A Review of the Holographic Relation between Linearized Gravity and the First Law of Entanglement Entropy

Rasmus Jaksland

November 30, 2017

Abstract

This thesis reviews the conjectured holographic relation between entanglement and gravity due to Mark van Raamsdonk and collaborators. It is accounted how the linearized Einstein equations both with and without matter in a $d + 1$-dimensional AdS background can be derived from the first law of entanglement entropy in a $d$-dimensional CFT. This derivation builds on the Ryu-Takayanagi formula that relates entanglement entropy for CFT subsystems to extremal surfaces in the AdS bulk. The relation between gravity and entanglement is also corroborated by a qualitative investigation of the duality between the thermofield double state and the maximally extended AdS/Schwarzschild black hole using the Bekenstein-Hawking formula. Furthermore, this qualitative argument is generalized to generic CFT states with a classical spacetime dual using the Ryu-Takayanagi.

The thesis also reviews the most relevant prerequisites for this holographic relation between gravity and entanglement: Anti-de Sitter spacetime, entanglement and entanglement entropy, gauge/gravity duality, the Ryu-Takayanagi formula, and linearized gravity.
Contents

1 Introduction 3
  1.1 Notation and conventions ........................................... 4

2 Anti-de Sitter Spacetime 6
  2.1 Global AdS ............................................................. 6
  2.2 Poincaré patch of AdS .............................................. 10
  2.3 Schwarzschild black holes ......................................... 11
  2.4 Asymptotically AdS black holes .................................. 16

3 Entanglement and Entropy 18
  3.1 The density matrix ................................................ 19
  3.2 Entanglement entropy .............................................. 23
  3.3 Entanglement in CFTs .............................................. 26

4 A Brief Introduction to Gauge/Gravity Duality 29
  4.1 $\mathcal{N} = 4$ Super Yang-Mills theory and type IIB string theory on $AdS_5 \times S^5$ 30
  4.2 Gauge/gravity duality: The weak form .............................. 34
  4.3 $CFT_d$ and $AdS_{d+1}$ ............................................ 40

5 The Ryu-Takayanagi Formula 43
  5.1 Relation to holography ............................................. 43
  5.2 Uses of the Ryu-Takayanagi formula .............................. 47
  5.3 Evidence for the Ryu-Takayanagi formula ......................... 48
Chapter 1

Introduction

Ever since its discovery by Maldacena in 1998, the AdS/CFT correspondence has intrigued researchers in quantum gravity. The correspondence entails a duality between theories with and without gravity, and recently, Mark van Raamsdonk and collaborators [1–4] have proposed that entanglement here plays a crucial role. The suggestion relies on a single entry from the AdS/CFT dictionary; the Ryu-Takayanagi formula [5] and its covariant generalization [6]. This formula relates entanglement entropy on the CFT side to areas of extremal co-dimension two surfaces on the AdS side, thereby relating an intrinsically quantum mechanical phenomena on the CFT side to the geometry of the bulk in the spacetime dual. This result, van Raamsdonk conjectures, proves an intimate relation between entanglement on the CFT side with spacetime itself on the AdS side. He argues “that the intrinsically quantum phenomenon of entanglement appears to be crucial for the emergence of classical spacetime geometry” [1, pp. 4–5].

A qualitative justification for this claim employs the example of a thermofield double state and its dual description as a maximally extended AdS-Schwarzschild black hole (the eternal black hole). In this case, the spacetime connectivity between the two exterior regions of the eternal black hole changes when the amount of entanglement between the two double states is changed. While this result may be established employing a holographic interpretation of the well known Bekenstein-Hawking formula, it may be generalized using the Ryu-Takayanagi formula to all quantum states with a classical spacetime dual. Thus, for such pairs this qualitative argument indicates that there is an intimate relation between entanglement in a CFT state and spacetime in the dual spacetime. The details of this qualitative argument will be the topic of chapter 7 of this thesis.
Chapter 8 will be dedicated to more rigorous and quantitative support for a relation between entanglement on the CFT side and gravity on the AdS side. More precisely, it will be demonstrated how the first law of entanglement entropy on the CFT side is equivalent to imposing the linearized Einstein equations without matter on the AdS side.

The linearized Einstein equations without matter are the $G_N \rightarrow 0$ limit of the Einstein equations, where $G_N$ is Newton’s constant. Going to the first subleading order in the $G_N$ expansion, bulk matter fields will have to be taken into account. Since these, when promoted to quantum fields, give rise to entanglement in the bulk, a correction to the Ryu-Takayanagi formula is required [7]. Chapter 9 will explore this correction and argue that the same entanglement constraint on the CFT side with this correction is equivalent to linearized Einstein equations including a source term.

Thus in summary, it is reviewed in this thesis how the Ryu-Takayanagi formula alone entails that the linearized Einstein equations both with and without matter are satisfied in a spacetime if the first law of entanglement entropy is satisfied in a dual CFT state. Before proceeding to this, however, chapters two through six will introduce relevant aspects of anti-de Sitter spacetime (chapter 2), entanglement and entanglement entropy (chapter 3), the AdS/CFT correspondence (chapter 4), the Ryu-Takayanagi formula (chapter 5), and linearized gravity (chapter 6).

1.1 Notation and conventions

Throughout this thesis, we will work in units where Planck’s constant, $\hbar$, Boltzmann’s constant, $k_B$, and the speed of light, $c$, to unity – $\hbar = c = k_B = 1$ – we will, however, keep the dependence on Newton’s constant, $G_N$, explicit since this will be instructive for later purposes. Adopting Einstein notation, we will assume an implicit sum over repeated indices. When ambiguities may occur, we will indicate the number of dimensions in which constants are defined by raised numbers in parenthesis. For instance, the notation $G_N^{(26)}$ signifies that this is Newton’s constant in 26-dimensional spacetime. Greek lower case indices starting with $\mu - \mu, \nu, \rho, \sigma$ – will be used for Minkowski spacetime coordinates, latin lower case indices starting from $i - i, j, k, l$ – will be used for Minkowski space coordinates. Latin lower case indices starting from $a - a, b, c, d, e, f$ – will be used for Poincaré coordinates of AdS spacetime and occasionally for coordinate independent relations. To avoid confusing in the numerical ordering of the Poincaré coordinates, we will use $t$ for the time component of tensors (for instance $A^t$) and $z$ for the AdS scaling coordinates component of tensors (for instance $A_z$). Thus, raised and lowered $t$ and $z$ are
not indices but tensor components. Most often, we will consider $d + 1$-dimensional AdS spacetime. Here $0 \leq a \leq d$, $0 \leq \mu \leq d - 1$, and $1 \leq i \leq d - 1$. We will use $D$ to denote the dimensionality of a spacetime. Thus, we will often find that $D = d + 1$. Finally, we will use capital indices starting with $A - A, B, C, D$ – for target space coordinates and greek lower case indices starting with $\alpha - \alpha, \beta$ – for world sheet coordinates.

Throughout, $g$ will denote the background metric, and a metric induced by $g$ on a surface (including the world sheet metric) will be denoted by $\gamma$. Unless stated otherwise, metrics will have Lorentzian signature and we adopt the sign convention $\text{diag}(-, +, \cdots, +)$. 
Chapter 2

Anti-de Sitter Spacetime

Throughout this thesis, we will be interested in so-called anti-de Sitter (AdS) spacetime: It is one of the cornerstones in gauge/gravity duality, it will serve as the background in which we will consider linearized gravity, and one of the central examples will deal with asymptotically AdS black holes. Consequently, we will subsequently spend some time introducing AdS spacetime and briefly asymptotically AdS black holes. To connect with the subsequent chapters, we introduce AdS in $d + 1$ dimensions. The material in this section is based on chapter 2 of [8] and chapter 5 of [9].

2.1 Global AdS

According to the theory of general relativity, spacetimes must satisfy the Einstein equations

$$G^E_{ab} = 8\pi G_N T_{ab}$$  \hspace{1cm} (2.1)

where $0 \leq a \leq d$, $T_{ab}$ is the energy momentum tensor and $G^E_{ab}$ is the Einstein tensor

$$G_{ab} \equiv R_{ab} - \frac{1}{2} R g_{ab} + \Lambda g_{ab}.$$  \hspace{1cm} (2.2)

Here $\Lambda$ is the cosmological constant, $R_{ab}$ is the Ricci tensor and $R$ is the Ricci scalar defined in terms of the Riemann tensor as

$$R_{ab} \equiv R_{abc}^c, \quad R \equiv g^{ab} R_{ab}.$$  \hspace{1cm} (2.3)

Minkowski spacetime is arguably the most well known solution to the vacuum Einstein equations without cosmological constant, i.e. Einstein equations for which $T_{ab} = 0$
and $\Lambda = 0$. The causal structure of this and other spacetimes can be depicted in a Penrose diagram, where the possibly infinite spacetime is brought into finite size nice by choosing an appropriate conformal factor. Null curves are always $\pm 45^\circ$ with vertical such that timelike directions are those directions bounded by the null curves. The Penrose diagram of Minkowski spacetime is depicted in figure 2.1.

Figure 2.1: Penrose diagram of $d$-dimensional Minkowski spacetime. A sphere $S^{d-2}$ in each point is implicit. $i^+$ denotes future timelike infinity, $i^-$ denotes past timelike infinity, $i^0$ denotes spacelike infinity, $\mathfrak{I}^+$ denotes future null infinity, and $\mathfrak{I}^-$ denotes past null infinity.

A characteristic feature of Minkowski spacetime is that it is a maximally symmetric spacetime in the sense that it has the maximal number of spacetime symmetries as given by Killing vector fields, $K_a$, that satisfy the equation

$$\nabla_a K_b + \nabla_b K_a = 0. \tag{2.4}$$

Here $\nabla_a$ is the covariant derivative, which is given in terms of the partial derivative, $\partial_a$, and the Christoffel connection, $\Gamma^c_{ab}$,

$$\nabla_a V_b = \partial_a V_b - \Gamma^c_{ab} V_c \tag{2.5}$$

where $\Gamma^c_{ab}$ may be expressed in terms of the metric as

$$\Gamma^c_{ab} = \frac{1}{2} g^{cd}(\partial_a g_{db} + \partial_b g_{da} - \partial_d g_{ab}). \tag{2.6}$$
One finds that the maximal number of symmetries and therefore Killing vector fields is \( \frac{1}{2}D(D+1) \) where \( D \) is the number of spacetime dimensions of the manifold \( (D = d + 1) \). Thus, any manifold with this many symmetries/Killing vector fields is maximally symmetric. In a maximally symmetric spacetime, the curvature must be the same everywhere; something that is trivially obeyed by Minkowski spacetime.

That curvature is the same everywhere entails that the Riemann tensor may be expressed in terms of the Ricci scalar, \( R \),

\[
R_{abcd} = \frac{R}{D(D-1)} (g_{ab}g_{cd} - g_{ad}g_{bc})
\]

(2.7)

For Minkowski spacetime, where \( R = 0 \) everywhere, this entails that the Riemann tensor vanishes everywhere. However, the relation indicates that we may consider two other maximally symmetric solutions to the Einstein equations in vacuum: One where \( R > 0 \) and one where \( R < 0 \). The latter type of spacetimes – the maximally symmetric spacetimes with \( R < 0 \) – will be our primary interest in the following. Such spacetimes are known as anti-de Sitter spacetimes and solve the Einstein equations with negative cosmological constant as may be seen from

\[
0 = G^E_{ab} = \left( R_{ab} - \frac{1}{2} R g_{ab} + \Lambda g_{ab} \right) = 2g^{ab} \left( R_{ab} - \frac{1}{2} R g_{ab} + \Lambda g_{ab} \right) = 2R - DR + 2\Lambda D
\]

(2.8)

from which it follows \( R = 2\Lambda D/(D-2) \). Thus, \( R \propto \Lambda \).

For convenience, we will consider anti-de Sitter spacetimes in \( (d+1) \) dimensions. This may be embedded in a \( (d+2) \) dimensional Minkowski spacetime, \( R^{d,2} \), with the metric

\[
ds^2 = -(dX^0)^2 + (dX^1)^2 + \cdots + (dX^d)^2 - (dX^{d+1})^2
\]

(2.9)

and given by the hypersurface subject to the restriction

\[-(X^0)^2 + \sum_{i=1}^d (X^i)^2 - (X^{d+1})^2 = -L^2.\]

(2.10)

This hypersurface may also be parametrized by

\[
X^0 = L \cosh(\rho) \cos(\tau) \\
X^i = L \Omega_i \sinh(\rho), \quad \text{for } i \in \{1, \ldots, d\} \\
X^{d+1} = L \cosh(\rho) \sin(\tau)
\]

(2.11)
where $\rho \in \mathbb{R}_+$, $\tau \in [0,2\pi]$, and $\Omega_i$ parametrizes a $(d-1)$-dimensional sphere, $S^{d-1}$, such that it satisfies $\sum_{i=1}^{d} \Omega_i^2 = 1$. These coordinates are known as global coordinates of $AdS_{d+1}$. In these global coordinates, the metric takes the form

$$ds^2 = L^2 \left( -\cosh^2(\rho) d\tau^2 + d\rho^2 + \sinh^2(\rho) d\Omega_{d-1}^2 \right).$$  \hspace{1cm} (2.12)$$

where $d\Omega_{d-1}^2$ is the standard metric on $S^{d-1}$. The vector field $\partial_\tau$ is a Killing vector field on the entire manifold and $\tau$ may therefore serve as a global time coordinate. It is worth noting that $\tau$ and therefore time is periodic in $2\pi$. Usually, we will mean by global anti-de Sitter spacetime, the universal covering of anti-de Sitter spacetime in global coordinates, where $\tau$ is unwrapped such that $\tau \in \mathbb{R}$.

For purposes below, we will be interested in the conformal boundary of anti-de Sitter spacetime. This is more easily studied if we replace $\rho$ by $\theta$ such that $\tan(\theta) = \sinh(\rho)$ which entails $\theta = [0, \frac{\pi}{2}]$. In these coordinates, the metric takes the form

$$ds^2 = \frac{L^2}{\cos^2(\theta)} \left( -d\tau^2 + d\theta^2 + \sin^2(\theta) d\Omega_{d-1}^2 \right).$$  \hspace{1cm} (2.13)$$

Since we are interested in a conformal boundary, the overall scaling factor $\frac{L^2}{\cos^2(\theta)}$ may be ignored and we see that for $\theta = \pm \frac{\pi}{2}$, we obtain $\partial AdS_{d+1}^{Global} = \mathbb{R} \times S^{d-1}$, where $\partial AdS_{d+1}^{Global}$ denotes the conformal boundary of (the universal covering of) global anti-de Sitter spacetime.

Figure 2.2 depicts the Penrose diagram for the universal covering of global anti-de Sitter spacetime; still suppressing the $\sin^2(\theta) d\Omega_{d-1}^2$ part of the metric. To include this, regard the Penrose diagram in figure 2.2 as a radial slice of a hypercylinder whose angular part is given by $\sin^2(\theta) d\Omega_{d-1}^2$. Thus, $\theta = 0$ is the center of the cylinder and $\theta = \frac{\pi}{2}$ is the boundary. Contrary to Minkowski spacetime, there is no spatial infinity since a light signal may travel to boundary, $\rho = \infty$, and back again in finite time. To see this, consider the metric (2.13) for a light ray (null curve) for which $ds^2 = 0$. Disregarding the factor $L^2/\cos(\theta)$ and the scaled sphere $\sin^2(\theta) d\Omega_{d-1}^2$, we have

$$0 = -d\tau^2 + d\theta^2 \Rightarrow d\tau = d\theta$$  \hspace{1cm} (2.14)$$

from which we see that light may travel to the boundary in finite time.
\[ \begin{align*}
\theta &= 0 \quad \theta = \frac{\pi}{2} \\
\rho &= 0 \quad \rho = \infty
\end{align*} \]

Figure 2.2: Penrose diagram of \( d + 1 \)-dimensional universal covering of global AdS spacetime. A contribution \( \sin^2(\theta)d\Omega_{d-1}^2 \) in each point is implicit. \( i^+ \) denotes future timelike infinity, \( i^- \) denotes past timelike infinity. Due to the causal structure of the spacetime, there is no spacelike, future null, or past null infinity.

### 2.2 Poincaré patch of AdS

Another parametrization of (2.10) that will be used repeatedly below is given in terms of the coordinates \( t \in \mathbb{R}, \bar{x} = (x^1, \ldots, x^{d-1}) \in \mathbb{R}^{d-1} \) and \( r \in \mathbb{R}_+ \). The parametrization then reads

\[
\begin{align*}
X^0 &= \frac{L^2}{2r} \left( 1 + \frac{r^2}{L^4} \left[ x^2 - t^2 + L^2 \right] \right), \\
X^i &= \frac{rx^i}{L}, \quad \text{for } i \in \{1, \ldots, d-1\}, \\
X^d &= \frac{L^2}{2r} \left( 1 + \frac{r^2}{L^4} \left[ x^2 - t^2 - L^2 \right] \right), \\
X^{d+1} &= \frac{rt}{L}.
\end{align*}
\]

(2.15)

These are only local coordinates since the restriction \( r > 0 \) entails that we only cover half of \( AdS_{d+1} \). This is the so-called Poincaré patch and the coordinates therefore are
known as Poincaré coordinates. The metric of this Poincaré patch of $AdS_{d+1}$ is
\[
ds^2 = \frac{L^2}{r^2} dr^2 + \frac{r^2}{L^2} \left( -dt^2 + d\vec{x}^2 \right).
\] (2.16)

In subsequent chapters, we will most often take it to be implicit that $AdS_{d+1}$ is the Poincaré patch of $AdS_{d+1}$.

We will also be interested in the conformal boundary of the Poincaré patch of $AdS_{d+1}$. This can be shown to be $d$-dimensional Minkowski spacetime i.e. $\mathbb{R}^{1,d-1}$, which obtains for $r \to \infty$. This is more readily seen, if we replace $r$ by defining $z = \frac{L^2}{r}$ such that the metric becomes
\[
ds^2 = \frac{L^2}{z^2} \left( -dt^2 + dz^2 + d\vec{x}^2 \right).
\] (2.17)

Here the conformal boundary is found at $z \to 0$ and ignoring the pre-factor we exactly get $d$-dimensional Minkowski spacetime.\footnote{More precisely stated, the metric $\tilde{d}s^2 = ds^2/z^2$ has the boundary $\mathbb{R}^{1,d-1}$ at $z \to 0$ and $\tilde{d}s^2$ is conformally equivalent to $ds^2$.} Thus, for any slice of the Poincaré patch with $z = \text{constant}$, the metric is that of Minkowski spacetime. The coordinate $z$ takes the form of a warpfactor.

For this reason, the Penrose diagram for the Poincaré path of $AdS_{d+1}$ is the same as that for $(d)$-dimensional Minkowski spacetime (figure 2.1). This may seem odd. Minkowski spacetime is after all very different from anti-de Sitter spacetime. The reason that the Poincaré patch can have the same Penrose diagram as Minkowski spacetime is that the Poincaré patch – as the name indicates – only covers a patch of AdS. This patch, it turns out, has the same causal structure as Minkowski spacetime in agreement with the claim the the Poincaré patch is conformally equivalent to Minkowski spacetime.

### 2.3 Schwarzschild black holes

Another interesting class of solutions to the Einstein equations are black hole solutions that feature event horizons. The so-called maximally extended asymptotically AdS-Schwarzschild black hole in global coordinates will be of particular interest since it will play a crucial role in developing the qualitative argument in favour of a relation between entanglement and spacetime.

However, we will begin with the well known Schwarzschild black hole. This is a spherically symmetric solution to the Einstein equations for $\Lambda = 0$. The metric may be
expressed as

\[ ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2d\Omega^2_{d-2} \]  \hspace{1cm} (2.18)

where \( d\Omega^2_{d-2} \) is the usual metric for \( S^{d-2} \), \( r \) is a radial coordinate and \( f(r) = 1 - \frac{2\mu}{r^{d-3}} \). The constant, \( \mu \) is given by

\[ \mu = \frac{8\pi G_N M}{(d-1)\text{Vol}(S^{d-1})} \]  \hspace{1cm} (2.19)

such that in 4 dimensions \( \mu_{d=4} = G_N M \) from which we obtain the well known expression for the Schwarzschild black hole in 4 dimensions

\[ ds^2 = -\left(1 - \frac{2G_NM}{r}\right)dt^2 + \left(1 - \frac{2G_NM}{r}\right)^{-1}dr^2 + r^2d\Omega^2_{d-2}. \]  \hspace{1cm} (2.20)

Since our primary interest is in the causal structure of black holes and this is the same in any number of dimensions, we will for convenience work in 4 dimensions in the following. Also, we will suppress the last part of the metric \( r^2d\Omega^2_{d-2} \), since this part of the metric merely contributes with a \((d-2)\)-dimensional sphere in all points in the conformal diagram of the spacetime and may be reinserted thus.

The metric of interest – the \((t,r)\) subspace – therefore takes the form

\[ ds^2 = -\left(1 - \frac{2G_NM}{r}\right)dt^2 + \left(1 - \frac{2G_NM}{r}\right)^{-1}dr^2. \]  \hspace{1cm} (2.21)

Two points stand out. One is \( r = 0 \), which is a curvature singularity and the other is \( r = 2G_NM \), which turns out to be a coordinate singularity. In general dimensions, the curvature singularity is still situated at \( r = 0 \) and the coordinate singularity (the black hole horizon) is situated at \( r = r_h \equiv (2\mu)^{\frac{1}{d-3}} \).

To investigate the causal structure, we want to consider the behaviour of light-cones and one finds that the slope of these – defined as \( \frac{dr}{dt} \) – are given by

\[ \frac{dr}{dt} = \pm \left(1 - \frac{2G_NM}{r}\right)^{-1} \]  \hspace{1cm} (2.22)

since the light-cone is spanned by the radial null curves \( ds^2 = 0 \). For \( r \to 2G_NM \), we see that \( \frac{dr}{dt} = 0 \); the light-cone closes up as we approach the event horizon. This, however, is merely a feature of these particular coordinates. A better choice of coordinates will resolve this and allow us to study the causal structure as we pass the event horizon. The problem that the light-cone closes up may be resolved, if we can find a coordinate, \( r^*(r) \), such that

\[ \frac{dr}{dt} = \pm r^*(r). \]  \hspace{1cm} (2.23)
This coordinate, known as the tortoise coordinate, must satisfy
\[ dr^* = \frac{dr}{1 - \frac{2G_NM}{r}} \] (2.24)
and integrating on both sides, one obtains
\[ r^* = r - 2G_NM + 2G_NM \ln \left( \frac{r}{2G_NM} - 1 \right). \] (2.25)
We see that the event horizon at \( r = 2G_NM \) is located, in the tortoise coordinate, at \( r^* \to -\infty \). This also entails that \( r^* \) is only defined for \( r \geq 2G_NM \). In the tortoise coordinates, the metric takes the form
\[ ds^2 = \left( 1 - \frac{2G_NM}{r^{(r^*)}} \right) (-dt^2 + dr^{*2}). \] (2.26)
In these coordinates, the metric is no longer singular at \( r = 2G_NM \), however, it is at the price that the event horizon is pushed to \( r^* \to -\infty \). This may be resolved by replacing the time coordinate with either one of the light-cone tortoise coordinates:
\[ v \equiv t + r^*, \quad u \equiv t - r^* \] (2.27)
where the infalling radial null geodesics have \( v = \) constant and outgoing radial null geodesics have \( u = \) constant. The causal structure of infalling objects may then be investigated with the coordinates \((v, r)\), where the metric takes the form
\[ ds^2 = -\left( 1 - \frac{2G_NM}{r} \right) dv^2 + dv dr + dr dv. \] (2.28)
The radial null curves, \( ds^2 = 0 \), satisfies
\[ \frac{dv}{dr} = \begin{cases} 0 & \text{if } r < 2G_NM, \\ 2 \left( 1 - \frac{2G_NM}{r} \right)^{-1} & \text{if } r \geq 2G_NM \end{cases} \] (2.29)
from which we see that the outgoing radial null curve simply tilts over as the object approaches and passes \( r = 2G_NM \) such that all future directed paths for \( r < 2G_NM \) goes in the direction of decreasing \( r \); there is no way out of the black hole once in.

However, we could also choose to consider the radial null curves in the coordinates \((u, t)\), where the metric takes the form
\[ ds^2 = -\left( 1 - \frac{2G_NM}{r} \right) du^2 - du dr + dr du. \] (2.30)
In this case, the radial null curves satisfies

$$\frac{du}{dr} = \begin{cases} -2 \left( 1 - \frac{2G_NM}{r} \right)^{-1} \\ 0 \end{cases}$$

(2.31)

where we again see that the light-cone tilts when \( r = 2G_NM \) is approached and passed.

In these coordinates, however, the situation is reversed. The horizon may be passed by following past directed curves such that for \( r < 2G_NM \) all such paths go in the direction of decreasing \( r \). This is a white hole; once you are out, there is no way back in.

In order to find coordinates that includes the inside of both the black and the white hole, one could propose to use \((u, v)\). This gives the metric

$$ds^2 = -\frac{1}{2} \left( 1 - \frac{2G_NM}{r} \right) \left( dvdu + dudv \right)$$

(2.32)

where \( r \) is related to \( u \) and \( v \) by

$$\frac{1}{2} (v - u) = r + 2G_NM \ln \left( \frac{r}{2G_CM} - 1 \right).$$

(2.33)

We see that the horizon at \( r = 2G_NM \) has once again been pushed to infinity and is found at either \( u = \infty \) or \( v = -\infty \). To bring the horizon to a finite value in \( v \) and \( u \), we define

$$v' = e^{\frac{v}{4GM}}$$

$$u' = -e^{-\frac{u}{4GM}}$$

(2.34)

and see that the horizon is now found at either \( v' = 0 \) or \( u' = 0 \) and that \( 0 < v' < \infty \), \(-\infty < u < 0 \). Since \( v \) and \( u \) are null coordinates so are \( v' \) and \( u' \). To express the metric in terms of one timelike coordinate and the rest spacetime, we introduce the coordinates known as Kruskal-Szekeres coordinates,

$$T = \frac{1}{2} (v' + u')$$

$$R = \frac{1}{2} (v' - u').$$

(2.35)

such that \(-\infty < T < \infty \) and \( 0 < R < \infty \). The metric becomes

$$ds^2 = \frac{32G^3_NM^3}{r} e^{-\frac{v}{4GM}} \left( -dT^2 + dR^2 \right)$$

(2.36)

and the horizon in these coordinates is found at \( T = R \) for \( T > 0 \) and at \( T = -R \) for \( T < 0 \). We can approach and pass the horizon following null curves given by \( T = \)
\( \pm R + \text{constant} \). Thus, these coordinates include both the inside and the outside of the white and black holes. The coordinates, however, may be analytically extended even further. In these coordinates, we can extend the range of \( R \) from \( 0 < R < \infty \) to \( -\infty < R < \infty \). The region where \( R < 0 \) mirrors the region \( R > 0 \) and therefore includes a second region exterior to the black and white holes. As it turns out, no choice of coordinates can extend this spacetime further and we therefore denote the metric (2.36) with coordinate range \( -\infty < T < \infty \) and \( -\infty < R < \infty \) as the maximally extended Schwarzschild black hole.

![Figure 2.3: Penrose diagram of the \( d \)-dimensional maximally extended Schwarzschild black hole geometry. A sphere \( S^{d-2} \) scaling as \( r^2 \) must be added to each point. Regions I and III cover regions that lie outside the horizon (dashed, \( r = r_h \equiv (2\mu)^{\frac{1}{d-3}} \)) of the black hole interior covered by region II. Region IV is the interior of a white hole.](image)

The causal structure of the maximally extended Schwarzschild black hole may be depicted in a Penrose diagram (see figure 2.3) where we use \( \tilde{u} = \arctan \left( \frac{u'}{\sqrt{2G_NM}} \right) \) and \( \tilde{v} = \arctan \left( \frac{v'}{\sqrt{2G_NM}} \right) \) to bring the coordinates into finite range. Still, we have suppressed the \( (d-2) \)-dimensional sphere, i.e. we consider the \( (t,r) \) subspace. The full spacetime may be obtained merely by including a \( (d-2) \)-dimensional sphere scaling as \( r^2 \) in each point in the diagram. As seen on figure 2.3, the Penrose diagram for the maximally extended Schwarzschild black hole shares with Minkowski space the same future null infinity, \( \mathcal{I}^+ \), past null infinity, \( \mathcal{I}^- \), and spacelike infinity, \( i^0 \), which verifies that the
maximally extended Schwarzschild black hole is indeed asymptotically Minkowski. One subtlety about the diagram is that future timelike infinity, $i^+$, and past timelike infinity, $i^-$, are distinct from the singularity at $r = 0$ despite the appearance to the contrary. There are many timelike paths that do not end on the singularity, but go to other values of $r$. From the diagram we see that no future directed timelike path goes from one of the exterior region to the other; these two regions are causally disconnected.

2.4 Asymptotically AdS black holes

As mentioned in the beginning of this chapter, what we are really interested in is the maximally extended AdS-Schwarzschild black hole. This is similar to the maximally extended Schwarzschild black hole developed above but differs in that it is asymptotically AdS rather than asymptotically flat. This entails that the function $f(r)$ in (2.18) takes the form

$$f(r) = 1 - \frac{2\mu}{r^{d-3}} + \frac{r^2}{L^2}.$$  (2.37)

![Penrose diagram of the maximally extended AdS-Schwarzschild black hole with a $d-1$ dimensional sphere over each point that scales as $r^2$. b) Depiction of two spacelike slices of the the eternal black hole ($T = 0$ and $T$ equals a positive constant) with one angular coordinate restored.](image)

Figure 2.4: a) Penrose diagram of the maximally extended AdS-Schwarzschild black hole with a $d-1$ dimensional sphere over each point that scales as $r^2$. b) Depiction of two spacelike slices of the the eternal black hole ($T = 0$ and $T$ equals a positive constant) with one angular coordinate restored.
Compared to the maximally extended Schwarzschild black hole, the causal structure therefore remains the same with the exception that the spacetime is asymptotically AdS. The Penrose diagram for the maximally extended AdS-Schwarzschild black hole is therefore similar to figure 2.3 with the exception that the future null infinity, the past null infinity and the spacelike infinity must be identical to that of the universal covering of global AdS spacetime. The Penrose diagram is seen in figure 2.4. Compared to figure 2.3, the asymptotic boundaries of regions I and III are now the universal covering of global AdS spacetime. Due to the nature of this asymptotic boundary, light signals may reach and return from the boundary in finite time as shown above. This has peculiar consequences for the maximally extended AdS-Schwarzschild black hole where \( r_h \gg L \).

If \( r_h \ll L \), we may neglect the term \( \frac{r^2}{L^2} \), and the maximally extended AdS-Schwarzschild black hole will have properties similar to that of the asymptotically flat maximally extended Schwarzschild black hole. Particularly, both will evaporate due to the emission of Hawking radiation. For \( r_h \gg L \), however, the maximally extended AdS-Schwarzschild black hole may be in an equilibrium with its Hawking radiation since the radiation can reach the asymptotic AdS boundary and return in finite time. Thus, if emission and reabsorption rates are equal then the maximally extended AdS-Schwarzschild black hole will not evaporate. For this reason, we denote such black holes as eternal black holes.

For the eternal black hole we find – as we did for the maximally extended Schwarzschild black hole – two causally disconnected regions, I and III. No signal can travel from one to the other. Still, we will describe the eternal black hole as a connected spacetime since a signal from the region I can intersect a signal from region III. The signals, however, can only intersect inside the black hole thereby precluding any causal connection between the two exterior regions I and III. The picture emerging is that of a wormhole so long as it is remembered that the wormhole is simply a two sided black hole, i.e. a black hole with two distinct exterior regions. That we may conceive of the eternal black hole as a wormhole is further signified by the depiction in figure 2.4b of \( T = 0 \) and a constant positive \( T \) slice of the spacetime.

These depiction more explicitly demonstrates features of the eternal black hole; most explicitly the scaling of the sphere \( S^{d-2} \) with \( r^2 \). Also, these depictions will prove useful in a qualitative assessment of the behavior of the spacetime when the area of the black hole horizon is decreased.
Chapter 3

Entanglement and Entropy

Entanglement is an inherently quantum mechanical phenomena. To see the peculiarity of this phenomena let us consider two discernible particles, 1 and 2, and two positions, A and B. Both particles behave in such a way that when measured it is found at site A half of the times and at site B the other half of the times. However, the two particles are never found at the same site simultaneously. Thus, when we measure and find particle 1 at site A, then a similar measurement at site B will find particle 2 there. Similarly, if particle 2 is found at site A, we can be sure to find particle 1 at site B.

We may initially attempt to ascribe a probability wave to each particle to account for the behaviour. Each particle is described by a wave with two sharp peaks at A and B and with an amplitude of 1/2 at each site, i.e. they are ascribed a probability of 1/2 for being at site A and 1/2 for being at site B. However, describing the particles with an individual probability wave cannot account for their behaviour never to be found at the same site. Generally, there is no way to describe the behaviour with a function only ranging over the sites A and B (usually this will be a three dimensional position space). This will inevitably leave out some information about the relation between the two particles. This is the information about their entanglement.

To account for their entanglement, i.e. for the effect that the particles are never found at the same site, we must instead describe the particles as living in configuration space. In configuration space there are four possible configurations; not sites since the states are no longer in position space. These may be denoted $AA, AB, BA$ and $BB$ where the first letter denotes the position of the first particle and the second letter denotes the position of the second particle. Associating probability to each of these configurations, it follows from the example that only $AB$ and $BA$ have a non-zero probability to be
occupied. More precise, the probability is $1/2$ for each. The lesson to learn is that it is necessary to describe the observed correlation of these particles in configuration space. It is not possible to describe it in position space as we are used to from classical physics.

For convenience we will subsequently use correlation between the spin state of two particles rather than a correlation between positions. We will assume that the particles are fermions and therefore have spin-$1/2$. Thus, each particle is either found to have spin up, $|\uparrow\rangle$, or spin down, $|\downarrow\rangle$. Calling the system composed of first particle, $A$, and the system composed of the second particle, $B$, their state-spaces is spanned by the states $\{|\uparrow\rangle_A, |\downarrow\rangle_A\}$ and $\{|\uparrow\rangle_B, |\downarrow\rangle_B\}$, respectively. The full system, therefore, is spanned by

$$\left\{|\uparrow\rangle_N, |\downarrow\rangle_N\right\} \otimes \left\{|\uparrow\rangle_B, |\downarrow\rangle_B\right\}$$

(3.1)

Now, the state of the full system, $|\Psi\rangle$, is such that if one particle is found to have spin up, then the other particle will be found to have spin down and vice versa. Again, we cannot account for this correlation if we describe each particle individually. While we can get the statistics right by assigning probability $1/2$ to each state for each particle, this does not account for the correlation. We cannot describe the full state as a product of $A$ and $B$ states.

The two systems (the two particles) are correlated, and $|\Psi\rangle$ therefore becomes:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left( |\uparrow\rangle_A |\downarrow\rangle_B + |\downarrow\rangle_A |\uparrow\rangle_B \right)$$

(3.2)

this is what is known as a Bell pair. This is the maximally entangled state consisting of two qubits.

### 3.1 The density matrix

The example demonstrates that two entangled particles cannot be described as two separable subsystems, but must be regarded as one whole. To introduce this notion of entanglement more formally, we will consider a quantum state in some Hilbert space, $|\Psi\rangle \in \mathcal{H}$ that can be decomposed into at least two spatially separated subsystems.\(^1\) This implies that $\mathcal{H}$ can be expressed as the tensor product of the Hilbert spaces of each subsystem. Decomposing $\mathcal{H}$ into two subsystems, we have

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$$

(3.3)

\(^1\) $|\Psi\rangle$ is a pure state of $\mathcal{H}$
where $\mathcal{H}_A$ and $\mathcal{H}_B$ are the Hilbert spaces of the two subsystems (denoted $A$ and $B$ respectively).

Given this decomposition of the Hilbert space, we may write a quantum state – a vector $|\Psi\rangle \in \mathcal{H}$ – in terms of basis states in the subsystem Hilbert spaces, $|\psi^A_i\rangle \in \mathcal{H}_A$ and $|\psi^B_j\rangle \in \mathcal{H}_B$:

$$|\Psi\rangle = \sum_{i,j} c_{i,j} |\psi^A_i\rangle \otimes |\psi^B_j\rangle$$

(3.4)

where $\sum_{i,j} |c_{i,j}|^2 = 1$.

Generally this is not a product state, i.e. a product of two pure states. Decomposing a system into two or more subsystems, it is generally not the case that we may express the full quantum state as a direct product of the states of the subsystems

$$|\Psi\rangle = \left(\sum_i c_i |\psi^A_i\rangle\right) \otimes \left(\sum_j c_j |\psi^B_j\rangle\right)$$

(3.5)

even though the Hamiltonian of either subsystem – due to locality – is independent of each other. When this equality does not hold, we will say that systems $A$ and $B$ are entangled.

This formal exposition sits well with the above account of the Bell pair. Here it was claimed that the spin of the particles were entangled, when we could not describe the state of the full system in terms of the states of two particles as spanned by $\{|\uparrow\rangle, |\downarrow\rangle\}$. Particularly, the outcome of measurements could not be accounted for by such a product state.

Generally, if $|\Psi\rangle$ is not a product state of $\mathcal{H}_A$ and $\mathcal{H}_B$, it follows that a pure state, $\sum_i c_i |\psi^A_i\rangle$, cannot give the same outcome as $|\Psi\rangle$ for a measurement on subsystem $A$. However, the outcome of such measurements can be reproduced by some ensemble of orthogonal states in $\mathcal{H}_A$, where each of these states, $\{|\psi^A_i\rangle\}$, are associated with a (classical) probability, $p_i$, of finding the system in that state when making the measurement on $A$.

Thus, the expectation value for the measurement on state $|\Psi\rangle$ agrees with the ensemble average of that measurement on this ensemble. Expressing this in terms of an operator, $O_A$, acting on subsystem $A$, this implies:

$$\langle O_A \rangle = \sum_i p_i \langle \psi^A_i | O_A | \psi^A_i \rangle$$

(3.6)

To prove that there is such a class, $|\psi^A_i\rangle$, of orthogonal states, consider the expectation
value of $O_A$ when the full state is $|\Psi\rangle$

$$
\langle O_A \rangle = \langle \Psi | O_A \otimes I | \Psi \rangle \\
= \left( \sum_{i,j} c_{i,j}^* \langle \psi_j^B | \otimes \langle \psi_i^A | \right) (\sum_{k,l} c_{k,l} | \psi_k^A \rangle \otimes | \psi_l^B \rangle) \\
= \sum_{i,j} \sum_{k,l} \delta_{jl} \langle \psi_i^A | c_{i,j}^* c_{k,l} O_A | \psi_k^A \rangle \\
= \sum_{i,k} \sum_j \sum_{n,m} c_{i,j}^* c_{k,j} \langle \psi_i^A | \psi_n^A \rangle \langle \psi_n^A | O_A | \psi_m^A \rangle \langle \psi_m^A | \psi_k^A \rangle \\
= \sum_{m} \langle \psi_m^A | \sum_{i,k} \sum_j c_{i,j}^* c_{k,j} \langle \psi_i^A | \psi_n^A \rangle \langle \psi_n^A | O_A | \psi_m^A \rangle \langle \psi_m^A | \psi_k^A \rangle \\
= \sum_{m} \langle \psi_m^A | O_A \rho_A | \psi_m^A \rangle \\
= \text{tr}(O_A \rho_A)
$$

where the last line is obtained by defining the operator

$$
\rho_A = \sum_{i,k} \sum_j c_{i,j}^* c_{k,j} \langle \psi_i^A | \psi_k^A \rangle \\
$$

and by the definition of the trace

$$
\text{tr}(O) = \sum_i \langle i | O | i \rangle .
$$

As seen, the operator $\rho_A$ also only acts on subsystem $A$ and it is known as the reduced density matrix. It may be obtained from the density matrix of the full system, $\rho \equiv |\Psi\rangle \langle \Psi|$, by tracing out the degrees of freedom outside of the subsystem. In the case where the full system is divided into two subsystems $A$ and $B$ this is

$$
\text{tr}_B(\rho) = \text{tr}_B(|\Psi\rangle \langle \Psi|) \\
= \sum_m \langle \psi_m^B | \left( \sum_{k,l} c_{k,l} | \psi_k^A \rangle \otimes | \psi_l^B \rangle \right) \left( \sum_{i,j} c_{i,j}^* \langle \psi_j^B | \otimes \langle \psi_i^A | \right) | \psi_m^B \rangle \\
= \sum_m \sum_{k,l} \sum_{i,j} \delta_{ml} \delta_{jm} c_{k,l} c_{i,j}^* | \psi_k^A \rangle \langle \psi_i^A | \\
= \sum_m \sum_{i,k} c_{k,i} c_{i,j}^* | \psi_k^A \rangle \langle \psi_i^A | \\
$$

21
which is equal to (3.8) above.

The density matrix is an hermitian operator, i.e. $\rho = \rho^\dagger$, and therefore it has real eigenvalues. Furthermore, the eigenvalues of the density matrix, $p_i$, can be shown to be non-negative. Thus

$$p_i \in \mathbb{R}^+ \quad (3.11)$$

The eigenvalues, $p_i$, are also sometimes known as the entanglement spectrum of the density matrix for a state $|\Psi\rangle$.

Using the associated eigenvectors we can express the reduced density matrix as

$$\rho_A = \sum p_i \frac{\phi_i^A}{\phi_i^A} \quad (3.12)$$

where $|\phi_i^A\rangle$ are the eigenvectors of $\rho_A$.

Inserting this into the expression for the expectation value of the operator $O_A$, we find

$$\langle O_A \rangle = \text{tr}(O_A \rho_A)$$

$$= \text{tr} \left( O_A \sum p_i \frac{\phi_i^A}{\phi_i^A} \right)$$

$$= \sum_{i,j} p_i \langle \phi_j^A | O_A | \phi_i^A \rangle \delta_{ij}$$

$$= \sum_i p_i \langle \phi_i^A | O_A | \phi_i^A \rangle \quad (3.13)$$

where we have used that the trace is the same in any orthonormal basis. This proves the claim that for the full system in a state $|\Psi\rangle$ there is an ensemble of orthogonal states in $\mathcal{H}_A$, $\{p_i, \phi_i^A\}$ that reproduces the expectation value of an operator, $O_A$, that only acts on subsystem $A$.

From the last line of (3.7), we notice that the reduced density matrix for subsystem $A$, $\rho_A$, contains all information about the state, $|\Psi\rangle$, in subsystem $A$. Concerning our current interest – entanglement – it is interesting to see how $\rho_A$ includes information about the entanglement between subsystem $A$ and the full system. Particularly, we find that if $\rho_A$ only has one non-zero eigenvalue then the subsystem $A$ is in a pure state and thereby not entangled with the rest of the system.

To see this, let us assume that only the eigenvalue $p_1$ of $\rho_A$ is non-zero. We then
have

\[ \langle O_A \rangle = \langle \Psi | O_A \otimes \mathbb{1} | \Psi \rangle \]
\[ = \text{tr}(O_A \rho_A) \]
\[ = \sum_i p_i \langle \phi_i^A | O_A | \phi_i^A \rangle \]
\[ = p_1 \langle \phi_1^A | O_A | \phi_1^A \rangle \] (3.14)

This implies that there is a single state of the Hilbert space for the subsystem – in this case \( | \phi_1^A \rangle \in \mathcal{H}_A \) – that includes all the information of full state in that subsystem; there is no classical uncertainty about the state of the subsystem. If the full system consists of two subsystems \( A \) and \( B \) and \( A \) is a pure state – in this case \( | \phi_1^A \rangle \) – then it follows that \( B \) is a pure state. Assuming that this pure state of \( B \) is \( | \theta_1^B \rangle \), we can thereby express the full state, \( | \Psi \rangle \), as

\[ | \Psi \rangle = | \phi_1^A \rangle \otimes | \theta_1^B \rangle \] (3.15)

This is recognized as a product state, and the two subsystems is therefore not entangled.

In summary, if the reduced density matrix for some subsystem has only one non-zero eigenvalue, then the subsystem is not entangled with the rest of the full system. The reduced density matrix contains the information about the entanglement between the subsystem and the rest of the full system.

### 3.2 Entanglement entropy

These properties makes it natural to use the reduced density matrix to quantify the entanglement between a subsystem and the rest of the system. For this purpose, we will define the entanglement entropy, \( S \), as the the von Neumann entropy associated with a density matrix \( \rho \) :

\[ S = - \text{tr}(\rho \log(\rho)). \] (3.16)

A number of properties makes the von Neumann entropy a good choice when quantifying entanglement. First, we observe that \( S \geq 0 \) which sits well with the fact that the notion of “anti-entanglement” is incomprehensible. To see that \( S \geq 0 \), remember that \( \rho \) must be diagonalized in the basis of its eigenvectors with diagonal elements equal to its eigenvalues, \( p_i \). Since \( 0 \leq p_i \leq 1 \), it follows that \( \log(p_i) \leq 0 \). Therefore, \( \text{tr}(\rho \log(\rho)) \leq 0 \) which implies that \( S \geq 0 \).

Another nice property is that the von Neumann entropy vanishes for a pure state. Thus, if there is no entanglement between a subsystem under consideration and the full
system, then the entanglement entropy of that system is zero. This follows, since it holds for the density matrix of a pure state that \( \rho^2 = \rho \). Thus
\[
\rho \log(\rho) = \rho \log(\rho^2) = 2 \rho_B \log(\rho)
\]
\[
\Rightarrow \rho \log(\rho_B) = 0
\]
\[
\Rightarrow S(\rho) = -\text{tr}(\rho \log(\rho)) = 0.
\]

The definiti

The von Neumann entropy also nicely reproduces the usual entropy formulas of well known ensembles. Taking the example of the microcanonical ensemble given by the combination of states and probabilities
\[
\{|E_i\}, p_i = \frac{1}{n}\quad (3.18)
\]
where \( E_i \in (E, E + dE) \) and \( n \) is the number of energy eigenstates in this interval. The entropy for the microcanonical ensemble is given as
\[
S_{\text{micro}} = \log(n) = \sum_i p_i \log(n)
\]
where the last equality is obtained by the identity
\[
\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \frac{1}{n} = 1
\]
which implies thinking of the total entropy as getting a contribution, \( p_i \log(n) \), from each state, \( E_i \), in the ensemble.

Now using the equality \( \log(n) = -\log\left(\frac{1}{n}\right) \) and that \( p_i = \frac{1}{n} \), we find
\[
S_{\text{micro}} = \sum_i p_i \log(n) = -\sum_i p_i \log(p_i) = -\text{tr}(\rho \log(\rho))
\]
where the last line follows since since \( p_i = \frac{1}{n} \) are the eigenvalues of \( \rho \) such that
\[
\rho = \begin{pmatrix}
\frac{1}{n} & 0 & \cdots & 0 \\
0 & \frac{1}{n} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \frac{1}{n}
\end{pmatrix}
\]
such that that taking the trace equals a sum over \( p_i \).

This is exactly the von Neumann entropy. However, it is worth noticing at this point that contrary to thermodynamic entropy which only is defined for equilibrium states,
entanglement entropy is defined for any state and we can therefore also consider dynamical changes in entanglement entropy. Something that will become relevant later on.

The von Neumann entropy also quantifies entanglement in agreement with our immediate expectations. To see this, consider the previous example of a bell pair system, but now with variable entanglement

$$|\Psi\rangle = \alpha |\uparrow\rangle_A |\downarrow\rangle_B + \beta |\downarrow\rangle_A |\uparrow\rangle_B$$  \hspace{1cm} (3.23)

where $|\uparrow\rangle_A$ is the spin state of subsystem A and $|\alpha|^2 + |\beta|^2 = 1$ due to normalisation.

To compute the entanglement entropy, i.e. the von Neumann entropy, for one of the subsystem A, one first has to find the reduced density matrix

$$\rho_A = \text{tr}_B (|\Psi\rangle \langle \Psi|)$$

$$= \sum_{i=\uparrow,\downarrow} \langle i|_B (\alpha |\uparrow\rangle_A |\downarrow\rangle_B + \beta |\downarrow\rangle_A |\uparrow\rangle_B) (\alpha^* \langle \uparrow|_A \langle \downarrow|_B + \beta^* \langle \downarrow|_A \langle \uparrow|_B) |i\rangle_B$$

$$= (\langle \uparrow|_B (\beta |\downarrow\rangle_A |\uparrow\rangle_B) (\beta^* \langle \downarrow|_A \langle \uparrow|_B) \rangle_B$$

$$+ \langle \downarrow|_B (\alpha |\uparrow\rangle_A |\downarrow\rangle_B) (\alpha^* \langle \uparrow|_A \langle \downarrow|_B + |\downarrow\rangle_B$$

$$= |\beta|^2 |\downarrow\rangle_A \langle \downarrow|_A + |\alpha|^2 |\uparrow\rangle_A \langle \uparrow|_A$$  \hspace{1cm} (3.24)

Having found the reduced density matrix, we can find an expression for the entanglement entropy

$$S_A = - \text{tr}(\rho_A \log \rho_A)$$

$$= - \left( |\beta|^2 \log(|\beta|^2) \right) - \left( |\alpha|^2 \log(|\alpha|^2) \right)$$

$$= - \left( (1 - |\alpha|^2) \log(1 - |\alpha|^2) \right) - \left( |\alpha|^2 \log(|\alpha|^2) \right)$$  \hspace{1cm} (3.25)

where the last line in obtained from the constraint $|\alpha|^2 + |\beta|^2 = 1$.

As seen from figure 3.2 the entanglement entropy is maximal for $|\alpha|^2 = |\beta|^2 = 0.5$. The entanglement entropy vanishes for $|\alpha|^2 = 1$ and $|\alpha|^2 = 0$ which corresponds to pure states of the subsystem A. This is exactly what is expected for a Bell pair state.

In general we find that if $S_A = 0$ then $|\psi_A\rangle$ is a pure state. If $S_A \neq 0$ then $|\psi_A\rangle$ is in a mixed state; the degrees of freedom in A are entangled with the degrees of freedom in the rest of the full system.

25
3.3 Entanglement in CFTs

Going to quantum field theories, some subtleties arise. A quantum field theory is defined on a $d$-dimensional Lorentzian spacetime, $\partial M$, thus to define a state of the system one must first choose a Cauchy (spacelike) slice, $\Sigma_{\partial M}$, of the spacetime that defines a moment of simultaneity.\textsuperscript{2} With respect to $\Sigma_{\partial M}$ we may define the Hilbert space of states $\mathcal{H}_{\Sigma_{\partial M}}$. For any state of the system, $|\Psi\rangle$, we have $|\Psi\rangle \in \mathcal{H}_{\Sigma_{\partial M}}$.

We may conceive of $\mathcal{H}_{\Sigma_{\partial M}}$ as a tensor product of the Hilbert spaces associated with each point and therefore as comprising of all the local degrees freedom in the quantum field theory. This construction is perhaps better understood as the continuum limit of a lattice system on $\Sigma_{\partial M}$ with some lattice spacing $\epsilon$. The degrees of freedom are located at the lattice sites and associated with the Hilbert space $\mathcal{H}_i$, where $i$ is some numbering of the lattice sites. The full Hilbert space is then given as a tensor product of the Hilbert space at each lattice site

$$\mathcal{H} = \otimes_i \mathcal{H}_i.$$  \hspace{1cm} (3.26)

The continuum limit is then obtained by sending $\epsilon \to 0$ and we can conceive of $\mathcal{H}_{\Sigma_{\partial M}}$ as constructed in this way.

\textsuperscript{2}The notation $\partial M$ and $\Sigma_{\partial M}$ is chosen for comply with later notation.
We want to interrogate about the entanglement entropy of some subregion \( B \subset \Sigma_{\partial M} \), i.e. about some quantitative measure of the entanglement between \( B \) and its compliment \( \overline{B} \). Here, the compliment of \( B \) is defined to satisfy \( B \cup \overline{B} = \Sigma_{\partial M} \). The entangling surface separating \( B \) and \( \overline{B} \), we denote \( \partial B \). Since a CFT is a local quantum field theory, there are specific degrees of freedom associated with specific spatial regions; again the analogy to a continuum limit of a lattice system is helpful. We can therefore regard the full quantum system as composed of two subsystems, \( Q_B \) and \( Q_{\overline{B}} \), associated with two spatially separated regions \( B \) and \( \overline{B} \). As a consequence we can decompose the Hilbert space of states of the full system as we did in the section above.\(^3\)

Having made this construction, one may obtain an expression for the density matrix operator \( \rho_B \). To evaluate the entanglement entropy one also needs to obtain an expression for the operator \( \log(\rho_B) \). However, taking the logarithm of a continuum operator – such as \( \rho_B \) in the context of a quantum field theory – proves to be technically complicated. In fact, it proves to be more tractable to evaluate the Rényi entropy

\[
S_B^{(q)} = \frac{1}{1-q} \log(\text{tr}(\rho_B^q))
\]

and then take the limit

\[
\lim_{q \to 1} S_B^{(q)} = S_B.
\]

The Rényi entropy may be evaluated using the so-called replica trick. It is beyond the scope of the current project to go into details of this procedure. The procedure is reviewed in [10], where it is also employed to obtain the entanglement entropy for 2-dimensional conformal field theories. These prove to be a special case, since they possess a Virasoro algebra with central charge \( c \). This simplifies the evaluation of entanglement entropy using the replica trick. Indeed, the evaluation of entanglement entropy proves to be impossible in all but the simplest higher dimensional cases.

In 2-dimensional conformal field theories, any spatial slice takes the form of a line. Define \( B \) to be the interval \(-a < x < a\) centred around the origin on. As shown in [10], the entanglement entropy for such a single interval is found to diverge, however, this is as expected since we, in a continuum CFT, will have unregulated short distance entanglement over the entangling surface \( \partial B \). To regulate this, we introduce the cutoff \( \epsilon \). With this cutoff, the entanglement entropy of a single interval centred around the origin is found to be

\[
S_B = \frac{c}{3} \log \left( \frac{2a}{\epsilon} \right).
\]
The simple expression for the entanglement entropy is a consequence of the Virasoro algebra, which is not realized in higher dimensional CFTs. In chapter 5, we will compare this result to an expression for the entanglement entropy of a single interval in a 2-dimensional conformal field theories as obtained by bulk-calculation, i.e. to an expression for $S_B$ obtained via the AdS side of the AdS/CFT correspondence. It is this correspondence – also known as the gauge/gravity duality – that we now turn to.
Chapter 4

A Brief Introduction to
Gauge/Gravity Duality

Ever since its conception two decades ago [11], the gauge/gravity duality has intrigued researchers in many branches of physics. The conjecture is that certain local field theories on a fixed spacetime background – i.e. without gravity – are dynamically equivalent to certain theories that include a gravitational interaction; they describe the same physics.

The most well known example of this duality is the one between conformal field theories\(^1\) in \(d\) dimensions (\(CFT_d\)) and anti-de Sitter spacetimes in \(d + 1\) dimensions (\(AdS_{d+1}\)), where the \(CFT_d\) is defined on a fixed spacetime that is identical to the \(d\)-dimensional asymptotic boundary of \(AdS_{d+1}\). For this reason, the gauge/gravity duality is also known as the AdS/CFT correspondence.

Initially, the duality may seem unlikely since one would generally expect the number of degrees of freedom in \(d + 1\) dimensions to be larger than the number of degrees of freedom in \(d\) dimensions. Thus, if the duality obtains, it suggests that the degrees of freedom in theories of gravity behave differently from those in local quantum field theories.

The first piece of evidence suggesting this to be the case is the holographic principle known from black hole thermodynamics. From statistical mechanics we know that entropy can be regarded as a measure for the degrees of freedom of a system. As shown in [12], for general relativity to be consistent with the laws of thermodynamics, the entropy of

\(^1\)Conformal field theories are quantum field theories that are conformally invariant i.e. invariant under dilations, \(x^\mu \rightarrow \lambda x^\mu\) where \(\lambda\) is a constant , and special conformal transformation, \(x^\mu \rightarrow \frac{x^\mu - ax^2}{1-2ax}\).
a black hole has to scale with the area of its event horizon, $A_{BH}$

$$S_{BH} = \frac{A_{BH}}{4G_N}$$  \hspace{1cm} (4.1)$$

Indeed, this can be shown to be the upper bound of the entropy in $d + 1$ dimensional Einstein gravity (the Bekenstein Bound) for a spatial volume bounded by the spatial surface with area $A$. Thus, in general relativity the entropy $S_{GR}$ for the $d$-dimensional spatial slice $\Sigma_{M^{d+1}}$ of the manifold $M^{d+1}$ scales as

$$S_{GR} \propto \text{Area}(\Sigma_{M^{d+1}}) \propto \text{Vol}(\Sigma_{\partial M^{d+1}})$$  \hspace{1cm} (4.2)$$

where $\Sigma_{\partial M^{d+1}}$ is the $d - 1$ dimensional boundary of $\Sigma_{M^{d+1}}$. Thus, gravity realises the holographic principle that the degrees of freedom in a volume scales with the surface area of that volume.

For a $d$-dimensional local quantum field theory, the entropy of the $d - 1$ dimensional spatial slice of a system scales with the volume of that spatial slice. Following the duality, a CFT state with the spacetime dual $M^{d+1}$ is defined on a spacetime identical to the asymptotic boundary of $M^{d+1}$, $\partial M^{d+1}$. Thus, the entropy, $S_{CFT}$ scales as

$$S_{CFT} \propto \text{Vol}(\Sigma_{\partial M^{d+1}})$$  \hspace{1cm} (4.3)$$

where $\Sigma_{\partial M^{d+1}}$ is a spacelike co-dimension two surface. This is in agreement with the result obtained for a gravitational theory under the assumption of the holographic principle. This, of course, is no proof of the conjectured duality between $CFT_d$ and $AdS_{d+1}$, however, it suggests that it is not absurd that a local quantum field theory in $d$ dimensions can have the same number of degrees of freedom as a gravitational theory in $d + 1$ dimensions.

4.1 $\mathcal{N} = 4$ Super Yang-Mills theory and type IIB string theory on $AdS_5 \times S^5$

Originally, the gauge/gravity duality was conceived by comparing an open string and a closed string perspective of a stack of $N$ D3-branes [11]. This lead to the conjecture that $\mathcal{N} = 4$ Super Yang-Mills theory (SYM) in four dimensional Minkowski spacetime is dual to type IIB string theory on the poincaré patch of the spacetime $AdS_5 \times S^5$.\footnote{For global AdS, the CFT dual is defined on $R \times S^3$.} $\mathcal{N} = 4$ SYM is a conformally invariant quantum field theory with gauge group $SU(N)$
and is characterised by the two free parameters $N$ – the rank of the gauge group – and Yang-Mills coupling constant, $g_{YM}$. This is referred to as the ‘CFT side’ of the duality. The string theory on $AdS_5 \times S^5$ is referred to as the ‘AdS side’ and is characterised by the string length $l_s = \sqrt{\alpha'}$, the string coupling constant $g_s$, and the curvature radius of $AdS_5$, $L$, which is also the curvature radius of $S^5$. As a free parameter, the latter two only occur as the ratio $l_s/L$ and are related – according to the AdS/CFT correspondence – to the free parameters on the CFT side in the following way

$$2g_{YM}^2 N = 2\lambda = \frac{L^4}{l_s^4} \quad (4.4)$$

where $\lambda = g_{YM}^2 N$ is known as the ’t Hooft coupling. The string coupling, $g_s$, and the Yang-Mills coupling, $g_{YM}$, are related by

$$g_{YM}^2 = 2\pi g_s. \quad (4.5)$$

Thus stated, the conjectured duality is one between a four and a ten dimensional theory. The holographic principle asserts that the number of degrees of freedom in a $d + 1$ dimensional gravity theory may be the same as the number of degrees of freedom in a $d$ dimensional gauge theory. This, however, still leaves five dimensions unaccounted for in the duality between four dimensional $\mathcal{N} = 4$ SYM and type IIB string theory on $AdS_5 \times S^5$. To explain how the number of degrees of freedom can match serves as a crude consistency check of the duality.

The details of the duality is such that the $CFT_4$ is defined on a spacetime identical to the asymptotic boundary of the dual spacetime; in this case the flat asymptotic boundary of the Poincaré patch of $AdS_5$. Therefore, the five dimensional version of the AdS side can be obtained by a Kaluza-Klein [13] reduction of the string theory on $S^5$. More precisely, $S^5$ can be expanded in a complete set of spherical harmonics such that any field on $AdS_5 \times S^5$ can be reduced to a tower of massive fields on $AdS_5$, where the masses are due to the origin of these fields in the expansion in spherical harmonics of $S^5$. If the holographic principle holds for this reduced theory in five dimensions, then it is consistent with it having the same number of degrees of freedom as a four dimensional gauge theory like $\mathcal{N} = 4$ Super Yang-Mills theory.

Again, this is no proof of the duality. Indeed, it turns out to be very hard to prove the duality for generic $g_{YM}$ and $\lambda$ (and thereby generic $g_s$ and $L/l_s$). A proof would require a full non-perturbative definition of quantized type IIB string theory on curved backgrounds, which we currently lack. This restricts most explicit calculations on the AdS side to the regime where string perturbation theory can be employed; the regime
where \( g_s \ll 1 \). If \( l_s/L \) is kept fixed, this reduces the AdS side to classical string theory in the sense that only tree level are taken into account, while the higher string genus expansions are disregarded. On the CFT side, this is equivalent to the limit where \( g_{YM} \ll 1 \) while \( \lambda = g_{YM}^2 N \) is kept fixed. Thus, \( N \to \infty \). This is the well known ’t Hooft limit, where only planar diagrams in the gauge theory contribute to leading order in \( 1/N \). Generally, for fixed \( \lambda \)

\[
\lambda = g_{YM}^2 N \Rightarrow \lambda = 2\pi g_s N \Rightarrow g_s \propto \frac{1}{N} \quad (4.6)
\]

Thus, an expansion in \( 1/N \) on the CFT side is equivalent to the string theory genus expansion in \( g_s \).

If \( l_s/L \to 0 \), the AdS side further simplifies. In this limit, the string length, \( l_s \), is much smaller than the curvature radius, \( L \) and therefore the strings may be approximated by point particles in the form of type IIB supergravity. This supergravity is defined on a weakly curved \( AdS_5 \times S^5 \) (since \( l_s \ll L \)) and is dual to \( \mathcal{N} = 4 \) SYM in its strong coupling limit, i.e. \( \lambda \gg 1 \). This limit is referred to as the weak form of the duality as opposed to the strong form, where \( l_s/L \) and therefore \( \lambda \) are kept fixed.

Even though a proof of the duality is not currently possible, a number of non-trivial tests and checks have been carried out. These have their outset in either the weak or the strong form of the duality and evaluates the same quantities on both sides; hitherto with perfect agreement. For instance, the symmetries of \( \mathcal{N} = 4 \) SYM form the supergroup \( PSU(2,2|4) \) and so does the type IIB string theory on \( AdS_5 \times S^5 \). Also, scattering amplitudes which are protected by supersymmetry as one goes from weak to strong coupling can be compared between the two sides and shown to be equal.

Assuming the weak form of the duality, tests are complicated by it being a weak/strong duality. This entails that one of the sides must be evaluated at strong coupling, which cannot be done at full generality on either side. However, observables independent of the coupling may still be compared and can be shown to be equal. Particularly, the classical fields of type IIB supergravity on \( AdS_5 \times S^5 \) maps to a subset of the operators in \( \mathcal{N} = 4 \) SYM.\(^3\)

Generally, there should be a one-to-one map between operators on the CFT side and fields on the AdS side which entails that the duality may be formulated as a correspondence between the generating functional for \( \mathcal{N} = 4 \) SYM and partition function of type IIB string theory. The generating functional for correlation functions of gauge theory

\(^3\)See [14] for a complete summary of the field content of \( \mathcal{N} = 2 \) supergravity in 10 dimensions.
operators, $\mathcal{O}(x)$, is obtained by perturbing the Lagrangian, $\mathcal{L}$, by a source field, $\phi_0(x)$

$$\mathcal{L} \rightarrow \mathcal{L} + \phi_0(x)\mathcal{O}(x).$$

(4.7)

The generating functional in Euclidean signature on the CFT side, $Z[\phi_0]_{CFT}$, thereby takes the form

$$Z[\phi_0]_{CFT} = N \int D\phi e^{-\int d^d x (\mathcal{L} + \phi_0(x)\mathcal{O}(x))} \equiv \left\langle \exp \left( \int d^d x \mathcal{O}(x) \phi_0(x) \right) \right\rangle$$

(4.8)

where $N$ is some normalization. Assume that $\phi_0(x)$ is the boundary value of the fluctuating bulk field $\phi(z, 0)$ such that

$$\lim_{z \to 0} z^{\Delta - d} \phi(x, z) = \phi_0(x)$$

(4.9)

where $\Delta$ is the dimension of $\mathcal{O}(x)$.\footnote{Observe that $\phi_0(x) \neq \phi(z = 0, x)$. Generally, $\phi(z = 0, x)$ is divergent.} One then finds that

$$\left\langle \exp \left( \int d^d x \mathcal{O}(x) \phi_0(x) \right) \right\rangle = Z_{\text{string}} \bigg|_{\lim_{z \to 0} z^{\Delta - d} \phi(x, z) = \phi_0(x)}$$

(4.10)

where $Z_{\text{string}} \bigg|_{\lim_{z \to 0} z^{\Delta - d} \phi(x, z) = \phi_0(x)}$ is the classical string theory partition function in Euclidean signature that takes the path integral over all field configurations for $\phi(z, x)$ which takes the value $\phi_0(x)$ at the asymptotic boundary of AdS.

One complication is that $Z_{\text{string}}$ is not explicitly known. Therefore, it is often necessary to consider only the limit where the string length is much smaller than the curvature radius, i.e. the weak form introduced above as opposed to the strong form in (4.10). In this limit, the leading contribution to $Z_{\text{string}}$ comes from the field $\tilde{\phi}(x, z)$ that solves the supergravity equations of motion subject to the condition

$$\lim_{z \to 0} z^{\Delta - d} \tilde{\phi}(x, z) = \phi_0(x).$$

(4.11)

One thereby obtains the relation

$$\left\langle \exp \left( \int d^d x \mathcal{O}(x) \phi_0(x) \right) \right\rangle = e^{-S_{\text{sugra}} \bigg|_{\lim_{z \to 0} z^{\Delta - d} \tilde{\phi}(x, z) = \phi_0(x)}}$$

(4.12)

where $S_{\text{sugra}} \bigg|_{\lim_{z \to 0} z^{\Delta - d} \tilde{\phi}(x, z) = \phi_0(x)}$ is the (renormalized) supergravity action.

If the source field of an operator in the CFT is known, then the relation (4.10) tells how operators (and their correlators) on the CFT side relate to fields (and field correlators) on the AdS side.
4.2 Gauge/gravity duality: The weak form

As the subsequent chapters will deal with the weak form of the gauge/gravity duality, we will develop that in more detail here. To obtain the weak form, two limits were taken on the AdS side: \( g_s \to 0 \) and \( l_s/L \to 0 \). On the CFT side, this corresponded to the limit \( N \to \infty \) and large \( \lambda \), where \( N \) is the rank of the gauge group of the theory and \( \lambda \) is the t’Hooft coupling. These limits were taken to avoid having quantum gravity and stringy corrections on the AdS side. This allows the approximation of the AdS side with the Einstein Hilbert action in AdS spacetime

\[
S_{EH} = \int d^{d+1}x\sqrt{g} \left( \frac{1}{16\pi G_N^{d+1}} \left( R + \frac{d(d-1)}{L^2} \right) + \mathcal{L}_M \right)
\]

(4.13)

where \(-\frac{d(d-1)}{L^2}\) is the cosmological constant in AdS spacetime and \(\mathcal{L}_M\) are matter terms.

With this approximation, the gauge/gravity duality can serve to relate general relativity in \( d + 1 \) dimensions to a \( d \)-dimensional conformal field theory. To motivate the approximation of the AdS side by general relativity, we will first argue how supergravity obtains as an approximation on the AdS side and then show how the Einstein Hilbert action is a special case of supergravity action.

Supergravity is an effective low-energy limit of superstring theory where only tree level interactions are included. Since energy – in units \( c = \hbar = 1 \) – scales like inverse length, this is the limit where the energy is much smaller than the inverse of characteristic length scale of the theory, which in the case of string theory is the string length, \( l_s \).

Thus, the low energy limit is equivalent to the limit where \( l_s \to 0 \). In supergravity, therefore, strings are approximated by point particles; stringy correction are ignored. This entails that only the massless closed string particles are included in the mode content of supergravity. To see this, remember that the mass spectrum of the particles in the closed string theory in Minkowski background is given by

\[
M^2 = \frac{4(N - 1)}{l_s^2}.
\]

(4.14)

where \( N = \sum_{n=1}^{\infty} \alpha_{-n} \alpha_n \) and \( \alpha_{-n} \) and \( \alpha_n \) are the annihilation and creation operators, respectively, of the closed string mode expansion. Thus, for Minkowski background, the massless modes – for which \( N = 1 \) – are independent of the string length, \( l_s \), whereas the massive modes scale as \( l_s^{-2} \). When \( l_s \to 0 \) the massive modes decouple such that only the massless modes contribute. Consequently, the low energy limit, and therefore supergravity on Minkowski background, only contains the massless modes of the closed
superstring. The argument generalizes to closed string theory of $AdS$ background (for more details, see [15]).

Among the field content of the closed string is a graviton, which we with some foresight will denote $g_{AB}(X)$, and a dilaton, $\Phi(X)$. Here, $X^A$ is the embedding function from the string world sheet to the target space. The graviton of course is interesting since it perturbs the background spacetime in which the string moves, and should therefore relate to the gravitational dynamics. However, also the fluctuations of the dilaton, it turns out, contribute to the metric perturbation as seen from the point of view of general relativity. To indicate how this obtains as well as how to derive the supergravity action, it is instructive to consider the graviton and dilaton parts of the string action. For convenience, we will work with the action for the closed bosonic string, however, the procedure generalizes to the full closed superstring action.

We will take $\alpha, \beta = 1, 2$ to be worldsheet indexes and $A, B = 0, 1, ... D - 1$ to be target space indexes, where $D$ is the dimension of the target space. We will further assume that the metric induced on the worldsheet by the background metric is $\gamma_{\alpha\beta}$, i.e. the background is not necessarily flat. The world sheet can be parametrized by the timelike coordinate $\xi^0$ and spacelike coordinate $\xi^1$. Since we will work in Euclidean signature, it is convenient to define the coordinate $\xi^2 = i \xi^0$ such that $\xi^2$ is the Wick rotation of the worldsheet time parameter. As usual, $\xi^1$ is the parameter for string extension; for the closed string $0 \leq \xi^1 \leq 2\pi$. The Euclidean, bosonic string action on the world sheet for the graviton and dilaton then takes the form

$$S_{g\Phi} = \frac{1}{4\pi} \int d^2 \xi \sqrt{\gamma} \left[ \frac{1}{l_s^2} \gamma^{\alpha\beta} g_{AB}(X) \partial_\alpha X^A \partial_\beta X^B + R^{(2)}(\Phi(X)) \right]$$

(4.15)

where $R^{(2)}$ is the 2-dimensional Ricci scalar on the world sheet.

The dilaton field may be decomposed into a constant part, $\langle \Phi \rangle$ (the vacuum expectation value of the dilation field) and fluctuations, $\tilde{\Phi}$, around $\langle \Phi \rangle$

$$\tilde{\Phi} = \Phi - \langle \Phi \rangle.$$  

(4.16)
We then find
\[
\frac{1}{4\pi} \int d^2\xi \sqrt{\gamma} R^{(2)} \Phi(X) = \frac{1}{4\pi} \int d^2\xi \sqrt{\gamma} R^{(2)} \Phi(X) + \chi \langle \Phi \rangle .
\]  
(4.17)

where \( \chi = \frac{1}{4\pi} \int d^2\xi \sqrt{\gamma} R^{(2)} \) such that \( \chi \) is identical to the Euler characteristic of the world sheet. As a consequence, it only depends on the topology of the world sheet
\[
\chi = -2(q - 1)
\]  
(4.18)

where \( q \) is the genus of the world sheet. One finds that \( \chi = 2 \) for a sphere \((q = 0)\), \( \chi = 0 \) for a torus \((q = 1)\) and that \( \chi \) is negative for higher genus, \( q \geq 2 \), surfaces. Separating the constant and fluctuating parts of the dilaton field, the action thereby takes the form
\[
S_{g\Phi} = S_{g\Phi} + \langle \Phi \rangle \chi
\]  
(4.19)

where the last term only depends on the world sheet topology.

The Euclidean generating functional associated with \( S_{g\Phi} \) takes the form as a path integral over world sheet metrics, \( \gamma_{\alpha\beta} \), of manifolds, \( \mathcal{M} \), and embedding functions, \( X^A \)
\[
Z_{g\Phi} = \int_{\mathcal{M}} D\gamma_{\alpha\beta} \int DX^A e^{-S_{g\Phi}}.
\]  
(4.20)

Since \( \langle \Phi \rangle \chi \) only depends on the topology, one may rewrite the generating functional as
\[
Z_{g\Phi} = \sum_{q=0}^{\infty} \left( e^{-\langle \Phi \rangle} \right)^{-2(q-1)} \int_{\mathcal{M}^q} D\gamma_{\alpha\beta} \int DX^A e^{-S_{g\Phi}}
\]  
(4.21)

where \( g_s \equiv e^{\langle \Phi \rangle} \); the constant part of the dilaton field serves as the string coupling constant. For \( g_s \ll 1 \), the first term in the sum dominates. In this limit, therefore, one only has to consider the path integral over world sheets that are topological spheres. This gives rise to the tree level diagrams known as Witten diagrams. In subsequent chapters, we will suppose this approximation to be valid.\(^9\)

It proves to be convenient initially to keep \( \Phi \) and only later decompose it in \( \langle \Phi \rangle \) and \( \tilde{\Phi} \). An effective Lagrangian obtained from perturbation theory in \( \Phi \) will take the form
\[
\mathcal{L}_{Eff} = \sum_{q=0}^{\infty} \mathcal{O} \left( e^{-2(q-1)\Phi} \right)
\]  
(4.22)

\(^9\)In fact the effective coupling is \( e^{\Phi(X)} \) which means that it must be small for all \( X \) in order for perturbation theory to be valid.
where terms $\propto e^{-2\Phi}$ are denoted as tree level, and higher order terms are string loop corrections.\footnote{These should not be confused with loop corrections due to an expansion in the string length, $l_s$.}

For general backgrounds, an observable calculated using $Z_{g\Phi}$ will diverge even at tree level. To regulate this divergence, we must introduce a cut-off, $\mu$, and the background fields will generally depend on this cut-off. To keep track of this dependence we introduce the functionals

$$\beta_{AB}^g = \mu \frac{\partial g_{AB}}{\partial \mu}, \quad \beta^{\Phi} = \mu \frac{\partial \Phi}{\partial \mu}. \quad (4.23)$$

An obvious problem with the background fields being dependent on the cut-off is that this threatens to break conformal invariance. Just like gauge invariance is a defining symmetry of Yang-Mills theory as a quantum theory, so is conformal invariance in string theory.

A good indicator as to whether conformal invariance is broken is to look at the expectation value of the trace of the energy momentum tensor, $\langle T^{\alpha\alpha} \rangle$, that must vanish. At each order of the expansion in $g_s$, the trace may be expressed in terms of the beta-functions. Not going into the details, one finds for the regulated action at tree level\footnote{For higher order string loop corrections see [16].}

$$\langle T^{\alpha\alpha} \rangle_{\text{Tree}} = \frac{\beta^{\Phi}}{12} R^{(2)} + \frac{1}{2l_s^2} \beta_{AB}^g g^{\alpha\beta} \partial_\alpha X^A \partial_\beta X^B. \quad (4.24)$$

For this trace to vanish, we see that

$$\beta_{AB}^g = \beta^{\Phi} = 0 \quad (4.25)$$

i.e. neither of the background fields can depend on the regulator, $\mu$. This proves to impose strict constraints on these background fields.

Even if string loop corrections are disregarded, $S_{g\Phi}$ is a non-linear two-dimensional quantum field theory action, it can generally only be studied in perturbation theory. The true dimensionless coupling for the action is $l_s/L$, where $L$ is the characteristic length of the background field $g_{AB}$: for instance the AdS curvature radius in AdS background. As seen, only the dimensionfull coupling, $l_s^2$, occurs explicitly in the action, while $L$ is hidden in $g_{AB}$. We may therefore expand $S_{G\Phi}$ in powers of $l_s^2$ just remembering that perturbation theory is only valid for $l_s/L \ll 1$ and not just $l_s \ll 1$.

Using this expansion, one finds

$$\beta^{\Phi} = D - 26 + \frac{3}{2} l_s^2 \left[ 4(\nabla \Phi)^2 - 4 \nabla_A \nabla^A \Phi - R^{(D)} \right] + \mathcal{O}(l_s^4) + \mathcal{O}(e^{2\Phi})$$

$$\beta_{AB}^g = l_s^2 \left( R^{(D)}_{AB} + 2 \nabla_A \nabla_B \Phi \right) + \mathcal{O}(l_s^4) + \mathcal{O}(e^{2\Phi}) \quad (4.26)$$
where we note that we are now considering the $D$-dimensional Ricci scalar and tensor. For $l_s/L \ll 1$ the terms $\mathcal{O}(l_s^4)$ → 0 and for $g_s \ll 1$ the string loop order corrections vanish. Imposing the constraint $\beta^i = 0$ (4.27), one finds in this limit (for $D = 26$) the following constraints in the background fields

\begin{align}
0 &= 4(\nabla \Phi)^2 - 4 \nabla_A \nabla^A \Phi - R^{(26)} \\
0 &= R^{(26)}_{AB} + 2 \nabla_A \nabla_B \Phi.
\end{align}

(4.27)

Notably, for constant $\Phi$ this reduces to

\begin{equation}
0 = R^{(D)}_{AB}.
\end{equation}

(4.28)

Thus as desired, in the limit $l_s/L \ll 1$ (corresponding to the low energy limit) which we consider here one obtains the Einstein equations without matter (4.28), when all other fields vanish (expect the dilaton field that can be a non-zero constant).

While the beta functions was obtained as the renormalization group flow for the regulated world sheet action, they can also be regarded as equations of motion for the background fields that the string propagates through. We may therefore change the perspective and search instead for a target space (and therefore 26-dimensional) action that has (4.27) as its equations of motion. Without going into the details, this action is

\begin{equation}
S_{Tree} = \frac{1}{2\kappa_0^2} \int d^{26}X \sqrt{-g} e^{-2\Phi} \left( R^{(26)} + 4 \partial_B \Phi \partial^B \Phi \right)
\end{equation}

(4.29)

where $\kappa_0$ is some constant. As expected from according to (4.22), we see an overall factor $e^{-2\Phi}$ which is consistent with the fact that we only included leading order in $e^{2\Phi}$ in the beta-functions, i.e. no string loop corrections.

$S_{Tree}$ is the low-energy effective action for the target space in the presence of a closed bosonic string. It is the low-energy action since (4.27) is valid only for in the limit $l_s \ll 1$ where higher order corrections vanishes. Loop corrections to this tree level action will be additional terms $\mathcal{O}(l_s^4)$.

Comparing $S_{Tree}$ to the Einstein Hilbert action there are some similarities, but also two important differences. First is the presence of the factor $e^{-2\Phi}$ and second is the positive sign of the kinetic term $4\partial_B \Phi \partial^B \Phi$. These, however, can be accommodated by combining the metric field, $g_{AB}$ with the fluctuating part of the dilaton field, $\tilde{\Phi}$

\begin{equation}
\tilde{g}_{AB} = e^{-4\Phi/(D-2)} g_{AB}.
\end{equation}

(4.30)

Associating this new metric with Ricci scalar, $\tilde{R}$ and taking $D = 26$ we can then rewrite the metric as

\begin{equation}
S_{Tree}^E = \frac{1}{2\kappa_0^2 (e^{(\Phi)})^2} \int d^{26}X \sqrt{-\tilde{g}} \left( \tilde{R}^{(26)} - \frac{1}{6} \partial_B \tilde{\Phi} \partial^B \tilde{\Phi} \right).
\end{equation}

(4.31)
This gives the same form as the Einstein Hilbert action in the presence of a field with kinetic term \( \frac{1}{6} \partial_B \tilde{\Phi} \partial^B \tilde{\Phi} \). Expressing the action like this is known as the Einstein perspective as opposed to the string perspective in \( S_{Tree} \). The constant factor can then be determined by comparison to the Einstein Hilbert action

\[
2\kappa_0^2 \left( e^{\langle \Phi \rangle} \right)^2 = 16\pi G_N^{(26)}
\]  

where \( G_N^{(26)} \) is Newton’s constant in 26 dimensions. Identifying, as above, \( g_s \equiv e^{\langle \Phi \rangle} \), one obtains

\[
\kappa_0^2 g_s^2 = 8\pi G_N^{(26)} \Rightarrow \kappa^2 = 8\pi G_N^{(26)}
\]  

where \( \kappa^2 = \kappa_0^2 g_s^2 \). Interestingly, \( G_N^{(26)} \propto g_s^2 \), which is the parameter of the genus expansion of the generating functional for the closed string, \( Z_g \Phi \). Defining the reduced Planck length in 26 dimensions \( l_p = 8\pi G_N^{(26)} \) and observing from dimensional analysis that \( \kappa^2 \propto l_s^2 g_s^2 \), we find, for the weak coupling regime, that \( g_s \ll 1 \Rightarrow l_p \ll l_s \). This supports the previous assertion that the regime \( g_s \ll 1 \) is the regime of classical gravity, if we assume that effects from quantum gravity only become significant when the length scale of the system approaches the Planck length.

Starting instead from the 10 dimensional type IIB string action, we can analogously obtain 10 dimensional type IIB supergravity as the low energy effective action if all string loop corrections are ignored. Since it is tree level in string loop corrections it involves an overall factor \( e^{-2\Phi} \) and includes the same terms derived above – the Ricci scalar and a kinetic term for the dilaton – as well as additional fields due to the additional field content of the type IIB superstring action; among them a self dual four form field. None of these, however, lead to terms in the action identifiable as a cosmological constant. Remarkably, the equations of motion for 10 dimensional supergravity still permits solutions where the non-compact spacetime is an anti-de Sitter; a spacetime with constant negative cosmological constant. This obtains, if the metric is \( AdS_5 \times S^5 \) and the field strength for the four form field adequately relates to the AdS curvature radius. Also, the dilaton field has to be constant, while the dynamics of the remaining supergravity fields decouple.

As asserted above, a Kaluza-Klein reduction of \( S^5 \) leads to a tower of massive modes in the 5 dimensional compactified theory on \( AdS_5 \). Thus, AdS spacetime can appear as a solution for higher dimensional supergravity after a Kaluza-Klein reduction of a compact internal space even if the higher dimensional supergravity contains no cosmological constant. A Kaluza-Klein reduction may therefore reduce this higher dimensional supergravity theory to a lower dimensional gravity theory with a cosmological constant.
Indeed, supergravity on $AdS_5 \times S^5$ may be compactified by a Kaluza-Klein reduction of $S^5$ to some gauged, 5 dimensional supergravity theory [13].\footnote{In gauged supergravity, the gravitinos couple to a gauge field with the consequence that the transformation law for the gravitinos includes a constant term. For the action to remain invariant under this transformation, a constant term – the cosmological constant – must be added to the action.} Such gauged supergravity theories does contain a cosmological constant and has $AdS$ spacetime as its (supersymmetric) vacuum [17].

The implications of this for the AdS/CFT correspondence is that in the low energy limit we may approximate the AdS side by the $d+1$-dimensional Einstein Hilbert action – and therefore the Einstein equations – with cosmological constant $\Lambda = -\frac{d(d-1)}{L^2}$, though we should note that this disregards contributions from the field strength of the four form field that, as argued above, does not decouple. Supposing the validity of this approximation, we can then finally provide holographic picture that will be the outset of the remainder of the thesis.

4.3 \textit{CFT}_d and $AdS_{d+1}$

The overall assumption is that Einstein gravity on $AdS_{d+1}$ is dual to some large $N$, large $\lambda$ CFT in $d$ dimensions, where $\lambda = 2g_{YM}^2 N$.

We will consider a one parameter family of CFT states, $|\Psi(\zeta)\rangle$ where $|\Psi(0)\rangle$ represents the CFT vacuum. We assume that $|\Psi(\zeta)\rangle$ lives in a $d$-dimensional Minkowski spacetime, $R^{d-1,1}$, such that $|\Psi(0)\rangle$ is dual to the $(d+1)$-dimensional, Poincaré patch of pure Anti-de Sitter spacetime with metric

$$ds^2 = \frac{L^2}{z^2} (dz^2 + dx^\mu dx^\nu)$$

where $L$ is a length scale that controls the negative curvature of bulk spacetime. The spacetime has a Minkowski boundary for $z \to 0$ which is in accordance with the general result of the AdS/CFT correspondence that the quantum state $|\Psi\rangle$ with a spacetime dual $M_\Psi$ is defined on the boundary of the dual spacetime, $\partial M_\Psi$. Often, we will be interested in arbitrary Cauchy surfaces $\Sigma_{\partial M_\Psi} \subset \partial M_\Psi$.

Generally, the spacetime dual of excited states $|\Psi\rangle$ can be represented by the metric

$$ds^2 = \frac{L^2}{z^2} (dz^2 + \Gamma_{\mu\nu}(x,z)dx^\mu dx^\nu)$$

where $\Gamma_{\mu\nu}$ is a parameter that controls the perturbation, i.e. how much the spacetime deviates from pure AdS. As seen, pure AdS corresponds to a spacetime where $\Gamma_{\mu\nu} = 0$. Thus, $\Gamma_{\mu\nu} \to 0$ for $\zeta \to 0$ whereas $\Gamma_{\mu\nu} \neq 0$ when $\zeta \neq 0$. 
For small enough $z$, one finds that $\Gamma_{\mu\nu}(x, z) = \eta_{\mu\nu} + z^d \bar{h}_{\mu\nu}(x, z)$ and we may therefore express the metric as

$$ds^2 = \frac{L^2}{z^2}(dz^2 + dx_\mu dx^\mu + z^d \bar{h}_{\mu\nu}(x, z)dx_\mu dx_\nu).$$

(4.36)

This is known as the Fefferman-Graham coordinates. The metric can be used when the gravitational excitations are near the asymptotic boundary, $z \to 0$. On the CFT side, this corresponds to the small excitations, i.e. to states $|\Psi(\zeta)\rangle$ with sufficiently small $\zeta$. These states are small perturbations of the CFT vacuum and the spacetime dual, $M(\zeta)$, can be represented as perturbations of pure AdS spacetime and will take the form of asymptotically AdS spacetimes described by (4.36).

Higher excited states will on the other hand be dual to AdS spacetimes with very different geometry that may even have different topology. For such spacetimes, we cannot assume that $l_s/L \ll 1$ which entails that the Einstein Hilbert action is no longer a valid description of the AdS side, but needs corrections of order $l_s^2$ and higher. This, therefore, takes us beyond the regime of classical (super)gravity. Intuitively, this follows since one has to take into account the effects of the massive string modes in high energy states.

An example of an excited CFT state with a non-trivial dual classical spacetime is the thermal state of a CFT on $d$-dimensional Minkowski spacetime that is dual to the planar Schwarzschild-AdS black hole in $d + 1$ dimensions

$$ds^2 = -f_M(r)dt^2 + \frac{dr^2}{f_M(r)} + \frac{r^2}{L^2}d\Omega_{d-2}^2$$

(4.37)

where $f_M(r) = \frac{r^2}{L^2} - \frac{\mu}{r^{d-2}}$. CFT states dual to black hole spacetimes will play an important role in chapter 7, however, our interest will be in the eternal black hole in $d + 1$ introduced in section 2.4. The eternal black hole is conjectured to be dual to a thermal double state of two CFTs defined on $\mathbb{R} \times S^{d-1}$ [18].

The two identical regions I and III of the eternal black hole (see figure 2.4) have asymptotic boundaries – denoted $A$ and $B$ in figure 2.4 – with spacetime $\mathbb{R} \times S^{d-1}$. These are also causally disconnected. In accord with the general result of the AdS/CFT correspondence, the CFT dual to the eternal black hole is defined on a spacetime identical

\[\text{A number of authors have recently questioned this conjectured duality (see [19–21]). Thus, in the assessment of the conclusions drawn from this duality, it is therefore worth keeping in mind that the duality remains disputed. The general claims about the relation between entanglement and spacetime is not directly disproved if this duality turns out to be false, the intuitive appeal of the qualitative argument will be somewhat damaged.}\]

\[\text{This follows since the eternal black hole spacetime is asymptotically pure global AdS.}\]
to this asymptotic boundary of the AdS spacetime, $A \cup B$.\textsuperscript{15} Thus, the full quantum system is comprised of two identical quantum subsystems, $Q_A$ and $Q_B$, that covers the local degrees of freedom of the CFTs living on the spacetimes identical to $A$ and $B$ respectively. Locality of the CFT requires that $Q_A$ and $Q_B$ cannot interact, i.e. their Hamiltonians must be uncoupled. However, they can still be correlated via entanglement.

More precisely, the two subsystems $Q_A$ and $Q_B$ can be associated with two identical Hilbert spaces of states, $\mathcal{H}_A$ and $\mathcal{H}_B$, that may be spanned by an orthogonal basis consisting of (again identical) energy eigenstates $\{|E^A_i\rangle\}$ and $\{|E^B_i\rangle\}$. The thermofield double state, $|\Psi\rangle$, that is the CFT state dual to the eternal black hole can then be expressed as

$$|\Psi\rangle = \frac{1}{\sqrt{Z(\beta)}} \sum_i e^{-\beta E_i/2} |E^A_i\rangle \otimes |E^B_i\rangle$$

(4.38)

where $\beta$ is the inverse temperature of one of the subsystems\textsuperscript{16} and $Z = \sum_i e^{-\beta E_i}$. Notably, the expression of the thermofield double state, (4.38), in terms of energy eigenstates explicitly unveils the local degrees of freedom of the two subsystems, $Q_A$ and $Q_B$, to be entangled through the weighted sum over states $\sum_i e^{-\beta E_i/2} |E^A_i\rangle \otimes |E^B_i\rangle$. This will be of interest later.

\textsuperscript{15}Note that for the eternal black hole this is not a contiguous spacetime, but instead a spacetime that consists of two disjoint copies of the same spacetime.

\textsuperscript{16}Since they are identical, the inverse temperature is the same in both subsystems.
Chapter 5

The Ryu-Takayanagi Formula

So far, we have only provided very superficial remarks about the quantitative relation between $d$-dimensional CFT and the $d + 1$-dimensional AdS spacetime. The reason for this is that the subsequent study solely relies on a single entry from the AdS/CFT dictionary; the Ryu-Takayanagi formula [5] and its covariant generalization [6]. This formula relates entanglement entropy on the CFT side to areas of extremal co-dimension two surfaces on the AdS side, thereby relating an intrinsically quantum mechanical phenomena on the CFT side to the geometry of the bulk in the spacetime dual. Using this relation alone, it can be shown that if the first law of entanglement entropy is satisfied in a CFT, then Einstein equations to linear order are satisfied in the spacetime dual.

5.1 Relation to holography

Above, we already mentioned the remarkable result from black hole thermodynamics due to Bekenstein and Hawking that relates the entropy of a black hole to its horizon area

$$S_{BH} = \frac{A_{BH}}{4G_N} \quad (5.1)$$

This gave rise to the holographic principle, which we used as a consistency check for the gauge/gravity duality. This entails a particular way of conceiving of the Bekenstein Hawking formula as a relation between some entropy of a CFT subsystem and the area of the black hole horizon; that is the area of a particular co-dimension two surface in the spacetime dual to that CFT. Thus, rather than being a statement about black hole thermodynamics, this holographic interpretation of the Bekenstein Hawking formula
regards it as a piece of the AdS/CFT dictionary. As a consequence, it is natural to speculate whether the Bekenstein Hawking formula signifies a more general relation between entropies of CFT subsystems and co-dimension two surfaces in the bulk of the spacetime dual of that CFT.

As it turns out, such a generalization does indeed exist. For spacetimes described by Einstein gravity with a dual CFT state, Ryu and Takayanagi conjectured that the entanglement entropy (defined as in (3.16)) of some spatial region on the CFT side is proportional to the area of the co-dimension two surface in the bulk of the spacetime dual that minimizes the area functional under certain conditions.

To put this more precisely, let us again consider quantum state, $|\Psi\rangle$, with the dual spacetime, $M_\Psi$. In accordance with the gauge/gravity duality, $|\Psi\rangle$ is defined in the asymptotic boundary of $M_\Psi$, $\partial M_\Psi$. Let $\Sigma_{\partial M_\Psi}$ be a Cauchy surface of $\partial M_\Psi$ and let $B$ be a subregion of the Cauchy slice such that $B \cup \overline{B} = \Sigma_{\partial M_\Psi}$, where $\overline{B}$ is the compliment of $B$. The entanglement entropy, $S_B$, of the quantum system living on $B$ is the entropy associated with entanglement between the quantum systems $Q_B$ and $\overline{Q_B}$ over entangling surface that coincides with the boundary of $B$, $\partial B$, i.e. the von Neumann entropy of the reduced density matrix $\rho_B = \text{tr}_{\overline{Q}_B} (|\Psi\rangle\langle\Psi|)$ (5.2)

where the degrees of freedom in $Q_{\overline{B}}$ are traced out of the full density matrix. Ryu and Takayanagi’s conjecture is that $S_B$ is proportional to the area of the co-dimension two surface, $\tilde{B}$, whose boundary is $\partial \tilde{B} \equiv \tilde{B}|_{\partial M_\Psi} = \partial B$ and which extremizes the area functional

$$A = \int d^d \sigma \sqrt{|\det(g_{ab})|}$$

where $g_{ab}$ is the metric of $M_\Psi$ induced on the surface $\tilde{B}$. If more surfaces extremizes (5.3), then the further requirement is imposed that $\tilde{B}$ must be the surface of least area that is homologous to $B$. Thus, $\tilde{B}$ must be the surface smoothly retractable to $B$ that minimizes the area functional. The Ryu-Takayanagi formula can then be expressed as

$$S_B = \frac{A(\tilde{B})}{4G_N^{d+1}}.$$  (5.4)

It is worth noting that the area of $\tilde{B}$ generally will be divergent. This is due to the infinite proper distance from any point in AdS spacetime to the boundary. This, however, is not a problem. As remarked in section 3.3, short distance entanglement over the entangling surface in continuum quantum field theories renders the entanglement entropy for any quantum subsystem divergent. A regulation of this short distance entanglement on the
Figure 5.1: For the purpose of illustration we here depict spherical space with time, $S^d \times R$, that is conformally equivalent to global AdS.
CFT side exactly corresponds to introducing a boundary cutoff, \( \epsilon \), such that \( z > \epsilon \). This cutoff then serves to regulate the area as given by (5.3).

Comparing the Ryu-Takayanagi formula with the Bekenstein-Hawking formula, the similarity is striking. This is as expected if the Ryu-Takayanagi formula is a generalization of the Bekenstein-Hawking formula in its holographic interpretation. Thus, to establish the Bekenstein-Hawking formula as a special case can serve as a first test of the Ryu-Takayanagi formula. If the entropy, \( S_{BH} \), in the Bekenstein Hawking formula is interpreted as an entanglement entropy, \( S_B \), over some entangling surface, \( \partial S_B \), then the black hole horizon must be interpreted as a co-dimension two surface, \( \tilde{B} \), that minimizes the area functional which has \( \partial B = \partial \tilde{B} \) and which is homologous to \( B \). The problem, therefore, consists in checking whether black hole horizons in general are extremal co-dimension two surfaces that are homologous to a region on the asymptotic boundary of the black hole spacetime.

For spacetimes with black holes situated in the bulk, the black hole horizon cannot coincide with a surface that ends on the boundary, since \( \partial \tilde{B}_{BH} = \emptyset \). Thus, if such bulk black hole horizons are to be identified with the extremal surfaces from the Ryu-Takayanagi formula, it must be for the special case where \( \partial B = \emptyset \). This is the limit where \( B \) covers the full spacetime boundary, i.e. \( \overline{B} = \emptyset \). Given these conditions, it is tempting to suppose that \( \tilde{B} \) is the empty set as well, however, since a black hole spacetime is not simply connected – it contains a singularity – the empty set is not homologous with \( B \). Instead, \( \tilde{B} \) must be a surface that enclose the singularity. Considering all such surfaces, the one with minimal area is the black hole horizon. Obviously, this immediately entails that the dual quantum state to a bulk black hole must be a thermal state. If \( B \) is such that its compliment, \( \overline{B} \), is empty, then the quantum state on \( B \) must be thermal in order for it to have non-zero entropy and thereby for it to have a black hole spacetime dual with non-zero horizon area. Thus, the Bekenstein-Hawking formula in its holographic interpretation may be viewed be a special case of the Ryu-Takayanagi formula.

Claiming this, however, it is necessary to make a qualifying remark about the type of entropy involved. If the full system is in a thermal state, it has an entropy and is therefore not a pure state. But knowing only this, the entropy cannot be designated as entanglement entropy; one simply does not know the origin of the entropy. For this reason, the type of entropy was left unqualified above. However for the eternal black hole, the thermal properties explicitly originate from entanglement effects and the entropy, therefore, is entanglement entropy. As stated in section 4.3, the eternal black hole is
dual to a thermofield double state

\[ |\Psi\rangle = \frac{1}{\sqrt{Z(\beta)}} \sum_i e^{-\beta E_i/2} |E_i\rangle \otimes |E_i\rangle. \]  

(5.5)

This is a pure state, i.e. the von Neumann entropy of the state $|\Psi\rangle$ is zero, but expressed as an entangled state of two subsystems. Tracing out the degrees of freedom of one of the subsystems, the other subsystem is in a thermal state. Knowing the full state, we can identify the thermal entropy of one of the subsystems with the entanglement entropy between the two subsystems. Either subsystem simply is in a thermal state, and consequently the extremal bulk co-dimension two surface whose area is proportional to this entanglement entropy must be identical with the black hole horizon following the argument above. Thus, when we consider the entropy of one of the subsystems in the eternal black hole and notes its origin, then this entropy can be identified as entanglement entropy. Therefore, the Bekenstein-Hawking formula applies to the eternal black hole as a special case of the Ryu-Takayanagi formula.

### 5.2 Uses of the Ryu-Takayanagi formula

As a practical tool, the Ryu-Takayanagi formula allows us to evaluate entanglement entropies in strongly coupled field theories by evaluating the area of $\tilde{B}$ in the weakly coupled dual gravity theory. However, we could also work the other way around. Knowing the entanglement entropy, we can determine the area of the co-dimension two surface, $\tilde{B}$. In itself, this area is not of much interest. It would be interesting, if one could reconstruct the entire dual spacetime from the entanglement entropy of the field theory. Knowing a single area of a co-dimension two surface in the bulk of this spacetime will not get us very far; many spacetimes will be compatible with this particular area. However, subdividing the Cauchy surface, $\Sigma_{\partial M^*}$, into different subsystems will give other entanglement entropies that can be related to different bulk surface areas.

In fact, the problem is highly overconstrained ensuring that a single spacetime at most – the spacetime dual of that CFT state – is compatible with all the areas in the bulk that corresponds to entanglement entropies on the CFT side. Thus, according to the Ryu-Takayanagi formula it is in principle possible to find the spacetime dual of any CFT simply by calculating entanglement entropies. It is worth noting that the entanglement entropies overconstrain the dual spacetime such that there, for a general CFT, will be no spacetime that is compatible with all the entropies. Consequently, only
a smaller subset of CFTs are expected to have dual descriptions as geometric spacetime theories.

However, finding the spacetime dual even for these CFTs faces severe technical difficulties both due to problems associated with the evaluation of entanglement entropies on the CFT side, and due to problems in solving the overconstrained problem on the gravity side. Also, the Ryu-Takayanagi formula is limited to first order in the bulk, and therefore correction will occur from quantum gravity if we go deeper into the bulk.

However, setting aside these technical difficulties and considering only first order perturbations in the bulk where the gravitational theory is well-described by classical Einstein gravity coupled to matter (without curvature couplings), the Ryu-Takayanagi formula suggests an intimate relation between entanglement in $d$ dimensions – an inherently quantum mechanical phenomena – and then the entire bulk spacetime in $d+1$ dimensions in the (rare) cases where gauge/gravity duality obtains.

It is exactly this relation that will be explored subsequently. However, prior to this some evidence for the Ryu-Takayanagi formula will be provided.

5.3 Evidence for the Ryu-Takayanagi formula

At its conception, the evidence of the Ryu-Takayanagi formula consisted in the a number of successful tests, where the extremal surface approach was employed to reproduce known results for entanglement entropies of CFT states. These quantitative comparisons are not very numerous due to the strong coupling of the CFT for $d > 2$. Thus, most of these comparisons are between $CFT_2$ and $AdS_3$.

One such example is the matching of the entanglement entropy for a single interval in a $CFT_2$ on $\mathbb{R}^{1,1}$. As we did in section 3.3, we will consider an interval with length $2a$ centred around the origin on the spacelike slice, $t = 0$, such that $B = \{x \in \mathbb{R} | -a < x < a\}$ and $\partial B = \{-a, a\}$. For a $CFT_2$ on $\mathbb{R}^{1,1}$, the spacetime dual is the Poincaré patch of pure AdS$_3$. Since the background is Poincaré-AdS, the metric induced on the spatial slice, $t = 0$, is $ds^2 = \frac{L^2}{z^2} (dx^2 + dz^2)$. Following the prescription of the Ryu-Takayanagi formula, the entaglement entropy of $B$ is equal to the area of a co-dimension two bulk surface, $\tilde{B}$, in the $t = 0$ plane parametrized by coordinates $x$ and $z$; see the metric (2.16). $\tilde{B}$ must be a curve that has $\tilde{B}|_{\partial M} = \partial B$, i.e. a curve whose endpoints at $z = 0$ are $x = -a$ and $x = a$. Furthermore, it must minimize the area functional (5.3).

We will choose the parametrize $B$ in terms of the boundary coordinate $x$. The metric
induced on $B$ – with the single component $g_{xx}$ – may then be expressed as

$$g_{xx} = \frac{L^2}{z^2} \left( \frac{\partial x}{\partial x} \frac{\partial x}{\partial x} + \frac{\partial z}{\partial x} \frac{\partial z}{\partial x} \right) = \frac{L^2}{z^2} \left( 1 + z'(x)^2 \right). \quad (5.6)$$

Since induced metric has only one component, we have $\det(g) = g_{xx}$. The area functional therefore takes the form

$$A = \int_{-a}^{a} dx \sqrt{g_{xx}} = \int_{-a}^{a} dx \frac{L}{z} \sqrt{1 + z'(x)^2}. \quad (5.7)$$

Defining $A = \frac{1}{z} \sqrt{1 + z'(x)^2}$, we may then find the extremum of the area functional using the Euler-Lagrange equation

$$0 = \frac{\partial A}{\partial z} - \frac{d}{dx} \frac{\partial A}{\partial z'} = -\frac{L \sqrt{1 + z'(x)^2}}{z^2} - \frac{d}{dx} \frac{L z'(x)}{z \sqrt{1 + z'(x)^2}} \quad (5.8)$$

The solution is $z = \sqrt{a^2 - x^2}$ – a semi-circle around the origin in the $zx$ place – as may be verified by inserting back into the numerator of Euler-Lagrange equation:

$$z z''(x) + z'(x)^2 + 1 = -\frac{a^2}{z^2} + \frac{x^2}{z^2} + 1 = 0 \quad (5.9)$$

where we have used $1 = \frac{a^2}{z^2} - x^2$. Thus, the co-dimension two surface, $\tilde{B}$, must satisfy $z^2 + x^2 = a^2$.

Inserting this into the expression for the area, we find

$$A = \int_{-a}^{a} dx \frac{L}{a^2 - x^2} \sqrt{1 + \frac{x^2}{a^2 - x^2}}. \quad (5.10)$$

As expected, this integration diverges. To regulate it, we require that $z > \epsilon$ for some small $\epsilon$. The integration then yields

$$A = 2L \ln \left( \frac{2a}{\epsilon} \right). \quad (5.11)$$

Inserting this into the Ryu-Takayanagi we find

$$S_B = \frac{A}{4G_N} = \frac{L}{2G_N} \ln \left( \frac{2a}{\epsilon} \right). \quad (5.12)$$

If we identify $c = \frac{3}{2} \frac{L}{4G_N}$ in accord with the $\text{AdS}_3/\text{CFT}_2$ dictionary, we get

$$S_B = \frac{c}{3} \ln \left( \frac{2a}{\epsilon} \right). \quad (5.13)$$
The same result as obtained by direct calculation in section 3.3. Besides direct agreement between the entanglement entropies in cases where the calculations are tractable on both the AdS and the CFT side, the Ryu-Takayanagi formula has also been shown to satisfy various known properties of entanglement entropy such as strong subadditivity. A partial derivation of the Ryu-Takayanagi formula for general spherical entangling surfaces was proposed in [22] and based on this a full derivation from the AdS/CFT dictionary was developed in [23]. Both derivations are beyond the scope of the current project. However, elements of [22] will be explored in later chapters since a number of aspects from this partial derivation of the Ryu-Takayanagi formula will prove useful for other purposes.
Chapter 6

Linearized Gravity in AdS

Background

In this section we introduce linearized gravity gravity first in a generic background and then in the specific case of an anti-de Sitter background.

6.1 Perturbing spacetime

Linearized gravity is a type of approximation scheme for Einstein gravity where the spacetime is characterized by some background metric, $g^0_{ab}$, plus a weak perturbation, $\delta g_{ab}$, that generates only small corrections to the background field such that the full metric field, $g_{ab}$, is given by

$$g_{ab} = g^0_{ab} + \delta g_{ab}$$  \hspace{1cm} (6.1)

where $\delta g_{ab}$ is assumed to be small compared to $g^0_{ab}$. Linearized gravity is then obtained by expanding Einstein's field equations to linear order in $\delta g_{ab}$. Since the perturbation, $\delta g_{ab}$, is small, we will assume that we can use the background metric to raise and lower indices.

For convenience, we will consider a one parameter family of metrics, $g_{ab}(\zeta)$, such that for each value of $\zeta$, $g_{ab}(\zeta)$ solves the Einstein equations and $g_{ab}(0) = g^0_{ab}$. Thus, the first order correction to the background metric – the perturbation $\delta g_{ab}$ – is given by

$$\left. \frac{\partial g_{ab}(\zeta)}{\partial \zeta} \right|_{\zeta=0} = \delta g_{ab}.$$  \hspace{1cm} (6.2)
The linearized Einstein equations can then similarly be obtained as
\[ \frac{\partial G_{ab}^E(\zeta)}{\partial \zeta} \bigg|_{\zeta=0} = 8\pi G_N \frac{\partial T_{ab}(\zeta)}{\partial \zeta} \bigg|_{\zeta=0}. \] (6.3)

These equations impose constraints on \( \delta g_{ab} \). To find them, we must establish what change is induced to linear order on the Einstein tensor by the linear order perturbation \( \delta g_{ab} \). For convenience, we will introduce the notation
\[ \frac{\partial G_{ab}^E(\zeta)}{\partial \zeta} \bigg|_{\zeta=0} \equiv \dot{G}_{ab}^E \] (6.4)
and may thereby rephrase the above as the task of finding an expression for \( \dot{G}_{ab}^E \) in terms of \( \delta g_{ab} \).

The Einstein tensor with cosmological constant is given by
\[ G_{ab}^E = R_{ab} - \frac{1}{2} R g_{ab} + \Lambda g_{ab} = \left( \delta^c_a \delta^d_b - \frac{1}{2} g^{cd} g_{ab} \right) R_{cd} + \Lambda g_{ab}. \] (6.5)

where \( R_{ab} \) is the Ricci tensor, \( R = g^{ab} R_{ab} \) is the Ricci scalar and \( \Lambda \) is the cosmological constant. Thus,
\[ \dot{G}_{ab}^E = \frac{\partial}{\partial \zeta} \bigg|_{\zeta=0} \left[ \left( \delta^c_a \delta^d_b - \frac{1}{2} g^{cd} g_{ab} \right) R_{cd} + \Lambda g_{ab} \right] \\
= \left( \delta^c_a \delta^d_b - \frac{1}{2} g^{cd} g_{ab} \right) \dot{R}_{cd} - \frac{1}{2} \left( g^{cd} \delta g_{ab} - \delta g^{cd} g_{ab} \right) R_{cd} + \Lambda \delta g_{ab} \] (6.6)

where we have defined \( \delta g_{ab} \equiv g^{ac} g^{bd} \delta g_{bd} \) from which it follows that \( \frac{\partial g_{ab}}{\partial \zeta} \bigg|_{\zeta=0} = -\delta g_{ab} \).\(^1\)

To find an expression for \( \dot{G}_{ab}^E \), one must therefore first find an expression for the perturbation of the Ricci tensor \( \dot{R}_{ab} \). The Ricci tensor is given as
\[ R_{ab} = R_{acb}^c \] (6.8)
where \( R_{acb}^c \) is the Riemann tensor. The Riemann tensor, in turn, depends on the metric connections:
\[ R_{abc}^d = -\partial_a \Gamma_{bc}^d + \partial_b \Gamma_{ac}^d - \Gamma_{ae}^d \Gamma_{bc}^e + \Gamma_{be}^d \Gamma_{ac}^e. \] (6.9)

\(^1\)To see this, consider
\[ \frac{\partial}{\partial \zeta} \delta^c_a = \frac{\partial}{\partial \zeta} g^{ac} g_{bc} \]
\[ 0 = g_{bc} \frac{\partial g^{ab}}{\partial \zeta} + g^{ac} \frac{\partial g_{bc}}{\partial \zeta} \]
\[ 0 = g^{cd} g_{bc} \frac{\partial g^{ab}}{\partial \zeta} + g^{cd} g^{ab} \delta g_{bc} \]
\[ 0 = \frac{\partial g^{ad}}{\partial \zeta} + \delta g^{ad} \] (6.7)
The linear order change in the Riemann tensor induced by the linear order perturbation, \( \delta g_{ab} \), may be expanded as

\[
\frac{\partial}{\partial \zeta} R_{d}^{abc} = - \frac{\partial}{\partial \zeta} \left( \partial_a \Gamma^d_{bc} - \partial_b \Gamma^d_{ac} + \Gamma^d_{ae} \Gamma^e_{bc} - \Gamma^d_{be} \Gamma^e_{ac} \right) = - \partial_a \dot{\Gamma}^d_{bc} + \partial_b \dot{\Gamma}^d_{ac} - \dot{\Gamma}^d_{ae} \Gamma^e_{bc} + \dot{\Gamma}^d_{be} \Gamma^e_{ac} - \partial_a \dot{\Gamma}^d_{bc} + \dot{\Gamma}^d_{be} \Gamma^e_{ac} + \dot{\Gamma}^d_{be} \Gamma^e_{ac}
\]  

(6.10)

where we have used that the partial derivative is independent of \( \zeta \). The last line follows since

\[
- \nabla_a \dot{\Gamma}^d_{bc} + \nabla_b \dot{\Gamma}^d_{ac} = - \partial_a \dot{\Gamma}^d_{bc} + \partial_b \dot{\Gamma}^d_{ac} - \dot{\Gamma}^d_{ae} \Gamma^e_{bc} + \dot{\Gamma}^d_{be} \Gamma^e_{ac} + \dot{\Gamma}^d_{be} \Gamma^e_{ac}.
\]  

(6.11)

Here we have used that the metric connection is torsion free: \( \Gamma^e_{ab} = \Gamma^e_{ba} \).

What remains is to find perturbation of the metric connection. To find an expression for this, consider initially the derivative with respect to \( \zeta \) of the vanishing covariant derivative of the metric

\[
0 = \frac{\partial}{\partial \zeta} \left( \partial_a g_{bc} - \Gamma^d_{ab} g_{dc} - \Gamma^d_{ac} g_{ad} \right) = - \Gamma^d_{ab} g_{dc} - \Gamma^d_{ac} g_{ad} + \nabla_a \delta g_{bc}.
\]  

(6.12)

Reorganizing this, we find

\[
\nabla_a \delta g_{bc} = \dot{\Gamma}^d_{ab} g_{dc} + \dot{\Gamma}^d_{ac} g_{ad}
\]  

(6.13)

Another constraint follows, since the metric connection must be torsion free in the first two indices. Thus,

\[
\dot{\Gamma}^c_{(ab)} = \dot{\Gamma}^c_{ab}.
\]  

(6.14)

Now, using first (6.13) and then that the metric connection must be torsion free, we find

\[
\dot{\Gamma}^c_{abc} = \nabla_a \delta g_{bc} - \dot{\Gamma}^c_{ac} = \nabla_a \delta g_{bc} - \dot{\Gamma}^c_{cab}.
\]  

(6.15)
Applying this two more times, we get
\[
\dot{\Gamma}^{abc} = \nabla_a \delta g_{bc} - \dot{\Gamma}_{cab} = \nabla_a \delta g_{bc} - \left( \nabla_c \delta g_{ab} - \dot{\Gamma}_{bca} \right) = \nabla_a \delta g_{bc} - \nabla_c \delta g_{ab} + \left( \nabla_b \delta g_{ca} - \dot{\Gamma}_{abc} \right).
\]  

(6.16)

The metric perturbation now occurs both on the left and on the right hand side above. Collecting these, dividing by two and raising the last index, we finally get
\[
\dot{\Gamma}^{c}_{ab} = \frac{1}{2} g^{cd} \left( \nabla_a \delta g_{bd} - \nabla_d \delta g_{ab} + \nabla_b \delta g_{da} \right).
\]  

(6.17)

As seen, this expression only depends on metric perturbation and covariant derivatives taken with respect to the background metric.

Inserting this expression, the perturbation of the Ricci tensor becomes
\[
\dot{R}^{c}_{ab} = \dot{R}^{c}_{ab} = -\nabla_a \dot{\Gamma}^{c}_{ab} + \nabla_c \dot{\Gamma}^{c}_{ab} = -\frac{1}{2} g^{cd} \nabla_a \left( \nabla_c \delta g_{bd} - \nabla_d \delta g_{cb} + \nabla_b \delta g_{cd} \right) + \frac{1}{2} g^{cd} \nabla_c \left( \nabla_a \delta g_{bd} - \nabla_d \delta g_{ab} + \nabla_b \delta g_{da} \right) = -\frac{1}{2} g^{cd} \left( R^{e}_{acdb} + R^{e}_{acdb} - \nabla_a \nabla_d \delta g_{cb} + \nabla_a \nabla_b \delta g_{cd} + \nabla_c \nabla_d \delta g_{ab} - \nabla_c \nabla_b \delta g_{da} \right) + \frac{1}{2} g^{cd} \nabla_c \left( \nabla_a \delta g_{bd} - \nabla_d \delta g_{ab} + \nabla_b \delta g_{da} \right) = -\frac{1}{2} g^{cd} \left( R^{e}_{acdb} + R^{e}_{acdb} - \nabla_a \nabla_d \delta g_{cb} + \nabla_a \nabla_b \delta g_{cd} + \nabla_c \nabla_d \delta g_{ab} - \nabla_c \nabla_b \delta g_{da} \right)
\]

(6.18)

where the last line follows from the identity
\[
\nabla_a \nabla_c \delta g_{bd} - \nabla_c \nabla_a \delta g_{bd} = R^{e}_{acb} \delta g_{ed} + R^{e}_{acdb} \delta g_{eb}.
\]  

(6.19)

Using the analogous identity
\[
\nabla_c \nabla_b \delta g_{da} - \nabla_b \nabla_c \delta g_{da} = R^{e}_{cba} \delta g_{de} + R^{e}_{cbda} \delta g_{ec}
\]

(6.20)

we can substitute for the last term above and act with $g^{cd}$ on all terms
\[
\dot{R}^{c}_{ab} = -\frac{1}{2} g^{cd} \left( R^{e}_{acdb} + R^{e}_{acdb} - \nabla_a \nabla_d \delta g_{cb} + \nabla_a \nabla_b \delta g_{cd} + \nabla_c \nabla_d \delta g_{ab} - \nabla_c \nabla_b \delta g_{da} \right) = -\frac{1}{2} \left( R^{e}_{a b} \delta g_{cd} + R^{e}_{a c} \delta g_{eb} - \nabla_a \nabla^e \delta g_{cb} + \nabla_a \nabla_b \delta g^{e}_{c} + \nabla_c \nabla^e \delta g_{ab} - \nabla_b \nabla^d \delta g_{da} - R^{e}_{e a} \delta g_{de} - R^{e}_{e b} \delta g_{de} \right) = -\frac{1}{2} \left( \nabla_c \nabla^e \delta g_{ab} - \nabla_a \nabla^c \delta g_{cb} - \nabla_b \nabla^e \delta g_{ca} + \nabla_a \nabla_b \delta g^{e}_{c} - R^{e}_{a b} \delta g_{bd} - R^{e}_{b d} \delta g_{da} + 2 R^{e}_{a b} \delta g_{dc} \right)
\]

(6.21)
where we have used $R^c_{eb} = R^e_b$, $R^c_{ac} = -R^c_{ca} = -R^d_e = R^d_e$, and also $R^c_{a b} \delta g_{d c} = R^c_{b a} \delta g_{c d}$.

6.2 Perturbing the Poincaré patch

The expression in terms of the Riemann and Ricci tensor is convenient since we will be interested in the special case where the background is the Poincaré patch of anti-de Sitter spacetime; a solution to Einstein’s equations with negative cosmological constant

$$\Lambda = -\frac{d(d - 1)}{2L^2}.$$ (6.22)

This is a maximally symmetric spacetime and the Riemann tensor may therefore be expressed as

$$R_{abcd} = -\frac{1}{L^2} (g_{ac} g_{bd} - g_{ad} g_{bc})$$ (6.23)

and the Ricci tensor takes the form

$$R_{ab} = -\frac{d}{L^2} g_{ab}.$$ (6.24)

Using the Poincaré coordinates, the metric tensor is diagonal, such that $g_{ab}$ and $R_{ab}$ vanishes for $a \neq b$. Consequently, the Riemann tensor contracted with the metric perturbation becomes:

$$R^c_{a b} \delta g_{c d} = g^{ce} g^{df} R_{aebf} \delta g_{cd}$$

$$= g^{ce} g^{df} (g_{ab} g_{ef} - g_{af} g_{eb}) - \frac{1}{L^2} \delta g_{cd}$$

$$= (\delta^c_f g^{df} g_{ab} - \delta^c_d \delta^f_a) - \frac{1}{L^2} \delta g_{cd}$$ (6.25)

and the Ricci tensor contracted with the metric perturbation becomes

$$R^c_a \delta g_{bc} = g^{cd} R_{ad} \delta g_{bc}$$

$$= g^{cd} \frac{-d}{L^2} g_{ad} \delta g_{bc}$$

$$= \frac{-d}{L^2} \delta^c_a \delta g_{bc}$$

$$= \frac{-d}{L^2} \delta g_{ba}.$$ (6.26)
Subsequently, we will primarily be interested in $G^E_{tt}$. Using the above we find

\[
\dot{G}^E_{tt} = \frac{1}{2} \left( -\delta g^{cd} g_{tt} + g^{cd} \delta g_{tt} \right) R_{cd} + \left( \delta_i^c \delta^d_l - \frac{1}{2} g^{cd} g_{tt} \right) \dot{R}_{cd} - \frac{d(d-1)}{2L^2} \delta g_{tt}
\]

\[
= \frac{1}{2} \left( -g^{ce} g^{df} \delta g_{ef} g_{tt} + g^{cd} \delta g_{tt} \right) \frac{d}{L^2} g_{cd}
\]

\[
+ \dot{R}_{tt} - \frac{1}{2} \left( \dot{R}_{tt} - \dot{R}_{zz} - \delta^{ij} \dot{R}_{ij} \right) - \frac{d}{2L^2} \delta g_{tt}
\]

\[
= \frac{d}{2L^2} \left( -\delta_i^c \delta^d_l g_{tt} + \delta^c_e - (d-1) \delta g_{tt} \right) + \frac{1}{2} \left( \dot{R}_{tt} + \dot{R}_{zz} + \delta^{ij} \dot{R}_{ij} \right) - \frac{d(d-1)}{2L^2} \delta g_{tt}
\]

\[
= \frac{d}{2L^2} \left( \delta^{ij} \delta g_{ij} - \delta g_{tt} \right) + \frac{1}{2} \delta_{ab} \dot{R}_{ab}
\]

(6.27)

where we have assumed radial gauge ($\delta g_{zz} = 0$) and used that $\delta^a_a = d + 1$.

What remains is to evaluate $\delta_{ab} \dot{R}_{ab}$:

\[
\dot{R}_{ab} = \delta_{ab} \left( -\frac{1}{2} \nabla_c \nabla_c \delta g_{ab} + \frac{1}{2} \nabla_a \nabla_c (\delta g_{bc} - \frac{1}{2} \delta g^d_d g_{bc}) \right)
\]

\[
+ \frac{1}{2} \nabla_b \nabla_c (\delta g_{ac} - \frac{1}{2} \delta g^d_d g_{ac}) + \frac{1}{2} R_a^c \delta g_{bc} + \frac{1}{2} R_b^c \delta g_{ac} - R_{c d} \delta g_{cd}
\]

\[
= \delta_{ab} \left( -\frac{1}{2} \nabla_c \nabla_c \delta g_{ab} + \nabla_a \nabla_c (\delta g_{bc} - \frac{1}{2} \delta g^d_d g_{bc}) + R_a^c \delta g_{bc} - R_c^d \delta g_{cd} \right)
\]

\[
= \frac{z^2}{L^2} \left[ -\frac{d}{z^2} \delta g_{tt} - \left( \partial_z^2 + \partial_i \partial^i - \frac{(d-5)}{z} \partial_z + \frac{d-4}{z^2} \right) \delta_{kl} \delta g_{kl} + \partial_i \partial^i \delta g_{ij} \right]
\]

(6.28)

where we have defined $\partial_i \equiv \delta_{ij} \partial^j$.

We will suppose that $\zeta$ is small such that the dual spacetimes can be represented using the Fefferman-Graham coordinates by the metric

\[
ds^2 = \frac{L^2}{z^2} \left( dz^2 + dx_\mu dx^\mu + z^d \tilde{h}_{\mu \nu} (x, z) dx^\mu dx^\nu \right)
\]

(6.29)

where we, in the last equality, have defined $h_{\mu \nu} \equiv \tilde{h}_{\mu \nu} L^2$. This is convenient since it permits us to write the perturbed metric on the form

\[
g_{ab} = g^0_{ab} + \delta g_{ab} = g_{ab}^{AdS} + z^{d-2} h_{ab}
\]

(6.30)

Here we have identified $\delta g_{ab} = z^{d-2} h_{ab}$ where $h_{ab}$ is equal to $h_{\mu \nu}$ with the additional components $h_{zz} = 0$ and $h_{az} = 0$. From this, we readily see that the Fefferman-Graham
coordinates presumes radial gauge for the perturbation $z^{d-2}h_{ab}$. Also, we have assumed that the background metric is the Poincaré patch of AdS:

$$
g_0^{ab} = \begin{pmatrix} -\frac{L^2}{z^2} & \frac{L^2}{z^2} \\ \frac{L^2}{z^2} & \ddots \\ & & & \frac{L^2}{z^2} \end{pmatrix}$$  

(6.31)

Inserting this into the expression for $\delta^{ab}\dot{R}_{ab}$, one finds

$$\delta^{ab}\dot{R}_{ab} = \frac{z^2}{L^2} \left[ -\frac{d}{z^2} z^{d-2} h_{tt} - \left( \frac{\partial_z^2}{z} + \partial_z \partial^i - \frac{(d-5)}{z} \partial_z + \frac{d-4}{z^2} \right) \delta^{kl} z^{d-2} h_{kl} + \partial^i \partial^j z^{d-2} h_{ij} \right]$$

\hspace{1cm} + \frac{z^2}{L^2} \left[ -\frac{d}{z^2} z^{d-2} h_{tt} - \frac{(d-2)}{z^2} - \frac{d-2}{z} \partial_z + \partial_z^2 + \partial_z \partial^i \right] \right] \delta^{kl} h_{kl} + \partial^i \partial^j h_{ij}$$

\hspace{1cm} = \frac{z^2}{L^2} \left[ -\frac{d}{z^2} z^{d-2} h_{tt} - \frac{(d-2)}{z^2} - \frac{d-2}{z} \partial_z + \partial_z^2 + \partial_z \partial^i \right] \right] \delta^{kl} h_{kl} + \partial^i \partial^j h_{ij}$$

\hspace{1cm} = \frac{z^2}{L^2} \left[ -\frac{d}{z^2} z^{d-2} h_{tt} - \frac{(d-2)}{z^2} - \frac{d-2}{z} \partial_z + \partial_z^2 + \partial_z \partial^i \right] \delta^{kl} h_{kl} + \partial^i \partial^j h_{ij}$$

\hspace{1cm} + \frac{z^2}{L^2} \left[ -\frac{d}{z^2} z^{d-2} h_{tt} - \frac{(d-2)}{z^2} - \frac{d-2}{z} \partial_z + \partial_z^2 + \partial_z \partial^i \right] \delta^{kl} h_{kl} + \partial^i \partial^j h_{ij}$$

Reinserting this into the expression for $\dot{G}^{E}_{tt}$ and again substituting $\delta g_{ab}$ with $z^{d-2}h_{ab}$, one finally gets

$$\dot{G}^{E}_{tt} = \frac{d}{2L^2} (\delta^{ij} z^{d-2} h_{ij} + z^{d-2} h_{tt})$$

\hspace{1cm} + \frac{z^2}{L^2} \left[ -\frac{d}{z^2} z^{d-2} h_{tt} - \frac{(d-2)}{z^2} - \frac{d-2}{z} \partial_z + \partial_z^2 + \partial_z \partial^i \right] \delta^{kl} h_{kl} + \partial^i \partial^j h_{ij}$$

\hspace{1cm} = \frac{z^2}{L^2} \left[ -\frac{d}{z^2} z^{d-2} h_{tt} - \frac{(d-2)}{z^2} - \frac{d-2}{z} \partial_z + \partial_z^2 + \partial_z \partial^i \right] \delta^{kl} h_{kl} + \partial^i \partial^j h_{ij}$$

\hspace{1cm} + \frac{z^2}{L^2} \left[ -\frac{d}{z^2} z^{d-2} h_{tt} - \frac{(d-2)}{z^2} - \frac{d-2}{z} \partial_z + \partial_z^2 + \partial_z \partial^i \right] \delta^{kl} h_{kl} + \partial^i \partial^j h_{ij}$$

\hspace{1cm} + \frac{z^2}{L^2} \left[ -\frac{d}{z^2} z^{d-2} h_{tt} - \frac{(d-2)}{z^2} - \frac{d-2}{z} \partial_z + \partial_z^2 + \partial_z \partial^i \right] \delta^{kl} h_{kl} + \partial^i \partial^j h_{ij}$$

The equation

$$\dot{G}^{E}_{tt} = 0$$  

(6.34)

is the tt-component of the linearized Einstein equations in vacuum in the presence of a negative cosmological constant $-\frac{d(d-1)}{2L^2}$.

As one may expect, $\dot{G}^{E}_{ab} = 0$ is also the equations of motion obtained to leading order for an expansion in $G_N$ of the Einstein-Hilbert action. The first subleading correction in $G_N$ turns on the coupling to matter fields with the inclusion of the source term $8\pi G_N \delta \langle T_{ab} \rangle$.

\hspace{1cm} \textit{57}
Chapter 7

Entanglement and Spacetime

An initial qualitative indication of the relation between gravity on the AdS side and entanglement on the CFT side may be obtained by considering the dual pair: the eternal black hole and the thermofield double state. Here the Bekenstein-Hawking formula interpreted holographically, i.e. as a special case of the Ryu-Takayanagi formula, can be used to monitor what happens to the black hole horizon in the eternal black hole when entanglement is added or removed between the double states in the thermofield double state. Thus, it serves as a probe of the relation between entanglement on the CFT side and some spacetime surface on the AdS side. While this special case may be studied employing a holographic interpretation of the Bekenstein-Hawking formula, general, holographic, classical spacetimes may similarly be studied using the Ryu-Takayanagi formula. Below, the special case of the thermofield double state/eternal black hole will be examined first, and we will then subsequently consider the generalization to arbitrary quantum states with a classical spacetime dual.

7.1 Entanglement and the eternal black hole

The thermofield double state is a particular state, $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, in which we can find the quantum system comprised of the subsystems $Q_A$ and $Q_B$. Interestingly, the expression of the thermofield double state, (4.38), in terms of energy eigenstates explicitly unveils how the degrees of freedom of the two subsystems are entangled through the weighted sum over states $\sum_i e^{-\beta E_i/2} |E_i^A\rangle \otimes |E_i^B\rangle$. More explicitly, we find that the
density matrix for one of the subsystems is given by

$$\rho_A = \text{tr}_B (|\Psi\rangle \langle \Psi|)$$

$$= \frac{1}{Z(\beta)} \sum_m \left( \sum_i e^{-\beta E_i/2} |E_i^A\rangle \otimes |E_i^B\rangle \right) \left( \sum_j e^{-\beta E_j/2} \langle E_j^A| \otimes \langle E_j^B| \right) |E_m^B\rangle$$

$$= \frac{1}{Z(\beta)} \sum_m \sum_i \sum_j e^{-\beta(E_j+E_i)/2} \delta_{jm} \delta_{im} |E_i^A\rangle \langle E_j^A|$$

$$= \frac{1}{Z(\beta)} \sum_m e^{-\beta E_m} |E_m^A\rangle \langle E_m^A|$$

(7.1)

from which it follows that

$$S_A = -\frac{1}{Z(\beta)} \sum_m e^{-\beta E_m} \log \left( \frac{1}{Z(\beta)} e^{-\beta E_m} \right)$$

$$= \sum_i e^{-\beta E_i} \sum_m e^{-\beta E_m} \left( \beta E_m + \log \left( \sum_j e^{-\beta E_j} \right) \right)$$

$$= \sum_i e^{-\beta E_i} \sum_m e^{-\beta E_m} \beta E_m + \sum_j \log \left( e^{-\beta E_j} \right)$$

(7.2)

where we have used $Z(\beta) = \sum_i e^{-\beta E_i}$. This will generally be non-zero except for the limit where $\beta \to \infty$, i.e. the limit where the temperature goes to zero. It is evident from this, how the inverse temperature $\beta$ controls the entanglement entropy and therefore the amount of entanglement between the degrees of freedom in the two subsystems.

From the reduced density matrix, $\rho_A$, we see that in the limit $\beta \to \infty$ only the ground state of the subsystem, $Q_A$, is occupied. This follows since all the energy eigenvalues other than the vacuum state, $E_i \neq 0$, are non-zero and therefore vanish in the limit $\beta \to \infty$ as a consequence of the factor $e^{-\beta E_i}$. Due to the symmetry of the thermofield double state, the same holds for the subsystem, $Q_B$. Thus, in the limit $\beta \to \infty$ the full system is the product of these two ground states

$$|\Phi\rangle = |E_0^A\rangle \otimes |E_0^B\rangle$$

(7.3)

i.e. $|\Psi\rangle \overset{\beta \to \infty}{=} |\Phi\rangle$.

Manifestly, $|\Phi\rangle$ is a product state and does not, as expected from the expression for $S_A$, contain entanglement between the degrees of freedom in $Q_A$ and $Q_B$. In this state, therefore, the two systems are completely uncorrelated. What is argued in [1] is that a first indication of the relation between entanglement and spacetime may be
obtained from the differences between the spacetime dual of a product state like $\Phi$ and an entangled state like $\Psi$.

The state $\Phi$ where there is no entanglement between $Q_A$ and $Q_B$ consists of the tensor product of two identical pure states; more specifically a product of the identical vacuum states of these subsystems, $|E_0\rangle$. If we suppose that a pure state is dual to some spacetime, then the product of two pure states is dual to the product of two such spacetimes. Thus, the full spacetime consists of the product of two completely uncorrelated spacetimes. It seems, therefore, reasonable to suppose that these two spacetimes are disconnected, i.e. no light signal travelling from one spacetime can intersect a light signal travelling from the other. From the conjectured duality between the thermofield double state and the eternal black hole, it follows that the spacetime dual of the thermofield double state, where $Q_A$ and $Q_B$ are entangled, is a connected spacetime, i.e. in the eternal black hole a light signal travelling from $A$ can intersect a light signal travelling from $B$. The comparison between $\Psi$ and $\Phi$ in the context of the thermofield double states suggests that entanglement between $Q_A$ and $Q_B$ is a necessary condition for connectivity between $A$ and $B$ in the dual eternal black hole.

This conclusion sits well with the Bekenstein-Hawking formula that relates the entropy of a black hole, $S_{BH}$, with its horizon area, $A_{BH}$ [12].

$$S_{BH} = \frac{A_{BH}}{4G} \quad (7.4)$$

The duality between the eternal black hole and the thermofield double state entails that the black hole entropy – regardless of its origin on the AdS side – is equal to the entanglement entropy of one of the subsystems in the thermofield double state [24]. Employing this relation, it is possible to monitor the horizon area as seen from either side of the eternal black hole, when entanglement between $Q_A$ and $Q_B$ is removed in the dual thermofield double state. Since the horizon area, $A_{BH}$, and entanglement entropy, $S_A$, is proportional, it follows that when entanglement is removed between $Q_A$ and $Q_B$, then the black hole horizon area decreases. Since the entanglement entropy is controlled by $\beta$, it follows that the black hole horizon area decreases when $\beta$ increases. Apparently, in the limit where all entanglement is removed between $Q_A$ and $Q_B$ (i.e. where $\beta \to \infty$), the entanglement entropy goes to zero and so does the horizon area. This limit is depicted in figure 7.1 for spatial slices $T = 0$ and $T$ equal to a positive constant, where $T$ is the timelike Kruskal-Szekeres coordinate.

The depiction suggests that the spacetime pinches such that it is only contiguous at $T = 0$. Thus, in the limit where $\beta \to \infty$ in the thermofield double state, the spacetime dual is connected by a single point. This does not exactly reproduce the
initial assumption that the product of two pure states, such as $|E_0\rangle \otimes |E_0\rangle$, are dual to two completely disconnected spacetimes. A singularity remains, whose interpretation is uncertain. This, however, should perhaps not worry us too much. Due to the critical minimum temperature for AdS-Schwarzschild black holes below which they cannot exist, it is unclear whether the above account of the limit $\beta \to \infty$ is correct. When $\beta \to \infty$, the temperature of the thermal state of $Q_A$ and $Q_B$ goes to zero, i.e. below the critical temperature. We should therefore be careful when taking the limit $\beta \to \infty$. Despite this complication, it is clear how removing (some) entanglement between the two subsystems of the thermofield double state, $Q_A$ and $Q_B$, decreases the connectivity in the dual eternal black hole spacetime. From this latter observation, van Raamsdonk – not considering the complication related the minimum temperature – concludes: “In this example, classical connectivity arises by entangling the degrees of freedom in the two components” [1, 2325, emphasis in original]. Entanglement between $Q_A$ and $Q_B$ is a necessary condition for spacetime connectivity between regions A and B in the eternal black hole.

7.2 Beyond the eternal black hole

While conceived in the context of the duality between the eternal black hole and the thermofield double state, the relation between entanglement and spacetime connectivity generalises to any quantum state with a classical spacetime dual.

Suppose again that $|\Psi\rangle$ is a generic quantum state with a classical spacetime dual $M_\Psi$ with boundary $\partial M_\Psi$ and $|\Psi\rangle$ is a state in the Hilbert space, $\mathcal{H}_{\partial M_\Psi}$, for a CFT.
defined on (a spatial slice of) $\partial M$. Now, divide the boundary $\partial M$ into two regions $B$ and $\overline{B}$, such that $B \cup \overline{B} = \partial M$ (see figure 5.1). Since a CFT is a local quantum field theory, there are specific degrees of freedom associated with specific spatial regions. We can therefore regard the full quantum system as composed of two subsystems, $Q_B$ and $Q_{\overline{B}}$, associated with two spatially separated regions $B$ and $\overline{B}$. As a consequence, we can decompose the Hilbert space of states of the full system as a tensor product of the Hilbert spaces of states of $Q_B$ and $Q_{\overline{B}}$:

$$\mathcal{H}_{\partial M} = \mathcal{H}_B \otimes \mathcal{H}_{\overline{B}}$$  \hfill (7.5)

$|\Psi\rangle$ can therefore be expressed as a sum over products of states $|\psi_i^B\rangle \in \mathcal{H}_B$ and $|\psi_j^{\overline{B}}\rangle \in \mathcal{H}_{\overline{B}}$:

$$|\Psi\rangle = \sum_{i,j} p_{i,j} |\psi_i^B\rangle \otimes |\psi_j^{\overline{B}}\rangle$$  \hfill (7.6)

This will generally not be a product state, i.e. a product of a state in $\mathcal{H}_B$ and $\mathcal{H}_{\overline{B}}$. Thus, the local degrees of freedom in $Q_B$ and $Q_{\overline{B}}$ will generally be entangled.

Again, assume that the spacetime dual of a product state $|\Phi\rangle = \left( \sum_i c_i |\psi_i^B\rangle \right) \otimes \left( \sum_j d_j |\psi_j^{\overline{B}}\rangle \right)$ is dual to two disconnected spacetimes. One then obtains the result that entanglement between $Q_B$ and $Q_{\overline{B}}$ in $|\Psi\rangle$ is a necessary condition for the dual spacetime $M_{\Psi}$ to be a connected spacetime. The duality between the thermofield double state and the eternal black hole is just a particular example of this.

We may monitor this more closely using the Ryu-Takayanagi formula:

$$S_B = \frac{A(\tilde{B})}{4G_N}$$  \hfill (7.8)

Again, it follows that changing the entanglement between $Q_B$ and $Q_{\overline{B}}$ changes the connectivity between the two corresponding regions in the spacetime dual.

When $|\Psi\rangle \rightarrow |\Phi\rangle$, the state of the full quantum system becomes a product of two pure states such that there is no entanglement between the local degrees of freedom in $Q_B$ and $Q_{\overline{B}}$. Thus, in this limit the entanglement entropy, $S_B$, goes to zero and so does the area of $\tilde{B}$ according to the Ryu-Takayanagi formula. More explicitly stated, in this limit the bulk metric changes such that the minimal area dividing the two asymptotic

\footnote{Some complications are involved in making such a decomposition in a gauge invariant way, but these will not be considered here.}
regions $B$ and $\overline{B}$ in the spacetime goes to zero; the spacetime dual of the quantum state pinches when $|\Psi\rangle \rightarrow |\Phi\rangle$. For the spatial surface $\Sigma_{\partial M_4}$, figure 7.2 depicts the limit where all entanglement is removed between $Q_B$ and $Q_{\overline{B}}$.

![Figure 7.2](image)

Figure 7.2: Depiction of the behavior of the spatial slice $\Sigma_{\partial M_4}$ when all entanglement is removed between $Q_B$ and $Q_{\overline{B}}$. Note that the quantum state is defined on a fixed spacetime identical to the asymptotic boundary of the spacetime dual. Thus, the change in $\overline{B}$ is solely due to changes in the bulk metric despite the appearance to the contrary.

Again, the limit where all entanglement is removed between the quantum subsystems does not exactly reproduce the expectation that the tensor product of two pure states is dual to a disconnected spacetime. The two regions $B$ and $\overline{B}$ remain connected by a single singular point that has no clear interpretation. Nevertheless, it evident how entanglement between $Q_B$ and $Q_{\overline{B}}$ is related to the connectivity between $B$ and $\overline{B}$ in the spacetime dual. Further support for this is found in an argument from the mutual information$^2$ between a point in $B$ and one in $\overline{B}$ to the effect that the proper distance in the dual spacetime between any two such points goes to infinity when the entanglement between $Q_B$ and $Q_{\overline{B}}$ goes to zero [25]. As summarised by van Raamsdonk, “the two regions of spacetime pull apart and pinch off from each other” [1, 2327]. In other words, the conclusion from section 7.1 extends even to spacetimes with a contiguous boundary.

If the AdS side and the CFT side in the AdS/CFT correspondence are regarded as two different ways to encode the same information about a physical system, then the above qualitative investigation of the relation between spacetime and entanglement in the AdS/CFT correspondence offers an indication of how some of this information is encoded in the two different descriptions. The aspect of the physical system described on the AdS side as spacetime geometry is encoded on the CFT side as entanglement.

$^2$Schematically, the mutual information, $I(A, B)$, can be defined as $I(A, B) = S(A) + S(B) - S(A \cup B)$ where $S(M)$ is the entanglement entropy between a region $M$ and the rest of the system [1].
structure. It is, however, important to notice that the argument merely demonstrates how entanglement on the CFT side is a necessary condition for spacetime connectivity on the AdS side. Indeed, bulk entanglement corrections to the Ryu-Takayanagi formula entails that entanglement is not sufficient for spacetime connectivity.
Chapter 8

Linearized Gravity from Entanglement

If the entanglement structure indeed can represent the same physics as bulk spacetime, there must be constraints on entanglement that corresponds to Einstein’s field equations. Finding such correspondences would provide more quantitative justification for the claim that the same physical phenomena can indeed be encoded either as entanglement structure on the CFT side or as bulk spacetime.

In the following, it will be shown that if perturbations of the CFT vacuum state satisfies a particular constraint on entanglement entropy then the spacetime dual of the perturbed state satisfies Einstein’s field equations expanded to linear order around pure AdS. Thus, there is a general constraint on entanglement entropy in the CFT state that corresponds to imposing Einstein’s field equations in the bulk of the spacetime dual.

The constraint imposed on the CFT side is the following:

$$dS = dE$$  \hspace{1cm} (8.1)

where $S$ is the entanglement entropy of a region in a CFT and $E$ is some energy associated with that region.

(8.1) is similar to the first law of thermodynamics and a derivation of Einstein equations from this relation sits well with previous attempts to derive these equations from thermodynamics [26]. However, contrary to the first law of thermodynamics (8.1) holds for arbitrary perturbations of the vacuum state and not only for thermal/equilibrium states as it is the case for the first law of thermodynamics.

To derive Einstein equation from (8.1), one must establish a translation between
this relation in the holographic CFT to quantities in the spacetime dual. For the entanglement entropy (LHS of (8.1)), this translation is obtained by the Ryu-Takayanagi formula:

$$S = \frac{A(\tilde{B})}{4G_N}.$$  \hspace{1cm} (8.2)

The change in energy (RHS of (8.1)) can be translated as an energy associated with the asymptotic behavior of the metric. This, together with the Ryu-Takayanagi formula allows the translation of (8.1) holding in the CFT into a relation that must hold in the spacetime dual, which turn out to be the linearized Einstein equations in vacuum.

The material of this chapter is based on [3, 4, 27]. No new results are derived here, however, a number of details are filled in which are not explicit elsewhere in the literature.

As in section 4.3, we will consider a one parameter family of CFT states, $|\Psi(\zeta)\rangle$ and assume that $|\Psi(\zeta)\rangle$ lives in a $d$-dimensional Minkowski spacetime, $R^{d-1,1}$, such that $|\Psi(0)\rangle$ has as its spacetime dual the $(d+1)$-dimensional pure Poincaré patch of Anti-de-Sitter spacetime. The metric of the dual spacetime therefore takes the form

$$ds^2 = \frac{L^2}{z^2} (dz^2 + dx_\mu dx^\mu)$$ \hspace{1cm} (8.3)

Again, the dual spacetime of the quantum state $|\Psi\rangle$ will be denoted $M_\Psi$. Its boundary will be denoted $\partial M_\Psi$ and arbitrary Cauchy surfaces on this boundary will be denoted $\Sigma_{\partial M_\Psi} \subset \partial M_\Psi$.

Within this framework, we first want to demonstrate that

$$\frac{d}{d\zeta} S_B = \frac{d}{d\zeta} E_{B}^{Hyp}$$ \hspace{1cm} (8.4)

holds for all small perturbations of the CFT vacuum state i.e. for all $|\Psi(\zeta)\rangle$ with small $\zeta$. Here, $S_B$ is the entanglement entropy of $|\Psi(\zeta)\rangle$ for a ball shaped region $B \subset \Sigma_{\partial M_\Psi}$ and $\frac{d}{d\zeta} S_B$ is the variation of this entropy compared to $|\Psi(0)\rangle$. $E_{B}^{Hyp}$ is the so-called hyperbolic energy of the region $B$ and $\frac{d}{d\zeta} E_{B}^{Hyp}$ is the variation of this energy with respect to $|\Psi(0)\rangle$.

Secondly, we want to demonstrate that if (8.4) holds for a quantum state then linearized Einstein equations hold in the spacetime dual of this quantum state. To do this, we will provide a holographic interpretation of $S_B$ and $E_{B}^{Hyp}$ which both translate as integrations over (parts of) $M(\zeta)$ in terms of the metric perturbation $h(x, z)$. Requiring an equality to first order between the holographic expressions of $\frac{d}{d\zeta} S_B$ and $\frac{d}{d\zeta} E_{B}^{Hyp}$ will then be shown to be equivalent to imposing Einstein’s equations to first order on $h(z, x)$.
In the following, it will be assumed that all states, $|\Psi(\zeta)\rangle$, are small perturbations of $|\Psi(0)\rangle$ and that their spacetime duals, $\mathcal{M}(\zeta)$, consequently are small perturbations of $M(0)$, i.e. of the Poincarè patch of pure AdS.

### 8.1 Entropy and the modular Hamiltonian

The entanglement entropy of a spatial region in a CFT is defined via the reduced density matrix. Again, we define for some CFT state $|\Psi\rangle$ the reduced density matrix associated with an arbitrary spatial region $B$ as:

$$
\rho_B = \text{tr}_B(|\Psi\rangle \langle \Psi|) \quad (8.5)
$$

where $\overline{B}$ is the compliment of $B$. The von Neumann entropy is

$$
S_B = -\text{tr}(\rho_B \log(\rho_B)) \quad (8.6)
$$

Now, consider the perturbed state $|\Psi(\zeta)\rangle$ such that $S_B$ is the entanglement entropy in the region $B$. Varying this with respect to $\zeta$, we find (to any order in $\zeta$)

$$
\frac{d}{d\zeta}S_B = \frac{d}{d\zeta} - \text{tr}(\rho_B \log(\rho_B)))
$$

$$
= -\text{tr} \left( \log(\rho_B) \frac{d}{d\zeta} \rho_B \right) - \text{tr} \left( \rho_B \frac{d}{d\zeta} \log(\rho_B) \right)
$$

$$
= -\text{tr} \left( \log(\rho_B) \frac{d}{d\zeta} \rho_B \right) - \text{tr} \left( \rho_B \frac{1}{\rho_B} \frac{d}{d\zeta} \rho_B \right)
$$

$$
= -\text{tr} \left( \log(\rho_B) \frac{d}{d\zeta} \rho_B \right) - \text{tr} \left( \frac{d}{d\zeta} \rho_B \right)
$$

$$
= -\text{tr} \left( \log(\rho_B) \frac{d}{d\zeta} \rho_B \right)
$$

where last line follows since the trace of density matrix must be unity.\(^1\) Defining the modular Hamiltonian of the region $B$: $H_B \equiv -\log(\rho_B(\zeta = 0))$ we get

$$
\frac{d}{d\zeta}S_B = \text{tr} \left( H_B \frac{d}{d\zeta} \rho_B \right) \quad (8.7)
$$

Here $H_B$ is independent of $\zeta$ and therefore we can take the derivative outside the trace

$$
\frac{d}{d\zeta} \text{tr}(H_B \rho_B) = \frac{d}{d\zeta} \langle H_B \rangle \quad (8.8)
$$

\(^1\)It has fixed normalization and therefore $\text{tr} \left( \frac{d}{d\zeta} \rho_B \right) = \frac{d}{d\zeta} \text{tr}(\rho_B) = 0.$
where the equality follows from the definition of the expectation value $\langle A \rangle = \text{tr}(\rho A)$.

Thereby, we obtain the general relation that holds for an arbitrary region $B \subset \Sigma_{\partial M}$ and for any perturbation:

$$\frac{d}{d\zeta} S_B = \frac{d}{d\zeta} \langle H_B \rangle$$  \hspace{1cm} (8.10)

Generally, the modular Hamiltonian, $H_B$, is not a local operator. Thus, it cannot in general be evaluated from the local fields of a quantum field theory. However, there are certain exception where $H_B$ is a local operator. One example is the special case where $B$ is a ball shaped region with radius $R$.\footnote{I.e. a particular spatial, co-dimension two surface of $R^{d-1,1}$.}

The domain of dependence for this region is defined as usual as consisting of those points for which all causal curves that passes through the point and also passes through $B$.

As shown by [22], there is a conformal mapping of this domain of dependence for a ball shaped region in Minkowski spacetime to a Rindler wedge of Minkowski spacetime. This is the part of Minkowski spacetime that is accessible for a uniformly boosted observer. Identifying $x = r \cosh(\eta)$ and $t = r \sinh(\eta)$, the Rindler wedge has the metric

$$ds^2 = dr^2 - r^2 d\eta^2$$  \hspace{1cm} (8.11)

where $r > 0$.

Generally, it holds that the density matrix for a half space (like the Rindler wedge) is thermal for the vacuum state of a Lorentz-invariant quantum field theory.\footnote{Unruh radiation is a well known effect of this.} To see this, consider first the Wick rotated (Euclidean) statement of the path integral for a quantum field theory where $t = i\tau$:

$$\langle \phi_1(\tau_1)|e^{-\beta H}|\phi_0(\tau_0) \rangle = N \int_{\phi(\tau_0)=\phi_0}^{\phi(\tau_1)=\phi_1} D\phi \ e^{-S_{\text{Eucl}}(\phi(\tau))}$$  \hspace{1cm} (8.12)

where $\phi_1$ and $\phi_0$ are respectively final and initial field configurations. This may be identified as density matrix of a thermal state $e^{-\beta H}/Z$ under the assumption $Z = N$ and $\tau_0 = 0$ and $\tau_1 = \beta$

$$\langle \phi_1|e^{-\beta H}|\phi_0 \rangle = N \int_{\phi(0)=\phi_0}^{\phi(\beta)=\phi_1} D\phi \ e^{-S_{\text{Eucl}}(\phi(\tau))}.$$  \hspace{1cm} (8.13)

The partition function, i.e the trace of the density matrix $e^{-\beta H}$, can be expressed as

$$Z(\beta) = \text{tr} \left( e^{-\beta H} \right) = \int_{\tau} \langle \phi(\tau)|e^{-\beta H}|\phi(\tau) \rangle.$$  \hspace{1cm} (8.14)
This is equivalent to a path integral over all periodic paths with period $\beta$ which we express as

$$Z(\beta) = \oint_{\phi(0) = \phi(\beta)} \mathcal{D}\phi \ e^{-S^{Eucl}(\phi(\tau))}. \quad (8.15)$$

We can obtain the reduced density matrix for a region $B$ by tracing out the degrees of freedom in the compliment of $B$, $B^c$

$$\langle \phi^B_1 | e^{-\beta H} | \phi^B_0 \rangle = \frac{1}{Z} \int_{\phi^B(\beta) = \phi^B_0} \int_{\phi^B(0) = \phi^B_0} \mathcal{D}\phi \ e^{-S^{Eucl}(\phi(\tau))}. \quad (8.16)$$

Now, note that for some state $|\Psi\rangle$

$$\lim_{\beta \to \infty} e^{-\beta H} |\Psi\rangle = N |E_0\rangle \quad (8.17)$$

where $|E_0\rangle$ is the vacuum state. This follows since all energy eigenstates other than the vacuum state are suppressed by a factor $e^{-\beta(E-E_0)}$. This entails that

$$\langle \phi^B_0 | NE_0 \rangle = \langle \phi_0 | \lim_{\beta \to \infty} e^{-\beta H} |\Psi\rangle = N \int_{\phi^B(\infty)} \mathcal{D}\phi \ e^{-S^{Eucl}(\phi(\tau))} \quad (8.18)$$

We then have

$$\langle \phi^B_1 | \rho^R_{vac} | \phi^B_0 \rangle = \langle \phi^B_1 | NE_0 \rangle \langle \phi^B_0 | NE_0 \rangle^* = N^2 \int_{\phi^B(0^-) = \phi^B_0} \mathcal{D}\phi \ e^{-S^{Eucl}(\phi(\tau))} \quad (8.19)$$

which follows since

$$\langle \phi^B_0 | NE_0 \rangle^* = \langle \Psi | \lim_{\beta \to \infty} e^{\beta H} |\phi_0\rangle \quad (8.20)$$

Consider the special case where we are interested in the density matrix for the Rindler wedge of the Minkowski vacuum, i.e. a half-space of a particular QFT vacuum state. First, define the region $R$ of Minkowski spacetime by $x^1 > 0$. Thus $R$ is a half-space.

We then have

$$\langle \phi^R_1 | \rho^R_{vac} | \phi^R_0 \rangle = N^2 \int_{\phi^R(\theta=2\pi) = \phi^R_0} \mathcal{D}\phi \ e^{-S^{Eucl}(\phi(\tau))} \quad (8.21)$$

where $\tau$ is still the wick rotated time coordinate. Change the coordinates to polar coordinates, $x^1 = r \cos(\theta)$ and $\tau = r \sin(\theta)$ such that the metric may be expressed as $ds^2 = dr^2 + r^2 d\theta^2$ (where we have suppressed the remaining direction $x^{i \neq 1}$). We then have

$$\langle \phi^R_1 | \rho^R_{vac} | \phi^R_0 \rangle = N^2 \int_{\phi^R(\theta=2\pi) = \phi^R_0} \mathcal{D}\phi \ e^{-S^{Eucl}(\phi(\tau))}. \quad (8.22)$$
Comparing to (8.16), we identify this as a density matrix over $Z = \mathbb{N}$ with $\beta = 2\pi$ and Hamiltonian, $H_\eta$, that generates evolution in the variable $\eta = i\theta$, i.e.

$$\langle \phi^R_1 | \rho^\text{vac}_R | \phi^R_0 \rangle =$$ (8.23)

In the coordinates, $(r, \eta)$, the metric takes the form $ds^2 = dr^2 - r^2 d\eta^2$ which we recognize as the metric of the Rindler wedge in Minkowski spacetime. $H_\eta$, therefore, is the generator of Rindler time (the boost generator of Minkowski spacetime) and $e^{-2\pi H_\eta}$ is a thermal density matrix of the Rindler wedge. Thus, combining the results (8.22) and (8.23), we find that reduced vacuum density matrix for the half space $R$ is equal to the thermal density matrix defined in terms of the boost generator with $\beta = 2\pi$, i.e.

$$\rho^\text{vac}_R = \frac{1}{Z} e^{-\beta H_\eta}.$$ (8.24)

In the Rindler wedge, $\eta$ is the time coordinate and classical time translations are generated by the vector field

$$d_\eta = xd_t + td_x.$$ (8.25)

In Minkowski space, time translations are simply generated by the vector field $d_t$ and the Hamiltonian, therefore, can be obtained from the energy density operator for Minkowski spacetime, $T_{tt}$, by

$$H = \int d^{d-1}x T_{tt}.$$ (8.26)

By the same procedure, one can obtain the Hamiltonian for the Rindler wedge, $H_\eta$, however, its expression as a sum of vector fields complicates things. The complication can be avoided by noting that the Hamiltonian must be the same for any Cauchy surface of the spacetime, we can therefore choose $t = 0$ such that the second term in (8.25) is zero. Thereby, $H_\eta$ can be expressed in terms of the full Minkowski spacetime energy density operator $T_{tt}$,

$$H_\eta = \int_{x > 0} d^{d-1}x \langle xT_{tt} \rangle$$ (8.27)

where $x > 0$ is required to ensure that we only include the right Rindler wedge. With this, we can obtain an expression for the density matrix for the ball shaped region using the inverse mapping from the Rindler wedge to the domain of dependence for $B$. Since we are considering a CFT state, the energy eigenstates, including the vacuum state, are invariant under this conformal mapping. Associating with this mapping the transformation $U$, on finds

$$\rho_B = U^\dagger \rho_\eta U$$

$$= \frac{1}{Z} e^{-2\pi U^\dagger H_\eta U}$$ (8.28)

70
where the last equality follows from (8.24). Now define $H_{B}^{\text{vac}} \equiv 2\pi \mathcal{U} H_{\eta} \mathcal{U} - \log(Z)$ such that $\rho_{B} = e^{-H_{B}^{\text{vac}}}$. From this it immediately follows that $H_{B}^{\text{vac}} = -\log(\rho_{B})$ and we may therefore identity $H_{B}^{\text{vac}}$ as the modular Hamiltonian and observe that (8.28) yields a direct expression for the modular Hamiltonian of a ball shaped region in terms of the known Hamiltonian for the Rindler wedge, $H_{\eta}$ and the transformation $\mathcal{U}$ from the Rindler wedge to the domain of dependence for a ball shaped region.

Using this, one finds

$$H_{B}^{\text{vac}} = 2\pi \int_{B} d^{d-1}x \frac{R^{2} - (\vec{x} - \vec{x}_{0})^{2}}{2R} T^{tt}(x)$$  \hspace{1cm} (8.29)$$

where, $T^{tt}$ is the energy density operator for the CFT and $(\vec{x} - \vec{x}_{0})^{2}$ serves as a squared radial coordinate centred in the center of the ball. For further details, see \cite{4, 22}.

We can then obtain the sought expression for the first order variation in $\zeta$ away from vacuum of the expectation value of modular Hamiltonian

$$\frac{d}{d\zeta} \bigg|_{\zeta=0} \langle H_{B} \rangle = 2\pi \int_{B} d^{d-1}x \frac{R^{2} - (\vec{x} - x_{0})^{2}}{2R} \frac{d}{d\zeta} \bigg|_{\zeta=0} \langle T^{tt}(x) \rangle$$  \hspace{1cm} (8.30)$$

Finally, making an identification between the expectation value of the modular Hamiltonian and the hyperbolic energy, $\langle H_{B} \rangle \equiv E_{B}^{\text{hyp}}$, the relation

$$\frac{d}{d\zeta} S_{B} = \frac{d}{d\zeta} E_{B}^{\text{hyp}}$$  \hspace{1cm} (8.31)$$

is obtained by substitution into (8.10).

### 8.2 Holographic interpretation of $\delta S_{B}$

To see what corresponds to

$$\frac{d}{d\zeta} S_{B} = \frac{d}{d\zeta} E_{B}^{\text{hyp}}$$  \hspace{1cm} (8.32)$$

in the spacetime dual, we must find an interpretation of $\frac{d}{d\zeta} S_{B}$ and $\frac{d}{d\zeta} E_{B}^{\text{hyp}}$ on the AdS side.

A holographic interpretation of entanglement entropy is already provided by the Ryu-Takayanagi formula (8.2). This involves the evaluation of the area of a minimal co-dimension two surface in the spacetime bulk that divides the boundary region $B$ and its compliment.
Generally, one can parametrize a co-dimension two surface as $X^a(\sigma)$ where $X$ is an embedding function. The area functional, $A(g, X)$, for such surfaces can then be expressed as

$$A(g, X) = \int d^{d-1}\sigma \sqrt{\det(\gamma_{\mu\nu})}$$  \hspace{1cm} (8.33)$$

where $\det(\gamma_{\mu\nu})$ is the determinant of the metric induced on this surface by the embedding function $X^a$ such that

$$\gamma_{\mu\nu} = g_{ab}(x) \frac{\partial X^a}{\partial \sigma^\mu} \frac{\partial X^b}{\partial \sigma^\nu}$$  \hspace{1cm} (8.34)$$

We will use the notation $A(g, X_{\text{ext}})$ for the extremal of the area functional, where $X_{\text{ext}}^a(\sigma)$ is the embedding function which extremizes the area.

Now, consider a small perturbation of the metric $g_{ab}(x)$

$$g_{ab} = g_{ab}^0 + \delta g_{ab}$$  \hspace{1cm} (8.35)$$

that will entail a small variation of the spacetime. In the original spacetime the area, $A$, was extremized by $X^0_{\text{ext}}$ but with the variation in spacetime due to the variation in $g$ another surface with embedding function, $X_{\text{ext}}$, will extremize $A$.

The variation in extremal surface area is then

$$\delta A(g, X_{\text{ext}}) = \frac{\delta A(g, X_{\text{ext}}^0)}{\delta g} \delta g + \frac{\delta A(g^0, X_{\text{ext}}) \delta X}{\delta g} \delta g$$  \hspace{1cm} (8.36)$$

Note that $X_{\text{ext}}^0$ extremizes the functional $A(g^0, X)$. This entails that

$$A(g^0, X_{\text{ext}}) = A(g^0, X_{\text{ext}}^0 + \delta X) = A(g^0, X_{\text{ext}}^0) + \mathcal{O}(\delta X^2)$$  \hspace{1cm} (8.37)$$

i.e. the first order variations of the unperturbed extremal surface with respect to variation in the embedding function vanishes.\(^4\) The second term in (8.36) can therefore be rewritten as

$$\frac{\delta A(g^0, X_{\text{ext}})}{\delta X} = \frac{\delta A(g^0, X_{\text{ext}}^0)}{\delta X} + \frac{\mathcal{O}(\delta X^2)}{\delta X} = \mathcal{O}(\delta X)$$  \hspace{1cm} (8.38)$$

We then observe that $\delta X$ is of order $\delta g$ from which it follows

$$\mathcal{O}(\delta X) \frac{\delta X}{\delta g} \delta g = \mathcal{O}(\delta g^2)$$  \hspace{1cm} (8.39)$$

We therefore find that

$$\delta A(g, X_{\text{ext}}) = \frac{\delta A(g, X_{\text{ext}}^0)}{\delta g} \delta g + \mathcal{O}(\delta g^2)$$  \hspace{1cm} (8.40)$$

\(^4\)For an extremum, the first order variation always vanishes.
Thus, to linear order in $\delta g$, the variation in $A(g, X_{\text{ext}})$ depends on the variation in $g$ while the embedding function, $X$, is held fixed.

The variation of $A(g, X_{\text{ext}})$ is controlled by the variation in the induced metric. To first order in $\delta g$, the variation in the induced metric is proportional to the variation in the metric since the first order variation of the embedding function vanishes. Thus, reintroducing the perturbation parameter $\zeta$, we can express the perturbation of the induced metric as

$$\frac{d}{d\zeta} \bigg|_{\zeta=0} \gamma(\zeta) = \delta \gamma$$

(8.41)

where we have suppressed the tensor notation and the dependence of $\gamma$ on $x$.

The variation to first order in $\zeta$ (i.e. to first order in $\delta g$) of the square-root of the induced metric, $\det(\gamma(\zeta))$, entering in the area functional (8.33) may then be expressed as

$$\frac{d}{d\zeta} \bigg|_{\zeta=0} \sqrt{\det(\gamma(\zeta))} = \frac{1}{2\sqrt{\det(\gamma(\zeta))}} \frac{d}{d\zeta} \bigg|_{\zeta=0} e^{\text{tr}(\ln(\gamma(\zeta)))}$$

(8.42)

where we have used the identity $\det(M) = \exp(\text{tr}(\ln(M)))$. The derivative becomes

$$\frac{d}{d\zeta} \bigg|_{\zeta=0} e^{\text{tr}(\ln(\gamma(\zeta)))} = \det(\gamma(0)) \frac{d}{d\zeta} \bigg|_{\zeta=0} \text{tr}(\ln(\gamma(\zeta)))$$

$$= \det(\gamma(0)) \text{tr}\left( \frac{d}{d\zeta} \bigg|_{\zeta=0} \ln(\gamma(\zeta)) \right)$$

$$= \det(\gamma(0)) \text{tr}\left( \gamma(0) \frac{d}{d\zeta} \bigg|_{\zeta=0} \gamma(\zeta) \right)$$

$$= \det(\gamma(0)) \text{tr}(\gamma(0) \delta \gamma)$$

$$= \det(\gamma(0)) \delta \gamma$$

(8.43)

where the last line reintroduces the tensor notation and replaces the trace with a sum over repeated indices. Thus we find the following expression for the variation of the square root

$$\frac{d}{d\zeta} \bigg|_{\zeta=0} \sqrt{\det(\gamma(\zeta))} = \frac{1}{2\sqrt{\det(\gamma(0))}} \det(\gamma(0)) \gamma^{\rho\lambda}_0 \delta \gamma_{\rho\lambda}$$

(8.44)

We can then insert this into the area functional, and get an explicit expression of the variation of the extremum of the area functional due to a first order variation in the metric

$$\delta A(g, X_{\text{ext}}) = \int d^{d-1}x \frac{1}{2} \sqrt{\det(\gamma(0))} \gamma^{\rho\lambda}_0 \delta \gamma_{\rho\lambda}.$$  

(8.45)

73
Now, we want again to consider the particular example where the perturbed metric $g_{ab}$ is given by

$$g_{ab} = g_{ab}^{\text{AdS}} + z^{d-2} h_{ab}. \quad (8.46)$$

The embedding function $X_{\text{ext}}$ of interest is the mapping from the co-dimension two surface, $\tilde{B}$, to the spacetime with metric $g_{ab}$, which extremizes the area of $\tilde{B}$, whose boundery at $z = 0$ is a ball shaped region, $B \subset \Sigma_{\partial M}$, with radius $R$ and which is defined on the constant time slice $t = 0$. In pure AdS, the bulk surface that extremizes the area of $\tilde{B}$ for a ball shaped boundary is

$$\tilde{x}^2 + z^2 = x_i x^i + z^2 = R^2 \quad (8.47)$$

where $x_i$ are the spatial Minkowski coordinates, i.e. the coordinates on the boundary $\Sigma_{\partial M}$.

Parametrizing this extremal surface using these boundary coordinates, such that $(t, x^i) = \sigma^\mu$, one finds the embedding function

$$X^a_{\text{ext}}(t, x^i) : \mathbb{R}^{d-1} \rightarrow \mathbb{R}^{d+1}, (t, x^i) \mapsto \begin{cases} X^0 = t_0 \\ X^1 = x^1 \\ \vdots \\ X^{d-1} = x^{d-1} \\ X^d = Z = \sqrt{R^2 - \tilde{x}^2} \end{cases} \quad (8.48)$$

Since we have chosen a constant time slice, we may disregard the time coordinate and consider only the spatial coordinates, $x^i$. The unperturbed induced metric, $\gamma^0_{ij}$, then takes the form

$$\gamma^0_{ij} = g_{ab} \frac{\partial X^a_{\text{ext}}}{\partial x^i} \frac{\partial X^b_{\text{ext}}}{\partial x^j} = \frac{L^2}{z^2} \sum_{k=1}^{d-1} \left( \frac{\partial X^k_{\text{ext}}}{\partial x^i} \frac{\partial X^k_{\text{ext}}}{\partial x^j} + \frac{\partial Z}{\partial x^i} \frac{\partial Z}{\partial x^j} \right). \quad (8.49)$$

where the last equality follows since $\partial t_0 / \partial t = 0$ and $g^0_{ab} = g_{ab}^{\text{AdS}}$ and therefore is diagonal with Lorentzian signature and elements $L^2 / z^2$. Now

$$\sum_{k=1}^{d-1} \frac{\partial X^k_{\text{ext}}}{\partial x^i} \frac{\partial X^k_{\text{ext}}}{\partial x^j} = \delta_{ij} \quad (8.50)$$

and

$$\frac{\partial Z}{\partial x^i} \frac{\partial Z}{\partial x^j} = \frac{x_i x_j}{z^2} \quad (8.51)$$
so we find that
\[
\gamma_{ij}^0 = \frac{L^2}{z^2} \left( \delta_{ij} + \frac{x_i x_j}{z^2} \right)
\] (8.52)
and
\[
\gamma_{0j}^j = \frac{z^2}{L^2} \left( \delta_{ij} - \frac{x^i x^j}{R^2} \right).
\] (8.53)
This can be verified using the relation
\[
\vec{x}^2 + z^2 = R^2 \Rightarrow \vec{x}^2 + z^2 + 1 = \frac{R^2}{z^2}
\] (8.54)
and the identity \(\gamma_{ij} \gamma^{jk} = \delta^k_i\):
\[
\gamma_{ij}^0 \gamma^{jk}_0 = \frac{L^2}{z^2} \left( \delta_{ij} + \frac{x_i x_j}{z^2} \right) \frac{z^2}{L^2} \left( \delta^{jk} - \frac{x^j x^k}{R^2} \right)
\]
\[
= \delta_{ij} \delta^{jk} + \delta^{ij} \frac{x_i x_j}{z^2} - \delta_{ij} \frac{x^j x^k}{R^2} - \frac{x_i x_j x^j x^k}{z^2 \ R^2}
\]
\[
= \delta_{ik}^k + \frac{x_i x^k}{z^2} - \frac{x_i x^k}{R^2} - \frac{x_i x_j x^j x^k}{z^2 \ R^2}
\]
\[
= \delta_{ik}^k + \frac{x_i x^k}{z^2} - \frac{x_i x^k}{R^2} - \frac{x_i x^k}{z^2} - \frac{x_i x^k}{R^2}
\]
\[
= \delta_{ik}^k
\] (8.55)
The square root of the determinant of the induced metric, \(\sqrt{\gamma} \), becomes
\[
\sqrt{\gamma^0} = \frac{L^{d-1}}{z^{d-1}} \sqrt{1 + \frac{\vec{x}^2}{z^2}} = \frac{L^{d-1}}{z^d} R
\] (8.56)
where the last equality follows from (8.54). The first order variation of the induced metric, \(\delta \gamma_{ij}\), is then given as
\[
\delta \gamma_{ij} = \delta g_{ab} \frac{\partial X^a_{\text{ext}}}{\partial x^i} \frac{\partial X^b_{\text{ext}}}{\partial x^j} = z^{d-2} h_{ab} \frac{\partial X^a_{\text{ext}}}{\partial x^i} \frac{\partial X^b_{\text{ext}}}{\partial x^j}
\]
\[
= z^{d-2} \left( h_{mn} \frac{\partial X^m_{\text{ext}}}{\partial x^i} \frac{\partial X^n_{\text{ext}}}{\partial x^j} + h_{z\mu} (\ldots) + h_{zz} (\ldots) \right)
\] (8.57)
where we have used the assumption (radial gauge) \(h_{z\mu} = h_{zz} = 0\) and that the embedding function, \(X^a_{\text{ext}}(x_i)\), is invariant under the first order variation of the the metric.
All of this can then be combined to obtain an expression for the first order variation around pure AdS in the area of the extremal surface with a ball shaped boundary

\[
\delta A^B_{\text{AdS}} = \int d^{d-1}x \frac{1}{2} \sqrt{\text{det} (\gamma_{\mu\nu}^0)} \gamma_0^{\rho\lambda} \delta \gamma_{\rho\lambda}
\]

\[
= \int d^{d-1}x \frac{1}{2} \left[ \frac{L^{d-1}}{z^d R} \right] \left[ \frac{z^2}{L^2} \left( \delta^{ij} - \frac{x^i x^j}{R^2} \right) \right] (z^{d-2} h_{ij})
\]

\[
= \frac{L^{d-3} R}{2} \int d^{d-1}x \left( \delta^{ij} - \frac{1}{R^2} (x^i - x^i_0)(x^j - x^j_0) \right) h_{ij}
\]

(8.58)

where the last line assumes that the center of the ball shaped region is \(x_0\). Using the Ryu-Takayanagi formula (5.4) we find

\[
\delta S_B = \frac{\delta A_{\text{ads}}}{4G_N} = L^{d-3} R \int_{|\vec{x} - \vec{x}_0| \leq R} d^{d-1}x \left( \delta^{ij} - \frac{1}{R^2} (x^i - x^i_0)(x^j - x^j_0) \right) h_{ij}
\]

(8.59)

where

\[
\delta S_B \equiv \frac{d}{d\zeta} \bigg|_{\zeta=0} S_B.
\]

(8.60)

Thus, for CFT state |\(\Psi\rangle\) that has pure AdS as its spacetime dual, a small variation in the entanglement entropy, \(\delta S_B\), of a ball shaped region \(B\) is proportional to the change in area of the extremal surface \(\tilde{B}\), \(\delta A_{\text{AdS}}\), due to the corresponding perturbation of the AdS metric.

### 8.3 Holographic interpretation of \(\delta E_B^{Hyp}\)

From

\[
H_B^{\text{vac}} = 2\pi \int_B d^{d-1}x \frac{R^2 - (\vec{x} - \vec{x}_0)^2}{2R} T^{tt}(x)
\]

we see that a holographic interpretation of the hyperbolic energy, \(E_B^{Hyp}\), merely requires an interpretation of the energy density operator, \(T^{tt}\), for the CFT.

To find this holographic interpretation, consider first an infinitesimal ball shaped region, \(B_{inf}\), on the boundary \(\Sigma_{\partial M}\) centred at \(x_0\). For such a infinitesimal ball, the expectation value of the energy density operator can be assumed to be the same throughout the ball. To leading order, we can therefore replace the function \(\frac{d}{d\zeta} \bigg|_{\zeta=0} \langle T^{tt}(x) \rangle\) with
its central value \( \frac{d}{d\zeta} \bigg|_{\zeta=0} \langle T^{tt}(x_0) \rangle \), where \( x_0 \) is the center. The variation in hyperbolic energy for such an infinitesimal ball, \( \frac{d}{d\zeta} \bigg|_{\zeta=0} E^{hyp}_{B_{m,f}} \), is given by (8.30) when taking the limit \( R \to 0 \)

\[
\frac{d}{d\zeta} \bigg|_{\zeta=0} E^{hyp}_{B_{m,f}} = \lim_{R \to 0} \left\{ 2\pi \int_{B} d^{d-1}x \frac{R^2 - (\bar{x} - \bar{x}_0)^2}{2R} \frac{d}{d\zeta} \bigg|_{\zeta=0} \langle T^{tt}(x) \rangle \right\} = \pi \frac{d}{d\zeta} \bigg|_{\zeta=0} \langle T^{tt}(x_0) \rangle
\]

\[
\Rightarrow \frac{2\pi R^d S_{d-2}}{d^2 - 1} \frac{d}{d\zeta} \bigg|_{\zeta=0} \langle T^{tt}(x_0) \rangle = \lim_{R \to 0} \delta S_B
\]

where \( S_{d-2} \) is the area of the unit ball of dimension \( d - 2 \) (surface area of the unit sphere in dimension \( d - 1 \)). From (8.4) it then follows that

\[
\frac{d}{d\zeta} \bigg|_{\zeta=0} E^{hyp}_{B_{m,f}} = \frac{d}{d\zeta} \bigg|_{\zeta=0} S_{B_{m,f}} = \lim_{R \to 0} \delta S_B
\]

Inserting the holographic interpretation for \( \delta S_B \) into (8.59) we find

\[
\frac{d}{d\zeta} \bigg|_{\zeta=0} \langle T^{tt}(x_0) \rangle = \frac{d^2 - 1}{2\pi S_{d-2}} \lim_{R \to 0} \left( \frac{1}{R^d} \delta S_B \right)
\]

To investigate the limit \( R \to 0 \), it is convenient to define coordinates that are constant in this limit

\[
\hat{x} = \frac{x^i - x_0^i}{R}, \quad \hat{z} = \frac{z}{R}
\]

With these coordinates we find the expression

\[
\delta S_B = \frac{L^{d-3}}{8G_N} \int_{\hat{x}^2 \leq 1} d^{d-1}\hat{x} R^{d-1} \langle \delta^{ij} - \hat{x}^i \hat{x}^j \rangle h_{ij}(x, z)
\]

and

\[
\frac{d}{d\zeta} \bigg|_{\zeta=0} \langle T^{tt}(x_0) \rangle = \frac{d^2 - 1}{2\pi S_{d-2}} \frac{L^{d-3}}{8G_N} \int_{\hat{x}^2 \leq 1} d^{d-1}\hat{x} (\delta^{ij} - \hat{x}^i \hat{x}^j) h_{ij}(x, z).
\]
The integrals are over symmetric intervals and therefore all terms where \( i \neq j \) will not contribute since
\[
\int_{-a}^{a} dx \int_{-b}^{b} dy \ (\delta_{xy} - xy) h_{xy}(x_0, z = 0) = h_{xy} \left( a^2 - (-a)^2 \right) \left( b^2 - (-b)^2 \right) = 0. \quad (8.68)
\]
We therefore have
\[
\frac{d}{d\zeta} \bigg|_{\zeta=0} \langle T^{tt}(x_0) \rangle = \frac{d^2 - 1}{2\pi S_{d-2}} \frac{L^{d-3}}{8G_N} h_{ii} \int_{\tilde{x}_1}^{\tilde{x}_2} d^{d-1} \hat{x} (1 - \hat{x}^2 \hat{x}^i). \quad (8.69)
\]
Since there is nothing particular about the numbering of the coordinates \( x^i \), the \( d - 1 \) integrals above must be equal
\[
\int_{\tilde{x}_2 \leq 1} d^{d-1} \hat{x} (1 - \hat{x}^2 \hat{x}^1) = \cdots = \int_{\tilde{x}_2 \leq 1} d^{d-1} \hat{x} (1 - \hat{x}^{d-1} \hat{x}^{d-1}) = S_{d-2} \frac{d}{d^2 - 1}. \quad (8.70)
\]
One thereby finds\(^5\)
\[
\frac{d}{d\zeta} \bigg|_{\zeta=0} \langle T^{tt}(x_0) \rangle = \frac{d^2 - 1}{2\pi S_{d-2}} \frac{L^{d-3}}{8G_N} \left( \frac{S_{d-2} d}{d^2 - 1} \delta_{ij} h_{ij} \right). \quad (8.71)
\]
Now nothing was assumed for the point \( x_0 \), so this holds for all \( x \). We can therefore substitute the first order variation of energy density, \( T^{tt}(x) \), in the CFT state with the metric perturbation at the boundary, \( \delta_{ij} h_{ij}(x) \) times some constant
\[
\frac{d}{d\zeta} \bigg|_{\zeta=0} \langle T^{tt}(x) \rangle \bigg|_{\delta E^H_{B}} = 2\pi \int_B d^{d-1} x \frac{R^2 - (\bar{x} - \bar{x}_0)^2}{2R} \frac{d}{d\zeta} \bigg|_{\zeta=0} \langle T^{tt}(x) \rangle
\]
\[
= \frac{L^{d-3} d}{16\pi G_N} \int_B d^{d-1} x \frac{R^2 - (\bar{x} - \bar{x}_0)^2}{R} \delta_{ij} h_{ij}(x, z = 0). \quad (8.72)
\]
This, thereby, provides a holographic interpretation of the hyperbolic energy for a ball shaped region on \( \Sigma_{M_\Psi} \). Below, we will define
\[
\frac{d}{d\zeta} \bigg|_{\zeta=0} E^H_{B} \equiv \delta E^H_{B} \quad (8.73)
\]
\(^5\)This is a rather remarkable result in its own right. Assuming only the Ryu-Takayanagi formula, we have derived the known relation \([28, 29]\) between the first order variation of the expectation value of the CFT stress energy tensor and the bulk metric perturbation \( h \).
8.4 Combining the holographic results

In the two previous sections we found holographic interpretations (and expressions) for \( \delta S_B \) and \( \delta E_{B}^{hyp} \). The relation \( \delta S_B = \delta E_{B}^{hyp} \) holds for any CFT quantum state. It will turn out that if and only if this equality holds for all disks in all Lorentz frames then the bulk metric satisfies the Einstein’s field equations to linear order in \( h \).

The two holographic interpretations are

\[
\delta E_{B}^{hyp} = \frac{L^{d-3}}{16 G_N R} \int_B d^{d-1}x (R^2 - (\vec{x} - \vec{x}_0)^2) \delta^{ij} h_{ij}(x, z = 0) \equiv \delta E_{B}^{Grav} \tag{8.74}
\]

and

\[
\delta S_B = \frac{L^{d-3}}{8 G_N R} \int_B d^{d-1}x \left( R^2 \delta^{ij} - (x^i x^j) \right) h_{ij} \equiv \delta S_{B}^{Grav} \tag{8.75}
\]

where we have introduced the shorthand \( \delta E_{B}^{Grav} \) and \( \delta S_{B}^{Grav} \) for the holographic interpretation of \( \delta E_{B}^{hyp} \) and \( \delta S_B \), respectively. Combining these and using the relation \( \delta S_B = \delta E_{B}^{hyp} \), we find

\[
\delta E_{B}^{Grav} = \delta S_{B}^{Grav}
\]

\[
\Rightarrow 2d \int_B d^{d-1}x \left( R^2 - (\vec{x} - \vec{x}_0)^2 \right) \delta^{ij} h_{ij}(x, z = 0) = \int_B d^{d-1}x \left( R^2 \delta^{ij} - x^i x^j \right) h_{ij}(x, z)
\]

(8.76)

a relation between the metric perturbation on the boundary and in the perturbation in the bulk. This bulk constraint was first shown to be equivalent to linearized Einstein equation in vacuum in [3], however, a more elegant derivation was developed in [4] inspired by the Iyer-Wald formalism [30]. Only the details of the latter derivation will be recounted here.

In [4], they show that the non-local constraint (8.76) is equal to the linearized Einstein equations in vacuum if there is a form, \( \chi_B(h) \), such that (off shell)

\[
\int_B \chi_B(h) = \delta E_{B}^{Grav}, \quad \int_B \chi_B(h) = \delta S_{B}^{Grav}
\]

(8.77)

and

\[
d\chi_B(h) = -\frac{1}{8\pi G_N} \xi^a_B G_{ab}^E \epsilon^b
\]

(8.78)

Here, \( d \) is the exterior derivative, \( G_{ab}^E \) are linearized Einstein equations, \( \xi^a_B \) is the Killing vector

\[
\xi_B = -\frac{2\pi}{R} (t - t_0) \left[ z \partial_z + (x^i - x_i^0) \partial_i \right] + \frac{\pi}{R} \left[ R^2 - z^2 - (x^i - x_i^0)^2 - (t - t_0)^2 \right] \partial_t.
\]

(8.79)
We see that all but the time component vanishes at time, $t = t_0$. Further, on the extremal surface $\tilde{B}$ that satisfies $(\vec{x} - \vec{x}_0)^2 + z^2 = R^2$ and $t = t_0$ also the time component, $\xi^t_B$, vanishes. Thus, $\tilde{B}$ is a bifurcation surface of the killing vector $\xi_B$. Finally, $\epsilon^b$ is the volume form of co-dimension one surfaces in the bulk

$$\epsilon^b = \frac{1}{d!} \epsilon^{bc_2 \ldots c_{d+1}} dx^c_2 \wedge \ldots \wedge dx^c_{d+1} \tag{8.80}$$

where $\epsilon_{a_1 \ldots a_{d+1}}$ is the antisymmetric tensor in $d + 1$ dimensions, for which we define $\epsilon_{xt^1 \ldots x^{d-1}} = \sqrt{-g}$. For a surface of constant time, we have $dt = 0$ and all component but $\epsilon^t$ vanishes. This follows since all terms of $\epsilon_{b \neq t}$ either include the antisymmetric tensor with repeated indices or wedge products that involve $dt$.

Defining the volume $\Pi$ as the volume bounded by $B \cup \tilde{B} \equiv \partial \Pi$, where $B$ and $\tilde{B}$ as above are spacelike surfaces at $t = t_0$. From that it follows

$$d\chi_B(h)|_\Pi = -\frac{1}{8\pi G_N} \xi^t_B G^E_{tt} \epsilon^t. \tag{8.81}$$

Furthermore, we require that

$$d\chi_B|_{\partial M} = 0 \tag{8.82}$$

where $\partial M$ is the boundary of AdS which entails that $d\chi_B|_B = 0$ since $B \subset \partial M$.

Assuming there is such a form, $\chi$, it can then be shown that if the non-local constraint $\delta S^\text{Grav}_B = \delta E^\text{Grav}_B$ is satisfied in the bulk, then Einstein’s field equations to linear order are satisfied in the bulk, i.e. the constraint implies that $\hat{G}^E_{ab} = 0$. First, relation $\delta S^\text{Grav}_B = \delta E^\text{Grav}_B$ implies that

$$0 = \delta S^\text{Grav}_B - \delta E^\text{Grav}_B = \int_B \chi_B - \int_B \chi_B = \int_{B \cup \tilde{B}} \chi_B = \int_{\Pi} d\chi_B \tag{8.83}$$

where the last equality follows from Stokes theorem. Next, we multiply by $R$, take the derivative with respect to $R$ and expand using Leibniz’ integration rule for forms [31]

$$0 = \frac{\partial}{\partial R} \int_{\Pi} R d\chi_B = \int_{\Pi} \tilde{v}_\partial (R d\chi_B) + \int_{\partial \Pi} \tilde{v}_\partial (R d\chi_B) + \int_{\Pi} \frac{\partial}{\partial R} R d\chi_B \tag{8.84}$$

where $\partial$ is the interior product and $\tilde{v}$ is a vector orthogonal to $R$, such that $\tilde{v}_\partial d\chi_B = d\chi_B(h) \cdot \hat{r}$, where $\hat{r}$ is a unit vector in the radial direction. It is an axiom of exterior derivatives that $d(d\omega) = 0$ for any differential form $d\omega$ of a smooth function $\omega$. Therefore, the first term in the expansion vanishes. Thus, inserting the expression (8.78) for $d\chi$ and using (8.81), we find

$$0 = -\frac{1}{8\pi G_N} \left( \int_{\partial \Pi} R \xi^t_B \hat{G}^E_{tt} \epsilon^t \cdot \hat{r} + \int_{\Pi} \frac{\partial}{\partial R} R \xi^t_B \hat{G}^E_{tt} \epsilon^t \right). \tag{8.85}$$
The first term vanishes, which can be seen by splitting the surface integral over \( \partial \Pi \) into an integral over \( B \) and one over \( \tilde{B} \) (valid since \( \partial \Pi = B \cup \tilde{B} \)). The integral over \( B \) vanishes due the requirement \( d\chi |_{\partial M} = 0 \). The integral over \( \tilde{B} \) vanishes since \( \tilde{B} \) is a bifurcation surface of \( \xi^t_B \) \( (\xi_B|_{\tilde{B}} = 0) \). Thus, only the second term above is non-vanishing. The derivative, restricted to the volume \( \Pi \) for which \( t = t_0 \), can be evaluated by inserting the time component of \( \xi^t_B \) as given in (8.79)

\[
\frac{\partial}{\partial R} R \xi^t_B \dot{G}_t^E e^t = \frac{\partial}{\partial R} R \frac{\pi}{R} \left[ R^2 - z^2 - (\vec{x} - \vec{x}_0)^2 \right] \dot{G}_t^E e^t = 2\pi R \dot{G}_t^E e^t. \tag{8.86}
\]

We then finally get

\[
0 = -\frac{R}{4G_N} \int_{\Pi} \dot{G}_t^E e^t \tag{8.87}
\]

which is satisfied only if \( \int_{\Pi} \dot{G}_t^E e^t = 0 \).

To see that this implies \( \dot{G}_t^E = 0 \), observe that (8.87) is of the form

\[
\int_{\Pi(R)} dA f(z, x) = \int_{\partial \Pi(R)} dA f(z, x)
\]

where \( dA \) is the area element on \( \partial \Pi(R) \). Expanding this integral, we find

\[
0 = \int_{\partial \Pi(R)} dA f(z, x) = \int_B dA f(z, x) + \int_{\tilde{B}} dA f(z, x) = \int_{\tilde{B}} dA f(z, x) \tag{8.89}
\]

where the last equality follows from (8.82). Now, we identify \( \int_{\tilde{B}} dA f(z, x) \) as the average moment of the function \( f(z, x) \) on \( \tilde{B} \) and find that this vanishes. Next, consider the derivative of \( \int_{\Pi(R)} d^{d-1} x f(z, x) \) with respect to \( x^i \):

\[
0 = \frac{\partial}{\partial x^i} \int_{\Pi(R)} d^{d-1} x f(z, x)
\]

\[
= \frac{\partial}{\partial R} \int_{\Pi(R)} d^{d-1} x f(z, x) \frac{\partial R}{\partial x^i}
\]

\[
= \int_{\partial \Pi(R)} dA x^i f(z, x)
\]

\[
= \int_{\tilde{B}} dA x^i f(z, x)
\]

where the last equality follows since \( 1/R \) only contributes with a constant factor. Again, we used Stokes theorem. We identify this as the first moment of \( f(x, z) \) on \( \tilde{B} \) in the \( x^i \) direction for vanishing average moment. As seen, this also vanishes.

Since the first order moment vanishes, it holds that

\[
0 = \int_{\Pi(R)} d^{d-1} x f(z, x) = \int_{\Pi(R)} d^{d-1} x f(z, x). \tag{8.91}
\]
Again, taking the derivative with respect to $x^i$ gives
\[
0 = \frac{\partial}{\partial x^i} \int_{\Pi(R)} d^{d-1}zd^i f(z, x) = \frac{\partial}{\partial R} \int_{\Pi(R)} d^{d-1}zd^i f(z, x) \frac{\partial R}{\partial x^i} = \int_B dA (x^i)^2 f(z, x). \tag{8.92}
\]
This, we recognize as the second moment of $f(x, z)$ on $\tilde{B}$ in the $x^i$ direction for vanishing average moment. Also this vanishes. We can continue this procedure and show that all moments of $f(x, z)$ on $\tilde{B}$ vanishes if $0 = \int_{\Pi(R)} dA f(z, x)$. Thus, $f(x, z) = 0$.

Going to our particular example $f(z, x) = -\frac{R}{4G_N} \cdot G_E(t)$ with
\[
0 = -\frac{R}{4G_N} \int_{\Pi} \dot{G}_E \epsilon^t \tag{8.93}
\]
it follows that $\dot{G}_E^{t} = 0$, since $R$, $G_N$ and $\epsilon^t$ are non-zero.

The argument so far has been carried assuming that the $\tilde{B}$ was a subset of a surface of constant time. Thus, we have assumed a frame of reference with four velocity, $u^\mu = (1, \vec{0})$. Going to a general frame of reference, one finds that the condition $\delta S_{Grav}^B = \delta E_{Grav}^B$ entails
\[
u^\mu u^\nu \dot{G}_E^{\mu \nu} = 0. \tag{8.94}
\]
As a crude test, we see immediately that this reproduces the result above for the four velocity $u^\mu = (1, \vec{0})$. If we assume that (8.94) holds for arbitrary four velocities $u^\mu$ then it follows that $\dot{G}_E^{\mu \nu} = 0$. Thus, the linearized Einstein equations are satisfied for the boundary coordinate directions. What remains is to show that $\dot{G}_E^{z \mu} = 0$ and $\dot{G}_E^{zz} = 0$. First note that $\dot{G}_E^{z \mu} = 0$ and $\dot{G}_E^{zz} = 0$ for $z = 0$ which follows from (8.82). These are analogous to initial value constraints that must be satisfied by $\dot{G}_E^{ab}$. Consider the general result of the initial value formulation of general relativity that if the spatial components of the Einstein equations are satisfied everywhere, then the initial value constraints given on some spatial slice of the spacetime are satisfied everywhere. This result follows from the Bianchi identity $\nabla^a G_E^{ab} = 0$ [32, 252-267]. However, this result may as well be applied to constraints at the surface $z = 0$. If the remaining components, $G_E^{\mu \nu}$, of the Einstein equations are satisfied in all of spacetime as well as the constraints at the surface $z = 0$ then these constraints are satisfied in all of spacetime. This entails that if $\dot{G}_E^{\mu \nu} = 0$ and $\dot{G}_E^{z \mu} = 0$ and $\dot{G}_E^{zz} = 0$ for $z = 0$ then $\dot{G}_E^{z \mu} = 0$ and $\dot{G}_E^{zz} = 0$ for all $z$. Thus, if $\delta S_{Grav}^B = \delta E_{Grav}^B$ is satisfied in the bulk and there is a form $\chi_B(h)$ that satisfies the stated constraints then it follows that linearized Einstein equations are satisfied in the bulk.
8.5 The form $\chi_B(h)$

What remains is to show that there is a form $\chi_B$ for which (8.77), (8.95), (8.78), and (8.82) holds. Let us make the ansatz

$$\chi_B = -\frac{1}{16\pi G_N} \left[ \frac{d}{d\zeta} \right]_{\zeta=0} \left( \nabla^a \xi^{b}_{B} \epsilon_{ab} + \xi^{b}_{B} \epsilon_{ab} (\nabla \delta g^{ac} - \nabla^a \delta g^{c}_{c}) \right)$$

(8.95)

where $\delta g^{c}_{c} = g^{bc} \delta g_{bc}$ and

$$\epsilon_{ab} = \frac{1}{(d-1)!} \epsilon_{abcd} \cdots \epsilon_{d+1} \right|_{d^c \cdots d^{d+1}} \nabla^a$$

(8.96)

This entails $\nabla_a \epsilon_{ab} = 0$ since the covariant derivative of the Levi-Civita tensor vanishes.

We may therefore expand the variation of $\nabla^a \xi^{b}_{B} \epsilon_{ab}$ with respect to the metric as

$$\frac{d}{d\zeta} \left|_{\zeta=0} \nabla^a \xi^{b}_{B} \epsilon_{ab} \right. = \frac{d}{d\zeta} \left|_{\zeta=0} \left( g^{ac} \nabla_c \xi^{b}_{B} \epsilon_{ab} \right) \right.$$  

(8.97)

where we have defined $\bar{\epsilon}_{ab} = \epsilon_{ab} / \sqrt{-g}$ and used

$$\bar{\epsilon}_{ab} \delta (\sqrt{-g}) = \frac{1}{2} \bar{\epsilon}_{ab} \sqrt{-g} g^{cd} \delta g_{cd} = \frac{1}{2} \epsilon_{ab} g^{cd} \delta g_{cd}. \quad (8.98)$$

Also, we have used the expression for the variation of the covariant derivative (6.17)

$$\delta \left( \nabla \xi^{b}_{B} \right) = \bar{\Gamma}^{b}_{abcd} \xi^{d}_{B} = \frac{1}{2} \epsilon_{ab} g^{cd} \left( \nabla_c \delta g_{de} + \nabla_d \delta g_{ce} - \nabla_e \delta g_{cd} \right) \xi^{d}_{B}$$

(8.99)

(8.97) may be reinserted into the expression for $\chi$:

$$\chi_B = -\frac{1}{16\pi G_N} \left[ \frac{1}{2} \epsilon_{ab} g^{ac} g^{de} \left( \nabla_c \delta g_{de} + \nabla_d \delta g_{ce} - \nabla_e \delta g_{cd} \right) \xi^{d}_{B} \right.$$  

(8.100)

$$+ \frac{1}{2} \epsilon_{ab} g^{ac} \delta g_{cd} \nabla_c \xi^{b}_{B} - \delta g^{ac} \epsilon_{ab} \nabla^c \xi^{b}_{B} + \xi^{b}_{B} \epsilon_{ab} (\nabla \delta g^{ac} - \nabla^a \delta g^{c}_{c}) \right]$$

We will consider $\chi_B$ restricted to the constant time surface $\Pi$, where all but the time component of $\xi_B$ vanishes. Thus, $\xi^{t}_{B} = \frac{2}{R} \left[ R^2 - z^2 - (x^4 - x^{d}_{0})^2 \right]$ which follows from the
expression (8.79). This entails that $\partial_z \xi^i_B = -\frac{2\pi z}{R}$ and $\partial_t \xi^i_B = -\delta_{ij} \frac{2\pi(x^j - x^j_0)}{R}$. One must also use that $\epsilon_{ij \mid \Pi} = \epsilon_{iz \mid \Pi} = \epsilon_{zz \mid \Pi} = 0$ and that $h_{it} = \delta^{ij} h_{ij}$. The latter can be verified by requiring that

$$D = g^{ab} g_{ab}$$

(8.101)

and

$$D = (g^{ab} + h^{ab})(g_{ab} + h_{ab}) = g^{ab} g_{ab} + 2g^{ab} h_{ab} + \mathcal{O}(h^2) \Rightarrow 2g^{ab} h_{ab} = -\mathcal{O}(h^2).$$

(8.102)

For small perturbation, the term $\mathcal{O}(h^2)$ vanishes from which it follows that $g^{ab} h_{ab} = 0$. We therefore have

$$0 = g^{ab} h_{ab} \Rightarrow h_{it} = \delta^{ij} h_{ij} + h_{zz}$$

(8.103)

where $h_{zz} = 0$ assuming radial gauge. Using these, we may evaluate $\chi_B$ on the surface $\Pi$:

$$\chi_B \mid \Pi = -\frac{1}{16\pi G_N} \left\{ \epsilon_z \left[ -\frac{2\pi z}{R} - \xi^i_B \partial_z \right] g^{jk} \delta g_{jk} \right. + \left. \epsilon_t \left[ -\delta_{il} \frac{2\pi(x^l - x^l_0)}{R} - \xi^i_B \partial_l \right] g^{jk} \delta g_{jk} - \epsilon_t \left[ -\delta_{lj} \frac{2\pi(x^l - x^l_0)}{R} - \xi^i_B \partial_j \right] g^{kj} \delta g_{ki} \right\}. \tag{8.104}$$

Again, we will consider the case where the metric is approximated by the Fefferman-Graham metric such that $\delta g_{ab} = z^{d-2} h_{ab}$. Using this expression for the metric perturbation and that $g^{ij} = \frac{z^d}{L^2} \delta^{ij}$, we find the final expression for $\chi_B$:

$$\chi_B \mid \Pi = -\frac{1}{16\pi G_N} \left\{ \epsilon_z \left[ -\frac{2\pi z}{R} - \xi^i_B \partial_z \right] \frac{z^d}{L^2} h_{jk} \right. + \left. \epsilon_t \left[ -\delta_{il} \frac{2\pi(x^l - x^l_0)}{R} - \xi^i_B \partial_l \right] \frac{z^d}{L^2} h_{jk} - \epsilon_t \left[ -\delta_{lj} \frac{2\pi(x^l - x^l_0)}{R} - \xi^i_B \partial_j \right] \frac{z^d}{L^2} h_{ki} \right\} \tag{8.105}$$
where we, in the last equality, have ensured that the indices on $\epsilon$ is lowered such that we can use definition (8.96). Note that the index on partial derivatives are still raised and lowered by the Kronecker-Delta, $\partial_i = \delta_{ij} \partial^j$.

With this expression we can now verify that

\[ \int_B \chi_B|\Pi = \delta E^{\text{grav}}_B \]  
(8.106)

\[ \int_B \chi_B|n = \delta S^{\text{grav}}_B \]  
(8.107)

\[ \chi_B|\Pi = -\frac{1}{8\pi G_N} \xi^t_B \dot{g}^E\epsilon^t. \]  
(8.108)

First, consider the integration over $B$. Here, $\epsilon_{zt} = 0$ since $dz = 0$ on $B$ which follows since $z = 0$ on $B$. Define

\[ u \equiv \frac{z^{d+2}}{16\pi G_N L^4} \left( \frac{2\pi z}{R} + \xi^t_B \frac{d}{z} + \xi^t_B \dot{z} \right) \delta^{jk} h_{jk} \]  
(8.109)

and observe that the integral of the $d-1$ form $u \epsilon_{zt}$ is

\[ \int_B u\epsilon_{zt} = \int_B d^{d-1}x \sqrt{|g|} u. \]  
(8.110)

Inserting this and the expression for $u$ into (8.110), we find

\[ \int_B u\epsilon_{zt} = \int_B d^{d-1}x \frac{L^{d+1}}{z^{d+1}} \frac{z^{d+2}}{16\pi G_N L^4} \left( \frac{2\pi z}{R} + \xi^t_B \frac{d}{z} + \xi^t_B \dot{z} \right) \delta^{jk} h_{jk} \]

\[ = \frac{L^{d-3}}{16\pi G_N} \int_B d^{d-1}x \frac{z^{d+2}}{z^{d+1}} \left( \frac{2\pi z}{R} + \xi^t_B \frac{d}{z} + \xi^t_B \dot{z} \right) \delta^{jk} h_{jk} \]

\[ = \frac{L^{d-3}d}{16\pi G_N} \int_B d^{d-1}x \xi^t_B \delta^{jk} h_{jk} \]  
(8.111)

where the last equality follows since $z = 0$ on $B$ and $\partial_z \delta^{jk} h_{jk}(x,0) = 0$. Inserting the expression for $\xi^t_B$ at $z = 0$ gives

\[ \int_B \chi_B|\Pi = \frac{L^{d-3}d}{16\pi G_N} \int_B d^{d-1}x \frac{\pi}{R} \left( R^2 - (\vec{x} - \vec{x}_0)^2 \right) \delta^{jk} h_{jk} \]

\[ = \frac{L^{d-3}d}{16\pi G_N R} \int_B d^{d-1}x \left( R^2 - (\vec{x} - \vec{x}_0)^2 \right) \delta^{jk} h_{jk} = \delta E^{\text{grav}}_B. \]  
(8.112)

Next, consider the integration over $\tilde{B}$. Since $\tilde{B}$ is a bifurcation surface of the killing vector $\xi_B$, it follows that $\xi^t_B = 0$ on $\tilde{B}$. Both $\epsilon_{zt}$ and $\epsilon_{zt}$ are non-vanishing, so we have

\[ \int_{\tilde{B}} \chi_{\tilde{B}}|\Pi = \int_{\tilde{B}} (u\epsilon_{zt} + v^t\epsilon_{zt}). \]  
(8.113)
Here we define the function $u$ as above but now restricted to the surface $\tilde{B}$:

$$u \equiv \frac{z^{d+2}}{16\pi G_N L^4} \left( \frac{2\pi z}{R} + \xi_B^i \frac{d}{z} + \xi_B \partial_z \right) \delta^{jk} h_{jk} = \frac{z^{d+2}}{16\pi G_N L^4} \frac{2\pi z}{R} \delta^{jk} h_{jk} \quad (8.114)$$

where the last equality follows since $\xi_B^i = 0$ on $\tilde{B}$. We also define in addition

$$v^i \equiv \frac{z^{d+2}}{16\pi G_N L^4} \left\{ \left[ \frac{2\pi(x^i - x_0^i)}{R} + \xi_B^i \partial^i \right] \delta^{jk} h_{jk} - \left[ \frac{2\pi(x^j + x_0^j)}{R} + \xi_B^i \partial^i \right] \delta^{jk} h_{kj} \right\}$$

$$= \frac{z^{d+2}}{16\pi G_N L^4} \left[ \frac{2\pi(x^i - x_0^i)}{R} \delta^{jk} h_{jk} - \frac{2\pi(x^j + x_0^j)}{R} \delta^{jk} h_{kj} \right]. \quad (8.115)$$

Again, we have

$$\int_{\tilde{B}} u \epsilon_{zt} = \int_{\tilde{B}} d^{d-1}x \sqrt{|g|} u. \quad (8.116)$$

where the integration now is over the surface $\tilde{B}$. For the integration over the $\epsilon_{ul}$-part, we have

$$\int_{\tilde{B}} v^i \epsilon_{ul} = \int_{\tilde{B}} \sqrt{|g|} v^i dz dx^1 \cdots dx^{i-1} dx^{i+1} \cdots dx^{d-1}$$

$$= - \int_{\tilde{B}} d^{d-1}x \frac{L^{d+1}}{z^{d+1}} v^i \frac{dz}{dx^i} \quad (8.117)$$

$$= \int_{\tilde{B}} d^{d-1}x \frac{L^{d+1}}{z^{d+1}} v^i \xi_{ji} \frac{(x^j - x_0^j)}{z}$$

where we have used the chain rule to replace $dz$ by $dx^i$ and $dz/dx^i = -\delta_{ji}(x^j - x_0^j)/z$.

For convenience, we will evaluate the integration over $u \epsilon_{zt}$ and $v^i \epsilon_{ul}$ separately:

$$\int_{\tilde{B}} u \epsilon_{zt} = \int_{\tilde{B}} \epsilon_{zt} \frac{z^{d+2}}{16\pi G_N L^4} \frac{2\pi z}{R} \delta^{jk} h_{jk}$$

$$= \int_{\tilde{B}} d^{d-1}x \frac{z^{d+2}}{16G_N RL^4} \frac{L_{d+1}}{z^{d+1}} \delta^{jk} h_{jk}$$

$$= \frac{L^{d-3}}{8G_N R} \int_{\tilde{B}} d^{d-1}x z^2 \delta^{jk} h_{jk}. \quad (8.118)$$
and

\[
\int_B v^i \epsilon_{it} = \int_B \frac{z^{d+2}}{16\pi G_N L^4} \left[ \frac{2\pi (x^i - x^i_0)}{R} \delta^{jk} h_{jk} - \frac{2\pi (x^j + x^j_0)}{R} \delta^{ik} h_{kj} \right] \epsilon_{it} 
= - \int_B d^{d-1} x \frac{z^{d+2}}{8G_N RL^4} \frac{L^{d+1}}{z^{d+1}} dx^i \left[ \frac{2\pi (x^i - x^i_0)}{R} \delta^{jk} h_{jk} - \frac{2\pi (x^j + x^j_0)}{R} \delta^{ik} h_{kj} \right] 
= \frac{L^{d-3}}{8G_N R} \int_B d^{d-1} x \frac{z (x^i - x^i_0)}{z} \left[ \frac{2\pi (x^i - x^i_0)}{R} \delta^{jk} h_{jk} - \frac{2\pi (x^j + x^j_0)}{R} h_{ij} \right] 
= \frac{L^{d-3}}{8G_N R} \int_B d^{d-1} x \left[ (x^i - x^i_0)^2 \delta^{jk} h_{jk} - (x^i - x^i_0)(x^j - x^j_0) h_{ij} \right].
\]

(8.119)

Putting these elements back together and using \( R^2 = z^2 + (\vec{x} - \vec{x}_0)^2 \), we finally obtain

\[
\int_B \chi_B |_{\Pi} = \int_B \epsilon_{zt} + \int_B v^i \epsilon_{it} 
= \frac{L^{d-3}}{8G_N R} \int_B d^{d-1} x \frac{z^{d+2}}{z^{d+1}} \frac{L^{d+1}}{z} \left[ \frac{2\pi (x^i - x^i_0)}{R} \delta^{jk} h_{jk} - \frac{2\pi (x^j + x^j_0)}{R} \delta^{ik} h_{kj} \right] 
= \frac{L^{d-3}}{8G_N R} \int_B d^{d-1} x \left[ R^2 \delta^{jk} h_{jk} - (x^i - x^i_0)(x^j - x^j_0) h_{ij} \right] = \delta S_B^{grav}. 
\]

(8.120)

Finally, consider the external derivative, \( d\chi |_{\Pi} \). Writing \( \chi |_{\Pi} \equiv u \epsilon_{zt} + v^i \epsilon_{it} \) and using that \( \epsilon_{tt} = 0 \), we can express the external derivative as

\[
d\chi_B |_{\Pi} = d(u \epsilon_{zt} + v \epsilon_{it}) = -d(\sqrt{-g} \epsilon_{zt} + \sqrt{-g} \epsilon_{ti}) = - \left( \frac{\partial u \sqrt{-g}}{\partial z} + \frac{\partial v^i \sqrt{-g}}{\partial x^i} \right) \tilde{\epsilon}_t 
\]

(8.121)

where we have defined \( \tilde{\epsilon}_{ab} = \epsilon_{ab} / \sqrt{-g} \). The minus sign arises since \( t \) must be the first rather than the second index of the Levi-Civita symbol inside \( \tilde{\epsilon}_t \), and the Levi-Civita is anti-symmetric. Using \( \sqrt{-g} = \frac{L^{d+1}}{\vec{x}^{d+1}} \) and the expressions for \( u \) and \( v^i \), we can evaluate
the two partial derivatives:

\[
\frac{\partial (u \sqrt{-g})}{\partial z} = \frac{\partial (u L^{d+1}_{\sqrt{-g}})}{\partial z} = \frac{\partial}{\partial z} \left( \frac{L^{d+1}}{2\pi L^d} \right) \left( \frac{2\pi z}{R} + \xi^i_B \frac{d}{z} + \xi^i_B \partial_z \right) \delta^{ijk} h_{jk} \\
= \frac{L^{d-3}}{16\pi G_N} \left( \frac{4\pi z}{R} + 2\frac{\pi z^2}{R} \partial_z - 2\pi z \frac{\partial z}{R} + \xi^i_B \partial_z + \xi^i_B \partial_z - 2\pi z \frac{\partial z}{R} \partial_z + \partial z \frac{\partial z}{R} \right) \delta^{ijk} h_{jk}
\]

where we have used that \( \partial_z \xi^i_B = -2\pi z/R \). The derivative \( \frac{\partial (v \sqrt{-g})}{\partial x^i} \) becomes

\[
\frac{\partial (v \sqrt{-g})}{\partial x^i} = \frac{\partial (v L^{d+1}_{\sqrt{-g}})}{\partial x^i} = \frac{\partial}{\partial x^i} \left( \frac{L^{d+1}}{2\pi L^d} \right) \left( \frac{2\pi (x^i - x^i_0)}{R} + \xi^i_B \partial^j \right) \delta^{ijk} h_{jk} \\
= \frac{L^{d-3}}{16\pi G_N} \left( \frac{2\pi}{R} \delta^{i} + \frac{2\pi (x^i - x^i_0)}{R} \partial_i \left( \partial_t \xi^i_B \right) \partial^j + \xi^i_B \partial^j \partial^k \delta^{jk} h_{jk} \right) \\
- \frac{L^{d-3}}{16\pi G_N} \left( \frac{2\pi}{R} + \frac{2\pi (x^i - x^i_0)}{R} \partial_i \left( \partial_t \xi^i_B \right) \partial^j + \xi^i_B \partial^j \partial^k \delta^{jk} h_{jk} \right)
\]

where we have used that \( \partial_t \xi^i_B = -2\pi (x^i - x^i_0)/R \) and \( \delta^i_i = d - 1 \).

Collecting the terms, we get:

\[
d\chi_B|_\Pi = - \left( \frac{\partial (u L^{d+1}_{\sqrt{-g}})}{\partial z} + \frac{\partial (v L^{d+1}_{\sqrt{-g}})}{\partial x^i} \right) \epsilon_t = \frac{L^{d-3}}{16\pi G_N} \left[ \frac{2\pi}{R} (2 - d) + \frac{\xi^i_B}{z} (d + 1) \partial_z + \xi^i_B \partial^2_z \right] \delta^{ijk} h_{jk}
\]

\[
+ \left( \frac{2\pi}{R} (d - 2) + \xi^i_B \partial_t \partial^j \delta^{ijk} h_{jk} - \xi^i_B \partial_t \partial^j \delta^{ijk} h_{jk} \right) \epsilon_t = \frac{L^{d-3}}{16\pi G_N} \left[ \frac{\xi^i_B}{z} (d + 1) \partial_z + \xi^i_B \partial^2_z + \xi^i_B \partial_t \partial^j \delta^{ijk} h_{jk} \right] \epsilon_t.
\]

88
To put the exterior derivative on the form $d\chi_B|\Pi = -\frac{1}{8\pi G_N}\xi_B^t \dot{G}^E_{tt} \epsilon^t$, we have to raise the index of $\bar{\epsilon}_t$ and replace it by $\epsilon^t$

$$\bar{\epsilon}_t = g^{tt} \frac{z^{d+1}}{L^{d+1}} \epsilon^t = -\frac{z^{d-1}}{L^{d-1}} \epsilon^t.$$  \hspace{1cm} (8.125)

Inserting this, we get

$$d\chi_B|\Pi = -\frac{z^d}{16\pi G_N L^2} \xi_B^t \left[ \left( \frac{d+1}{z} \partial_z + \partial_z^2 + \partial_i \partial^i \right) \delta^{jk} h_{jk} + \partial_i \partial^i \delta^{jk} h_{kj} \right] \epsilon^t$$

$$= -\frac{1}{8\pi G_N} \xi_B^t \dot{G}^E_{tt} \epsilon^t$$  \hspace{1cm} (8.126)

where

$$\dot{G}^E_{tt} = \frac{z^d}{2L^2} \left[ \left( \frac{d+1}{z} \partial_z + \partial_z^2 + \partial_i \partial^i \right) \delta^{jk} h_{jk} + \partial_i \partial^i \delta^{jk} h_{kj} \right]$$  \hspace{1cm} (8.127)

in agreement with the result found for the linearized Einstein equations in AdS in section 6.2. Thus, we see that the linearized Einstein equations in vacuum follows from the holographic interpretation of the first law of entanglement entropy.
Chapter 9

Beyond Linearized Gravity

So far, we have on the AdS side considered the limit \( G_N \to 0 \) of general relativity. In this limit, the source term appearing in the Einstein equations, \( 8\pi G_N T_{ab} \), is suppressed since it is of order \( G_N \). Thus, to go beyond vacuum solutions we must include also the first subleading order in \( G_N \) in the expansion of the Einstein equations. The limit \( G_N \to 0 \) is – following the identification (4.33) – equivalent to the limit \( g_s \to 0 \). According to the AdS/CFT dictionary, \( g_s \propto \frac{1}{N} \) (4.6) and therefore the first subleading order in \( G_N \) corresponds to an expansion on the CFT side that includes the first subleading order in \( \frac{1}{N} \). In the previous chapter, the CFT side relation

\[
\frac{d}{d\zeta} S_B = \frac{d}{d\zeta} H_{B^p}
\]

(9.1)

was derived in full generality, i.e. to any order in \( \zeta \). This relation therefore holds also to subleading orders in \( \frac{1}{N} \).

Our interest here will be to see what this first law of entanglement entropy corresponds to on the AdS side, if we include the first subleading order in \( \frac{1}{N} \). We will find that this gives rise to a source term in linearized Einstein gravity exactly corresponding to the source term, \( 8\pi G_N T_{ab} \), in the Einstein Equations. This result was first obtained in [33] which the present chapter is based on.

9.1 \( \frac{1}{N} \) corrections to the Ryu-Takayanagi formula

As already indicated, once we include subleading orders in \( \frac{1}{N} \) on the CFT side, we will have to consider matter fields on the AdS side. To first subleading order in \( G_N \), we can approximate the AdS side as a semi-classical gravity, i.e. by a classical spacetime
$M$ and a quantum matter field living on that spacetime, $|\Phi\rangle_{\text{Bulk}}$ such that the dual of some quantum state $|\Psi\rangle$ is the pair $(M, |\Phi\rangle_{\text{Bulk}})$. The approximation remains semi-classical since we assume that the spacetime $M$ is a classical spacetime, i.e. it cannot be a superposition of spacetimes, which supposedly would be allowed in a full theory of quantum gravity.

Going to subleading order in $\frac{1}{N}$ or equivalently $G_N$ induces corrections to the Ryu-Takayanagi formula. More precisely, it is conjectured in [7] that subleading order corrections to the Ryu-Takayanagi formula is due to bulk entanglement over the surface $\tilde{B}$, that is the entanglement between the degrees of freedom in the region $\Pi$ – the region bounded by $B \cup \tilde{B}$ – and the rest of the system. The correction takes the form of the entanglement entropy of the quantum field on $\Pi$

$$S_B(|\Psi\rangle) = A(\tilde{B}) + S_\Pi(|\Phi\rangle_{\text{Bulk}})$$

(9.2)

where we have signified that the CFT entanglement entropy, $S_B$, is a function of the CFT state $|\Psi\rangle$ and that the bulk entanglement entropy, $S_\Pi$, is a function of the state of the bulk field, $|\Phi\rangle_{\text{Bulk}}$. Just as we motivated the Ryu-Takayangi formula as a generalization of the Bekenstein-Hawking formula, we can motivate this correction to the Ryu-Takayanagi formula by continuing the analogy to the case of black holes where corrections are also supposed to arise due to entanglement over the black hole horizon. It is essentially this result that has been generalized in [7] to arbitrary surfaces that minimizes the area functional.

Subsequently, we will simply assume that this corrected Ryu-Takayanagi formula is sound for quantum states with a semi-classical spacetime dual. With this formula, we can – in analogy to the previous chapter – obtain holographic interpretations of $S_B$ and $E_{B,Hyp}$. Again, a holographic interpretation of $S_B$ is immediately provided now by the corrected Ryu-Takayanagi formula. The variation $\delta S_B$ therefore consists of the already derived variation of the area of the surface $\tilde{B}$ (8.59) and the variation of the bulk entanglement entropy over the surface $\tilde{B}$. The variation of the latter, $\delta S_\Pi(|\Phi\rangle_{\text{Bulk}})$, in principle originates in two sources: variation of the surface $\tilde{B}$ and variation of the bulk field $\delta |\Phi\rangle_{\text{Bulk}}$. The former, however, vanishes to first order. This follows since entanglement entropy – just like the area of the minimal surface $\tilde{B}$ – is extremized by the embedding function $X_{\text{ext}}^a(x_i)$ and the first order variation always vanishes for such an extremum. What remains, therefore, is to find an expression for the variation of $S_\Pi(|\Phi\rangle_{\text{Bulk}})$ due to the variation of the bulk field $\delta |\Phi\rangle_{\text{Bulk}}$.
9.2 A holographic interpretation of $\delta S_B$ and $\delta E^{HyP}_B$

In analogy to the entanglement entropy of a quantum subsystem on the CFT side, the bulk entanglement entropy, $S_{\Pi}(\ket{\Phi}_{Bulk})$, is given in terms of the reduced density matrix, $\rho^B_{\Pi}$, for the region $\Pi$

$$S_{\Pi}(\ket{\Phi}_{Bulk}) = - \text{tr}\left(\rho^B_{\Pi} \log(\rho^B_{\Pi})\right). \quad (9.3)$$

Depending only on the variation, $\delta \ket{\Phi}_{Bulk}$, we may express the variation of the bulk entanglement entropy in terms of the variation of the $\delta S_{\Pi}(\ket{\Phi}_{Bulk})$ by an argument parallel to that of (8.7)

$$\delta S_{\Pi}(\ket{\Phi}_{Bulk}) = - \text{tr}\left(\delta \rho^B_{\Pi} \log(\rho^B_{\Pi})\right) \quad (9.4)$$

where we have again used that the trace of the density matrix must be unity. Defining the bulk modular Hamiltonian

$$H^B_{\Pi} \equiv - \log(\rho^B_{\Pi}) \quad (9.5)$$

we find

$$\delta S_{\Pi}(\ket{\Phi}_{Bulk}) = \text{tr}\left(\delta \rho^B_{\Pi} H^B_{\Pi}\right)$$

$$= \delta \text{tr}\left(\rho^B_{\Pi} H^B_{\Pi}\right)$$

$$= \delta \braket{H^B_{\Pi}} \quad (9.6)$$

where the last equality follows since $H^B_{\Pi}$ is independent of the variation and from the definition of the expectation value.

Similarly to the modular Hamiltonian on the CFT side, the bulk modular Hamiltonian cannot in general be expressed in terms of local operators. However, for the particular case where the full spacetime is AdS spacetime and we are interested in the modular Hamiltonian of the region $\Pi$, the modular Hamiltonian can be expressed as an integral over the bulk energy momentum tensor

$$H^B_{\Pi} = \int_{\Pi} \xi^a_B \mathcal{T}^{ab}_{Bulk} \epsilon^b \quad (9.7)$$

where $\xi^b_B$ and $\epsilon^a$ are given as in (8.79) and (8.96). The derivation of this expression is analogous to the derivation in section 8.1. Further details can be found in [33]. With this expression for $H^B_{\Pi}$ we find that

$$\delta S_{\Pi}(\ket{\Phi}_{Bulk}) = \delta \braket{H^B_{\Pi}} = \int_{\Pi} \xi^a_B \delta \braket{\mathcal{T}^{ab}_{Bulk}} \epsilon^b. \quad (9.8)$$
For the hyperbolic energy, we found in section 8.3 the first order holographic interpretation
\[
\delta E_{B}^{Hyp} = 2\pi \int_{B} dt^{d-1} x R^{2} \frac{r^{2}}{2R} \delta \left\langle T^{tt}(x, z = 0) \right\rangle. \tag{9.9}
\]
This was obtained by considering an infinitesimal ball on the boundary; subleading corrections to \( \delta E_{B}^{Hyp} \) should therefore be related to the energy-momentum tensor of the bulk region associated with this infinitesimal boundary ball. Thus, if we assume that the expectation value of bulk energy-momentum tensor, \( \left\langle T_{ab}^{Bulk} \right\rangle \), vanishes on the boundary, it follows that the correction to \( \delta E_{B}^{Hyp} \) vanishes to first subleading order in \( \frac{1}{N} \).

### 9.3 Linearized gravity coupled to matter

In the previous chapter, we found (8.83)
\[
\delta S_{B}^{Grav} - \delta E_{B}^{Grav} = \int d\chi_{B}(h) \tag{9.10}
\]
where \( d\chi_{B} = -\frac{1}{8\pi G_N} \xi_{B}^{a} \hat{G}_{ab}^{E} \epsilon^{b} \). Including the correction to \( S_{B}^{Grav} \), we have the following relation to first subleading order in \( G_N \) in the bulk
\[
\delta S_{B}^{Grav} - \delta E_{B}^{Grav} = -\int_{\Pi} \frac{1}{8\pi G_N} \xi_{B}^{a} \hat{G}_{ab}^{E} \epsilon^{b} + \int_{\Pi} \xi_{B}^{a} \delta \left\langle T_{ab}^{Bulk} \right\rangle \epsilon^{b} \tag{9.11}
\]
\[
= -\frac{1}{8\pi G_N} \int_{\Pi} \xi_{B}^{a} \left( \hat{G}_{ab}^{E} - 8\pi G_N \delta \left\langle T_{ab}^{Bulk} \right\rangle \right) \epsilon^{b}.
\]
Now, imposing as a constraint the holographic interpretation of the first law of entanglement entropy \( \delta S_{B}^{Grav} = \delta E_{B}^{Grav} \), we find
\[
0 = -\frac{1}{8\pi G_N} \int_{\Pi} \xi_{B}^{a} \left( \hat{G}_{ab}^{E} - 8\pi G_N \delta \left\langle T_{ab}^{Bulk} \right\rangle \right) \epsilon^{b}. \tag{9.12}
\]
An argument analogous to that in section (8.4) implies that this constraint is satisfied only if \( \hat{G}_{ab}^{E} - 8\pi G_N \delta \left\langle T_{ab}^{Bulk} \right\rangle = 0 \), i.e. if the linearized Einstein equations with source term \( 8\pi G_N \delta \left\langle T_{ab}^{Bulk} \right\rangle \) are satisfied in the bulk.

This, therefore, demonstrates that imposing the constraint \( \delta S_{B} = \delta E_{B}^{Hyp} \) to subleading order in \( \frac{1}{N} \) on a quantum state with a semi-classical asymptotically AdS spacetime dual is equivalent to imposing the Einstein equations on that spacetime to subleading order in \( G_N \) in the bulk, i.e. the spacetime satisfies the equations
\[
\hat{G}_{ab}^{E} = 8\pi G_N \delta \left\langle T_{ab}^{Bulk} \right\rangle. \tag{9.13}
\]
The constraint, \( dS_B = dE_B^{Hyp} \), we argued holds to any order in \( \frac{1}{N} \). This may trigger the consideration whether it is equivalent to the Einstein equations to all orders. In order for \( \delta \langle T^{ Bulk}_{ab} \rangle \) to occur in (9.13), it must be the case that \( \langle T^{Bulk}_{ab} \rangle \) is present in the general spacetime constraint equivalent the general CFT constraint \( dS_B = dE_B^{Hyp} \).

Prima facie, one could think that there were other operators, \( O^{Bulk}_{ab} \), in this general spacetime constraint. The idea would be that the source term could be of the form \( 8\pi G_N \langle T^{Bulk}_{ab} + O^{Bulk}_{ab} \rangle \). As argued in [33], a rather simple observation renders such a construction unlikely. From this general source term, we must obtain \( 8\pi G_N \delta \langle T^{Bulk}_{ab} \rangle \) to subleading order in \( G_N \), i.e. \( \delta \langle O^{Bulk}_{ab} \rangle \) must vanish to this order. This would obtain if the operator \( O^{Bulk}_{ab} \) was such that it annihilated the bulk vacuum state. However, any expansion in terms of creation and annihilation operators of a local quantum field theory operator will include terms without annihilation operators – in other words, only with creation operators – and therefore, the operator \( O^{Bulk}_{ab} \) cannot be a local quantum field theory operator since \( \delta \langle O^{Bulk}_{ab} \rangle = 0 \). There may be more subtle ways to make this construction, however, the argument indicates that the full semi-classical Einstein equations,

\[
G^E_{ab} = 8\pi G_N \langle T^{Bulk}_{ab} \rangle
\]

(9.14)

are the general constraints on the spacetime equivalent to CFT constraint \( dS_B = dE_B^{Hyp} \).

A more indirect argument to this effect is the observation that entanglement may occur between all types of degrees of freedom. All degrees of freedom in a subsystem contribute to the entanglement entropy and are therefore taken into account when the entanglement entropy for the subsystem is evaluated. In this way, entanglement is universal. As is well known, the Einstein equations signifies that all stress energy are sources of gravity. Regardless of the origin of the stress energy it sources the gravitational field. As they write in [33]: “we can say that the universality of the gravitational interaction comes directly from the universality of entanglement”. This, we may argue, is why the full semi-classical Einstein equations should be equivalent to some constraint on entanglement in the dual quantum state.

It is worth remarking that even if we can obtain the full semi-classical Einstein equations from \( dS_B = dE_B^{Hyp} \), there still is an important restriction to the domain of application of this result essentially due to the fact that the spacetime is still classical. This is explicitly seen from the fact that the Einstein tensor (with cosmological constant) \( G^E_{ab} \), on the LHS of (9.14), is not a quantum operator. Thus, (9.14) as a whole is not
an operator relation; which can be regarded as the reason why we have to take the expectation value of $T_{ab}^{Bulk}$. A more subtle point leading to the same conclusion is that the Ryu-Takayanagi formula is ill-defined if the geometry under consideration is a superposition of spacetimes. Since the whole machinery above is founded on the Ryu-Takayanagi formula, it follows that we can at most obtain semi-classical gravity using it as our holographic dictionary. Thus, unless further correction are made to the Ryu-Takayanago formula, we cannot go beyond this domain of validity.
Chapter 10

Conclusion

As demonstrated in chapter 8 and 9, the linearized Einstein equations both with and without matter in a $d + 1$-dimensional AdS background can be derived from the first law of entanglement entropy in a $d$-dimensional CFT. Furthermore, it was speculated that the full semi-classical Einstein gravity follows from imposing to first law of entanglement to all order in $\frac{1}{N}$.

The result is quite remarkable. Gravity in $d + 1$ dimensions may be encoded in an intrinsically quantum mechanical structure – entanglement structure – in a $d$-dimensional CFT. The dynamics described by the Einstein’s field equations can be reinterpreted as a first law like dynamical constraint on entanglement. But however remarkable this may be, the import of the result is initially less significant. The holographic relation between gravity and entanglement does not in the form developed here offer new insights the overarching problem of contemporary theoretical physics: Quantum gravity. By offering the alternative interpretation of semi-classical gravity as entanglement structure, we are offered an alternative description of a theory that we already know well and which arguably is less than full theory of quantum gravity.\footnote{It should be noted that there are those who argue for the possibility of semi-classical gravity on the fundamental level. See for instance \cite{34} for a recent discussion.} For this and other reasons, there are continued research on the holographic relation between entanglement and gravity. It is beyond the scope of the present project to account for this research in detail, however, we will here just mention a few of the recent developments.

In \cite{35}, the holographic relation between entanglement and gravity is rigorously extended to non-linear gravity by showing that the second order perturbation of Einstein’s field equations follows from the first law of entanglement entropy. Again, the Ryu-
Takayanagi formula is employed as the only ingredient of the AdS/CFT dictionary. This may prove relevant, since a holographic relation between entanglement entropy and an emergent spacetime structure similar to the Ryu-Takayanagi formula has recently been recovered in the context of loop quantum gravity [36]. This may suggest an independence from the AdS/CFT correspondence (and therefore from string theory) for a holographic prescription like the Ryu-Takayanagi formula. Obtaining the Ryu-Takayanagi in loop quantum gravity is driven by the relation between loop quantum gravity and tensor networks. Interesting, recent research has also explored the links from the relation between entanglement and gravity studied here to (random) tensor networks [37, 38].

In the chapters above, the Ryu-Takayanagi formula played a crucial role in the derivation of the holographic relation between gravity and entanglement. In [39] it is argued that one can obtain the same result using only the Bekenstein-Hawking formula under the assumption that the entanglement entropy is maximized for small fixed volume balls.

A research program relevant for classical general relativity is the conjectured relation between the Fisher information defined in terms of relative entropy for CFT subsystems and canonical energy in a dual AdS spacetime. It has been shown in [40] that the Fisher information is positive definite which entails that the corresponding canonical energy is positive definite as well. This in turn suggests a new positive energy theorem in general relativity.

There are also work that explores other constraints on entanglement in CFTs and how these translate holographically as constraints in the bulk. For instance, mutual entanglement information is considered in [41] and strong subadditivity (as well as mutual information) is considered in [42].

Acknowledgements

This project is the result of many insightful discussion with more people than I can list here. However, I would in particular like thank my supervisor Troels Harmark for his valuable comments on earlier drafts of this work. Honourable mention also goes to my most frequent discussion partners Bjarke Nielsen and Niels Linnemann who have helped me sort out the numerous questions that arose during the genesis of this project.
Bibliography

[1] M. Van Raamsdonk, “Building up spacetime with quantum entanglement,” General Relativity and Gravitation, vol. 42, no. 10, pp. 2323–2329, 2010.

[2] M. Van Raamsdonk, “A patchwork description of dual spacetimes in AdS/CFT,” Classical and Quantum Gravity, vol. 28, no. 6, p. 065002, 2011.

[3] N. Lashkari, M. B. McDermott, and M. Van Raamsdonk, “Gravitational dynamics from entanglement “thermodynamics”,” Journal of High Energy Physics, vol. 2014, no. 4, p. 195, 2014.

[4] T. Faulkner, M. Guica, T. Hartman, R. C. Myers, and M. Van Raamsdonk, “Gravitation from entanglement in holographic CFTs,” Journal of High Energy Physics, vol. 2014, no. 3, p. 51, 2014.

[5] S. Ryu and T. Takayanagi, “Holographic Derivation of Entanglement Entropy from the anti-de Sitter Space/Conformal Field Theory Correspondence,” Phys. Rev. Lett., vol. 96, no. 18, p. 181602, 2006.

[6] V. E. Hubeny, T. Takayanagi, and M. Rangamani, “A covariant holographic entanglement entropy proposal,” Journal of High Energy Physics, vol. 0707, no. 062, 2007.

[7] T. Faulkner, A. Lewkowycz, and J. Maldacena, “Quantum corrections to holographic entanglement entropy,” Journal of High Energy Physics, vol. 2013, no. 11, p. 74, 2013.

[8] M. Ammon and J. Erdmenger, Gauge/Gravity Duality: Foundations and Applications. Cambridge: Cambridge University Press, 2015.

[9] S. Carroll, Spacetime and Geometry: An Introduction to General Relativity. Harlow: Pearson Education Limited, 2014.
[10] P. Calabrese and J. Cardy, “Entanglement entropy and quantum field theory,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2004, no. 06, p. P06002, 2004.

[11] J. Maldacena, “The Large-N Limit of Superconformal Field Theories and Supergravity,” *International Journal of Theoretical Physics*, vol. 38, no. 4, pp. 1113–1133, 1999.

[12] J. D. Bekenstein, “Black Holes and Entropy,” *Physical Review D*, vol. 7, no. 8, pp. 2333–2346, 1973.

[13] J. H. Schwarz, “Covariant field equations of chiral N = 2 D = 10 supergravity,” *Nuclear Physics B*, vol. 226, no. 2, pp. 269–288, 1983.

[14] P. Howe and P. West, “The complete N =2, d = 10 supergravity,” *Nuclear Physics B*, vol. 238, no. 1, pp. 181–220, 1984.

[15] D. Berenstein, J. Maldacena, and H. Nastase, “Strings in flat space and pp waves from N= 4 Super Yang Mills,” *Journal of High Energy Physics*, vol. 2002, no. 04, p. 013, 2002.

[16] C. Callan, C. Lovelace, C. Nappi, and S. Yost, “String loop corrections to beta functions,” *Nuclear Physics B*, vol. 288, pp. 525–550, 1987.

[17] M. Günaydin, L. Romans, and N. Warner, “Compact and non-compact gauged supergravity theories in five dimensions,” *Nuclear Physics B*, vol. 272, no. 3, pp. 598–646, 1986.

[18] J. Maldacena, “Eternal black holes in anti-de Sitter,” *Journal of High Energy Physics*, vol. 2003, no. 04, p. 021, 2003.

[19] D. Marolf and A. C. Wall, “Eternal black holes and superselection in AdS/CFT,” *Classical and Quantum Gravity*, vol. 30, no. 2, p. 025001, 2013.

[20] S. G. Avery and B. D. Chowdhury, “No Holography for Eternal AdS Black Holes.” 2013.

[21] S. D. Mathur, “What is the dual of two entangled CFTs?.” 2014.

[22] H. Casini, M. Huerta, and R. C. Myers, “Towards a derivation of holographic entanglement entropy,” *Journal of High Energy Physics*, vol. 2011, no. 5, p. 36, 2011.
[23] A. Lewkowycz and J. Maldacena, “Generalized gravitational entropy,” *Journal of High Energy Physics*, vol. 2013, no. 8, p. 90, 2013.

[24] R. Emparan, “Black hole entropy as entanglement entropy: a holographic derivation,” *Journal of High Energy Physics*, vol. 2006, no. 06, p. 012, 2006.

[25] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, “Area Laws in Quantum Systems: Mutual Information and Correlations,” *Physical Review Letters*, vol. 100, no. 7, p. 070502, 2008.

[26] T. Jacobson, “Thermodynamics of Spacetime: The Einstein Equation of State,” *Physical Review Letters*, vol. 75, no. 7, pp. 1260–1263, 1995.

[27] M. Van Raamsdonk, “Lectures on Gravity and Entanglement,” https://arxiv.org/abs/1609.00026, 2016.

[28] V. Balasubramanian and P. Kraus, “A Stress Tensor for Anti-de Sitter Gravity,” *Communications in Mathematical Physics*, vol. 208, no. 2, pp. 413–428, 1999.

[29] K. Skenderis, “Asymptotically Anti-de Sitter Spacetimes and their Stress Energy Tensor,” *International Journal of Modern Physics A*, vol. 16, no. 05, pp. 740–749, 2001.

[30] V. Iyer and R. M. Wald, “Some properties of the Noether charge and a proposal for dynamical black hole entropy,” *Physical Review D*, vol. 50, pp. 846–864, July 1994.

[31] H. Flanders, “Differentiation Under the Integral Sign,” *The American Mathematical Monthly*, vol. 80, no. 6, pp. 615–627, 1973.

[32] R. M. Wald, *General Relativity*. Chicago: Chicago University Press, 1984.

[33] B. Swingle and M. Van Raamsdonk, “Universality of Gravity from Entanglement,” https://arxiv.org/abs/arXiv:1405.2933v1, 2014.

[34] A. Tilloy and L. Diósi, “Sourcing semiclassical gravity from spontaneously localized quantum matter,” *Physical Review D*, vol. 93, no. 2, p. 024026, 2016.

[35] T. Faulkner, F. M. Haehl, E. Hijano, O. Parrikar, C. Rabideau, and M. Van Raamsdonk, “Nonlinear gravity from entanglement in conformal field theories,” *Journal of High Energy Physics*, vol. 2017, no. 8, p. 57, 2017.
[36] M. Han and L.-Y. Hung, “Loop quantum gravity, exact holographic mapping, and holographic entanglement entropy,” Physical Review D, vol. 95, no. 2, p. 024011, 2017.

[37] B. Swingle, “Entanglement renormalization and holography,” Phys. Rev. D, vol. 86, no. 6, p. 065007, 2012.

[38] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter, and Z. Yang, “Holographic duality from random tensor networks,” Journal of High Energy Physics, vol. 2016, no. 11, p. 9, 2016.

[39] T. Jacobson, “Entanglement Equilibrium and the Einstein Equation,” Physical Review Letters, vol. 116, no. 20, p. 201101, 2016.

[40] N. Lashkari and M. Van Raamsdonk, “Canonical energy is quantum Fisher information,” Journal of High Energy Physics, vol. 2016, no. 4, p. 153, 2016.

[41] P. Hayden, M. Headrick, and A. Maloney, “Holographic mutual information is monogamous,” Physical Review D, vol. 87, no. 4, p. 046003, 2013.

[42] N. Bao, S. Nezami, H. Ooguri, B. Stoica, J. Sully, and M. Walter, “The holographic entropy cone,” Journal of High Energy Physics, vol. 2015, no. 9, p. 130, 2015.