Holographic Entanglement Entropy: A way to Information Paradox

Hao Yang

Supervised by Prof Jerome Gauntlett

Department of Physics

Imperial college London

Submitted in part fulfilment of the requirements for the degree of

Master of Science in Quantum Fields and Fundamental Forces of Imperial College London

September 18, 2023

Abstract

Black hole information loss paradox has been a famous open question in theoretical physics for a long time. A breakthrough based on holographic entanglement entropy, called the "island rule", was proposed in 2019, which recovers the Page curve by considering the entanglement between the black hole and its radiation.

In this dissertation, we introduce the idea of holographic entanglement entropy through some simple examples in $\mathrm{AdS}_3/\mathrm{CFT}_2$. Then we review some important milestones of information loss, including Page curve and black hole complementarity, which leads to the so-called "central dogma". In the last part, we will see how holographic entanglement entropy can be applied to the information loss problem, and how replica wormhole supports the island rule.

Acknowledgement

I would like to thank my supervisor, Professor Jerome Gauntlett, for introducing me such an interesting topic, and also for the weekly meetings which are really enlightening.

I would thank one year of study at Imperial College I learned so much. As an ex-mathematical student, it was a challenge to get used to physicists' thinking way. Thanks to my friends, Steven Hsia, Jieming Lin, Zihan Wang and Zhihao Zhang, it was enjoyable to learn from them. Although here may not be a good place for any philosophical discussion, this experience inspired many thoughts of mine about the relationship and the differences between mathematics and physics. After some painful hesitancy, I finally confirm the areas that I want to work on. Here I would like to show my special appreciation to Qixuan Xu, who provided mush help with my PhD application.

Finally, I would appreciate my parents' unconditional support, though they can hardly understand what I am doing. I am grateful for the companionship of my mother. I could have finished my dissertation much earlier without her.

Contents

Introduction

The origin of the black hole information paradox is back to 1973 when J. A. Wheeler [29] first used the term – "No-hair" to illustrate a property of classical black holes, which states that a stationary black hole has no property other than a finite number of parameters. This statement implies all the information of the matters forming and infalling a classical black hole will completely lost. There was no problem until 1975 when Stephen Hawking, to support the idea of black hole thermodynamics [6], introduced the quantum field theory to prove a black hole can really radiate as a thermal object [21]. The radiation that the black hole releases will finally result in the evaporation of the black hole. According to his calculation [22], the black hole will break quantum information. More precisely, no matter what the ingoing state is (pure state or mixed state), the outgoing state is a mixed state. This would certainly break the unitary of the time evolution which is an axiom of quantum mechanics.

Since the original paradox appeared, there have been many variants proposed. A famous one is the "black hole cloning" which break the "No-cloning Theorem". To solve this problem, Susskind et al. ([42], [41]) claim that the information does not only cross through the horizon but is also reflected on the horizon, and more importantly no-one can see the information and its copy at the same time. They argued that quantum mechanics can only be described locally and individually, and this conjecture is historically called the Black Hole Complementarity (BHC). However, a more tricky problem called "Firewall Paradox" (or "Almheiri-Marolf-Polchinski-Sully (AMPS) Paradox") [5] was mentioned 10 years later, which found that the paradox will still be even for a single infalling observer. As well as the discussion of the Firewall paradox is increasing ([27], [44], [34], [35]), the information paradox becomes a hot point again.

In 1993, Don N. Page [33] assumed the unitary of black hole evaporation, and produced the so-called Page curve. This curve indicates a different result from Hawking's calculation, in which the entropy of the Hawking radiation returns to zero finally and hence the information is preserved. Page curve shows us a possible way to solve the information paradox. If one can obtain the Page curve without the assumption of the unitary, the information paradox can be solved.

During the end of the twentieth century, there was another story. The study of string theory led Juan Maldacena [28] to discover the relation between the gauge theory and the quantum gravity theory, which we call the AdS/CFT duality now. Maldacena's remarkable discovery has inspired many theorists to find more evidence to support this idea. In 2006, Shinsei Ryu and Tadashi Takayanagi [40] proposed a formula to calculate the entanglement entropy in a gravitational description, which is called the Holographic Entanglement Entropy (HEE). This formula has later proven to be valid ([26], [15], [32]) and has some improved versions ([24], [18], [17]). As well as holographic entanglement entropy strongly supports AdS/CFT duality, some people find its potential application in information paradox.

A breakthrough appeared in 2019 when Geoffrey Penington [37] stated, by considering the entanglement between the black hole and its radiation, that the contribution to the entropy of the Hawking radiation includes not only the region outside the black hole but also a region inside. This is the so-called "island rule" which explains why the fine-grained entropy of the radiation will return to zero finally and hence the information is preserved.

The main points of this dissertation will focus on introducing holographic entanglement entropy, and how this idea will be helpful to the black hole information loss paradox.

Chapter 1

Preliminary

In this chapter, we will briefly review some background required for the discussion of the HEE and black hole information paradox. This chapter mainly contains three parts. The first part is about black hole thermodynamics. We can find more details about that in [29] or [12]. In the second part, we review some ideas from quantum information. A famous textbook by Nielsen and Chuang in quantum information and quantum computing [31] could be a good reference. The last and the most important part is about AdS/CFT correspondence. There are many useful materials, such as [30], [14] and [25].

1.1 Black Holes

Soon after Albert Einstein discovered his famous equation of general relativity in 1915, Karl Schwarzschild found the first (static vacuum) solution of Einstein's field equation, with a hyper-surface dividing the spacetime into two causally-independent regions. This hyper-surface is called the **event horizon**, of which the internal area is called the **black hole**. The black hole could be the most remarkable creatures in general relativity and has become one of the most popular research objects in theoretical physics.

1.1.1 Kerr-Newman black hole

A Kerr-Newman (KN) black hole is a charged, rotating vacuum solution of the Einstein-Maxwell equations. The metric can be given in Boyer-Lindquist coordinates

$$ds^2 = -\frac{\Delta - a^2\sin^2\theta}{\Sigma}dt^2 - 2a\frac{\sin^2\theta(r^2+a^2-\Delta)}{\Sigma}dtd\phi + \frac{(r^2+a^2)^2 - \Delta a^2\sin^2\theta}{\Sigma}\sin^2\theta d\phi^2 + \frac{\Sigma}{\Delta}dr^2 + \Sigma d\theta^2 \tag{1.1}$$

where

$$\Sigma = r^2 + a^2\cos^2\theta, \qquad \Delta = r^2 - 2Mr + a^2 + Q^2, \qquad a = \frac{J}{M}, \tag{1.2}$$

As shown above, we see there are actually only three parameters to determine the metric – the mass M, the angular momentum J and the electric charge Q [1] .

No-hair Theorem [2]

No-hair theorem states that all stationary solutions of the Einstein-Maxwell equations can be determined by only three independent parameters – mass, angular momentum and electric charge, and all other properties are uniquely determined by these three parameters. One example we have already seen is the KN solution (1.1). This theorem implies that the information of the matter forming or infalling into a black hole will completely disappear, which could be the beginning of the black hole information loss problem. However, it is worth noting that this is a classical result.

1.1.2 Black hole thermodynamics

In 1973, Bardeen, Carter and Hawking [6] found the four laws of black hole mechanics, purely derived from the classical Einstein's field equations:

[1] Sometime there can be four or more finite numbers of parameters (including the magnetic charge), but we only consider three here for simplicity.

[2] Precisely speaking, it is not a theorem but a conjecture since there is still no rigorous mathematical proof for a general case.

- **The Zeroth Law**: Assuming the dominant energy condition, the surface gravity κ is constant over the event horizon of a stationary black hole.

- **The First Law**: The mass M, the area of the event horizon A, the angular momentum J and the electric charge Q satisfy

$$\delta M = \frac{\kappa}{8\pi}\delta A + \Omega_H \delta J + \Phi_H \delta Q. \tag{1.3}$$

  where κ is the surface gravity, Ω_H is the angular velocity and Φ_H is the electric potential.

- **The Second Law**: Assuming the weak energy condition and cosmic censorship, the area of the event horizon A never decreases:

$$\Delta A \geq 0 \tag{1.4}$$

- **The Third Law**: It is impossible to reduce κ to zero by a finite sequence of operations.

We may compere the laws of black hole mechanics with the ordinary laws of thermodynamics [20]:

- **The Zeroth Law**: The Temperature T is constant throughout a system in thermal equilibrium.

- **The First Law**: The internal energy E, the entropy S, the angular momentum J and the total electric charge Q satisfy

$$\delta E = T\delta S + \Omega \delta J + \Phi \delta Q \tag{1.5}$$

  where T is the temperature, Ω is the angular velocity and Φ is electric potential.

- **The Second Law**: In any physical process, the entropy cannot decrease

$$\Delta S \geq 0 \tag{1.6}$$

- **The Third Law**: It is impossible to reduce T to zero by a finite sequence of operations.

We can see the amazing similarity between the laws of black hole dynamics and ordinary thermodynamics, which implies somehow a black hole may behave like a thermal object. It should not be difficult to find the following corresponding relations:

$$T \propto \kappa \qquad \text{and} \qquad S \propto A. \tag{1.7}$$

However, people did not believe this result because there was no evidence showing a classical black hole could radiate just like other thermal objects. In 1975, Hawking [21] showed that black hole can really create and emit particles by quantum mechanical effects, which we call the **Hawking radiation** in the later time. He also found the specific expression of the temperature of the black hole:

$$T_H = \frac{\hbar\kappa}{2\pi k_B c} \tag{1.8}$$

and the (Bekenstein-Hawking) entropy S_{BH} of the black hole:

$$S_{\mathrm{BH}} = \frac{Ac^3}{4G\hbar} \tag{1.9}$$

Generalized second law

Bekenstein [7] mentioned that as a full spacetime containing two parts – the black hole and the outer region, the entropy of the whole system should be

$$S_{\rm gen} = S_{\rm BH} + S_{outside} = \frac{Ac^3}{4G\hbar} + S_{outside} \tag{1.10}$$

This is called the **generalized entropy** and should satisfy the generalized second law:

$$\Delta S_{\rm gen} \geq 0 \tag{1.11}$$

We will later see the idea of generalized entropy inspires the improvements of the holographic entanglement entropy formula.

1.2 Quantum Information

It should be clarified what we mean information before talking about information loss. The information of a quantum system is usually described by an **ensemble** (or a **statistical mixture**) of quantum states, $\{p_i, |\psi_i\rangle\}$, which is a collection of all possible (orthonormal) quantum states that the system can adopt, together with the corresponding probability. If the state of the system is known exactly, e.g. the state is $|\Psi\rangle$ with the corresponding probability $p = 1$ and the probability to adopt other state is zero, we say this system is in a **pure state**; otherwise, the system is in a **mixed state**.

We may act an arbitrary operator $\hat{U}$ on this quantum system,

$$|\psi\rangle \to |\psi'\rangle = \hat{U}\,|\psi\rangle \tag{1.12}$$

then we measure this system again and obtain a new ensemble $\{p'_i, |\psi'_i\rangle\}$. If the probabilities $p_i =$

p'_i, we say the information is preserved under the operator $\hat{U}$, i.e. the probability distribution does not change.

1.2.1 Density Operator

Everyone should be familiar with quantum mechanics described by the state space as above. However, there is an alternate formulation based on the **density operator** (or **density matrix**). For a general ensemble $\{p_i, |\psi_i\rangle\}$, the density operator is defined as

$$\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i|. \tag{1.13}$$

For a system in pure state $|\Psi\rangle$, we can write the density operator as $\rho = |\Psi\rangle \langle\Psi|$. We should note that the eigenvalue of the density operator is the probability,

$$\rho |\psi_i\rangle = p_i |\psi_i\rangle \tag{1.14}$$

and this is the motivation of the definition of the density matrix. A way to characterize pure or mixed states is to calculate $\mathrm{Tr}\rho^2$:

$$\mathrm{Tr}\rho^2 = \sum_i p_i^2 \begin{cases} = (\mathrm{Tr}\rho)^2 = 1 & \text{for pure states} \\ < (\mathrm{Tr}\rho)^2 = 1 & \text{for mixed states} \end{cases} \tag{1.15}$$

The average of any observable $\hat{O}$ can be given in terms of density operator:

$$
\begin{aligned}
\langle \hat{O} \rangle &= \sum_i p_i \langle \psi^i | \hat{O} | \psi^i \rangle \\
&= \sum_{ijk} p_i \langle \psi^i | \psi^j \rangle \langle \psi^j | \hat{O} | \psi^k \rangle \langle \psi^k | \psi^i \rangle \\
&= \sum_{ijk} p_i \langle \psi^k | \psi^i \rangle \langle \psi^i | \psi^j \rangle \langle \psi^j | \hat{O} | \psi^k \rangle \\
&= \sum_{jk} \langle \psi^k | \rho | \psi^j \rangle \langle \psi^j | \hat{O} | \psi^k \rangle \\
&= \sum_{jk} \rho_{kj} \hat{O}_{jk} = \mathrm{Tr}(\rho \hat{O})
\end{aligned}
\tag{1.16}
$$

Let $\{p_i, \psi_i\}$ be a quantum ensemble and an arbitrary transformation $\hat{U}$ acting on it. The density operator transforms as

$$
\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i| \rightarrow \rho' = \sum_i p_i \hat{U} |\psi_i\rangle \langle \psi_i| \hat{U}^\dagger = \hat{U} \rho \hat{U}^\dagger \tag{1.17}
$$

To preserve the information of system, we must have

$$
\langle \psi_i | \rho | \psi_i \rangle = p_i \equiv p_i' = \langle \psi_i' | \rho' | \psi_i' \rangle = \langle \psi_i | \hat{U}^\dagger \hat{U} \rho \hat{U}^\dagger \hat{U} | \psi_i \rangle \quad \Rightarrow \quad \hat{U}^\dagger \hat{U} = \mathbb{1} \tag{1.18}
$$

Hence, information preservation assumes the unitary of the transformation, which is an important postulate of quantum mechanics. If a black hole can be seen as a quantum system from the outside, the loss of information and the unitary of quantum mechanics cannot be both true. This is the reason why people think there is a paradox.

1.2.2 Quantum Statistics

Recall the Gibbs entropy formula from classical statistical mechanics:

$$S_G = -\sum_i p_i \ln p_i \tag{1.19}$$

From (1.14), we can easily extend the definition of entropy into quantum statistics by canonical quantization,

$$S_{\text{vN}}(\rho) = -\text{Tr}(\rho \ln \rho) = -\sum_i p_i \ln p_i \tag{1.20}$$

where p_i are the eigenvalues of the density matrix ρ. This is called the **von Neumann entropy** (or the **quantum entropy**). Here are some important properties of von Neumann entropy:

i. Lower bound: $S_{\text{vN}}(\rho) \geq 0$ and $S_{\text{vN}}(\rho) = 0$ if and only if ρ is pure.

ii. Upper bound: $S_{\text{vN}}(\rho) \leq \ln N$ and $S_{\text{vN}}(\rho) = \ln N$ if and only if ρ is totally disordered (or totally mixed), i.e. $\rho = \frac{1}{N}\mathbb{1}_{N\times N}$.

iii. $S_{\text{vN}}(\rho \otimes \sigma) = S_{\text{vN}}(\rho) + S_{\text{vN}}(\sigma)$.

iv. $S_{\text{vN}}(\hat{U}\rho\hat{U}^\dagger) = S_{\text{vN}}(\rho)$ where $\hat{U}$ is unitary.

v. For non-negative numbers $\{\lambda_i\}$ with $\sum_i \lambda_i = 1$, we have $S_{\text{vN}}(\sum_i \lambda_i \rho_i) \geq \sum_i \lambda_i S_{\text{vN}}(\rho_i)$.

There are two important inequalities of the von Neumann entropy:

(1) Subadditivity:

$$|S(A) - S(B)| \leq S(A \cup B) \leq S(A) + S(B) \tag{1.21}$$

(2) Strong subadditivity:

$$S(A \cup B \cup C) + S(B) \leq S(A \cup B) + S(B \cup C) \tag{1.22}$$

These two inequalities are rather complicated to prove if using typical quantum information methods, but can be easily proven by holographic entanglement entropy. We will see this in the next chapter.

The von Neumann entropy is not easy to calculate in general because of the logarithmic function, hence we usually use the **n-th Renyi entropy**

$$S^{(n)}(\rho) = \frac{1}{1-n}\ln[\mathrm{Tr}(\rho^n)] = \frac{1}{1-n}\ln\left[\sum_i p_i^n\right] \tag{1.23}$$

to evaluate von Neumann entropy. By the analytic continuation $\mathrm{Re}(n) > 1$, we take the limit

$$\begin{aligned}\lim_{n\to 1}\frac{1}{1-n}\ln\left[\sum_i p_i^n\right] &= -\frac{\partial}{\partial n}\ln\left[\sum_i p_i^n\right]\Bigg|_{n=1} \\ &= -\frac{1}{\sum_i p_i^n}\left[\sum_i p_i^n \ln p_i\right]\Bigg|_{n=1} \\ &= -\sum_i p_i \ln p_i = -\mathrm{Tr}(\rho\ln\rho)\end{aligned} \tag{1.24}$$

where we use $\sum_i p_i = 1$ at the last line. Consequently, we have

$$S_{\mathrm{vN}}(\rho) = \lim_{n\to 1} S^{(n)}(\rho) \tag{1.25}$$

This strategy is called the **replica trick** since we make n copies for the density matrix.

Density matrix for canonical ensembles

We calculate the density matrix of canonical ensembles for later use. Assume the canonical ensemble reaches thermodynamic equilibrium, then the entropy S_{vN} is maximized with definite internal energy. We use the method of Lagrange multipliers to find the density matrix:

$$0 = \delta S_{\mathrm{vN}} = \sum_k \delta\rho_{kk}\ln\rho_{kk} + \sum_k \rho_{kk}\delta\ln\rho_{kk} \tag{1.26}$$

There are two constraints:

(1) Fixed Internal energy:

$$\langle H \rangle = \mathrm{Tr}(\rho H) = \sum_k \rho_{kk} E_k \implies \delta \langle H \rangle = \sum_k \delta \rho_{kk} E_k = 0 \tag{1.27}$$

(2) Normalization condition:

$$\mathrm{Tr}\rho = \sum_k \rho_{kk} = 1 \implies \delta(\mathrm{Tr}\rho) = \sum_k \delta \rho_{kk} = 0 \tag{1.28}$$

Introducing Lagrange multipliers β and γ, we define the Lagrange function:

$$\mathcal{L} = S_{\mathrm{vN}} + \beta \langle H \rangle + \gamma \mathrm{Tr}\rho \tag{1.29}$$

then we obtain

$$\begin{aligned} \delta\mathcal{L} = \sum_k \delta\rho_{kk} \left[(\ln \rho_{kk} + 1) + \beta E_k + \gamma \right] = 0 &\implies \rho_{kk} = \exp(-\beta E_k - \gamma - 1) \\ &\implies \rho_{kk} = \frac{\exp(-\beta E_k)}{\sum_i \exp(-\beta E_i)} \end{aligned} \tag{1.30}$$

Recalling the partition function $Z = \sum_k \exp(-\beta E_k) = \mathrm{Tr}(e^{-\beta H})$, we can write the density matrix as

$$\rho = \frac{1}{Z} e^{-\beta H} \tag{1.31}$$

We may find that this density matrix is very similar to the time evolution operator e^{-iHt}, which inspires us to regard β as the imaginary time it. It is just like Wick rotation! If the density matrix has the form of (1.31), we call the system is in a **thermal state**.

Thermal Correlation function

Recall the correlation function defined by

$$\langle \hat{O}(\tau, x)\hat{O}(0,0)\rangle = \mathrm{Tr}\left[\rho \hat{O}(\tau, x)\hat{O}(0,0)\right] \tag{1.32}$$

Assume $\tau > 0$ and $-\beta < \tau < \beta$, then the thermal correlatio function

$$\begin{aligned} G_\beta(\tau, x) &= -\mathrm{Tr}[\rho T_E \hat{O}(\tau, x)\hat{O}(0,0)] \\ &= -\frac{1}{Z}\mathrm{Tr}[e^{-\beta H}\hat{O}(\tau, x)\hat{O}(0,0)] \\ &= -\frac{1}{Z}\mathrm{Tr}[\hat{O}(0,0)e^{-\beta H}\hat{O}(\tau, x)] \\ &= -\frac{1}{Z}\mathrm{Tr}[\hat{O}(0,0)e^{-\beta H}\hat{O}(\tau, x)e^{\beta H}e^{-\beta H}] \\ &= -\frac{1}{Z}\mathrm{Tr}[e^{-\beta H}\hat{O}(0,0)\hat{O}(\tau-\beta, x)] \\ &= -\frac{1}{Z}\mathrm{Tr}[e^{-\beta H}\hat{O}(\tau-\beta, x)\hat{O}(0,0)] = G_\beta(\tau-\beta, x) \end{aligned} \tag{1.33}$$

where T_E is the Euclidean time ordering. We can see the thermal correlation function has a period of the imaginary time β. This is a characteristic for thermal states, and its temperature is given by $T = 1/\beta$.

1.2.3 Entanglement

Entanglement is an essential difference between classical and quantum statistics. A classical ensemble is just a collection of states, and there is no interaction between these states. A quantum ensemble is more than that so that we cannot simply consider these states separately. To better understand the idea of the entanglement, we consider a 2-qubit system as a Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. The basis of this Hilbert space is given by[3]

$$\{|0\rangle_A |0\rangle_B , |0\rangle_A |1\rangle_B , |1\rangle_A |0\rangle_B , |1\rangle_A |1\rangle_B\}$$

Then, we consider two states as follow:

$$\begin{aligned} |\alpha\rangle =& \frac{1}{\sqrt{2}} (|0\rangle_A |1\rangle_B + |1\rangle_A |0\rangle_B) \\ |\beta\rangle =& \frac{1}{\sqrt{2}} (|0\rangle_A |0\rangle_B + |0\rangle_A |1\rangle_B) = |0\rangle_A \otimes \frac{1}{\sqrt{2}} (|0\rangle_B + |1\rangle_B) \end{aligned} \tag{1.34}$$

As shown in (1.34), we cannot write $|\alpha\rangle$ as a product state, which means these two sub-systems A and B are **entangled**. Meanwhile, we can represent $|\beta\rangle$ as a tensor product, which is not entangled.

Because of the entanglement, given a state in the full system $\mathcal{H}$, it is not always possible to find a state in the sub-system which gives the same measurements. To get the information of the sub-system, we define the **reduced density matrix** for sub-system A

$$\rho_A = \mathrm{Tr}_B \rho \tag{1.35}$$

where Tr_B is called a **partial trace** over $\mathcal{H}_B$ meaning taking trace only for states in $\mathcal{H}_B$. Moreover, the von Neumann entropy with respect to the reduced density matrix ρ_A

$$S_A \equiv S(\rho_A) = -\mathrm{Tr}(\rho_A \ln \rho_A) \tag{1.36}$$

is called the **entanglement entropy** of the sub-system A.

Again, we take $|\alpha\rangle$ as an example. The reduced density matrix for sub-system A is given

[3]We use $|a\rangle_A |b\rangle_B = |a\rangle_A \otimes |b\rangle_B$

by

$$\rho_A = \mathrm{Tr}_B\rho = \frac{1}{2}\left(|0\rangle_A\langle 0|_A + |1\rangle_A\langle 1|_A\right) = \frac{1}{2}\mathbb{1}_{2\times 2} \tag{1.37}$$

which is totally mixed. We can see that although the total system is in a pure state, the subsystem A is in a (maximal) disordered state. It tells us that complete knowledge of the total system doesn't lead to complete knowledge of the subsystem. This is the biggest difference between classical and quantum system, and it comes from entanglement.

Moreover, we calculate the entanglement entropy

$$S_A = -\mathrm{Tr}(\rho_A \ln \rho_A) = \ln 2 = -\mathrm{Tr}(\rho_B \ln \rho_B) = S_B \tag{1.38}$$

which means there are 2 qubits are entangled. Hence, we can use the entanglement entropy to measure entanglement.

For pure state, e.g. $\rho = |\alpha\rangle\langle\alpha|$, the von Neumann entropy is zero $S_{\mathrm{vN}}(\rho) = 0$ and the entanglement entropy is non-zero, $S_A = S_B = \ln 2 \neq 0$. Hence, the entanglement entropy provides a good measurement of entanglement for pure states, and the information is stored in entanglement between A and B. Moreover, as we have seen from the above example, we can prove that the entanglement entropy $S_A = S_B$ for pure states.

For mixed state, the entanglement entropy is no longer a good measure of entanglement since it mixes quantum and classical correlations. For example, we consider a separable system $\tilde{\rho} = \rho_A \otimes \rho_B$ which is obviously not entangled. However, we can find that the entanglement entropy is still non-zero $S_A = S_B \neq 0$ but the von Neumann entropy of the total system $S_{\mathrm{vN}}(\rho) = S(A \cup B) = S_A + S_B$. Instead, we use **mutual information** to measure entanglement, which is defined by

$$I(A;B) = S(A) + S(B) - S(A \cup B) \tag{1.39}$$

For pure state, we can find $I(A;B) \neq 0$. For non-entangled mixed states, e.g. $\tilde{\rho} = \rho_A \otimes \rho_B$, the mutual entanglement is zero $\tilde{I}(A;B) = 0$ and there is no (classical and quantum) correlation.

1.2.4 Fine-grained entropy vs Coarse-grained entropy

We have seen that the von Neumann entropy or the entanglement entropy is invariant under the unitary time evolution and does not satisfy the second law of thermodynamics. Obviously, it is not the ordinary thermal entropy. This is a motivation to distinguish two kinds of entropy:

- **Fine-grained entropy**: It is just another name of von Neumann entropy, which is invariant under the unitary evolution.

- **Coarse-grained entropy**: When we don't know exactly the density matrix ρ of the system, but only know the measurement of a subset of physical observable O_i. Then we find all possible density matrices $\tilde{\rho}$ such that $\langle O_i \rangle = \mathrm{Tr}(\tilde{\rho} O_i)$. The maximum von Neumann entropy $S_{\mathrm{vN}}(\tilde{\rho})$ over all $\tilde{\rho}$ is chosen to be the coarse-grained entropy. The coarse-grained entropy obeys the second law of thermodynamics

According to these definitions, it is easy to find

$$S_{\text{fine}} \leq S_{\text{coarse}} \tag{1.40}$$

because we use $\tilde{\rho}$ as a candidate of ρ which provides a upper bound.

In summary, the coarse-grained entropy is a statistical quantity, which is like a cheap version of the fine-grained entropy. The ordinary thermodynamics entropy is coarse-grained, since we fix the internal energy or the volume and then maximize the entropy. Beikenstein-Hawking entropy is also a coarse-grained entropy because it increases as time. Now a natural question is what is the fine-grained entropy for a black hole? As we will see later, it is given by the quantum extremal surface.

1.3 Conformal Field Theory

Roughly speaking, a Conformal Field Theory (CFT) is a quantum field theory invariant under conformal transformations, containing Poincaré, dilation and special conformal transformations.

1.3.1 Conformal Group

Consider the metric tensor $g_{\mu\nu}$ in a d-dimensional spacetime. A conformal transformation of the coordinates is an invertible mapping $x \to x'$ such that

$$g'_{\mu\nu}(x') = \Omega^2(x) g_{\mu\nu}(x) \tag{1.41}$$

for some non-zero function $\Omega(x)$. In other words, a conformal transformation is a local dilatation. As we can check, conformal transformations form a group that we call the **conformal group**. One special case is when $\Omega(x) \equiv \pm 1$, the conformal group reduces into the Poincaré group.

For simplicity, we assume the spacetime is flat, i.e. $g_{\mu\nu} = \eta_{\mu\nu} = \mathrm{diag}(-1, 1, \ldots, 1)$. It can be proved that, for $d > 2$, the most general infinitesimal conformal transformation $x^\mu \to x'^\mu = x^\mu + \epsilon^\mu(x)$ obeys [14]

$$\epsilon^\mu(x) = a^\mu + \alpha x^\mu + M^\mu{}_\nu x^\nu + 2(x \cdot b)x^\mu - b^\mu x^2 \tag{1.42}$$

where we can identify each term into a kind of transformations (See Table 1.1).

We can redefine these generators (See Table 1.1) as following:

$$\begin{array}{ll} J_{\mu\nu} = M_{\mu\nu} & J_{(d+1)d} = D \\ J_{(d+1)\mu} = \frac{1}{2}(P_\mu - K_\mu) & J_{d\mu} = \frac{1}{2}(P_\mu + K_\mu) \end{array} \tag{1.43}$$

	Transformations	Generators
(translation)	$x'^\mu = x^\mu + a^\mu$	$P_\mu = -i\partial_\mu$
(dilation)	$x'^\mu = \alpha x^\mu$	$D = -ix^\mu\partial_\mu$
(rotation)	$x'^\mu = M^\mu{}_\nu x^\mu$	$M_{\mu\nu} = i(x_\mu\partial_\nu - x_\nu\partial_\mu)$
(SCT)	$x'^\mu = \frac{x^\mu - b^\mu x^2}{1-2b\cdot x+b^2x^2}$	$K_\mu = -i(2x_\mu x^\nu\partial_\nu - x^2\partial_\mu)$

Table 1.1: Conformal transformations and the generators

then the new generators obey the **Conformal algebra**:

$$[J_{ab}, J_{cd}] = i(\tilde{\eta}_{ad}J_{bc} + \tilde{\eta}_{bc}J_{ad} - \tilde{\eta}_{ac}J_{bd} - \tilde{\eta}_{bd}J_{ac}) \tag{1.44}$$

where $\tilde{\eta}_{ab} = \mathrm{diag}(-1,-1,1,\dots,1)$. As we can check, the conformal algebra is isomorphic to $\mathfrak{so}(2,d)$.

1.3.2 Primary operators

According to the conformal algebra, we can see that P_μ and K_μ can be considered as the raising and lowing operators with respect to D.

$$[D, P_\mu] = -iP_\mu, \qquad [D, K_\mu] = iK_\mu \tag{1.45}$$

It is natural to define the ground state of D,

$$K_\mu \left|\phi_0\right\rangle = 0 \tag{1.46}$$

which is annihilated by K_μ. The ground state is not unique, and we call all these states the **primary states**. We can also define the **primary operator** $\Phi(x)$ such that

$$\Phi(x)\left|0\right\rangle = \left|\phi_0(x)\right\rangle \tag{1.47}$$

According to the definition of the primary state, the primary operator should obeys

$$[D, \Phi(0)] = -i\Delta\Phi(0), \qquad [K_\mu, \Phi(0)] = 0 \tag{1.48}$$

where Δ is called the **scaling dimension** of the operator Φ.

1.3.3 Correlation function

The conformal symmetry provides a powerful restriction on quantum fields. As a consequence, the form of the correlation function in CFT is rather simple,

$$\langle \Phi_1(\lambda x_1) \cdots \Phi_n(\lambda x_n) \rangle = \lambda^{-\Delta_1 - \Delta_2 - \cdots - \Delta_n} \langle \Phi_1(x_1) \cdots \Phi_n(x_n) \rangle \tag{1.49}$$

Specially the two-point function has the form,

$$\langle \Phi_1(x_1)\Phi_2(x_2) \rangle = \frac{c_{12}}{|x_1 - x_2|^{2\Delta}}, \quad \Delta_1 = \Delta_2 \equiv \Delta \tag{1.50}$$

where c_{12} is constant and it must vanish if $\Delta_1 \neq \Delta_2$.

1.4 Anti-de Sitter Spacetime

A (d+1)-dimensional Anti-de Sitter (AdS) Spacetime[4] is a maximally symmetric space with an negative constant curvature, which is a solution of Einstein's equations with negative cosmological constant:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = -\Lambda g_{\mu\nu} \tag{1.51}$$

where $\Lambda = -\frac{d(d-1)}{2\ell^2}$ and ℓ is called the **AdS radius**.

A nice way to think of AdS_{d+1} is to embed it into a (d+2)-dimensional Minkowski spacetime

[4] We consider the (d+1)-dimensional case here for later convenience.

M_{d+2} as a hyperboloid sub-manifold:

$$T_1^2 + T_2^2 - X^i X^i = \ell^2, \qquad i = 1, \dots, d. \tag{1.52}$$

An important point is that the Minkowski spacetime M_{d+2} has the signature $(2, d)$, which means it has two time-like coordinates rather than one. Hence, the metric for M_{d+2} is

$$d\tilde{s}^2 = -dT_1^2 - dT_2^2 + dX^i dX^i \tag{1.53}$$

We could also see that AdS_{d+1} has an isometry group, $SO(2, d)$, from the symmetry of the Minkowski spacetime M_{d+2}, which has the same symmetry with CFT_d. It is an evidence of holographic principle.

To obtain the induced metric for AdS_{d+1}, it would be better using the global coordinates as follow:

$$\begin{cases} T_1 = \ell \cosh\rho \cos\tau \\ T_2 = \ell \cosh\rho \sin\tau \\ X^i = \ell \sinh\rho \Omega_i \end{cases} \tag{1.54}$$

where Ω_i are the parameters for a (d-1)-sphere, i.e. $\sum_i \Omega_i^2 = 1$. By taking (1.54) into (1.53), it is not hard to find the **global metric** for AdS_{d+1}:

$$ds^2 = \ell^2(-\cosh^2\rho d\tau^2 + d\rho^2 + \sinh^2\rho d\Omega_{d-1}^2) \tag{1.55}$$

where $d\Omega_{d-1}^2$ is the metric of (d-1)-Sphere and $d\Omega_{d-1} = d\theta^2 + \sin^2\theta d\Omega_{d-2}^2$. This metric has only one time-like coordinate as expected, and we should also note that the time coordinate τ has a natural period of 2π, but we still define τ in real number, $\tau \in \mathbb{R}$.

For better understanding of AdS spacetime, we introduce some other useful coordinate systems for AdS Spacetime:

(1) **Static coordinates**: Introducing $r = \ell \sinh\rho$ and $t = \ell\tau$ in (1.55), we have

$$ds^2 = -f dt^2 + \frac{1}{f} dr^2 + r^2 d\Omega_{d-1}^2 \tag{1.56}$$

where $f = 1 + \frac{r^2}{\ell^2}$. Its components are independent of time, and hence it is static. It also has a similar structure with the Schwarschild metric. Since it covers the whole spacetime, it is also called the global coordinates sometimes.

(2) **Conformal coordinates**: Introducing $\tan\theta = \ell \sinh\rho$, we have

$$ds^2 = \frac{\ell^2}{\cos^2\theta}\left(-d\tau^2 + d\theta^2 + \sin^2\theta d\Omega_{d-1}^2\right) \tag{1.57}$$

where $\frac{\ell^2}{\cos^2\theta}$ is a conformal factor. It is easy to see that this metric is conformal related to the metric of a cylinder $\mathbb{R} \times S^d$, and hence has the same causal structure.

(3) **Poincaré coordinates**: Introducing the new coordinate system:

$$\begin{aligned} t &= \ell \frac{\sin\tau}{\cos\tau - \Omega_d \sin\rho}, \\ z &= \ell \frac{\cos\rho}{\cos\tau - \Omega_d \sin\rho}, \\ x^i &= \ell \frac{\Omega_i \sin\rho}{\cos\tau - \Omega_d \sin\rho}, \end{aligned} \tag{1.58}$$

then we have the metric

$$ds^2 = \frac{\ell^2}{z^2}(-dt^2 + dx^i dx^i + dz^2) \tag{1.59}$$

which only covers a small region of the AdS spacetime, called the **Poincaré patch**. This patch is bounded by a causal diamond, so we call it "Poincaré".

In many cases, we prefer to consider an **asymptotically AdS spacetime** with a conformal boundary. While approaching its boundary, it has the same geometry with the AdS geometry.

1.5 AdS/CFT Duality

The AdS/CFT correspondence could be one of the most famous examples of the holographic principle [8]. This duality was first discovered by Juan Maldacena [28] when he studied the low-energy limit of brane systems in string theory. The exact description hasn't been found yet, but a general one can be given by following [19]:

Statement. *Any relativistic conformal field theory on $\mathbb{R} \times S^{d-1}$ can be interpreted as a theory of quantum gravity in an asymptotically $AdS_{d+1} \times M$ spacetime. Here M is some compact manifold that may or may not be trivial.*

More roughly speaking, the CFT is living on the boundary of the AdS spacetime. Now a natural question is how to relate these two descriptions. The answer is using the "dictionary". There are two main dictionaries. One is the **BDHW dictionary** (or **extrapolate dictionary**), of which the basic idea is that we can "push" an operator in AdS to its boundary to obtain an operator in CFT. The other one is called the **GKP-W dictionary**, which can be simply interpreted as

$$Z_{\text{CFT}} = Z_{\text{AdS+fields}} \tag{1.60}$$

where the partition function of the CFT should be equal to the one of the AdS plus the fields on it. We should notice that the AdS/CFT duality is telling there are two descriptions for the same things, the AdS description or the CFT one. We cannot use both descriptions at the same time!

Chapter 2

Entanglement Entropy in AdS/CFT

In this chapter, we will introduce the general methods to calculate the entanglement entropy in CFT and the holographic entanglement entropy. More precisely, we will focus on examples in AdS_3/CFT_2. More detailed discussions in entanglement entropy and quantum field theory can be found in [10],[11] and [13]. For higher dimensional cases of the holographic entanglement entropy, a textbook by Rangamani and Takayanagi [38] would be very helpful, so would a shorter review by Ryu and Takayanagi [39]. We will also use the path integral representation for operators, which has been introduced in Appendix A. It would be very helpful if the reader views Appendix A before this chapter.

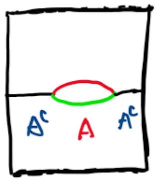
2.1 Entanglement Entropy in Conformal Field Theory

A general calculation of the entanglement entropy is based on the replica trick (1.25) and here we will use the following:

$$S_A = -\mathrm{Tr}(\rho_A \ln \rho_A) = -\frac{\partial}{\partial n}\mathrm{Tr}(\rho_A^n)\bigg|_{n=1} \tag{2.1}$$

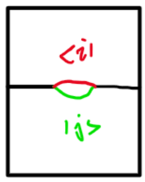
So our first task is to evaluate $\mathrm{Tr}(\rho_A^n)$. According to Appendix A, the reduced density matrix over a sub-system A can be represented by

$$\rho_A = \mathrm{Tr}_{A^c}\rho = \tag{2.2}$$



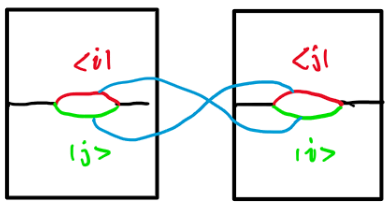
where we glue the cut of the region other than A together and leave the region A open. Its element can be obtained by restricting the boundaries,

$$(\rho_A)_{ij} = \langle i|\rho_A|j\rangle = \tag{2.3}$$



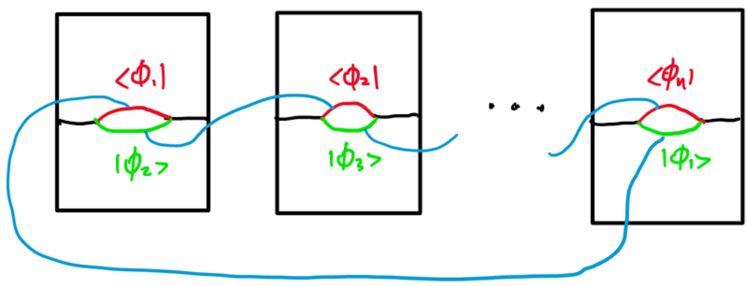
Hence, the traces of ρ_A^n are given as the partition function $Z_n(A)$ on a n-sheeted Riemann surface $\mathcal{R}_n$:

$$Z_2(A) \equiv \mathrm{Tr}\rho_A^2 = \sum_{i,j} \langle i|\rho_A|j\rangle \langle j|\rho_A|i\rangle =$$



$$\cdots \tag{2.4}$$

$$Z_n(A) \equiv \mathrm{Tr}\rho_A^n =$$



where we glue the boundaries of the same index by taking traces (See the blue lines). As a result, $Z_n(A)$ is given by path integral on n-sheeted Riemann surface $\mathcal{R}_n$ with boundary conditions

$$\phi_m(\tau = 0^-, x) = \phi_{m+1}(\tau = 0^+, x), \quad \phi_n(\tau = 0^-, x) = \phi_1(\tau = 0^+, x), \quad 1 \le m \le n \tag{2.5}$$

where $x \in [u, v]$. Then for a normalized reduced density matrix ρ_A, we have $\mathrm{Tr}\rho_A^n = \frac{Z_n(A)}{Z^n}$. Specially for $n = 1$, we have $\mathrm{Tr}\rho_A = 1$. Hence, the replica trick gives

$$S_A = -\lim_{n\to 1} \frac{\partial}{\partial n} \frac{Z_n(A)}{Z^n} \tag{2.6}$$

It could be very complicated to calculate $Z_n(A)$ on such a n-sheeted Riemann surface $\mathcal{R}_n$. Hence, the next step is to simplify the calculation by taking a conformal transformation that maps the n-sheeted Riemann surface $\mathcal{R}_n$ to a single Riemann surface $\mathbb{C}$, such that

$$\begin{aligned} Z_n(A) &= \int [D\phi]_{\mathcal{R}_n} \exp\left[-\int_{\mathcal{R}_n} dx d\tau \mathcal{L}[\phi(\tau, x)]\right] \\ \equiv Z_{\mathbb{C}}(A) &= \int [D\phi_1 D\phi_2 \cdots D\phi_n]_{\mathbb{C}} \exp\left[-\int_{\mathbb{C}} dx d\tau (\mathcal{L}[\phi_1(\tau, x)] + \cdots + \mathcal{L}[\phi_n(\tau, x)])\right] \end{aligned} \tag{2.7}$$

with boundary conditions

$$\phi_i(\tau = 0^+, x) = \phi_{i+1}(\tau = o^-, x), \quad x \in [u, v], \quad i = 1, \ldots, n \tag{2.8}$$

where the original fields ϕ on each sheet are mapped onto a single surface $\mathbb{C}$ and hence we have n fields. By considering these n fields together as a field with n components on a complex plane $\mathbb{C}$, we can finally represent $Z_n(A)$ in terms of some correlation functions (See [10] and [11] for details) and then the entanglement entropy can be calculated.

In general, the calculation of entanglement entropy in CFT is still complicated and we cannot obtain an exact result in many cases, even if we follow the previous steps. It is common to use numerical methods to calculate them, but we do have precise results for some simple cases. For example, we consider the entanglement entropy for the vacuum state in CFT_2 at a single interval. Suppose A is a single interval of length L_A. Then the entanglement entropy is given exactly by

$$S_A = \frac{c}{3} \log \frac{L_A}{\epsilon} \tag{2.9}$$

where c is the central charge of the CFT and ϵ is the UV cut-off because this entropy is actually divergent. We will later see this result can be also obtained from the geometry of the dual AdS spacetime.

2.2 Holographic Entanglement Entropy

2.2.1 HRT Formula and Quantum Extremal Surface

In 2006, after checking in many numerical computations, Shinsei Ryu and Tadashi Takayanagi [40] made an argument that the entanglement entropy in a (d+1)-dimensional CFT can be obtained from the area of d-dimensional minimal surfaces in AdS_{d+2}, in a way similar to the Bekenstein-Hawking formula for black hole entropy. Soon later, Hubeny, Rangamani and Takayanagi extended the time-independent case that Ryu and Takayanagi considered into the time-dependent one [24], which is summarized as the so-called Hubeny-Rangamani-Takayanagi (HRT) formula (2.10).

Consider A a spatial sub-region of the boundary of some asymptotically-AdS geometry, where the CFT is living (See Fig 2.1). The entanglement entropy for the sub-system A is given by

$$S_A = \frac{\mathrm{Area}(\gamma_A)}{4G} \tag{2.10}$$

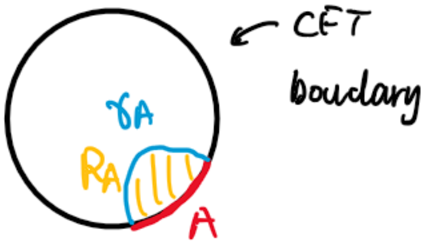


Figure 2.1: A spatial sub-region A on the conformal boundary of some asymptotically AdS geometry. γ_A is a the RT-surface. $\mathcal{R}_A$ is the region bounded by A and γ_A.

with minimal surface γ_A that extremes the area. As we can see, this formula (2.10) is very similar with the Bekenstein-Hawking formula (1.9). The surface γ_A is called **RT-surface** with the following requirements:

- co-dimension two, i.e. γ_A has two fewer dimensions than AdS space-time,
- space-like surface with external area in AdS,
- anchored to the AdS boundary, i.e. $\partial\gamma_A = \partial A$,
- homologous (continuous deformable) to A,
- If there are several external surfaces, pick the one with minimal area.

The HRT formula is original a hypothesis and then proven to be valid in many cases (See [26], [15], [32]).

Some improvements of HEE formula

After the HRT formula was proposed, people made a sequence of improvements. In 2013, Lewkowycz and Maldacena [26] added the contribution of the fields in the area $\mathcal{R}_A$, where $\mathcal{R}_A$ is the area bounded by a region A and a surface γ_A (See Fig 2.1),

$$S(A) = \frac{\text{Area}(\gamma_A)}{4G} + S_{\text{field}}(\mathcal{R}_A) \tag{2.11}$$

with minimal surface γ_A that extremes its area. This formula provides a quantum corrected version to HRT formula, and after that Faulkner, Lewkowycz and Maldacena [18] argue it in a higher order.

Engelhardt and Wall [17] improved it further, which is valid at arbitrary order.

$$S(A) = \frac{\text{Area}(\gamma_A)}{4G} + S_{\text{field}}(\mathcal{R}_A) \tag{2.12}$$

with minimal surface γ_A that extremes $\left[\frac{\text{Area}(\gamma_A)}{4G} + S_{\text{field}}(\mathcal{R}_A)\right]$. We should notice that the surface γ_A in (2.12) is different from the one in (2.10) and (2.11), because they extremes different objects. This improved surface (2.12) is called the **quantum extremal surface** (Quantum Extremal Surface (QES)) and the corresponding entropy (2.12) is the **generalized entropy** of $\mathcal{R}_A$. It is a generalization of the generalized entropy of black hole.

In later discussion, we will assume the stationary RT-formula (2.10), and use it to calculate HEE for some examples.

How to find external surface in AdS$_{d+1}$?

Following the HEE formula (2.10), a natural question is how to find the RT-surface γ_A in AdS$_{d+1}$? Suppose the induced metric of γ_A is given by

$$ds_\gamma^2 = h_{\mu\nu} dx^\mu dx^\nu, \quad \mu, \nu = 0, 1, \ldots, d-2 \tag{2.13}$$

then of which the area is

$$\text{Area}(\gamma_A) = \int \sqrt{h} d^{d-1}x \tag{2.14}$$

where h is the determinant of the metric $h_{\mu\nu}$, i.e. $h = \det h_{\mu\nu}$.

Assume the metric of the asymptotic AdS_{d+1} is given by

$$ds^2 = g_{AB} dX^A dX^B, A, B = 0, \ldots, d \tag{2.15}$$

and $X(x^\mu)$ is the extremal surface. we can write the induced metric as

$$ds_\gamma^2 = g_{AB} \frac{\partial X^A}{\partial x^\mu} \frac{\partial X^B}{\partial x^\nu} dx^\mu dx^\nu = h_{\mu\nu} dx^\mu dx^\nu \tag{2.16}$$

So the area is a functional of $X(x^\mu)$, and we can use Euler-Lagrange equations with boundary condition $\partial\gamma_A = \partial A$ to find the external surface. Now we have a general method to find the RT-surface. Let us see the following examples.

2.2.2 Example: Vacuum state in (1+1)-dimensional CFT in $\mathbb{R}^{1,1}$

Consider a 2D CFT in vacuum and a region A be an internal of length L_A. Without the loss of generality, we choose

$$x \in \left[-\frac{L_A}{2}, -\frac{L_A}{2}\right]. \tag{2.17}$$

It is dual to Poincaré AdS_3,

$$ds^2 = \frac{\ell_{\mathrm{AdS}}^2}{z^2}(-dt^2 + dx^2 + dz^2) \tag{2.18}$$

with the boundary at $z = 0$. In the static case, there is a natural time coordinate t and hence γ_A will live in a Cauchy slice (i.e. fixed-t slice) according to the time-translation invariance. Since the vacuum state is static, we can use the time translation invariance and choose a fixed-t surface, namely $t = 0$. Then we have the reduced metric

$$ds^2 = \frac{\ell_{\mathrm{AdS}}^2}{z^2}(dx^2 + dz^2) \tag{2.19}$$

The Area (or Length in this case) of the RT-surface is given by

$$\text{Length}(\gamma_A) = \int_{\gamma_A} ds = \int_{\gamma_A} \frac{\ell_{\text{AdS}}}{z}\sqrt{dx^2 + dz^2} = \ell_{\text{AdS}} \int_{-L_A/2}^{L_A/2} dx \frac{1}{z}\sqrt{1 + z'^2} \tag{2.20}$$

where $z' = \frac{dz}{dx}$. As mentioned in Section 2.2.1, the length of γ_A is a functional of $z(x)$ and we use the method of variation to find the curve $z(x)$ extremize Length(γ_A). Let the Lagrange function $\mathcal{L} = \frac{1}{z}\sqrt{1+z'^2}$, the the Euler-Lagrange equation is

$$0 = \frac{\partial \mathcal{L}}{\partial z} - \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial z'} = -\frac{1}{z^2}\sqrt{1+z'^2} - \frac{d}{dx}\frac{z'}{z\sqrt{1+z'^2}} \tag{2.21}$$

By solving (2.21), we find the geodesic is actually a semi-circle shown in Fig 2.2, which satisfies

$$z^2 + x^2 = (\frac{L_A}{2})^2 \tag{2.22}$$

Taking (2.22) back to (2.20), we will find this integral is divergent. Hence, we need to introduce

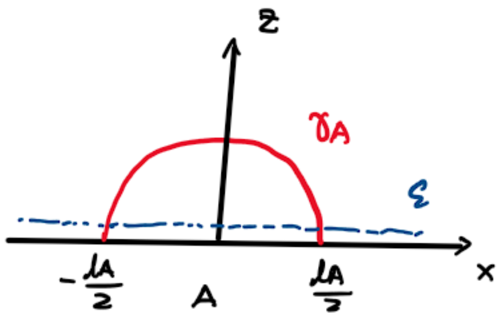


Figure 2.2: The CFT is living on a spatial line, and we choose a sub-region $A = [-L_A/2, L_A/2]$. The RT-surface γ_A is a semi-circle (See red curve). Since the length of γ_A is divergent, we have to set a cut-off (See blue line).

a UV cutoff ϵ and obtain

$$\text{Length}(\gamma_A) = 2\ell_{\text{AdS}} \log(\frac{L_A}{\epsilon}) \tag{2.23}$$

Hence, the holographic entanglement entropy is given by

$$S_A = \frac{\text{Length}(\gamma_A)}{4G} = \frac{\ell_{\text{AdS}}}{2G}\log(\frac{L_A}{\epsilon}) \tag{2.24}$$

Recalling the fact in $\text{AdS}_3/\text{CFT}_2$ the the central charge for the CFT is given by

$$c = \frac{3\ell_{\text{AdS}}}{2G} \tag{2.25}$$

We can write the HEE as following,

$$S_A = \frac{c}{3}\log\frac{L_A}{\epsilon} \tag{2.26}$$

and this agrees with the result in the CFT case (2.9).

2.2.3 Holographic proof of sub-additivity and strong sub-additivity

As mentioned before, the proofs of the subadditivity and strong subadditivity are rather difficult in quantum mechanics, but they are quite easy and obvious from the viewing of holographic principle! Firstly, we prove the subadditivity (1.21) and again we consider the lowest dimension case shown in Fig 2.3. As we can see, the length of the black curve is shorter than the sum of the lengths of the red and green curves, since the black curve is a geodesic. According to the RT formula (2.10), it is easy to show

$$S(A\cup B) \leq S(A) + S(B) \tag{2.27}$$

and then the subadditivity (1.21) is proven. Similarly, we can prove the strong subadditivity (1.22) from Fig 2.4. As we can see, the green curve is shorter than the black one, as well as the

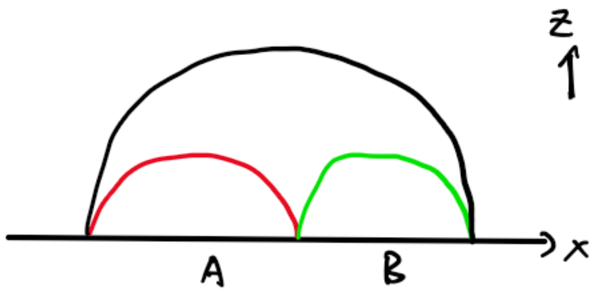


Figure 2.3: Holographic proof of strong subadditivity: According to the RT formula, the black curve represents $S_{A\cup B}$, the red one is S_A and the green one is S_B.

red one is shorter than the blue one. Hence, we have

$$\text{Length}(\gamma_B)+\text{Length}(\gamma_{ABC}) = \text{red}+\text{green} \leq \text{blue}+\text{black} = \text{Length}(\gamma_{AB})+\text{Length}(\gamma_{BC}) \quad (2.28)$$

and the strong subadditivity is well proven. This elegant proof was first provided by Headrick

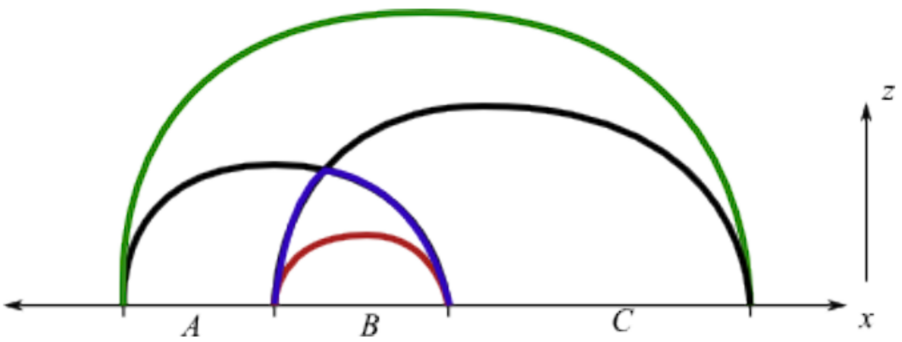


Figure 2.4: Holographic proof of strong subadditivity [20]: the green curve is S_{ABC}, the black curve plus the blue curve is $S_{AB} + S_{BC}$, and the red curve is S_B

and Takayangi in 2007 [23], which values the holographic principle as a powerful tool.

2.2.4 Holographic Entanglement Entropy at finite size

Consider a CFT$_2$ on $\mathbb{R} \times S^1_L$ describing the vacuum state on compact space, $x \in S^1_L$, where L is the perimeter of the circle. This is dual to static AdS$_3$ (See Fig 2.5) with the metric,

$$ds^2 = -(1+\frac{r^2}{\ell^2_{\text{AdS}}})dt^2 + \frac{1}{1+\frac{r^2}{\ell^2_{\text{AdS}}}}dr^2 + r^2 d\varphi^2 \quad (2.29)$$

We consider the region A with $\varphi \in [-\varphi_A, \varphi_A]$ and choose the Cauchy surface $\Sigma : t = 0$ (See

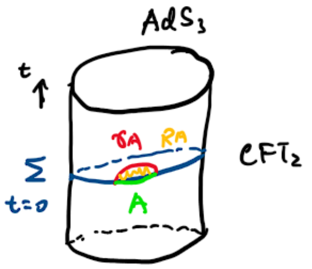


Figure 2.5: $\mathrm{AdS}_3/\mathrm{CFT}_2$

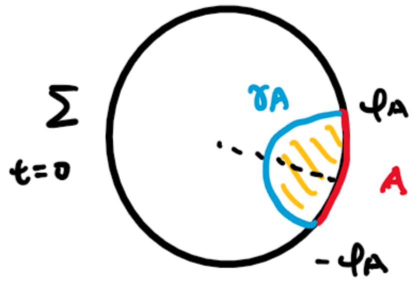


Figure 2.6: Cauchy slice $t = 0$ in $\mathrm{AdS}_3/\mathrm{CFT}_2$

Fig 2.6). Following the general method mentioned above, we can find that

$$r(\varphi) = \ell_{\mathrm{AdS}} \left(\frac{\cos^2 \varphi}{\cos^2 \varphi_A} - 1 \right)^{-\frac{1}{2}} \tag{2.30}$$

is geodesic at $t = 0$ and the holographic entanglement entropy for this case is

$$S_A = \frac{c}{3} \log \left(\frac{L}{\pi \epsilon} \sin \left(\frac{\pi L_A}{L} \right) \right) \tag{2.31}$$

where $c = \frac{3\ell_{\mathrm{AdS}}}{2G}$ is the central charge of the CFT and L_A is length of A.

If we consider the non-compact limit, i.e. L is sufficiently large such that $L_A \ll L$, we have

$$S_A = \frac{c}{3} \log \left(\frac{L}{\pi \epsilon} \sin \left(\frac{\pi L_A}{L} \right) \right) \approx \frac{c}{3} \log \left(\frac{L}{\pi \epsilon} \cdot \frac{\pi L_A}{L} \right) = \frac{c}{3} \log \left(\frac{L_A}{\epsilon} \right) \tag{2.32}$$

which reproduces the result for the vacuum CFT_2 we have already obtained above (2.26).

2.2.5 Holographic Entanglement Entropy at finite temperature

We move to the case that the CFT is defined at an infinite size but at a finite temperature. For example, we consider a CFT_2 on $S^1_\beta \times \mathbb{R}$, where S^1_β shows that it has a period β of the imaginary time. Hence, it indeed describes a thermal state on non-compact space $x \in \mathbb{R}$ and is dual to the planar Bañados-Teitelboim-Zanelli (BTZ) black hole with metric:

$$ds^2 = -\frac{r^2 - r_+^2}{\ell_{\text{AdS}}^2}dt^2 + \frac{1}{r^2 - r_+^2}dr^2 + \frac{r^2}{\ell_{\text{AdS}}^2}dx^2 \tag{2.33}$$

Again, we consider a sub-region $A = [-L_A/2, L_A/2]$. Following a similar process with Section 2.2.2, We can find the extremal surface satisfies

$$\frac{dr}{dx} = \frac{r}{\ell_{\text{AdS}}^2}\sqrt{(r^2 - r_+^2)\left(\frac{r^2}{r_*^2} - 1\right)}, \quad r_* = r_+ \coth\left(\frac{L_A}{2}r_+\right) \tag{2.34}$$

where $x \in (-L_A/2, L_A/2)$. The HEE can be finally obtained by

$$S_A = \frac{c}{3}\log\left(\frac{\beta}{\pi\epsilon}\sinh\left(\frac{\pi L_A}{\beta}\right)\right) \tag{2.35}$$

where $\beta = \frac{2\pi\ell_{\text{AdS}}^2}{r_+}$ and $c = \frac{3\ell_{\text{AdS}}}{2G}$.

In the low temperature limit, $\beta \to \infty$, we have

$$S_A = \frac{c}{3}\log\left(\frac{\beta}{\pi\epsilon}\sinh\left(\frac{\pi L_A}{\beta}\right)\right) \approx \frac{c}{3}\log\left(\frac{\beta}{\pi\epsilon}\cdot\frac{\pi L_A}{\beta}\right) = \frac{c}{3}\log\left(\frac{L_A}{\epsilon}\right) \tag{2.36}$$

which comes bake to the result in the non-compact vacuum state (2.26).

2.2.6 Holographic Entanglement Entropy at finite size and finite temperature

Now we consider a more complex case with both finite size and finite temperature, e.g. a CFT_2 on a torus $S_\beta^1 \times S_L^1$. It describes thermal state on S_L^1 and is dual to global BTZ black hole. The metric of the global BTZ black hole is given by

$$ds^2 = -\frac{r^2 - r_+^2}{\ell_{\text{AdS}}} dt^2 + \frac{\ell_{\text{AdS}}}{r^2 - r_+^2} dr^2 + r^2 d\varphi^2 \tag{2.37}$$

which has a horizon at $r = r^+$. Without loss of generality, we set $\ell_{\text{AdS}} = 1$ and the radius of the boundary circle $R = 1$, hence the perimeter $L = 2\pi R = 2\pi$.

Since the metric is static, we choose a Cauchy slice as usual, e.g. $t = 0$. Then, the metric reduces to

$$ds^2 = \frac{1}{r^2 - r_+^2} dr^2 + r^2 d\varphi^2 \tag{2.38}$$

However, we should find that there will be two different homotopy types for the extremal surface sharing boundary with A (shown in Fig 2.7) because of the existence of the BTZ black hole.

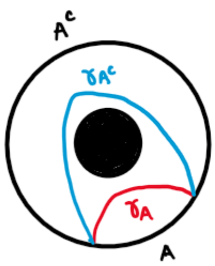


Figure 2.7: Two homotopy types of the extremal surfaces (See red and blue curves) because of the existence of the black hole.

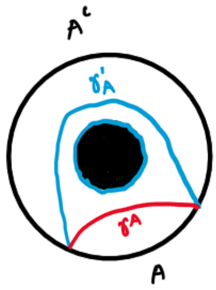


Figure 2.8: Two kinds of extremal surfaces (See red and blue curves) homologous to A. Note the blue one is disconnected.

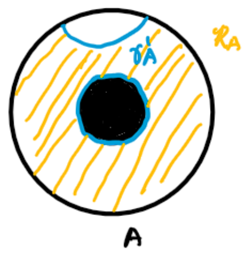


Figure 2.9: When A is large compered with the whole system, $\mathcal{R}_A$ covers almost every parts except the black hole.

As we can see from Fig 2.7, γ_A is homologous to A but γ_{A^c} is not. Alternatively, we consider a disconnected surface (See Fig 2.8)

$$\gamma'_A = \gamma_{A^c} \cup \mathcal{H} \tag{2.39}$$

where $\mathcal{H}$ is the horizon of the BTZ black hole. We can verify that γ'_A is homologous to A since $\partial\mathcal{R}_A = \gamma'_A \cup A$ (See Fig 2.9). Hence, γ'_A satisfies all the conditions of the RT-surface. According to the rule of the RT-surface, we should choose the minimal surface,

$$S_A = \min \left\{ \frac{\text{Area}(\gamma_A)}{4G}, \frac{\text{Area}(\gamma'_A)}{4G} \right\} \tag{2.40}$$

Hence, there are two cases worthwhile to consider:

I. If A is sufficiently small compared to the whole system (See Fig 2.8) such that the length of γ_A is smaller than of γ'_A, then γ_A is chosen to be the RT-surface, of which the function is given by

$$r(\varphi) = r_+ \left(1 - \frac{\cosh^2(r_+\varphi)}{\cosh^2(r_+\varphi_A)}\right)^{-\frac{1}{2}} \tag{2.41}$$

and the HEE is

$$S_A = \frac{\text{Length}(\gamma_A)}{4G} = \frac{c}{3}\log\left[\frac{\beta}{\pi\epsilon}\sinh\left(\frac{L\varphi_A}{\beta}\right)\right] = \frac{c}{3}\log\left[\frac{\beta}{\pi\epsilon}\sinh\left(\frac{2\pi R\varphi_A}{\beta}\right)\right] \tag{2.42}$$

If taking the limit $R \to \infty$ while keep the length of A, $L_A = 2R\varphi_A$, fixed, then we get

$$S_A = \frac{c}{3}\log\left[\frac{\beta}{\pi\epsilon}\sinh\left(\frac{\pi L_A}{\beta}\right)\right] \tag{2.43}$$

which gives the infinite size limit and returns to the above result at non-compact size but at a finite temperature (2.35).

II. If A is sufficiently large compared to the whole system (See Fig 2.9) such that the length of γ_A is larger than of γ'_A, then γ'_A is chosen to be the RT-surface. The HEE can be verified to be

$$S_A = \frac{c}{3}\log\left[\frac{\beta}{\pi\epsilon}\sinh\left(\frac{R}{\beta}(\pi - \varphi_A)\right)\right] + \frac{c}{3}\pi r_+ \tag{2.44}$$

where γ_{A^c} contributes to the first term and the horizon $\mathcal{H}$ contributes to the second term

[1].

As the size of the region A increasing, there is a critical point φ_A^* that transitions between Case I (2.42) and Case II (2.44), which is explicitly given by

$$\varphi_A^*(r_+) = \frac{1}{r_+}\coth^{-1}(2\coth(\pi r_+) - 1) \tag{2.45}$$

and depends on the value of r_+.

We can plot the graph of the holographic entanglement entropy changing with the size of the sub-region (shown in Fig 2.10). For whom is familiar with information loss problem, this

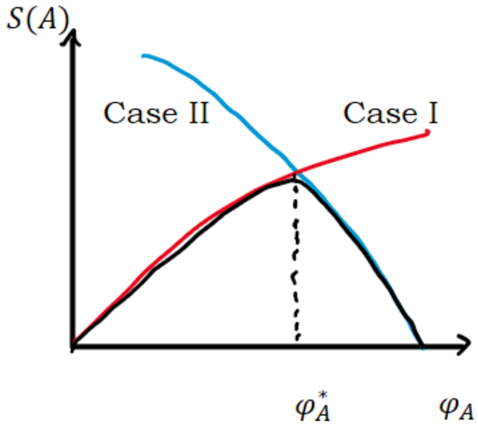


Figure 2.10: The blue curve shows the entropy by the connected extremal surface, namely small A. The red curve shows the entropy by the disconnected extremal surface, namely large A. S_A has a critical point at $\varphi_A = \varphi_A^*$ because of the choice of the RT-surface.

curve is very similar to the Page curve (See Fig 3.1). This is one of the motivations of the island rule (3.12).

In the case of $\varphi_A = \pi$, the region A actually covers the entire CFT, and then we have $S_A = S_{\mathrm{CFT}}^{thermal}$. On the other hand, the γ_{A^c} contribution just vanishes in this case and only the Bekenstein-Hawking contribution remains $S_A = S_{\mathrm{BH}}$. Hence, a natural outcome gives

$$S_{\mathrm{BH}} = S_A = S_{\mathrm{CFT}}^{thermal} \tag{2.46}$$

[1]The second term is just the Bekenstein-Hawking entropy of the BTZ black hole.

Chapter 3

Black Hole Information Paradox

Hawking showed that a pure state black hole [1] will evaluate into a thermal state, which breaks the unitarity of quantum mechanics [22]. This is the main contradiction of the black hole information paradox. In this chapter, we will review some important milestones in information paradox. Then, we will introduce some latest discoveries about the island rule and replica wormholes.

3.1 Unitary Evaporation

First, we review the whole evolution of a black hole. At the beginning, the black hole is formed from the gravitational collapse. The collapse happens so fast that the area of the horizon increases very rapidly, during which there is almost no radiation. Once the black hole appears, it will emit thermal radiation, and its mass gradually decreases over time. The original pure state is split into two parts, the black hole and its radiation. Obviously, the entropy of radiation is non-zero because of the entanglement. Typically, after waiting for a long time, the black hole will evaporate completely and the Hawking radiation is fully emitted.

According to Hawking's calculation, the black hole evaporates as well as more and more

[1] Here we mean a black hole from the collapse of a pure state and we will only consider this kind of black hole in this chapter.

radiation is emitted. Hence, the entropy of the radiation increases until the black hole evaporates completely (shown as the blue curve in Fig 3.1), and we cannot read information from such a mixed state. In 1993, by assuming the unitary of the evaporation, Page [33] obtained a different curve (See the black curve in Fig 3.1).

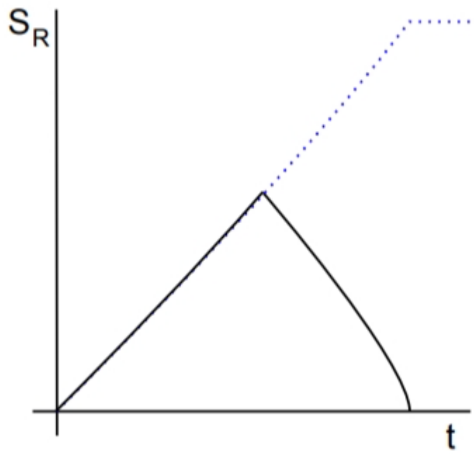


Figure 3.1: [] Shown as the blue curve, Hawking indicated that the entropy of the Hawking radiation increases as time and reaches a maximum when the black hole fully evaporates. On the other hand, the unitary of the evaporation leads to the entropy increasing but then decreasing back to zero finally (shown as the black curve). This critical time is called the **Page time**.

To derive the Page curve, we introduce a useful theorem first.

Theorem (Page Theorem)**.** *For any bipartite system $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$, we have*

$$\int d\mathcal{U} \left\| \rho_A - \frac{\mathbb{1}}{|A|} \right\|_1 \leq \sqrt{\frac{|A|^2 - 1}{|A||B| + 1}} \tag{3.1}$$

where $d\mathcal{U}$ is the Haar measure, $|\cdot|$ is the dimension of the Hilbert space and $\|\cdot\|_1$ is L_1 norm defined by $\|M\|_1 = Tr\sqrt{M^\dagger M}$.

An important corollary of the Page theorem is following.

Corollary. *If $1 < |A| \ll |B|$, then we have*

$$\int d\mathcal{U} \left\| \rho_A - \frac{\mathbb{1}}{|A|} \right\|_1 \approx \sqrt{\frac{|A|}{|B|}} \approx 0 \tag{3.2}$$

In other words, if the degrees of the freedom of the sub-system A is relatively small enough, A is very closed to a micro canonical ensemble.

At early time when the black hole just forms, there is almost no radiation, hence we have the degrees of freedom of the radiation is much smaller than the black hole, $|\text{Rad}| \ll |\text{BH}|$. According to the above corollary, we obtain

$$S_{fine}(\text{Rad}) \approx \ln|\text{Rad}| \equiv S_{coarse}(\text{Rad}) \propto tT \tag{3.3}$$

Also because of $\ln|\text{Rad}| \ll \ln|\text{BH}|$, we have

$$S_{coarse}(\text{Rad}) \ll S_{coarse}(\text{BH}) \tag{3.4}$$

As the evaporation process proceeds, the mass of the black hole gradually decreases, as well as the Hawking radiation increases. Hence, there will be a moment we call the **Page time** when $S_{coarse}(\text{Rad}) \approx S_{coarse}(\text{BH})$.

Consider the moment long after the Page time, we have $|\text{Rad}| \gg |\text{BH}|$, then again by the corollary, the fine-grained entropy of the black hole is given by

$$S_{fine}(\text{BH}) \approx S_{coarse}(\text{BH}) = \frac{A}{4G} \approx 0 \tag{3.5}$$

since the black hole almost evaporates completely. For a pure state black hole, we should also have

$$S_{fine}(\text{Rad}) = S_{fine}(\text{BH}) \approx 0 \tag{3.6}$$

Hence, the fine-grained entropy of the radiation is back to zero, and we obtain the Page curve. The importance of the Page curve is that it points us a practical way to solve the information paradox. If we can obtain Page curve without the assumption of the unitary, we solve the paradox.

3.2 Complementarity, Firewall and Central Dogma

3.2.1 Black Hole Complementarity and Firewall

As mentioned in Section 1.2, if we assume the unitary of the evaporation, the information of the in-falling objects will be preserved in the Hawking radiation. However, this assumption will cause a contradiction with the quantum "no-cloning" theorem.

Theorem ("No-cloning" theorem)**.** *Consider a Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. There is no unitary operator such that*

$$U \left|\psi\right\rangle_A \left|0\right\rangle_B \left|\phi\right\rangle_C = \left|\psi\right\rangle_A \left|\psi\right\rangle_B \left|\phi\right\rangle_C \tag{3.7}$$

Briefly speaking, the "no-cloning" theorem tells us that it is impossible to copy the quantum information. The only way to send the quantum information is to lose it. Let us see an example to explaining why black hole results quantum cloning.

Consider an observer Alice (denoted by "A") on a "nice"[2] slice Σ_1, carrying some quantum information, jumps into the black hole. In a later time (shown as the "nice" slice Σ_2 in Fig 3.2), another observer Bob (denoted by "B") can recover Alice's information from the radiation. However, we may notice that Alice's information will appear on Σ_2 at both A' and B if we assume the evaporation to be unitary. This is what we call the **black hole cloning**, and it is conflict with the "no-cloning" theorem.

Someone may think the horizon splits spacetime into two causal independent regions, so no observer can see both copies at the same time. This could be right. If Bob does want to see Alice's copy, he has to jump into the black hole as well. We wish he is also lucky enough to

[2]Here "nice" means small curvatures such that the semi-classical approximation is valid.

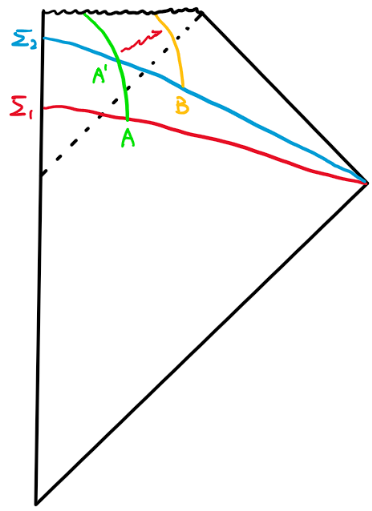


Figure 3.2: A on a "nice" slice Σ_1 is infalling into the black hole, carrying some quantum information. After a long time, Another observer B on "nice" slice Σ_2 can recover the information of A from the Hawking radiation, as well as the copy of the information A' appears on the same Cauchy slice

find Alice, or receive the message that Alice sent to him (See the red arrow in Fig 3.2), before he collides the singularity. As a result, he will obtain two copies of the information and breaks "no-cloning" theorem again. However, according to Susskind et al. [42], Alice will never have enough time to send her message. She must send the message within the Planck time, and it is impossible!

To solve this problem, Susskind, Thorlacius and Uglum [42] argued that no single observer can see both A' and B, based on the following three postulates:

- **Postulate 1**: The process of formation and evaporation of a black hole, as viewed by a distant observer, can be described entirely within the context of standard quantum theory. In particular, there exists a unitary S-matrix which describes the evolution from infalling matter to outgoing Hawking-like radiation.

- **Postulate 2**: Outside the stretched horizon of a massive black hole, physics can be described to good approximation by a set of semi-classical field equations.

- **Postulate 3**: To a distant observer, a black hole appears to be a quantum system with discrete energy levels. The dimension of the subspace of the states describing a black hole of mass M is the exponential of the Bekenstein entropy $S(M)$.

We may notice this is a fact that unitary quantum mechanics evolution on "nice slice" is contradicted with "unitary" of the black hole evaporation, whether it can be observed or not. However, we can give an new interpretation. If nobody can see both copies on a "nice" slice, why should such a slice exist? As a result, Susskind et al. ([42], [41]) proposed the so-called **black hole complementarity**, which states that quantum mechanics only need to describe the experience of individual observer, who are appropriately restricted by causality in what they do.

We may compare it to Bohr's complementarity principle or Heisenberg uncertainty principle, which are essential parts of the consistency of quantum mechanics. Black hole complementarity could also be an essential part of the consistency of some quantum gravity theory.

However, like what happened in the past, not every people satisfy with this interpretation. In 2013, Almheiri, Marolf, Polchinski and Sully [5] made an argument against the black hole complementarity, which is now called the **AMPS paradox** or **"Firewall problem**. They introduced the fourth postulate:

Postulate 4: *A freely infalling observer will experience nothing out of ordinary when crossing the horizon.*

which comes from the Rindler approximation near the horizon. We can find for an infalling observer, the horizon is nothing but just Minkowski spacetime. The basic idea of the AMPS paradox is that the fourth postulate is contradicted with the first three postulates of the BHC (especially the first postulate), and this will lead to the entanglement-monogamy problem. After all, the "Firewall" paradox is still an open question, and it is not directly related to our main topic so we will not talk about it in detail here.

3.2.2 Central Dogma

The three postulates of the BHC are summarized into one single idea called the **central dogma** [3], which states as follow:

Statement (Central Dogma). *As seen from the outside, a black hole can be described in terms of a quantum system with Area/$(4G_N)$ degrees of freedom, which evolves unitarily under time evolution.*

That is, we can replace a black hole with some spacetime around (up to some cut-off surface) with a quantum system, which evolves unitarily. The central dogma is still a controversial topic till now. For example, Unruh and Wald [43] think the first postulate can be given up with but the fourth postulate can be adopted. We should emphasize that the information loss problem is a paradox only if the central dogma is true.

3.3 The entropy of Hawking Radiation

3.3.1 Island Rule

According to the discussions we have already had, it should be an natural idea to use the holographic entanglement entropy formula (2.12) to calculate the fine-grained entropy of the black hole ([1], [37], [4]),

$$S_{fine}(\mathrm{BH}) = \min_X \left\{ \mathrm{ext}_X \left[\frac{\mathrm{Area}(X)}{4G} + S_{\text{field}}(\Sigma_X) \right] \right\} \tag{3.8}$$

where X is the QES and can be in the black hole interior, Σ_X is the region between the QES and some UV cut-off surface (See Fig 3.8). According to this formula (3.8), the fine-grained entropy of the black hole relies on its interior. Although some black holes are seen the same from the outside, the different black hole has a different interior and hence have a different fine-grained entropy. We find that there are at least two extremal surfaces for a pure state black hole. One is near the horizon $X \approx \mathcal{H}$ and the other is near the singularity $X \approx \varnothing$. Hence, the Bekenstein-Hawking entropy $S_{coarse}(\mathrm{BH})$ may not be equal to $S_{fine}(\mathrm{BH})$ since the horizon may not minimize $S_{fine}(\mathrm{BH})$, and this will result the Page curve for the black hole. Let us do a simple check.

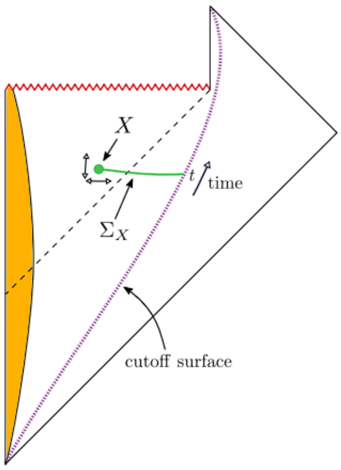


Figure 3.3: [3] The fine-grained entropy of the black hole depends on the green region Σ_X between the QES X and some cut-off surface.

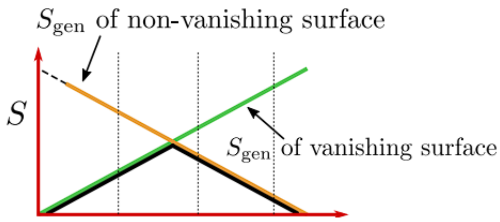


Figure 3.4: [3] The generalised entropy of the vanishing surface $X \approx \varnothing$ increases (See green curve), as well as the one of the non-vanishing surface decreases (See yellow curve). By considering the minimum, we obtain the Page curve for the black hole (See black curve).

At early time when the black hole just forms, the QES should be the one near the singularity $X \approx \varnothing$ for a pure state black hole. Since the area of X nearly vanishes and there is no radiation emitted yet, we have the fine-grained entropy is close to zero,

$$S_{fine}(\mathrm{BH}) = S_{\text{gen}}(X \approx \varnothing) \approx S_{\text{field}}(\Sigma_X) \approx 0 \tag{3.9}$$

While more radiation is escaping out of the cut-off surface, the entropy $S_{\text{field}}(\Sigma_X)$ is increasing because of the entanglement with the radiation within the cut-off surface (shown as the green line in Fig 3.4).

At the ending stage of black hole evaporation, the QES is near the horizon $X \approx \mathcal{H}$. Since the QES is very close to the horizon and the region Σ_X becomes significantly small, the contribution of the fields on Σ_X almost vanishes. Hence, we have

$$S_{fine}(\text{BH}) = S_{\text{gen}}(X \approx \mathcal{H}) \approx \frac{\text{Area}(\mathcal{H})}{4G} \approx S_{coarse}(\text{BH}) \tag{3.10}$$

As the evaporation is processing, the area of the horizon decreases, and hence the coarse-grained entropy of the black hole deceases (shown as the yellow line in Fig 3.4). After all, the fine-grained entropy of the black hole is chosen to be the minimum of these two extremal surfaces,

$$S_{fine}(\text{BH}) = \min\{S_{\text{gen}}(X \approx \varnothing), S_{\text{gen}}(X \approx \mathcal{H})\} \tag{3.11}$$

Hence, it satisfies the Page curve (shown as the black curve in Fig 3.4). However, the story is not complete yet because what we need is the Page curve for the Hawking radiation.

Since there is either the black hole or the radiation in the spacetime, a natural guess is that all regions other than Σ_X contribute to the fine-grained entropy of the radiation, and these regions may be disconnected [3]. Hence, we give a similar formula for the Hawking radiation (See also Fig 3.5):

$$S_{fine}(\text{Rad}) = \min_X \left\{ \text{ext}_X \left[\frac{\text{Area}(X)}{4G} + S_{\text{field}}[\Sigma_{\text{Rad}} \cup \Sigma_{\text{Island}}] \right] \right\} \tag{3.12}$$

where X is the extremal surface that extremes $\left[\frac{\text{Area}(X)}{4G} + S_{\text{field}}[\Sigma_{\text{Rad}} \cup \Sigma_{\text{Island}}]\right]$. The region Σ_{Island} is called the **island**. Moreover, we should differ this with the coarse-grained entropy of radiation,

$$S_{coarse}(\text{Rad}) = S_{\text{field}}(\Sigma_{\text{Rad}}) \tag{3.13}$$

[3] We have seen this case before in the example of the global BTZ black hole (See Section 2.2.6).

Again, we check the fine-grained entropy for the radiation follows the Page curve.

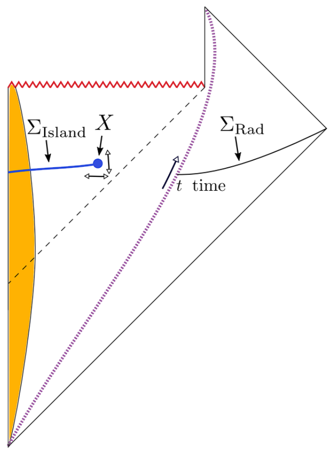


Figure 3.5: [3] There are two disconnected regions contributing to $S_{fine}(\text{Rad})$. The blue one is in the black hole interior called the island, the other (See black curve) is out of some cut-off surface.

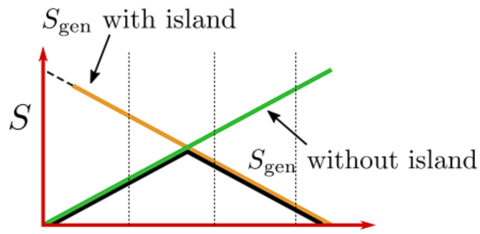


Figure 3.6: [3]

Similarly to our discussion for the evaporating black hole, we have $X \approx \varnothing$ at early time, and obviously the island vanishes as well $\Sigma_{\text{Island}} \approx \varnothing$. In this case, the fine-grained entropy for the radiation is approximately equal to the thermal entropy outside the cut-off surface,

$$S_{fine}(\text{Rad}) \approx S_{coarse}(\text{Rad}) \tag{3.14}$$

and again it increases over time since the radiation escapes out of the cut-off surface (See the green line in Fig 3.6). At later time when the horizon minimize the fine-grained entropy, i.e. $X \approx \mathcal{H}$ and the black hole almost evaporates fully $\text{Area}(\mathcal{H}) \approx 0$, we have

$$S_{fine}(\text{Rad}) \approx S_{\text{field}}[\Sigma_{\text{Rad}} \cup \Sigma_{\text{Island}}] \tag{3.15}$$

and we can prove that $S_{\text{field}}[\Sigma_{\text{Rad}} \cup \Sigma_{\text{Island}}] \approx 0$ since the island provides a purification of the outgoing radiation (See the yellow line in Fig 3.6). Consequently, we recovers the Page curve for the Hawking radiation (shown as the black curve in Fig 3.6).

3.3.2 Entanglement Wedge Reconstruction

Now we need to make some clarification on the central dogma (3.2.2). What do we mean the degrees of freedom of the black hole here? We should emphasize that the degrees of freedom of the black hole are not equal to those seen from the outside. An evidence comes from the island rule.

We notice the the fine-grained entropy formula for the black hole (3.8) only depends on the region Σ_X (See Fig 3.3). This means that all the information of the black hole is stored in the region Σ_X. Hence, not every part of the black hole interior contributes to the degrees of the freedom of the black hole, but to the degrees of the freedom of the radiation, i.e. the island belongs to the the degrees of the freedom of the radiation.

Moreover, we can prove that $S_{\text{field}}(\Sigma_1) = S_{\text{field}}(\Sigma_2)$ if Σ_1 and Σ_2 are causally related (or they are in the same causal wedge). Hence, the information is actually stored in a area called the **entanglement wedge** (See Fig 3.7) since two causal related Cauchy surfaces carry the same information.

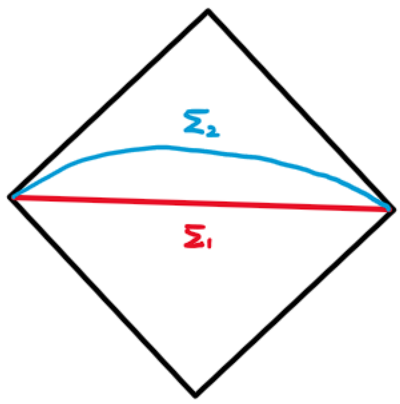
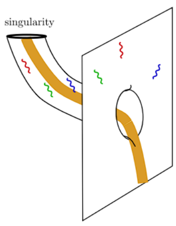


Figure 3.7: Σ_1 and Σ_2 are two partial Cauchy surfaces, sharing the same domain of dependence. They carry the same information which is stored in this causal diamond called the entanglement wedge

A more precise statement is called the "entanglement wedge reconstruction hypothesis" [3], which says that if we have a relatively small number of qubits in an unknown state but located inside the entanglement wedge of the black hole, then by preforming operations on the black hole degrees of freedom, we can read off the state of the qubits.

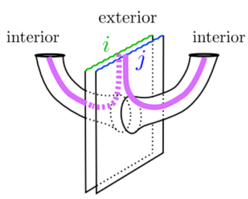
3.3.3 Replica wormholes

After the island rule is discovered, a derivation similar to the replica trick (1.25) explains further why the island rule is reasonable ([36], [2]). First, we consider the evaporating black hole that is topologically equivalent to the following pure state,

$$|\Psi\rangle =$$ 

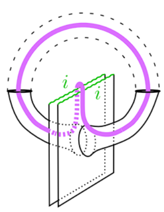
(3.16)

The entry of the density matrix $\rho = |\Psi\rangle\langle\Psi|$ is given by adding boundary states on the exterior,

$$\rho_{ij} = \langle i|\Psi\rangle\langle\Psi|j\rangle =$$ 

(3.17)

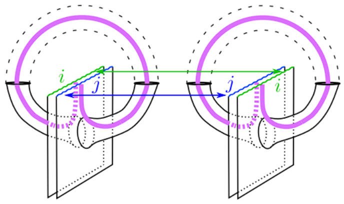
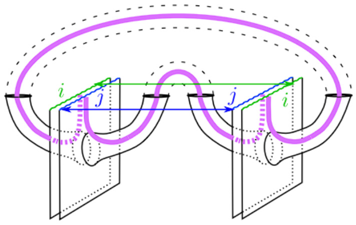
By taking the trace of the density matrix, we glues the boundaries.

$$\mathrm{Tr}\rho = \sum_i \langle i|\Psi\rangle\langle\Psi|i\rangle =$$ 

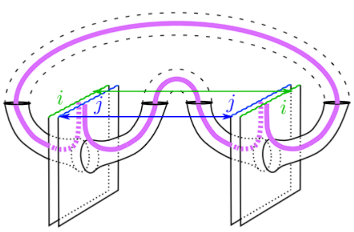
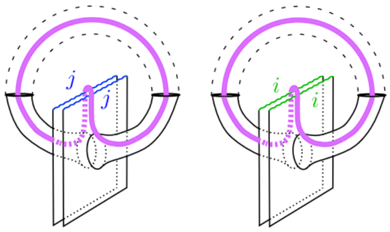
(3.18)

However, when we consider $\mathrm{Tr}(\rho^2)$, there are two cases for connecting the interiors:

$$\mathrm{Tr}(\rho^2) = \sum_{i,j} \langle i|\Psi\rangle \langle \Psi|j\rangle \langle j|\Psi\rangle \langle \Psi|i\rangle$$

$=$  $+$  $+ \cdots$ (3.19)

where the first term is called the **Hawking saddle** and the second term is called the **replica wormhole saddle**. We can find that the replica wormhole saddle is just $(\mathrm{Tr}\rho)^2$,

 $=$  (3.20)

If the Hawking saddle is dominant, we have $\mathrm{Tr}(\rho^2) < (\mathrm{Tr}\rho)^2$. Hence, the radiation is in a mixed state. If the replica wormhole saddle is dominant, we have $\mathrm{Tr}(\rho^2) \approx (\mathrm{Tr}\rho)^2$, which is satisfied only for pure states. So the radiation is in a pure state. In this way, we explain why the state of the radiation will return pure at the ending stage of the evaporation, and this derivation strongly supports the island rule.

Chapter 4

Discussion

4.1 Conclusion

We have introduced what is the holographic entanglement entropy and seen many examples. Then we discussed the black hole as a quantum system. Next, we applied the idea of holographic entanglement entropy to the information paradox and used the island rule to recover the Page curve of the Hawking radiation.

After all the discussion we have had, we may be able to solve one question now. What is the missing part in Hawking's original calculation? The answer is the black hole interior. Hawking only considered the exterior and obtained the coarse-grained entropy of the radiation. Because of the disorder from the entanglement with a part of the black hole's interior, the entropy seen from the outside is non-zero. Therefore, the information seems to be lost. In other words, some part of the information is stored within the black hole interior, namely the island. By including the contribution of the island, we obtain all the information again.

4.2 Open Questions

Although the information paradox seems to be solved, it is actually far from that. First, our analysis is still qualitative. We know the information is preserved, but we don't know the details of this process and how to decode the information from Hawking radiation. For example, how to choose the cut-off surface? What is the exact state of the black hole? Second, since we know there is entanglement between the black hole and its radiation, is there any choice to break the entanglement? How can we do so and what if we break the entanglement?

There are some even more fundamental problems with this solution. We haven't had an exact description for AdS/CFT correspondence. How much can we believe in the AdS/CFT tool? Moreover, there is no experimental evidence for string theory till now. A more depressing fact is that we have not observed Hawking radiation from any experiment so far. The whole story seems to be imaginary. After all, our theory of quantum mechanics is still incomplete. We should really ask ourselves. Do we really have the solution? Or even do we really have a paradox?

4.3 To the Infinite

The last part is for sci-fi fans. Here we assume the information paradox is well solved. We know exactly what happened during the evaporation of the black hole and how to decode the information from the radiation. We may ask is there any potential application of the black hole? One of the main topics of this dissertation is the quantum entanglement. Based on this principle, we can at least improve two kinds of technologies – quantum encryption and quantum computer.

We already know quantum entanglement plays an important role in these areas. One of the technological difficulties is how to stabilize the entangled states. As we have seen, there exists entanglement between the black hole and its radiation. Suppose this entanglement is rather stable because of its geometric nature. First, we know that we cannot decode the information from the radiation until waiting for a period of time called the "scrambling time". It is sufficient to imagine some kind of quantum encryption based on this idea. A similar idea is provided in the latest paper by Brakerski [9]. Second, we can use (micro) black holes to build quantum computers. A recent paper also talk about this possibility [16]. A black hole is also a good storage of information because if its storage capacity. A simple calculation shows that

$$\text{Area} \sim S \sim \log N \tag{4.1}$$

where N is the degrees of the freedom (we may simply consider it as the number of bits). This relation indicates that we only need a relatively small space to store an extremely large number of data. Finally, this may be the answer to many questions, such as the Fermi paradox. It is possible that the information supporting the existence of advanced extraterrestrial life is very common in our universe. However, it is encoded by some technology based on black holes and we cannot read it, so we simply think they don't exist!

Certainly, creating micro black holes is not an easy task, at least for now, but we believe crazy technology is born from crazy theory!

Appendix A

Path integral representations for states and operators

We introduce a method to represent quantum states or operators by path integral. This appendix is mainly based on the lecture notes by Thomas Hartman [20]. In this dissertation, this representation is used in the entanglement entropy in CFT (See Section 2.1) and replica wormholes (See Section 3.3.3).

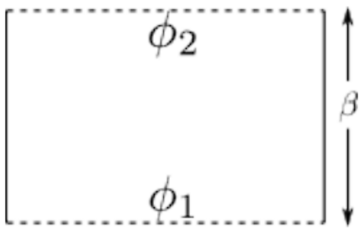
A.1 Transition amplitudes

Transition amplitudes measure the probability of a system transforming from one state to another. A transition amplitude under evolution by $e^{-\beta H}$ can be defined by a Euclidean path integral:

$$\langle \phi_2 | e^{-\beta H} | \phi_1 \rangle = \int_{\phi(\tau=0)=\phi_1}^{\phi(\tau=\beta)=\phi_2} D\phi e^{S_E[\phi]} \tag{A.1}$$

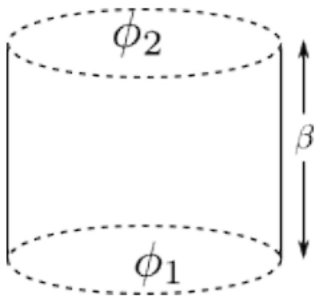
which we can regard as a ordinary integral over some manifold with boundary conditions at the time coordinate, and this integral depends on the topology of the manifold. For example, if the space is a line (or a plane for higher dimensions), the path integral is defined on $\mathbb{R}^{d-1} \times L_\beta$, where L_β is an internal of length β,

$$\langle\phi_2|e^{-\beta H}|\phi_1\rangle = \tag{A.2}$$



or similarly, if the space is a circle (or a sphere for higher dimensions), the path integral is defined on $S^{d-1} \times L_\beta$,

$$\langle\phi_2|e^{-\beta H}|\phi_1\rangle = \tag{A.3}$$



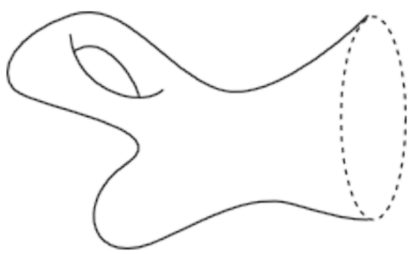
A.2 Wave functional

A wave function can be defined as a path integral with an open cut

$$|\Psi\rangle = |\phi_1(\tau)\rangle = e^{-\tau H}\,|\phi_1\rangle = \int_{\phi(\tau=0)=\phi_1}^{\phi(\tau=\beta)=??} D\phi e^{S_E[\phi]} = \tag{A.4}$$

More generally, we can consider any path integral with an open cut Σ as a quantum state on Σ, and the topology can be rather complex. For example,

$$|X\rangle = \tag{A.5}$$



A.2.1 Ground State

Consider a state $|A\rangle$ expanded in terms of energy eigenstates:

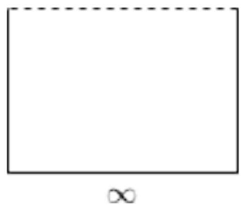
$$|A\rangle = \sum_n a_n |n\rangle , \qquad H|n\rangle = E_n |n\rangle \tag{A.6}$$

Then taking $|A\rangle$ to evolve over a long Euclidean time (i.e. $\tau \to \infty$), we have

$$\lim_{\tau\to\infty} e^{-\tau H} |A\rangle = \lim_{\tau\to\infty} \sum_n a_n e^{-\tau E_n} |n\rangle \approx e^{-\tau E_0} a_0 |0\rangle \quad \Rightarrow \quad |0\rangle \propto \lim_{\tau\to\infty} e^{-\tau H} |A\rangle \tag{A.7}$$

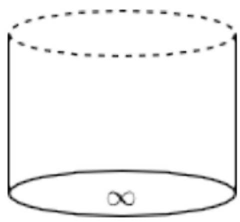
According to (A.4), the ground state on a line can be represented as

$$|0\rangle_{line} \propto \int_{\phi(\tau=-\infty)=0}^{\phi(\tau=0)=??} D\phi e^{-S_E} = \tag{A.8}$$



which is a path integral on a semi-infinite plane, with an open cut at the edge. Similarly, for the ground state on a circle,

$$|0\rangle_{circle} = \tag{A.9}$$

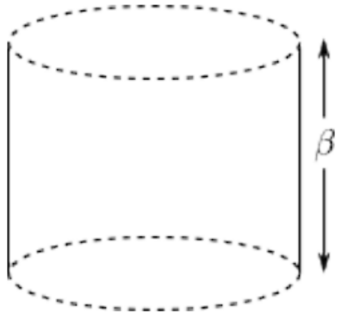


which is a semi-infinite Euclidean cylinder.

A.3 Density Operator

Recalling the definition of the density operator (1.13), we find it is defined by ket states and bra states, and hence can be represented as a path integral with two open cuts. For example, we consider a pure state on a circle, $\rho = |\Psi\rangle \langle\Psi|$,

$$\rho = e^{-\beta H} = \tag{A.10}$$

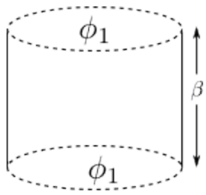


where we glue the boundaries of the ket state and the bra state and leave two open cuts.

A.4 Thermal Partition Function

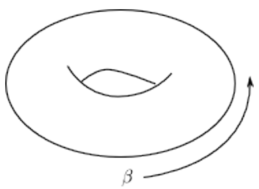
For a thermal state at temperature $T = 1/\beta$, the thermal partition function is given by

$$Z(\beta) = \mathrm{Tr} e^{-\beta H} = \sum_i \langle\phi_i|e^{-\beta H}|\phi_i\rangle = \sum_i \tag{A.11}$$



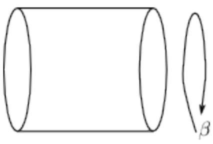
By summing over i, we need to glue the top and the bottom of the cylinder together. As a result, the thermal partition function on a circle is equivalent to a path integral on a torus:

$$Z(\beta)_{circle} = \tag{A.12}$$



Similarly, the thermal partition function on a line is equivalent to a path integral on an infinitely long cylinder of period β:

$$Z(\beta)_{line} = \tag{A.13}$$



List of Acronyms

QFT Quantum Field Theory

HEE Holographic Entanglement Entropy

CFT Conformal Field Theory

AdS Anti-de Sitter

BTZ Bañados-Teitelboim-Zanelli

KN Kerr-Newman

QES Quantum Extremal Surface

HRT Hubeny-Rangamani-Takayanagi

BHC Black Hole Complementarity

AMPS Almheiri-Marolf-Polchinski-Sully

Bibliography

[1] Ahmed Almheiri, Netta Engelhardt, Donald Marolf, and Henry Maxfield. The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole. *Journal of High Energy Physics*, 2019(12):63, December 2019. arXiv:1905.08762 [gr-qc, physics:hep-th].

[2] Ahmed Almheiri, Thomas Hartman, Juan Maldacena, Edgar Shaghoulian, and Amirhossein Tajdini. Replica Wormholes and the Entropy of Hawking Radiation. *Journal of High Energy Physics*, 2020(5):13, May 2020. arXiv:1911.12333 [hep-th].

[3] Ahmed Almheiri, Thomas Hartman, Juan Maldacena, Edgar Shaghoulian, and Amirhossein Tajdini. The entropy of Hawking radiation. *Reviews of Modern Physics*, 93(3):035002, July 2021. arXiv:2006.06872 [gr-qc, physics:hep-th].

[4] Ahmed Almheiri, Raghu Mahajan, Juan Maldacena, and Ying Zhao. The Page curve of Hawking radiation from semiclassical geometry. *Journal of High Energy Physics*, 2020(3):149, March 2020. arXiv:1908.10996 [gr-qc, physics:hep-th].

[5] Ahmed Almheiri, Donald Marolf, Joseph Polchinski, and James Sully. Black Holes: Complementarity or Firewalls? *Journal of High Energy Physics*, 2013(2):62, February 2013. arXiv:1207.3123 [gr-qc, physics:hep-th].

[6] J. M. Bardeen, B. Carter, and S. W. Hawking. The four laws of black hole mechanics. *Communications in Mathematical Physics*, 31(2):161–170, June 1973.

[7] Jacob D. Bekenstein. Generalized second law of thermodynamics in black-hole physics. *Physical Review D*, 9(12):3292–3300, June 1974. Publisher: American Physical Society.

[8] Raphael Bousso. The holographic principle. *Reviews of Modern Physics*, 74(3):825–874, August 2002. arXiv:hep-th/0203101.

[9] Zvika Brakerski. Black-Hole Radiation Decoding is Quantum Cryptography, May 2023. arXiv:2211.05491 [quant-ph].

[10] Pasquale Calabrese and John Cardy. Entanglement Entropy and Quantum Field Theory. *Journal of Statistical Mechanics: Theory and Experiment*, 2004(06):P06002, June 2004. arXiv:hep-th/0405152.

[11] Pasquale Calabrese and John Cardy. Entanglement entropy and conformal field theory. *Journal of Physics A: Mathematical and Theoretical*, 42(50):504005, December 2009. arXiv:0905.4013 [cond-mat, physics:hep-th, physics:quant-ph].

[12] Sean M. Carroll. *Spacetime and geometry: an introduction to general relativity.* Pearson Education, Harlow, pearson new international edition edition, 2014.

[13] H. Casini and M. Huerta. Entanglement entropy in free quantum field theory. *Journal of Physics A: Mathematical and Theoretical*, 42(50):504007, December 2009. arXiv:0905.2562 [hep-th].

[14] Philippe Di Francesco, Pierre Mathieu, and David Sénéchal. *Conformal Field Theory.* Graduate Texts in Contemporary Physics. Springer, New York, NY, 1997.

[15] Xi Dong. The gravity dual of Rényi entropy. *Nature Communications*, 7(1):12472, August 2016. Number: 1 Publisher: Nature Publishing Group.

[16] Gia Dvali and Zaza N. Osmanov. Black holes as tools for quantum computing by ad-

vanced extraterrestrial civilizations, January 2023. arXiv:2301.09575 [astro-ph, physics:gr-qc, physics:hep-th, physics:physics, physics:quant-ph].

[17] Netta Engelhardt and Aron C. Wall. Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime. *Journal of High Energy Physics*, 2015(1):73, January 2015. arXiv:1408.3203 [gr-qc, physics:hep-th].

[18] Thomas Faulkner, Aitor Lewkowycz, and Juan Maldacena. Quantum corrections to holographic entanglement entropy. *Journal of High Energy Physics*, 2013(11):74, November 2013. arXiv:1307.2892 [hep-th].

[19] Daniel Harlow. TASI Lectures on the Emergence of the Bulk in AdS/CFT, July 2018. arXiv:1802.01040 [gr-qc, physics:hep-th, physics:quant-ph].

[20] Thomas Hartman. Lectures on Quantum Gravity and Black Holes.

[21] S. W. Hawking. Particle creation by black holes. *Communications in Mathematical Physics*, 43(3):199–220, August 1975.

[22] S. W. Hawking. Breakdown of predictability in gravitational collapse. *Physical Review D*, 14(10):2460–2473, November 1976. Publisher: American Physical Society.

[23] Matthew Headrick and Tadashi Takayanagi. A holographic proof of the strong subadditivity of entanglement entropy. *Physical Review D*, 76(10):106013, November 2007. arXiv:0704.3719 [cond-mat, physics:hep-th, physics:quant-ph].

[24] Veronika E. Hubeny, Mukund Rangamani, and Tadashi Takayanagi. A Covariant Holographic Entanglement Entropy Proposal. *Journal of High Energy Physics*, 2007(07):062–062, July 2007. arXiv:0705.0016 [cond-mat, physics:gr-qc, physics:hep-th].

[25] Jared Kaplan. Lectures on AdS/CFT from the Bottom Up.

[26] Aitor Lewkowycz and Juan Maldacena. Generalized gravitational entropy. *Journal of High Energy Physics*, 2013(8):90, August 2013. arXiv:1304.4926 [hep-th].

[27] Juan Maldacena and Leonard Susskind. Cool horizons for entangled black holes. *Fortschritte der Physik*, 61(9):781–811, September 2013. arXiv:1306.0533 [hep-th].

[28] Juan M. Maldacena. The Large N Limit of Superconformal Field Theories and Supergravity. *International Journal of Theoretical Physics*, 38(4):1113–1133, 1999. arXiv:hep-th/9711200.

[29] Charles W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W. H. Freeman, San Francisco, 1973.

[30] Makoto Natsuume. *AdS/CFT Duality User Guide*, volume 903 of *Lecture Notes in Physics*. Springer Japan, Tokyo, 2015.

[31] Michael A. Nielsen. *Quantum computation and quantum information*. Cambridge University Press, Cambridge ;, 10th anniversary ed. edition, 2010.

[32] Tatsuma Nishioka. Entanglement entropy: holography and renormalization group. *Reviews of Modern Physics*, 90(3):035007, September 2018. arXiv:1801.10352 [hep-th].

[33] Don N. Page. Information in Black Hole Radiation, August 1993. arXiv:hep-th/9306083.

[34] Kyriakos Papadodimas and Suvrat Raju. The Black Hole Interior in AdS/CFT and the Information Paradox. *Physical Review Letters*, 112(5):051301, February 2014. arXiv:1310.6334 [gr-qc, physics:hep-ph, physics:hep-th, physics:quant-ph].

[35] Kyriakos Papadodimas and Suvrat Raju. State-Dependent Bulk-Boundary Maps and Black Hole Complementarity. *Physical Review D*, 89(8):086010, April 2014. arXiv:1310.6335 [gr-qc, physics:hep-ph, physics:hep-th, physics:quant-ph].

[36] Geoff Penington, Stephen H. Shenker, Douglas Stanford, and Zhenbin Yang. Replica wormholes and the black hole interior, April 2020. arXiv:1911.11977 [gr-qc, physics:hep-th, physics:quant-ph].

[37] Geoffrey Penington. Entanglement Wedge Reconstruction and the Information Paradox, August 2020. arXiv:1905.08255 [gr-qc, physics:hep-th, physics:quant-ph].

[38] Mukund Rangamani and Tadashi Takayanagi. *Holographic Entanglement Entropy*, volume 931 of *Lecture Notes in Physics*. Springer International Publishing, Cham, 2017.

[39] Shinsei Ryu and Tadashi Takayanagi. Aspects of Holographic Entanglement Entropy. *Journal of High Energy Physics*, 2006(08):045–045, August 2006. arXiv:hep-th/0605073.

[40] Shinsei Ryu and Tadashi Takayanagi. Holographic Derivation of Entanglement Entropy from AdS/CFT. *Physical Review Letters*, 96(18):181602, May 2006. arXiv:hep-th/0603001.

[41] L. Susskind and L. Thorlacius. Gedanken Experiments involving Black Holes. *Physical Review D*, 49(2):966–974, January 1994. arXiv:hep-th/9308100.

[42] L. Susskind, L. Thorlacius, and J. Uglum. The Stretched Horizon and Black Hole Complementarity. *Physical Review D*, 48(8):3743–3761, October 1993. arXiv:hep-th/9306069.

[43] William G. Unruh and Robert M. Wald. Information loss. *Reports on Progress in Physics*, 80(9):092002, July 2017. Publisher: IOP Publishing.

[44] Mark Van Raamsdonk. Evaporating Firewalls. *Journal of High Energy Physics*, 2014(11):38, November 2014. arXiv:1307.1796 [gr-qc, physics:hep-th].