Facing black hole crises

Black hole thermodynamics, the information paradox, and non-local effects

Cheng, P.

Publication date
2021

Document Version
Final published version

FACING BLACK HOLE CRISES

Black hole thermodynamics, the information paradox, and non-local effects

Peng Cheng

Facing Black Hole Crises

Black hole thermodynamics, the information paradox, and non-local effects

Peng Cheng

This research was performed at the Institute for Theoretical Physics Amsterdam (ITFA) of the University of Amsterdam (UvA) and was supported by the Chinese Scholarship Council (CSC).

UNIVERSITEIT VAN AMSTERDAM

ISBN: 978-94-641-9321-3

Cover: three-fold wormhole.

Facing Black Hole Crises

Black hole thermodynamics, the information paradox, and non-local effects

Academisch Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties
ingestelde commissie,
in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 20 oktober 2021, te 14.00 uur

door

Peng Cheng

geboren te Dongying

Promotiecommissie

Promotor

prof. dr. J. de Boer	Universiteit van Amsterdam

Copromotor

dr. D.M. Hofman	Universiteit van Amsterdam

Overige leden

prof. dr. K.E. Schalm	Universiteit Leiden
prof. dr. E.A. Bergshoeff	Rijksuniversiteit Groningen
prof. dr. E.P. Verlinde	Universiteit van Amsterdam
dr. A. Castro Anich	Universiteit van Amsterdam
dr. B.W. Freivogel	Universiteit van Amsterdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

To My Grandma

Preface

Misfortune might be a blessing in disguise, and vice versa.

– Laozi, Tao Te Ching (4th century BC)

The essence of the Taoism philosophy is the attitude towards the world, and the so-called balance between Yin and Yang is actually about a peaceful state of mind. The title of this thesis is about black hole crises. According to Laozi: misfortune might be a blessing in disguise, and vice versa; crises can always be chances to better theories in physics history. Don't you remember lord Kelvin's nineteenth-century clouds?

This thesis is devoted to a better understanding of the topics related to gravity and black hole physics. Quantising gravity is a longstanding problem in theoretical physics, and the black hole information paradox is one of the first severe conflicts we have met between general relativity and quantum mechanics. In the current thesis, we will start an adventure tour on the black hole information paradox, phase transitions of black holes, infrared structure of gauge theory and gravity, and emergent gravity. It is an adventure because we will face many crises that we do not completely understand, but it will eventually worth it because we may gain much more than we expected.

Publications

This thesis is based on the following publications:

- **Peng Cheng**, Jan de Boer, and Diego Hofman,
  Phase transitions of black holes in boxes.
  In preparation.

  Presented in chapter 3. The candidate participated in all the conceptual discussions, carried out most of the calculations, and wrote the draft of the whole paper.

- **Peng Cheng** and Yang An,
  Soft black hole information paradox: Page curve from Maxwell soft hair of a black hole,
  Phys. Rev. D **103** (2021) 126020 [arXiv: 2012.14864 [hep-th]].

  Presented in chapter 4. The candidate participated in all the conceptual discussions, carried out all the calculations, and wrote the draft of the whole paper.

- Yang An and **Peng Cheng**,
  Realize emergent gravity to generic situations,
  Eur. Phys. J. C **81**, 789 (2021) [arXiv: 2004.14059 [hep-th]].

  Presented in chapter 5. The candidate participated in most of the conceptual discussions, verified some calculations, and helped revise the final version of the paper.

Other publications by the author:

- **Peng Cheng**,
  Evaporating black holes and late-stage loss of soft hair, [arXiv: 2108.10177 [hep-th]].

- **Peng Cheng**, Shao-Wen Wei and Yu-Xiao Liu,
  Critical phenomena in the extended phase space of Kerr-Newman-AdS black holes,
  Phys. Rev. D **94** (2016) [arXiv: 1603.08694 [gr-qc]].

- Shao-Wen Wei, **Peng Cheng** and Yu-Xiao Liu,
  Analytical and exact critical phenomena of d-dimensional singly spinning Kerr-AdS black holes,
  Phys. Rev. D **93** (2016) [arXiv: 1510.00085 [gr-qc]].

- Shao-Wen Wei, **Peng Cheng**, Yi Zhong and Xiang-Nan Zhou,
  Shadow of noncommutative geometry inspired black hole,
  J. Cosmol. Astropart. Phys. **2015** (2015) 004 [arXiv: 1501.06298 [gr-qc]].

Contents

1 Introduction

Appreciate the beauty of the universe;
analyse the logic of everything.

– Zhuangzi, The Land Under The Heaven (369-286 BC)

The discovery of black hole thermodynamics has radical influence on our understanding of general relativity and its relationship with quantum mechanics. Can a black hole be described as a quantum system with micro-states that have entropy equals the Bekenstein-Hawking entropy? Does a black hole system evolve unitarily under time evolution? Questions like those have led to vast disputes and arguments. And until now, we have not reached a general agreement about the nature of quantum gravity and black holes. We still have many different theories and opinions toward the quantum structure of black hole; it does not make sense that "there are a thousand Hamlets in a thousand people's eyes"[1] in theoretical physics.

The so-called *central dogma* [1] represents a nice expectation from theoretical physicists, who do not want to give up the idea that the whole world is quantum mechanical. However, everything in the world has a price to pay; one may need to exchange what he thinks is the most valuable thing with non-local effects, breaking of energy conditions, or other weird phenomena.

[1] This is an Asian proverb about different people have different interpretations of the same thing. I learned it from my primary school curriculum but only realised it isn't a European proverb very recently.

All those confusions and disputes are not all bad, and we should regard them as stepping-stones to a better understanding of the subject or even a better theory that explains all the confusions and unifies all different opinions. Actually, we have already learned a lot from the crises related to black hole physics. For example, the holography principle is largely inspired by attempting to find a microscopic description of the Bekenstein-Hawking entropy. In addition, many ideas related to entanglement entropy, quantum information, hydrodynamics, condensed matter theory, and quantum field theory are originated from black hole physics.

Recently, with the understanding of island prescription, as a PhD student, I witnessed how theoretical physicists learn from crises, digest the lessons from the new insights, and then develop new concepts and theories to understand more. It is very impressive to experience and my fortune to watch lots of new ideas springing up like bamboo shoots after a spring rain.

Outline of the thesis

This thesis is an overview of the research I performed at the Institute for Theoretical Physics, University of Amsterdam. Moreover, it mainly contains the author's understanding of the topics related to the black hole information paradox, black hole thermodynamics, infrared structure, and other related topics. The outline of the thesis is as follows.

In chapter 2, we review and discuss some relevant concepts and basic ideas that will be important in the main context of the thesis. First of all, in subsection 2.1.1, we provide some background material about black hole thermodynamics, including thermal field theory, entanglement entropy, Euclidean path integral, et cetera. Then we move to discuss black hole phase transitions in subsection 2.1.2. We review topics such as the Hawking-Page phase transition, black hole chemistry, and the holographic dual of those phase transitions in the corresponding field theory. Black hole thermodynamics leads to two main confusions of black hole physics: the black hole micro-states and the black hole information paradox. Section 2.2 discusses different approaches to understand black hole micro-states from string theory, holography, loop quantum gravity, and other theories. Then, we also explain how the conformal field theory method explains the universality of Bekenstein-Hawking entropy and the logic behind the hidden conformal symmetry on black hole background. In section 2.3, we describe some basic ideas and issues with the black hole information paradox. First, we provide a review of the firewall paradox, quantum computa-

tion, and other concepts. Then section 2.3.2 reviews the recent developments related to the island prescription, quantum extremal surfaces, double holography, replica wormholes, the thermo-mixed double state, and other exciting ideas. The material in this chapter is organised in a pedagogical way.

In chapter 3, we put a U(1) gauge theory on different backgrounds and analyse the behaviour of the partition function of the system in different temperature limits. The partition function is evaluated by the Euclidean path integral. Firstly, we put the gauge theory between two flat parallel plates and analyse the different behaviours of the system, as the temperature varies in section 3.2. We perform canonical analysis in the flat case to see the physical degrees of freedom in the phase space and integrate over the physical degrees of freedom in the path integral. The flat case results can serve as a frame of reference for black hole calculations. Then we put the whole system onto a black hole contained in a box in section 3.3. There are also two boundaries: the stretched horizon and the boundary of the box. We find that there are two different Bekenstein-Hawking-like entropies in the finite-temperature black hole case. The part from bulk fluctuation modes is very similar to the entropy in 't Hooft's "brick wall" model [2] and can be interpreted as the entanglement across the stretched horizon [3]. There is an extra Bekenstein-Hawking-like entropy that only appears in the very-low-temperature limit, which persists in the (near)-extremal case. This contribution comes from the radius zero modes ϕ and the boundary stretched Wilson lines W, who have localisation on $\partial_\tau \phi = \partial_\tau W = 0$ in the Euclidean path integral at low temperature. The chapter also includes some further discussions.

Then in chapter 4, we use Maxwell soft hair to derive the Page curve. We firstly review concepts like the infrared structure of gauge theory and gravity, Yoshida's decoupling theorem, Petz map to construct interior operators in section 4.2. Then we derive an effective action for soft hair modes by treating the would-be-gauge degrees of freedom as a transition function between the near-horizon and infrared region. Next, we evaluated the size of the phase space of Maxwell soft hair. Finally, in section 4.4, we get a Page curve of the black hole by comparing the rates of Hawking radiation and soft hair measurement. There are also some speculative discussions related to the black hole information paradox in this chapter.

The chapter 5 discusses emergent gravity. We firstly review the Casini-Bekenstein bound in section 5.2. The variation of Casini-Bekenstein bound naturally reproduces Newton's second law in Rindler space and local gravitational force

for Schwarzschild black hole. In Section 5.3, we further develop the techniques to derive inertial force utilising the entanglement first law. Then we discussed some concepts related to the holographic interpretation, extremal surface, complexity tendency, and other related ideas.

In chapter 6, we summarise the main topics and results discussed in this thesis. What's more, we conclude the thesis with outlooks and some further research directions.

Those chapters are followed by a bibliography, a dutch summary, and acknowledgements.

Conventions

In this thesis, we adopt the natural unit $c = k = \hbar = 1$ unless otherwise specified. When reviewing the micro-states of black hole, we keep the reduced Planck's constant $\hbar$ explicitly in chapter 2. Throughout the thesis, the pronoun "we" is used because most of the work is based on collaboration.

2 Black Hole Crises

This chapter contains an introduction to some basic concepts and background material relevant to this thesis. We are mainly aiming to provide a general review of black hole physics, puzzles and exciting developments. Especially, we also discuss the recent developments related to the black hole information paradox (BHIP) and new understanding of holographic duality.

In section 2.1, we review black hole thermodynamics and the ideas related to the holographic principle. Section 2.2 offers different approaches that are trying to understand black hole micro-states, like string theory, loop quantum gravity, holography, and so on. Then we discuss how to look for hidden conformal symmetries in a black hole system, and how the conformal field theory method explains the universality of the Bekenstein-Hawking entropy. In section 2.3, we discuss the basic concepts of BHIP and the island prescription.

2.1 Black Hole Thermodynamics

What would happen if an object is dropped into a black hole? The black hole absorbs this amount of energy and reaches a new equilibrium. The so-called "four laws of black hole mechanism" that capture the above physics were discovered in the 1970s [4–11], for example, the first law

$$dM = \frac{\kappa}{8\pi G_N} d\mathrm{A} + \Omega dJ + \Phi dQ \tag{2.1}$$

where κ is the surface gravity of the black hole, and A is the area of the horizon. J, Q, Ω, and Ψ are other parameters of a Kerr-Newman black hole. Although strongly reminiscent of the laws of thermodynamics, it was hard to believe a classical black hole has a temperature and radiates particles at that time. Only when Hawking confirmed Hawking radiation using quantum field theory on a black hole background in [12], people, including Hawking himself, start to believe the black hole is much more interesting than a "dark star" [13]. The first law of black hole thermodynamics can be written as

$$dE = TdS + ... \tag{2.2}$$

where we identify the mass of black hole M as the internal energy E, the area of the horizon as the entropy S, and the surface gravity κ as the temperature T. Now, the black hole is equipped with entropy and temperature

$$S = \frac{\mathrm{A}}{4\hbar G_N}, \qquad T = \frac{\hbar\kappa}{2\pi}. \tag{2.3}$$

This discovery is like opening Pandora's box: tons of confusions and inspirations come out of it.

2.1.1 The temperature of a black hole

Let us use the Euclidean periodicity trick to derive the temperature of a black hole. It is well-known that the finite temperature quantum field theory (QFT) is periodic in imaginary time

$$t \sim t + i\beta, \tag{2.4}$$

with inverse temperature $\beta = 1/T$. This can be seen from the path integral representation of the thermal field theory [14]. For example, the inverse

temperature is encoded in the thermal Green's function

$$\begin{aligned} G_\beta(\tau, x) &= -\frac{1}{Z}\text{tr}\ [e^{-\beta H}\mathcal{O}(\tau, x)\mathcal{O}(0,0)] \\ &= -\frac{1}{Z}\text{tr}\ [e^{-\beta H}e^{\beta H}\mathcal{O}(0,0)e^{-\beta H}\mathcal{O}(\tau, x)] \\ &= G_\beta(\tau - \beta, x)\,, \end{aligned} \tag{2.5}$$

where $\tau = it$ is the Euclidean time. Let us now focus on our 4-dimensional black hole system

$$ds^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega_2^2\,, \tag{2.6}$$

with

$$f(r) = (1 - \frac{r_s}{r}), \qquad r_s = 2G_N M\,. \tag{2.7}$$

The near-horizon region of the above Schwarzschild black hole metric can be written as a 2-dimensional Rindler space times S^2, whose metric can be expressed as

$$ds^2 \approx -\rho^2 d\eta^2 + d\rho^2 + r_s^2 d\Omega_2^2\,, \tag{2.8}$$

with

$$\eta = \kappa t = \frac{t}{2r_s}\,, \qquad \rho = \frac{2\sqrt{r - r_s}}{\sqrt{f'(r_s)}}\,. \tag{2.9}$$

The Rindler spacetime describes uniformly accelerating observers in Minkowski spacetime, and the Unruh effect [15] claims that the Minkowski vacuum appears to be a thermal state with the temperature proportional to the acceleration of the Rindler observer. The Hawking temperature is essentially the Unruh temperature because of the *equivalence principle*. The temperature can be seen from the following argument. The above near-horizon metric with imaginary time $\theta = i\eta$, can be written as

$$ds^2 \approx \rho^2 d\theta^2 + d\rho^2 + r_s^2 d\Omega_2^2\,, \tag{2.10}$$

which has a conical singularity at $\rho = 0$ unless $\theta \sim \theta + 2\pi$. The smooth horizon condition enforces a periodicity of the Euclidean time of the black hole system. Then, the temperature for an observer sitting at $r = \infty$ is

$$T = \frac{\kappa}{2\pi} = \frac{1}{8\pi G_N M}\,. \tag{2.11}$$

The above argument is very hand-waving, and we can do better by looking at a thermal system more carefully. It turns out any thermal density matrix can be

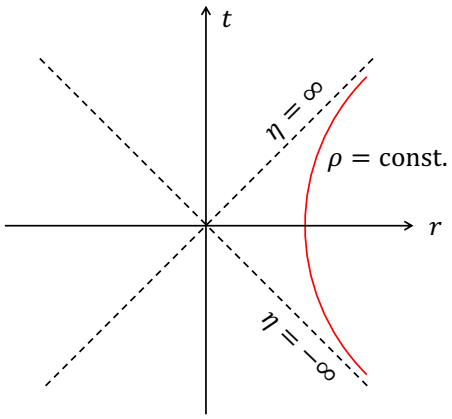


Figure 2.1: The near-horizon region of a Schwarzschild black hole can be regarded as 2-dimensional Rindler spacetime cross S^2.

obtained by tracing out part of a pure system, and the emergence of the Unruh effect can be interpreted as our ignorance of the region behind the Rindler horizon, as shown in figure 2.1. Similarly, in the black hole background, the Hartle-Hawking thermal state can be gotten by tracing out the region behind the horizon. Let us see this argument more explicitly below.

Here, we review the path integral representation of QFT firstly [16]. The transition amplitude for a scalar field theory can be represented as

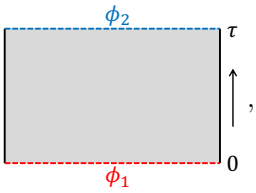
$$\langle\phi_2(t_2)|\phi_1(t_1)\rangle = \langle\phi_2|\, e^{-iH(t_2-t_1)}\,|\phi_1\rangle = \int_{\phi(t_1)=\phi_1}^{\phi(t_2)=\phi_2} \mathcal{D}\phi\; e^{iS}\,. \tag{2.12}$$

Writing it with Euclidean time $\tau = it$, the transition amplitude can be expressed as

$$\langle\phi_2(\tau)|\phi_1(0)\rangle = \langle\phi_2|\, e^{-\tau H}\,|\phi_1\rangle = \int_{\phi(0)=\phi_1}^{\phi(\tau)=\phi_2} \mathcal{D}\phi\; e^{-S_E}\,, \tag{2.13}$$

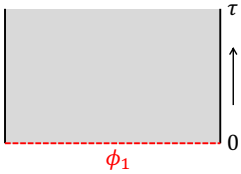
with the Euclidean action S_E. The above path integral can be illustrated diagrammatically as follows

$$\int_{\phi(0)=\phi_1}^{\phi(\tau)=\phi_2} \mathcal{D}\phi\; e^{-S_E} = \tag{2.14}$$



with path integral over the grey region from ϕ_1 to ϕ_2. Here ϕ_1 and ϕ_2 serve as boundary conditions on a fixed Cauchy surface. Correspondingly, the wave function can also be expressed as

$$|\Psi\rangle = e^{-\tau H}\,|\phi_1\rangle = \int_{\phi(0)=\phi_1}^{??} \mathcal{D}\phi\; e^{-S_E} = \tag{2.15}$$



which can be measured by the overlap with state $\langle\phi_2|$ as $\Psi[\phi_2] = \langle\phi_2|\Psi\rangle$.

The ground state can be gotten by evolving any state $|X\rangle$ by ∞ amount of time. Decomposing the state $|X\rangle$ on energy eigenstates $H\,|n\rangle = E_n\,|n\rangle$, we have

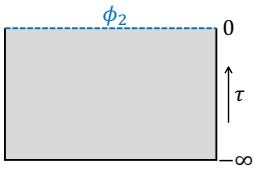
$$|X\rangle = \sum_n x_n\,|n\rangle\ . \tag{2.16}$$

Then, the ground state can be obtained by

$$\lim_{\tau\to\infty} e^{-\tau H}\,|X\rangle = \lim_{\tau\to\infty} \sum_n e^{-\tau E_n} x_n\,|n\rangle \to |0\rangle\ , \tag{2.17}$$

which is because all the high energy eigenstates with non-zero E_n are largely suppressed by $e^{-\tau E_n}$ as $\tau \to \infty$. Diagrammatically, we can depict the vacuum wave functional as

$$\langle \phi_2 | 0 \rangle = \int_{\phi(-\infty)}^{\phi(0)=\phi_2} \mathcal{D}\phi \; e^{-S_E} = \tag{2.18}$$



What's more, it is straightforward to see that the vacuum-to-vacuum amplitude can be expressed as

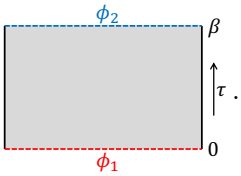
$$\langle 0 | 0 \rangle = \int_{\phi(-\infty)}^{\phi(\infty)} \mathcal{D}\phi \; e^{-S_E} = \tag{2.19}$$

Here we simply glue two diagrams shown in (2.18) together, which can be seen by inserting the identity

$$\langle 0 | 0 \rangle = \sum_{\phi} \langle 0 | \phi \rangle \, \langle \phi | 0 \rangle \; . \tag{2.20}$$

The density matrix is an operator that takes a bra and a ket, and then spits out a number. We can depict the thermal density matrix as

$$\langle \phi_2 | \, \rho \, | \phi_1 \rangle = \langle \phi_2 | \, e^{-\beta H} \, | \phi_1 \rangle = \tag{2.21}$$



We can get the thermal partition function by taking trace of the density matrix ρ, which can be represented as

$$Z(\beta) = \text{tr } e^{-\beta H} = \sum_{\phi} \langle \phi | \, e^{-\beta H} \, | \phi \rangle = \quad \beta \quad . \tag{2.22}$$

We can also insert some operators in the bulk to calculate correlation functions, and then the above picture (2.22) perfectly explains why the thermal Green's function shown in equation (2.5) has periodicity β. All the above diagrams can also be generalised to the situations that ϕs are defined on boundaries with different topology. And there can also be handles and holes in the bulk, like some "hand" $|H\rangle$ state in the bulk

$$\langle H | \phi_2 \rangle \quad = \quad , \tag{2.23}$$

Now, let us consider the Rindler space shown in figure 2.1, and derive the Unruh effect from the Minkowski vacuum. The density matrix of a 1+1-dimensional Minkowski vacuum can be written as

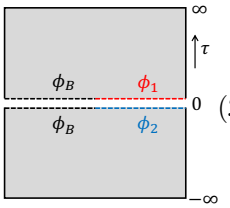
$$\rho = |0\rangle_M \langle 0| \ . \tag{2.24}$$

Because the presence of the Rindler horizon, we divide the 1-dimensional spatial line into two parts, the $r > 0$ region A and the $r < 0$ region B. The reduced density matrix in region A can be written as

$$\rho_A = \text{tr}_B \, |0\rangle_M \langle 0| \ , \tag{2.25}$$

whose path integral representation is

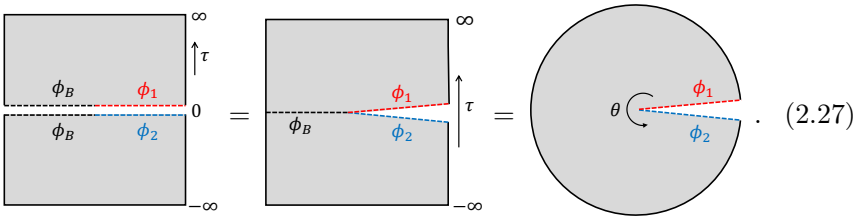
$$\langle\phi_2|\,\rho_A\,|\phi_1\rangle = \sum_{\phi_B} \langle\phi_2|\otimes\langle\phi_B|0\rangle_M\,\langle 0|\phi_B\rangle\otimes|\phi_1\rangle = \sum_{\phi_B}$$



(2.26)

The trace over fields ϕ_B in region B glues the B region together. With a different foliation, the above density matrix can be written as a thermal state with inverse temperature $\beta' = 2\pi$

$$\sum_{\phi_B}$$



. (2.27)

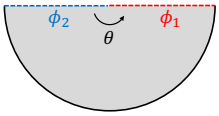
Here we treat the θ direction in metric (2.10) as the Euclidean time, and the above density matrix can be written as

$$\langle\phi_2|\,\rho_A\,|\phi_1\rangle = \langle\phi_2|\,e^{-2\pi H_\eta}\,|\phi_1\rangle\ , \tag{2.28}$$

with the Rindler Hamiltonian H_η. Now, we can identify the inverse temperature of Rindler space as $\beta' = 2\pi$. We will revisit the "Pacman" figure shown in equation (2.27) later in section 2.3. Now, we have derived the Unruh density matrix by tracing out the B region of Minkowski spacetime.

Besides the density matrix, the Minkowski vacuum itself can be treated by a different point of view. The transition amplitude between Minkowski vacuum and $\langle\phi_2|\otimes\langle\phi_1|$ can be represented as

$$\langle\phi_2| \otimes \langle\phi_1|0\rangle_M = \quad = \langle\phi_1| e^{-\frac{\beta}{2}H_\eta} |\phi_2\rangle \ , \tag{2.29}$$



which can be further written as

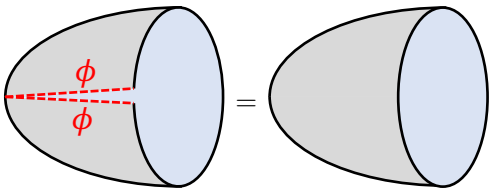
$$\langle\phi_1| e^{-\frac{\beta}{2}H_\eta} |\phi_2\rangle = \sum_n e^{-\frac{\beta}{2}E_n} \langle\phi_1|n_1\rangle \otimes \langle\phi_2|n_2\rangle^* \ . \tag{2.30}$$

$|n\rangle$ is the eigenstates of Rindler Hamiltonian H_η on the left and right sides, with $H_\eta |n\rangle = E_n |n\rangle$. The * represents CPT transformation. Comparing (2.29) and (2.30), we can conclude that the Minkowski vacuum is a purification of Rindler Hilbert space

$$|0\rangle_M = \sum_n e^{-\frac{\beta}{2}H_\eta} |n_2\rangle^* \otimes |n_1\rangle \ , \tag{2.31}$$

which is called *thermofield double* (TFD) state [17]. We have shown the origin of the thermal spectrum for Rindler observer restricted to the Rindler patch. The same argument can be generalised to black hole cases, which can be used to explain the thermal spectrum of the Hartle-Hawking state. The disk partition function obtained from tracing the Pacman figure is replaced by the so-called "cigar" geometry because of the curvature of the black hole. The black hole partition function can be represented as

$$Z_{BH} = \sum_\phi \quad = \quad . \tag{2.32}$$



The reason why we spend so much time on the thermal density matrix will be clear in section 2.3, where we will use path integral as the main tool to deal with concepts in the BHIP.

2.1.2 Black hole phase transitions

Now we can write the Euclidean path integral more explicitly. Let us focus on a $(d+1)$-dimensional Einstein-Hilbert theory with a negative cosmological constant Λ, i.e. Anti-de Sitter (AdS) spacetime, and denote the Euclidean action as S_E. The classical saddles, which are solutions of the equations of motion, are denoted as $\bar{g}^{(i)}_{\mu\nu}$, where the superscript (i) labels different possible saddles in the gravitational path integral. The path integral can be expressed as

$$Z(\beta) \approx \sum_i e^{-\frac{1}{\hbar G_N} S_E[\bar{g}^{(i)}_{\mu\nu}]} \times \det \left[\frac{1}{\hbar G_N} \frac{\delta^2 S_E}{\delta h^{(i)}_{\mu\nu} \delta h^{(i)}_{\mu\nu}} \Big|_{g^{(i)}_{\mu\nu} = \bar{g}^{(i)}_{\mu\nu}} \right]^{-1/2} \times Z_{QFT} \,, \quad (2.33)$$

where we have written the coupling constant G_N and the reduced Planck's constant $\hbar$ explicitly. $h^{(i)}_{\mu\nu}$ are the linearised gravitational fluctuations around the corresponding saddles $\bar{g}^{(i)}_{\mu\nu}$

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu} \,. \quad (2.34)$$

Higher-order fluctuations are ignored here. To leading approximation, we only need to count the contributions from classical saddles

$$Z(\beta) \approx \sum_i e^{-\frac{1}{\hbar G_N} S_E[\bar{g}^{(i)}_{\mu\nu}]} \,, \quad (2.35)$$

which will give us the classical black hole thermodynamics. We will come back to the other parts of (2.33) in chapter 3.

For now, let us only focus on the classical saddles. A concept of free energy related to the above path integral can be defined as

$$F = -T \ln Z \,, \quad (2.36)$$

where the Boltzmann constant k_B is set to be 1. The free energy is a good character of the dominant classical saddle point, and the saddle that has the lowest free energy plays the dominant role in the path integral.

Now we have turned the problem of finding classical saddles of Einstein's gravity into a thermodynamics problem. The phase transition of the thermodynamic system implies the transition of dominant saddles in the gravity theory. The thermodynamics system with only a few parameters is much easier to study than the non-linear Einstein's equation.

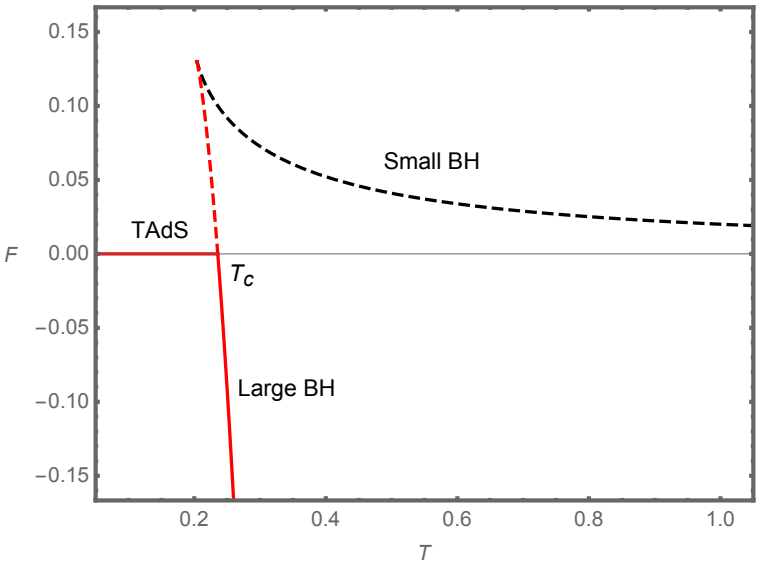


Figure 2.2: The Hawking-Page phase transition. The phase with smaller free energy is preferred. At a very low temperature, the thermal AdS is the only solution of the theory. The black hole saddles start to show up at a certain temperature, and the large black hole with a positive specific heat is always the favourable phase compare to the small black hole phase. At $T = T_c$, there is a first-order phase transition between the thermal AdS and the large black hole phase. At high temperature $T > T_c$, the large black hole is the preferred phase.

Example: the Hawking-Page phase transition

Obviously, there are at least two sets of solutions for the Euclidean action: Euclidean black hole solution and thermal AdS (TAdS) solution. The partition function (2.35) can be written as

$$Z \approx \exp\left[-\frac{1}{\hbar G_N} S_E[\bar{g}^{BH}_{\mu\nu}]\right] + \exp\left[-\frac{1}{\hbar G_N} S_E[\bar{g}^{TAdS}_{\mu\nu}]\right] + \dots . \tag{2.37}$$

In modified gravities, there might be other saddles we missed, but here we are relatively safe because of Birkhoff's theorem [18]. Writing the G_N and $\hbar$ explicitly, the action is of order $1/(\hbar G_N)$, and we have $1/(\hbar G_N) \to \infty$ in the classical limit; thus, the saddles are very steep. The above sum is exponentially dominated by the larger $S_E[\bar{g}^{(i)}_{\mu\nu}]$, and a small difference of S_E can result in a large difference in the partition function.

The classical action $S_E[\bar{g}^{(i)}_{\mu\nu}]$ can be easily evaluated by including Gibbons-Hawking boundary term and counter terms. Ignoring the constant Casimir energy, the free energy of the TAdS can be calculated as zero, and the free energy of the black hole is [19–21]

$$F_{BH} = -T \ln Z[\bar{g}^{BH}_{\mu\nu}] = \frac{\Omega_{d-1}}{16\pi\hbar G_N}(r_s^{d-2} - \frac{r_s^d}{l^2}) \tag{2.38}$$

where l is the radius of AdS spacetime, and Ω_{d-1} is the volume of unit sphere S^{d-1}. Note that the specific heat is positive as we can see from figure 2.2, so we conclude that for

$$T < T_c = \frac{d-1}{2\pi l} \tag{2.39}$$

we have

$$F_{BH} > F_{TAdS}\,, \tag{2.40}$$

the TAdS spacetime is the preferred saddle. At high temperature $T > T_c$, the finite temperature black hole is the dominant saddle, as explained in figure 2.2. The above phase transition between TAdS and finite black hole is known as the Hawking-Page phase transition [22].

Black hole chemistry

Recent interesting developments of black hole thermodynamics are mainly related to the so-call *black hole chemistry* [23–25], where the cosmological constant Λ is regarded as the pressure P of the system. Lots of dazzling phase structures and phase transitions were discovered soon in the so-called extended phase space.

The idea of treating the cosmological constant Λ as a dynamical variable was first proposed by Teitelboim and Brown in papers [26, 27], and then developed in [21, 28–31]. The phase transition between small and large black holes [32–35] was precisely identified with the van der Waals (vdW) phase transition of the liquid-gas system in [36–39], by regarding the cosmological constant as the thermodynamic pressure and its conjugate quantity as the volume. There are many new features in the extended phase space. For example, the equation of state can be used for comparison with ordinary thermodynamic systems [39], and the thermodynamic volume was shown to satisfy the reverse isoperimetric inequality [40]. Moreover, one important reason to consider black hole chemistry is that one can obtain a more fundamental theory that admits the variation of the cosmological constant [36].

Let us explain the basic idea and some crucial features of the black hole chemistry, using a $(d+1)$ dimensional charged black hole. If the cosmological constant Λ is regarded as the thermodynamic pressure of the system

$$P = -\frac{\Lambda}{8\pi G_N} = \frac{d(d-1)}{16\pi G_N l^2}\,, \tag{2.41}$$

the first law of the black hole thermodynamics is modified to be

$$dM = TdS + VdP + \Phi dQ\,. \tag{2.42}$$

V is the thermodynamic volume defined as the conjugate of P [37, 38, 40]

$$V = \left(\frac{\partial M}{\partial P}\right)_{S,Q}\,. \tag{2.43}$$

Now, the mass of the black hole $M(S,P,Q)$ is no longer interpreted as the internal energy, but enthalpy of the system [36]. The enthalpy is related to the internal energy $U(S,Q,V)$ by a Legendre transformation

$$M = U + PV\,. \tag{2.44}$$

This means that, the mass of an asymptotically AdS black hole is regarded as the total energy to create a black hole and place it in a box with pressure P.

The Smarr formula [41], which can be obtained by Euler's scaling argument, is also modified. The modified Smarr relation with the PV term included, is read as [36]

$$\frac{d-2}{d-1}M = TS + \frac{d-2}{d-1}\Phi Q - \frac{2}{d-1}PV\,, \tag{2.45}$$

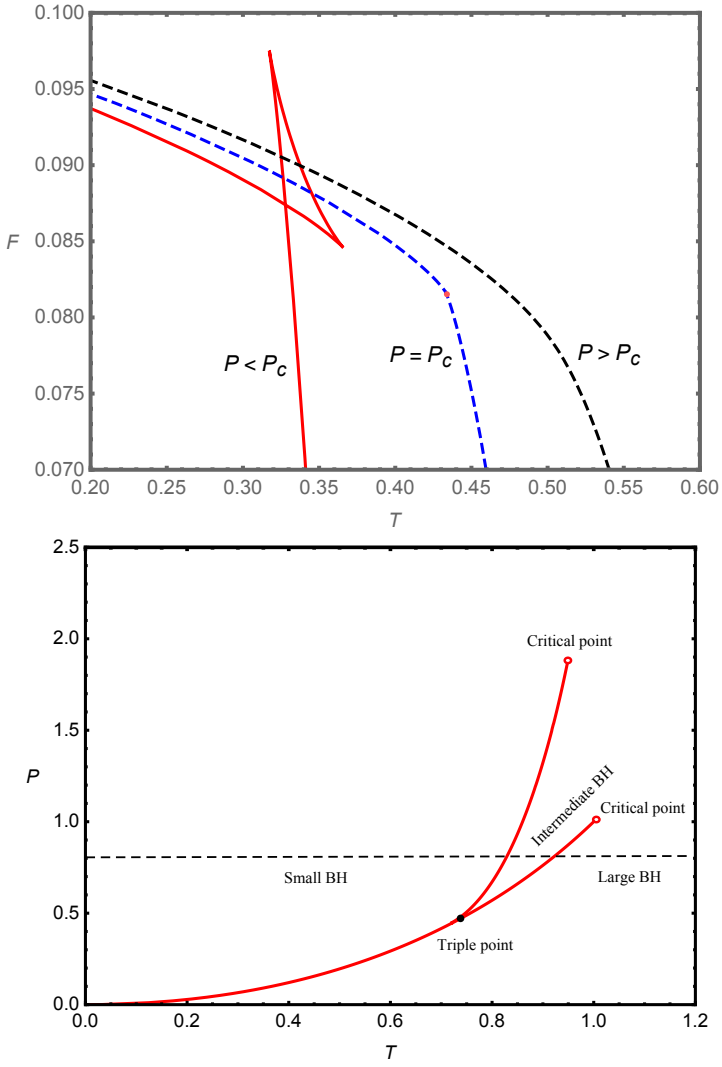


Figure 2.3: **First panel:** Free energy of RN black holes. At the critical point $P = P_c$, there is a first-order phase transition. For $P < P_c$, the free energy exhibits swallowtail behaviour, and we have a second-order phase transition. For high-pressure case, there is only one phase and no phase transition. **Second panel:** Triple phase transition for 6-dimensional Kerr-AdS black holes. The phase transition is reminiscent of the solid/liquid/gas phase transitions with a tri-critical point.

which is consistent with the relation of the conserved charges from the gravitational calculation.

Now, let us take a 4-dimensional Reissner-Nordstrom-AdS (RN-AdS) metric

$$\begin{aligned} ds^2 &= -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega_2^2 \,, \\ f(r) &= 1 - \frac{2G_N M}{r} + \frac{G_N Q^2}{r^2} + \frac{r^2}{l^2} \,, \end{aligned} \tag{2.46}$$

as an example and work out its phase structure. The conserved charges are read as

$$\begin{aligned} M = \frac{1}{2G_N}\left(r_+ + \frac{G_N Q}{r_+} + \frac{r_+^3}{l^2}\right) \,, \quad S = \frac{\pi r_+^2}{G_N} \,, \quad \Phi = \frac{G_N Q}{r_+} \,, \\ T = \frac{1}{4\pi r_+}\left(1 - \frac{G_N Q^2}{r_+^2} + \frac{3r_+^2}{l^2}\right) \,, \quad V = \frac{4\pi r_+^3}{3} \,, \end{aligned} \tag{2.47}$$

where r_+ is the outer horizon of the black hole. The modified Smarr relation (2.45)

$$M = 2TS + \Phi Q - 2PV \,. \tag{2.48}$$

is checked. The free energy of the system can also be expressed as

$$F = M - TS = \frac{l^2 r_+^2 - r_+^4 + 3G_N Q^2 l^2}{4G_N \; l^2 r_+} \,, \tag{2.49}$$

which is illustrated in the first panel of figure 2.3, where we can see the swallowtail behaviour. There is a critical value for P_c

$$P_c = \frac{1}{96 G_N Q^2} \,, \tag{2.50}$$

where we have a first-order phase transition between small and large black hole. Below the critical pressure, the free energy exhibit the swallowtail behaviour, which is a sign of second-order phase transition. It is interesting to look closely at the critical point. At the critical point, the critical ratio

$$\frac{P_c v_c}{T_c} = \frac{3}{8} \,, \tag{2.51}$$

is a universal constant and independent of the parameters of the black hole, where v_c is the diameter of the system. We also have the standard mean field theory exponents [38]

$$\alpha = 1 \,, \quad \beta = \frac{1}{2} \,, \quad \gamma = 1 \,, \quad \delta = 3 \,. \tag{2.52}$$

The ratio and exponents are exactly matched with the vdW fluid. The same critical exponents and scaling laws near the critical point, the same oscillatory behaviour of the isothermal line, and the swallowtail behaviour of the free energy imply that black hole systems are just behaving like our daily life material, which can be regarded as a piece of strong evidence for the central dogma of black hole. Many works demonstrated that there exists an extensive small-large black hole phase transition in the AdS black hole systems [42–62]. Moreover, it also exhibits some interesting multicritical phenomena, such as the reentrant phase transition and triple point [63–69]. An example of the triple phase transition is illustrated in the second panel of figure 2.3. It has been proposed that the black hole phase transitions can also be detected by astronomical observations, such as quasi-normal modes, photon orbits [70–72]

A further analogy between the vdW fluid and AdS black hole was made in paper [73], where a new concept, the number density of black hole molecules, was introduced. The difference in the number density between the small and large black hole phases naturally provides us with an order parameter to describe the phase transition. When an AdS black hole system crosses the coexistence curve, there is a non-vanishing latent heat, just like an ordinary thermodynamic system. It was also revealed that there exist interactions between black hole molecules. These provide a new perspective for the phase transition of AdS black holes.

Holographic dual description

The story told above is purely a gravitational story. The terms like free energy and thermodynamics are used because we are describing the classical saddle points in the gravitational path integral, as described at the beginning of section 2.1.2. There is also a nice field theory counterpart of the Hawking-Page phase transition [74–77], because of the AdS/CFT correspondence. The semi-classical gravity theory has a large-N $SU(N)$ gauge theory duality, and the Hawking-Page phase transition is dual to the confined/deconfined phase transition of the $SU(N)$ gauge theory.

Note that the free energy of the black hole phase is proportional to N^2, according to the dictionary

$$F_{BH} \propto \frac{1}{G_N} \propto N^2 \,. \tag{2.53}$$

So, the Hawking-Page phase transition is a transition between a phase with very small free energy and a phase with free energy proportional to N^2, as

shown in figure 2.2. Naively, if we look at the thermal partition function

$$Z = \int dE\ e^{-\beta E} D(E) \tag{2.54}$$

with density of states $D(E)$, the states with small energy contribute significantly because of the factor $e^{-\beta E}$ in (2.54), unless the high energy states have a very large density of state

$$D(E) \propto e^{N^2}\ . \tag{2.55}$$

This indeed is the case because the adjoint degrees of freedom contribute on the same footing at high temperature. The phase with $E \propto N^2$ may overwhelm the other phase if the temperature is high enough.

So we can conclude that there is a phase transition in the field theory between the confined and deconfined phase that corresponds to the Hawking-Page phase transition. Moreover, at low temperature (large β), the dominant contribution is from the states with energy $E \propto N^0$, which is the confined phase. At high temperature (small β), the states with energy $E \propto N^2$ overwhelm the other states because of the large density of states, which is the deconfined phase of the field theory. The confined phase corresponds to the TAdS phase in the gravity side, and the deconfined phase corresponds to the black hole phase.

Of course, one may expect to see a similar rich phase structure of field theory corresponding to the black hole chemistry discussed above. Does the large-N field theory that has a holographic dual also exhibits some vdW type or triple phase transitions in some instances? It is very interesting to understand, and there indeed are lots of attempts in this direction [78–83]. The difficulty lies in treating the cosmological constant Λ as a variable means the number of colours should also be regarded as variable in the field theory. One may need to add a chemical potential for colour [83]. Nevertheless, those ideas need further understanding.

Black hole thermodynamics is a fascinating topic, and it is very puzzling at the same time. There are two obvious paradoxes directly from black hole thermodynamics, and we do not even have good answers for them. The first one is that is there a statistical interpretation of black hole thermodynamics? To answer this question requires a good understanding of the micro-states of a black hole, and we need to find exponent of S_{BH} micro-states to explain the thermodynamic entropy of the black hole. The second puzzle is about the unitary evolution of a black hole system when Hawking radiation is considered.

Black hole evaporates because of the Hawking effect. Whether the evaporation process respect unitary evolution of quantum mechanics or not is closely related to the so-called black hole information paradox (BHIP). A quantum theory of gravity is still far away, and we do not even completely understand the semi-classical black hole. We call this the *Black Hole Crisis* because there are so many puzzles and paradoxes related to the black hole physics. We will try to review the understanding of black hole micro-state in the next section and review the recent development related to the BHIP in section 2.3.

2.2 Black Hole Micro-states

In this section, we will give a brief introduction of selected approaches to black hole statistical mechanics. We have discussed that the black hole system has thermodynamics, which might reflect the properties of the underlying microscopic states. If so, the Bekenstein-Hawking entropy

$$S = \frac{\text{Area}}{4\hbar G_N} \tag{2.56}$$

should be explained by the number of micro-states. We will explain the basic ideas of some approaches trying to reproduce the above entropy (2.56) in section 2.2.1 and spend more time on the hidden conformal symmetry near the black hole horizon in section 2.2.2. Many different research programs are trying to give a quantum theory of gravity, and they all regard successfully reproducing the Bekenstein-Hawking entropy as a sign of victory. Despite completely different start points and structures of those theories, they all declare that their theories can give out exactly the same Bekenstein-Hawking entropy. Isn't it also a problem to have so many different explanations of black hole micro-states? What are the common structures lie behind those theories need further understanding.

2.2.1 Different approaches

Here we would like to provide some examples that can successfully reproduce the Bekenstein-Hawking entropy. It is not realistic to cover all different approaches and to describe too many details of each theory in this subsection. The point we want to make here is that one of the main puzzles related to black hole micro-states is that there are different theories with different underlying microstructures all successfully reproduced the same Bekenstein-Hawking entropy, while there can only be one correct answer (and its variant).

"The" right quantum gravity theory needs to persuade other approaches that they are counting the wrong micro-states despite they can give out the right Bekenstein-Hawking entropy.

String theory was the first theory to give rise to a microscopic derivation of the Bekenstein-Hawking entropy [84]. The basic idea is the follows. A certain class of 5-dimensional supersymmetric BPS black holes in string theory can be interpreted as bound states of D_1 and D_5 branes wrapped on $T^4 \times S^1$, with momentum along S^1. On the one hand, the macroscopic Bekentein-Hawking entropy can be expressed in terms of the charges of the black hole solution. On the other hand, the microscopic entropy of the weakly bounded D_1-D_5 branes with strings stretched between them, can be evaluated. The brane calculation done in weak coupling can be extrapolated to strong coupling because of the presence of supersymmetry. It turns out that the entropy of the D_1-D_5 bound system matches the macroscopic Bekenstein-Hawking entropy. Soon, the calculation was extended to more general cases like non-extremal black holes [85–87], and was used in arguments of Hawking radiation [88]. The Fuzzball proposal [89–91] can also be regarded as a generalisation of explaining black hole micro-states using D_1-D_5 states.

AdS/CFT correspondence is always a good tool to understand bulk gravity theory. For example, in (2+1) dimensions, the Banados-Teitelboim-Zanelli black hole has a 2-dimensional conformal field theory (CFT) dual description, and the entropy can be produced by the CFT, whose density of states is determined by the central charge c [92–94]. Moreover, Ryu, Takayanagi, and collaborators (RT/HRT) [95, 96] have proposed the holographic entanglement entropy formula by embedding the field theory into a higher dimensional AdS spacetime. And soon after the RT formula was proposed, Emparan has shown it can be used to obtain the Bekenstein-Hawking entropy [97].

Loop quantum gravity (LQG) was also used to understand the microscopic degrees of freedom of black hole. There are many different proposals, and the author by no means an expert in this area. Here we just use the area operator in the spin network as an example to show that LQG can also reproduce the Bekentein-Hawking entropy. For a given surface Σ, the eigenvalue of the area operator $\hat{A}_\Sigma$ on spin networks can be expressed as

$$A_\Sigma = 8\pi\gamma\, G_N \sum_j \sqrt{j(j+1)}\,. \tag{2.57}$$

The sum j is over the edges across Σ, and γ is the Barbero-Immirzi (BI)

parameter [98–100]. Choosing the isolated horizon of a black hole and counting the number of spin network states, with suitable BI parameter γ, eventually reproduces the Bekentein-Hawking entropy [101–106].

What's more, there are lots of other proposals. For example, Hawking, Perry and Strominger have proposed the soft hair of a black hole can be responsible for the black hole entropy [107–113], induced gravity was shown successfully reproduced the desired entropy [114–116], quantum fields living very close to the horizon and correlated with the outside environment also have an entropy proportional to the area [2, 3, 117–120], and so on.

2.2.2 CFT description of the black hole entropy

Here is a general comment on most of the above approaches trying to understand the black hole microstates. It is not hard to put some so-called "microscopic structures" near the horizon and naively claim that the microscopic structure is stored in the unit of the Plank area l_p^2. Then, the number of states living on (or near) the horizon is just

$$\Omega \propto \exp \frac{\text{Area}}{l_p^2}\,, \tag{2.58}$$

which gives out an entropy

$$S \propto \frac{\text{Area}}{\hbar G_N}\,. \tag{2.59}$$

For most of the approaches, a more tricky part is the definition of concepts like "stretched horizon" and "isolated horizon" where one can put microscopic structures on.

It turns out the universal coefficient 1/4 in the Bekentein-Hawking entropy (2.56) is the hardest part to reproduce. Where is the coefficient 1/4 come from? One possible answer is: symmetry. As we have already seen, if there is hidden conformal symmetry near the horizon, and the density of states of two-dimensional CFT is controlled by the Cardy formula [121, 122]. The universality of the Cardy formula might be the right candidate to explain the universality of the coefficient 1/4 in (2.56). Hidden conformal symmetry near the horizon is extended studied in the literature [93, 94, 123–130], and here we follow Carlip [131–133] and are going to explain how to use four steps to build a CFT description.

1. Impose boundary conditions on appropriate boundaries. As the first step, we need to specify the boundary conditions. Usually, we have two

boundaries, the AdS boundary and the black hole horizon. It is a bit tricky to treat the horizon as a boundary. In the path integral formalism, the boundary conditions amount to split the spacetime along the horizon, and the full integration is reduced to two separate integrals, each with suitable boundary conditions. The full path integral can be reproduced by glueing the two parts and integrating over boundary conditions [134–138]. Still, defining boundary conditions on a null surface is difficult; the strategy is firstly imposing boundary conditions on the stretched horizon [139] and then taking the "real" horizon limit.

2. Find the algebra of diffeomorphism in the presence of boundary conditions. There are standard methods of deriving the algebra of diffeomorphism, for example, covariant phase space formalism [140–144]. For a gravitational system with first-class constraints $\mathcal{H}_\mu$, we have the generator associated with the diffeomorphism ξ^μ

$$H[\xi] = \int_\Sigma d^d x \; \xi^\mu \mathcal{H}_\mu \,. \tag{2.60}$$

defined by integrating over a Cauchy surface Σ. The Poisson brackets can define the algebra for the generators

$$\{H[\xi], H][\eta]\} = H[\{\xi, \eta\}] \,. \tag{2.61}$$

In the presence of boundaries, the situation is a bit complicated, and the above generators and Poisson brackets should be modified accordingly [145–151]. It turns out we need to add an extra boundary term to the generators

$$\bar{H}[\xi] = H[\xi] + B[\xi] = \int_\Sigma d^d x \; \xi^\mu \mathcal{H}_\mu + \int_{\partial\Sigma} d^{d-1} x \; \xi^\mu \mathcal{B}_\mu \,. \tag{2.62}$$

$B[\xi]$ is determined by the boundary conditions. The modified Poisson bracket becomes

$$\{\bar{H}[\xi], \bar{H}[\eta]\} = \bar{H}[\{\xi, \eta\}] + K[\xi, \eta] \,, \tag{2.63}$$

where $K[\xi, \eta]$ is a central term. The central term $K[\xi, \eta]$ is essential in getting the CFT description.

3. Look for a preferred sub-algebra Diff S^1. The central term $K[\xi, \eta]$ depends on the specific boundary conditions and can be complicated to calculate. However, if one can find a sub-algebra diffeomorphism of a circle such that

$$\{\xi, \eta\} = \xi\eta' - \eta\xi' \,, \tag{2.64}$$

the central extension of this sub-algebra is more or less fixed [152], which is read as

$$\{\bar{H}[\xi], \bar{H}[\eta]\} = \bar{H}[\{\xi, \eta\}] + \frac{c}{48\pi} \int d\theta (\xi' \eta'' - \eta' \xi'') \,, \tag{2.65}$$

which is the Virasoro algebra with central charge c. The Virasoro algebra is the fundamental symmetry algebra of a 2-dimensional CFT and thus can be dealt with with CFT methods. The existence of the sub-algebra Diff S^1 is not guaranteed, but certainly not surprising in lots of cases. The boundary of AdS_3 is a cylinder, so it is not hard to find a Diff S^1. Then it is natural to generalise everything to the geometries contains AdS×trivial geometry, like near-horizon-extremal-Kerr (NHEK) geometry [153]. We can also find the Diff S^1 in the null Killing vector of a generic black hole [154].

4. Use CFT methods to calculate the entropy. The entropy is completely determined by the Cardy formula [121, 122]. Moreover, the Cardy formula gives out universal entropy independent of the details of the theory, and can successfully reproduce the Benkenstein-Hawking entropy for BTZ black hole [93, 94], extremal Kerr black hole [125], and more general black holes [132]. Besides Virasoro algebra, Bondi-Metzner-Sachs algebra is also used as an engine in some recent papers [133, 155].

Okay, for now let us agree that the hidden conformal symmetry can explain the universal coefficient of Bekenstein-Hawking entropy, but what are those symmetry arguments tell us about the micro-states of black holes? Goldstone's theorem [156–158] claims that if a global symmetry is spontaneously broken, the massless Goldstone boson corresponds to each broken symmetry generator and can be viewed as an excitation along with the "would-be symmetry". The diffeomorphism of the spacetime is also broken because of the presence of boundary conditions. Although, being gauge symmetry, Goldstone's theorem remains true in this case. Moreover, the "would-be-diffeo" or "would-be-gauge" degrees of freedom become physical at the boundary and should be responsible for the symmetry calculations here. This is also coincident with the recent discussions related to edge modes [119, 136, 159].

2.3 The Black Hole Information Paradox and the Island Prescription

The second crisis because a black hole is a thermal system is the BHIP problem. This section provides some basic concepts and recent developments related

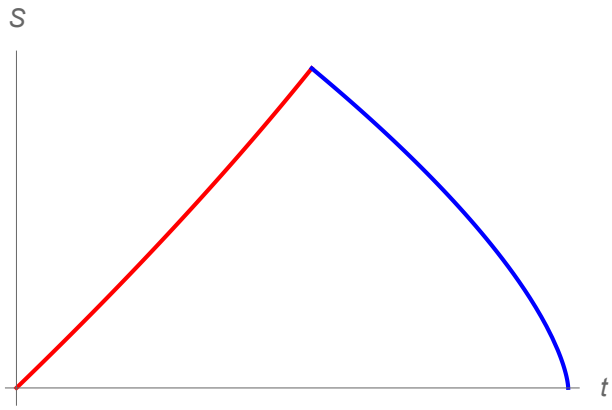


Figure 2.4: Page curve: the behaviour of the entropy of radiation and black hole during the evaporation.

to the BHIP. The fact that a black hole has thermodynamics and statistical mechanics would naturally lead us to the so-called *central dogma* [1]:

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.

However, the unitary evolution of the black hole suffers many setbacks along the way, which will be introduced in this section.

Naively, the most straightforward argument of the BHIP is that if we start with a black hole collapsed from a pure state, the black hole evaporates, and we are left with a bunch of Hawking radiation in a thermal state. The evolution from a pure state to a thermal state is not unitary, thus violating the central dogma or the basic principle of quantum mechanics (QM). But what should a unitary evolution look like? Don Page [160, 161] suggested the fine-grained entropy of a black hole and Hawking radiation should follow the so-called *Page curve* to be consistent with the unitarity. The Page curve is shown in figure 4.7. A derivation of the Page curve from a microscopic point of view can be regarded as a proof of the unitary evolution of the black hole and a solution of the BHIP in some sense.

A more careful analysis of the paradox is looking closely at the Hawking radiation process and seeing what contradiction we get during the evaporation. A "nice slice" Σ is a spacelike surface that can be stretched into the black

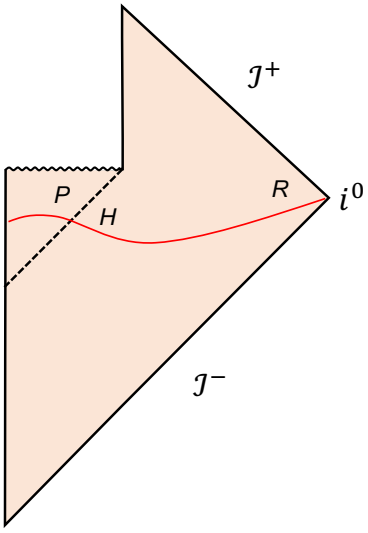


Figure 2.5: Penrose diagram of an evaporating black hole.

hole interior. Let us say Alice with her diary jumped into the black hole a scrambling time ago. Supposing the Hawking radiation carry the information of the diary, Hayden-Preskill protocol [162] tells us that Bob, sitting outside the horizon on the nice slice Σ, can get the information of the diary by collecting the radiation. Alice herself will come to the same nice slice, so we have two copies of the diary on Σ. This seems not consistent with the "no-cloning" theorem [163, 164] of QM, which claims that it's not possible to create an identical copy of an unknown quantum state. The *black hole complementarity* [165–168] tries to persuade us to forget about the nice slice. As far as those two copies of the diary are separated in different causal diamonds and can not meet each other or observed by the same observer, we are fine.

2.3.1 Firewall

The contradiction was made more concrete and sharp by Almheiri, Marolf, Polchinski, and Sully (AMPS) [169] in 2012. Let us consider a late time black hole shown in figure 2.5. The early radiation is denoted as R, the late Hawking radiation as H, and P is the partner of Hawking radiation H. On the one hand, for a unitary evaporation following the Page curve, H is supposed to purify R; thus, H and R are entangled. On the other hand, the near-horizon region of the black hole is Minkowski vacuum, thus H and P are maximally

entangled. H both entangled with R and P is a violation of the *monogamy of entanglement* [170], which claims that H can only has one partner to entangle with. If unitarity requires H and R are entangled, then H must "divorce" with P, which means the near-horizon region can not be Minkowski vacuum anymore. This means that the horizon is no longer flat and smooth, and the infalling observer will encounter *firewall* at the horizon.

Here is a short proof of entanglement monogamy in equations. Let us denote R_H as the part that maximally entangled with H, so we can write the entropy of H and R_H as

$$S_{HR_H} = 0\,. \tag{2.66}$$

The mutual information between P and HR_H is defined as

$$I_{P,HR_H} = S_P + S_{HR_H} - S_{HPR_H} = S_P - S_{HPR_H}\,. \tag{2.67}$$

The triangle inequality implies that

$$\begin{aligned} S_{HPR_H} &\leq S_P + S_{HR_H} = S_P & (2.68)\\ S_{HPR_H} &\geq |S_P - S_{HR_H}| = S_P\,; & (2.69) \end{aligned}$$

thus we can conclude that $S_{HPR_H} = S_P$. Thus, we have the mutual information

$$I_{P,HR_H} = S_P - S_{HPR_H} = 0, \tag{2.70}$$

i.e.,

$$\rho_{P,HR_H} = \rho_P \otimes \rho_{HR_H} \tag{2.71}$$

which means that there is no entanglement between P and HR_H, of course, the entanglement between H and P is also lost. We can also assume H and P are maximally entangled and then get the conclusion that H can not purify R_H; thus, the entropy of radiation increase all the time.

Now, we have seen the contradiction between the equivalent principle and unitary evolution of black hole systems, sharpened by the AMPS argument. Many attempts were made after the AMPS. Here we would like to provide argument related to the "decoding task" [171–175].

Soon after the AMPS, Harlow and Hayden proposed a way to avoid firewall [171]. The basic logic is very similar to the black hole complementarity. We were saying that H and R_H are maximally entangled, but R_H is a small part of the early radiation R. As far as no one can distil R_H out of R, we are fine. The quantum computation argument claims that the task of distilling

R_H out of R is exponentially hard, and the distilling time can be much longer than the evaporating time of the black hole. In such a sense, H and R are decoupled, and the smooth horizon is saved. Aaronson [172] further simplified the Harlow-Hayden decoding task and proved that the decoding task is as hard as inverting a one-way function.

Yoshida has a very similar idea and realises the decoupling between H and R by including the Hilbert space of observer $\mathcal{H}_O$. The basic idea can be summarised as Yoshida's decoupling theorem [173]: if the dimension of $\mathcal{H}_O$ is much larger than the Hilbert space of late radiation $\mathcal{H}_H$, i.e., $d_O \gg d_H$, early and late radiation are decoupled, and one can reconstruct the Hawking partner P without using the early radiation R. Pasterski and Verlinde [175] then legitimise the inclusion of observer Hilbert space by gravitational dressing and claim that the observer degrees of freedom can be identified with the soft hair degrees of freedom of a black hole system. The soft hair introduces an observer-dependent firewall, and the infalling observer will never knock into any AMPS firewall before reaching the singularity. The construction of P can be performed by the Petz map [176, 177]. We will encounter this idea again in chapter 4.

Note that the above argument related to decoupling between H and R, rely on the dimension of the Hilbert space of H is much smaller than the Hilbert space of black hole or radiation, which is approximately true right after the Page time. However, at the very end of the evaporation, the von Neumann entropy of radiation and black hole are both order $\mathcal{O}(1)$; the decoupling theorem then faces a challenge. Let us say that the radiation R is only left with two qubits, distil one-qubit R_H out of R such that $S_{H,R_H} = 0$ is not a hard problem anymore. Then the firewall shows up again at the end of the evaporation. If the firewall shows up once on the horizon, there would be a firewall all over the horizon because of the general covariance of the general relativity. A similar agreement applies to Yoshida's decoupling theorem, when the dimension of soft hair Hilbert space $d_S = d_O \gg d_H$ is not held anymore. We call this as *alopecia paradox*, as a pun on the annoyance of losing "soft hair". We will discuss this problem later and try to give a possible solution of the alopecia paradox.

2.3.2 The island prescription

The recent progress on the BHIP was initiated by [178, 179] in 2019. The new prescription of computing the entropy of the Hawking radiation was called

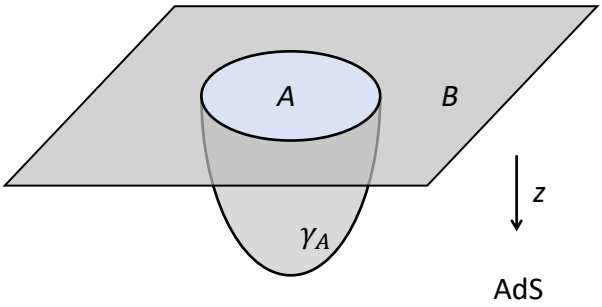


Figure 2.6: Holographic entanglement entropy. The entanglement entropy between regions A and B is calculated by embedding the spacetime to a higher dimensional AdS spacetime. The entanglement entropy is proportional to the area of the minimal surface γ_A.

"Island rule" later in [180]. This section provides basic ideas related to the island prescription.

Let us first look at the *quantum extremal surface* (QES). Entanglement entropy is a measurement of entanglement between two systems A and B. Inspired by AdS/CFT, the holographic entanglement formula, shown in figure 2.6, proposed a holographic way of calculating boundary entanglement entropy by embedding the spacetime into a higher dimensional AdS spacetime, known as RT/HRT formula [95, 96]. The entanglement entropy can be calculated by finding the minimal area surface γ_A in a higher dimensional AdS spacetime as

$$S_A = \min\left[\frac{\text{Area}(\gamma_A)}{4G_N}\right] . \tag{2.72}$$

γ_A is called RT surface. The RT/HRT formula is refined and generalised to more general situations [181–184], where the surface γ_A is chosen such that the generalised entropy

$$S_{gen} = \frac{\text{Area}(\gamma_A)}{4G_N} + S_{QFT}(\Sigma_\gamma) , \tag{2.73}$$

takes the minimal value [184]. Σ_γ is the region bounded by A and γ_A. Now the entanglement entropy (2.72) can be written as

$$S_A = \min\left[\frac{\text{Area}(\gamma_A)}{4G_N} + S_{QFT}(\Sigma_\gamma)\right] . \tag{2.74}$$

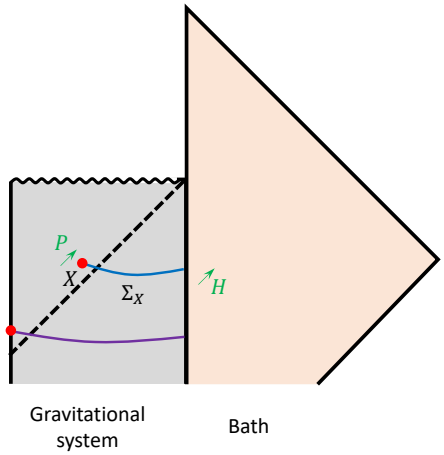


Figure 2.7: Extremal surfaces of an evaporating black hole. At the beginning of the evaporation, the extremal surface is vanishing. While after the Page time, the extremal surface changed to a surface behind the horizon. The Page curve can be obtained because of the transition.

The subtlety arises because the minimal surface means it minimises the generalised entropy in spacelike directions and maximises the generalised entropy in timelike direction; in such a sense, the surface is called *maximin surface* by Wall [185]. Appreciating this point, the von Neumann entropy can be obtained by firstly choosing the surfaces that extremise the generalised entropy and then selecting the one that gives rise to the minimal entropy value. Writing it in equations, the entropy can be expressed as

$$S = \min_X \left[\text{ext}_X \left(\frac{\text{Area}(X)}{4G_N} + S_{\text{semi-cl}}(\Sigma_X) \right) \right] , \tag{2.75}$$

where X is the extremal surface, and $S_{\text{semi-cl}}(\Sigma_X)$ is the von Neumann entropy for quantum fields living in Σ_X. The above formula can be derived by path integral prescription [181], and are generalised to non-holographic systems [178, 179, 186, 187].

Now, we can see how the QES gives out the Page curve of an evaporation black hole. The definition of the generalised entropy is not relying on holography, and thus it can be defined in any region of the bulk spacetime. For example, let us consider a black hole system coupled with a reservoir where Hawking

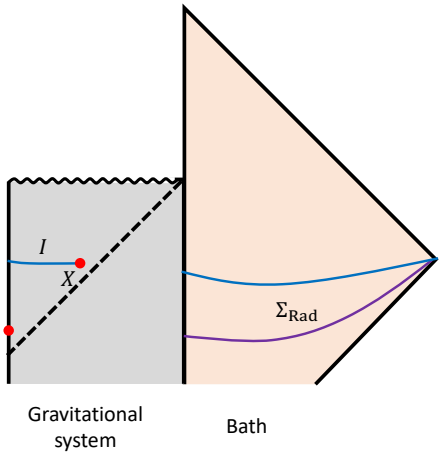


Figure 2.8: Island rule for computing the fine-grained entropy of radiation. At the late time, we need to include the island region inside the horizon when calculating the entropy of radiation.

radiation can pass through the boundary freely, as shown in figure 2.7. The extremal surface is a co-dimensional two surface, thus are shown as red dots in figure 2.7.

There are at least two QES in the above setup. Firstly, the vanishing surface is always a possible choice for X. When the extremal surface is vanishing, there is no contribution from the area of X, and the generalise entropy equal to the von Neumann entropy of quantum fields living on the purple surface shown in figure 2.7,

$$S_{BH} \approx S_{\text{semi-cl}}(\Sigma_{\emptyset})\,, \tag{2.76}$$

which counts the entanglement between the Hawking radiation and its partner.

There is also another extremal surface X that extremises the generalised entropy. As shown in figure 2.7, if the area of X connected with the blue Σ_X shrinks, quanta P can be included in Σ_X; thus the area contribution $\text{Area}(X)/(4G_N)$ decreases, and the semi-classical entropy $S_{\text{semi-cl}}(\Sigma_X)$ increases. When the amounts of entropy increased and decreased from both sides are equal to each other, the surface X is an extremal surface. It can be shown that the extremal surface is not far behind the horizon [178, 179]. By excluding most of the black hole interior, the semi-classical entropy is not comparable

with the area of X. So, the fine-grained entropy of the black hole approximately equals the area of the horizon,

$$S_{BH} \approx \frac{\mathrm{A}_{hor}}{4G_N}\,, \tag{2.77}$$

which is decreasing as the black hole evaporates.

Now we have shown that there are at least two extremal surfaces in computing the fine-grained entropy of a black hole, and we need to compare those two cases and only pick the minimal one. There are not too many Hawking radiation at the beginning of the evaporation, so

$$S_{\text{semi-cl}}(\Sigma_\emptyset) < \frac{\mathrm{A}_{hor}}{4G_N}\,, \tag{2.78}$$

and the extremal surface is the vanishing one. The entropy of the black hole goes as Hawking's calculation. After the Page time t_{page}, the number of particles escaping from the horizon is larger than the coarse-grained entropy of the black hole, so we have

$$S_{\text{semi-cl}}(\Sigma_\emptyset) > \frac{\mathrm{A}_{hor}}{4G_N}\,, \tag{2.79}$$

which means the extremal surface is changed to the second case, and the entropy of the black hole decrease as the black hole evaporates. The above entropy exactly matches the Page curve of an evaporating black hole [160, 161], as shown in figure 4.7. Thus, we can conclude that the transition between the two extremal surfaces precisely gives out the Page curve of an evaporating black hole.

How about the entropy of radiation? The most direct idea is to use equation (2.74) to calculate the fine-grained entropy of radiation and see if the entropy follows the Page curve or not. It is not clear to the author how to directly use (2.74) to get a disconnected Σ_X giving rise to the Page curve of radiation. Here we can try a different angle. It turns out assuming unitarity, the above argument is equivalent to the "island rule" for computing the entropy of radiation. A derivation of the island rule from a microscopic point of view can be regarded as a proof of unitary evolution and Page curve for radiation.

Assuming unitarity, the combined system of the black hole and the bath is a pure state, which means that the fine-grained entropy of radiation should be given by the following prescription

$$S_{\text{rad}} = \min_X \left[\text{ext}_X \left(\frac{\text{Area}(X)}{4G_N} + S_{\text{semi-cl}}(\Sigma_{\text{rad}} \cup \Sigma_{\text{I}}) \right) \right]\,, \tag{2.80}$$

where the island region is shown in figure 2.8. The extremal surface X is the same for the black hole argument. With the above prescription, the entropies (2.74) and (2.80) implies the combination of gravitational system and bath is a pure state.

The derivation of island rule is a different story, and we will introduce the basic idea later. For now, let us look at what the island prescription means. Certainly, the island rule means the Page curve for the radiation because it is already hinted at in our assumption. At the beginning of the evaporation, X is an vanishing surface, and we do not need to worry about any island, which means the entropy is the same as Hawking's calculation. After Page time, the extremal surface behind the horizon gives the minimal generalised entropy, and the entropy of the radiation more or less equals the area of the island. This is the similar argument as the black hole case, and the entropy of radiation exactly follows the Page curve shown in figure 4.7.

At first glance, the island rule is rather strange. The fine-grained entropy of radiation has a large difference from the semi-classical result after Page time:

$$S_{\text{rad}} = S_{\text{semi-cl}}(\Sigma_{\text{rad}}) + \text{large corrections coming from non-local effects.}$$

The non-local effects are because of the entanglement, which is the essence of ER=EPR [188]. Now, after t_{Page}, the entanglement wedge of radiation also includes the interior of the horizon, so one can in principle construct any operator inside the horizon just by looking at the radiation, which is consistent with the Hayden-Preskill protocol. The reconstruction [189, 190] can be performed by the Petz map.

2.3.3 Lessons from the island prescription

Claire is an ordinary people who live in a Minkowski spacetime. Her job is collecting some thermal radiation received from an unknown company. Basically, what she needs to do is just counting the clicks on the machine and writing down the numbers. She usually submits a monthly report to her boss to report the von Neumann entropy of the particles she collected. It is an easy job and well-paid. However, she quitted her job in May 2019, and no one in her company knows why.

The proposal of island prescription has profound influences on our current understanding of gravity, quantum information, and holography.

First of all, the derivation of island rule from a holographic point of view brings

us the idea of *double holography* [180], which can also be used to understand high dimensional black holes [191–193]. The basic idea is that the entropy of the radiation can also be holographically calculated by embedding the system into an even higher dimensional AdS spacetime, as shown in figure 2.9. The gravitational part can be described by Planck brane extended into the bulk in the Randall-Sundrum brane-world model [194, 195]. The solid black line in figure 2.9 corresponds to the gravitational system shown in figure 2.8, and the orange line corresponds to the bath. If no gravitational system is presented, the extremal surface is more or less the purple line shown in the first panel of figure 2.9. and it indeed is the case at the beginning of the evaporation. So before t_{page}, no island is needed, and the entropy is more or less the semi-clasical entropy of the radiation. However, after the Page time, the extremal surface that ends on the Planck brane, as shown in the second panel of figure 2.9, has the smaller generalised entropy, which is the case we need to include the island.

Replica trick and Euclidean path integral

In papers [190, 196], the authors proposed to use replica trick and Euclidean path integral to derive the island rule without using holography. Here we mainly follow [190] in this subsection. Let us use $|\Psi\rangle$ to represent the whole system, and we have

$$|\Psi\rangle = \frac{1}{\sqrt{kZ_1}} \sum_{i=1}^{k} |\psi_i\rangle_B |i\rangle_R \,, \tag{2.81}$$

where $|\psi_i\rangle_B$ denotes the black hole states, $|i\rangle_R$ denotes radiation, and the label i labels the entanglement between the radiation and black hole. The normalisation factor is chosen such that under

$$\langle\psi_i|\psi_j\rangle_B = \delta_{ij} Z_1 \,, \qquad \langle i|j\rangle_R = \delta_{ij} \,. \tag{2.82}$$

the amplitude

$$\sum_{\Psi} \langle\Psi|\Psi\rangle = 1 \,. \tag{2.83}$$

The density matrix of radiation can be obtained by tracing out the black hole system

$$\rho_R = \sum_m \langle\psi_m|\Psi\rangle \langle\Psi|\psi_m\rangle = \frac{1}{k} \sum_{i,j} |i\rangle_R \langle j| \langle\psi_j|\psi_i\rangle_B \tag{2.84}$$

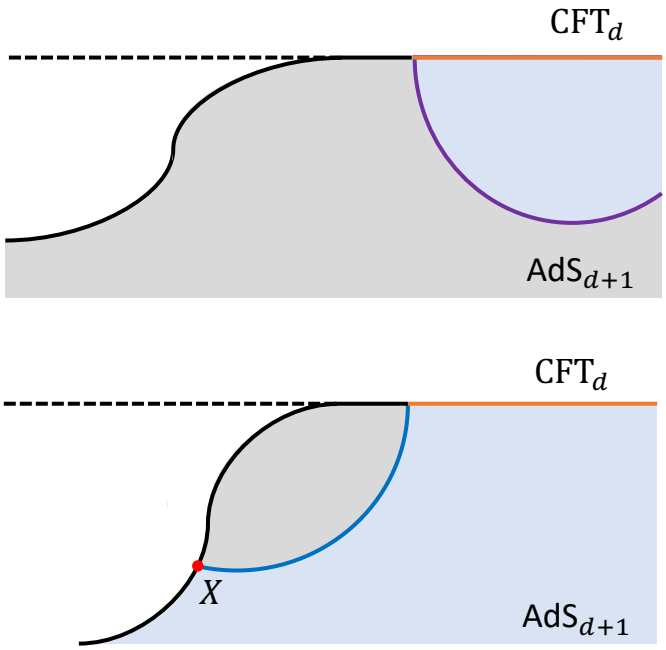


Figure 2.9: Double holography model for calculating the entropy of radiation. The orange line corresponds to the bath region in figure 2.8, and the solid black line corresponds to the gravitational region. The system is embedded into a higher dimensional AdS spacetime, and there is a transition of bulk extremal surface as shown in the two panels. At the beginning of the evaporation, the extremal surface calculating the entropy of the bath is shown in the first panel, and no island is needed. After the Page time, the extremal surface that ends on the Planck brane gives out smaller generalised entropy. So, we need to include the island in calculating the entropy of radiation, as shown in the second panel.

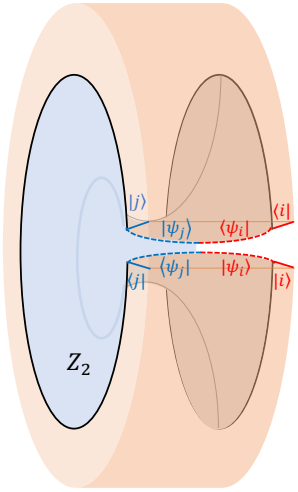


Figure 2.10: Replica wormhole saddle in calculating the purity of radiation (2.86). The picture is a diagrammatic representation of $\sum_{i,j} \langle\psi_i|\psi_i\rangle \langle\psi_j|\psi_j\rangle$. There is a wormhole geometry connecting the gravitational part, and the final result is $\sum_{i,j} \langle\psi_i|\psi_i\rangle \langle\psi_j|\psi_j\rangle = k^2 Z_2$.

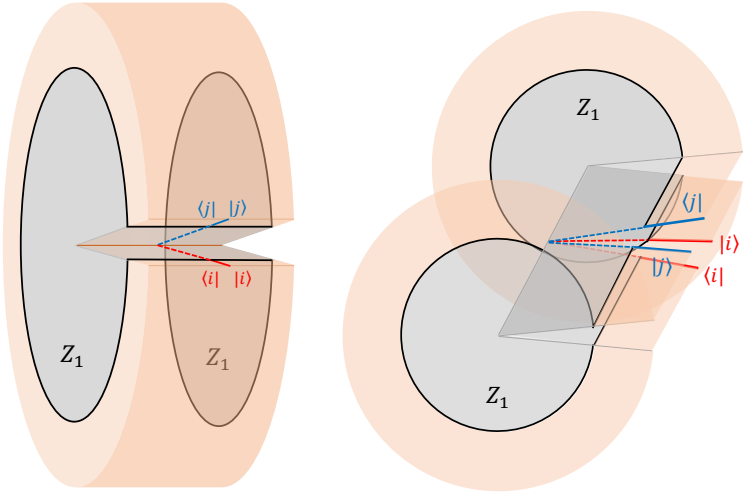
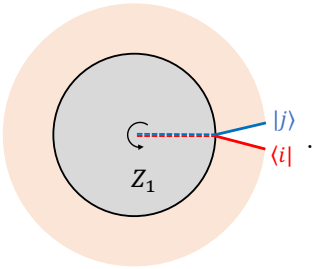


Figure 2.11: Hawking saddle in calculating the purity of radiation (2.86). The corresponding result is $\sum_{i,j} \langle\psi_i|\psi_j\rangle \langle\psi_j|\psi_i\rangle = k^2 Z_1^2$. The second picture is the same result shown in a different angle in case the first one is not clear.

which can be represented as

$\rho_R =$ 

. (2.85)

The density matrix is illustrated diagrammatically in the same way as we discussed in section 2.1.

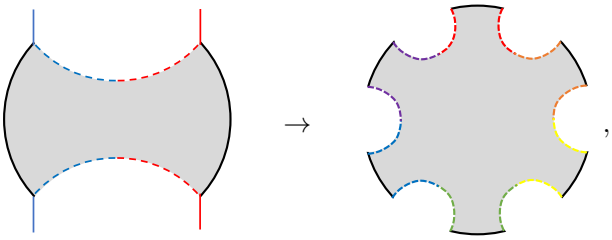
Instead of directly use the density matrix to calculate the entropy, one can also look at the purity, where we have

$$\begin{aligned}
\mathrm{tr}[\rho_R^2] &= \frac{1}{k^2 Z_1^2} \sum_l \sum_{i,j} \sum_{m,n} \langle l|i\rangle \langle j| \langle \psi_j|\psi_i\rangle |m\rangle \langle n| \langle \psi_n|\psi_m\rangle |l\rangle \\
&= \frac{1}{k^2 Z_1^2} \sum_{i,j} \sum_{m,n} \langle n| \langle \psi_n|\psi_m\rangle |i\rangle \langle j| \langle \psi_j|\psi_i\rangle |m\rangle \\
&= \frac{1}{k^2 Z_1^2} \sum_{i,j} \sum_{m,n} \langle \psi_n|\psi_m\rangle \langle \psi_j|\psi_i\rangle (\delta_{ni}\delta_{jm} + \delta_{ji}\delta_{nm}) \\
&= \frac{1}{k^2 Z_1^2} \sum_{i,j} [\langle \psi_i|\psi_j\rangle \langle \psi_j|\psi_i\rangle] + \frac{1}{k^2 Z_1^2} \sum_{i,j} [\langle \psi_i|\psi_i\rangle \langle \psi_j|\psi_j\rangle] \\
&= \frac{k Z_1^2 + k^2 Z_2}{k^2 Z_1^2}\,,
\end{aligned} \tag{2.86}$$

where Z_2 represent the *replica wormhole* shown in figure 2.10. The first part of equation (2.86) is illustrated in figure 2.11. Now, the purity can be expressed as

$$\mathrm{tr}[\rho_R^2] = \frac{k Z_1^2 + k^2 Z_2}{k^2 Z_1^2} = \frac{1}{k} + \frac{Z_2}{Z_1^2}\,. \tag{2.87}$$

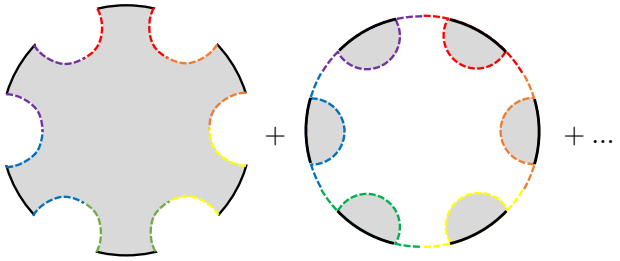
The n-th Renyi entropy $\mathrm{tr}[\rho_R^n]$ can be calculated by the same logic, which also contains fully connected geometry, fully disconnected geometry, and other geometries between them. The fully connected geometry can be diagrammatically represented by generalising the 2-fold replica wormhole geometry as



(2.88)

where the dashed lines with the same colour are traced over. Here the 2-fold wormhole on the left part of (2.88) is a flattened version of the grey region shown in figure 2.10, and one can get figure 2.10 by gluing the dashed lines with the same colour together. Similarly, gluing the dashed lines in the right part of (2.88), we get n-fold replica wormhole geometry. All in all, the n-th Renyi entropy can be expressed as

$\mathrm{tr}[\rho_R^n] =$



(2.89)

The von Neumann entropy can be obtained by analytically continue n to 1, which can be written as

$$S_R = -\mathrm{tr}(\rho_R \ln \rho_R) = -\lim_{n\to 1} \frac{1}{n-1} \ln \mathrm{tr}[\rho_R^n] \,. \tag{2.90}$$

So von Neumann entropy is essentially determined by $\mathrm{tr}[\rho_R^n]$. Let us see when different saddles dominate during the evaporation of the black hole. For the purity $\mathrm{tr}[\rho_R^2]$, when the number of Hawking radiation k is very small, the Hawking saddle that contributes as $1/k$ is the dominant contribution, and we have

$$\mathrm{tr}[\rho_R^2] \approx \frac{1}{k} \,. \tag{2.91}$$

At the late period of the radiation, when k is very large, we have

$$\mathrm{tr}[\rho_R^2] \approx \frac{Z_2}{Z_1^2}\,. \tag{2.92}$$

To get an impression of what Z_2/Z_1^2 looks like, one can use a 2-dimensional gravity theory as a toy model to calculate Z_2 and Z_1. In two-dimensional Jackiw-Teitelboim gravity [197, 198], the gravity path integral only contributes topologically. So we have

$$\frac{Z_2}{Z_1^2} = e^{-S_0} \tag{2.93}$$

where S_0 is the Bekenstein-Hawking entropy. The corresponding entropy of Z_2/Z_1^2 in 2 dimensional is the Bekenstein-Hawking entropy. Then, the corresponding entropy of the radiation calculated from the replica wormhole saddle is proportional to the area of the horizon. The Page curve can be obtained by the transition between the Hawking saddle and the wormhole saddle.

For the n-th Renyi entropy, there are more different saddles as we vary k, and the saddles with more disconnected components dominate at the early stage when k is relatively small. The connected geometries play more important roles at the late stage of evaporation. Again, we can add all the different saddles together, and the transitions between different saddles lead to an entropy of radiation that consistent with unitarity [190]. Let us still call the curve that illustrating the entropy as "Page curve", although there are many transitions than a single phase transition in the originally proposed Page curve in figure 4.7.

Note that there is a further puzzle. If we look closely at the amplitude

$$\langle \psi_i | \psi_j \rangle_B = \delta_{ij} Z_1\,, \tag{2.94}$$

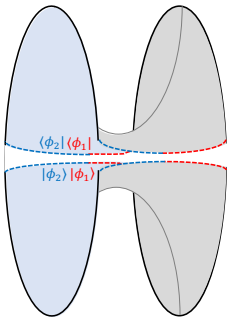
$$|\langle \psi_i | \psi_j \rangle_B|^2 = \delta_{ij} Z_1^2 + Z_2\,. \tag{2.95}$$

We have (2.94) squared does not equal (2.95). To solve the tension, one might need to think the gravitational path integral result as averaged over same ensemble [199], and the second term in (2.95) is the variance. This discovery also cause a wide discussion about ensemble average [200–202], baby universe [203–207] and other new understandings, which we will not cover in this thesis.

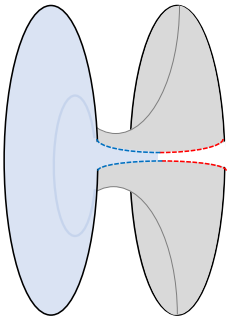
2.3.4 The replica wormhole as a transition amplitude between different vacua

In this subsection, let us put the entropy of radiation and Page curve aside, and think about what the replica wormhole saddle, shown in figure 2.10, means in gravity. Looking at the replica wormhole, we find that we have lost our favourite TFD state (2.29) as discussed in section 2.1. If one insist that the density matrix of the replica wormhole is gotten from tracing over some TFD-like state, we can think the density matrix ρ_{WH}, with "WH" representing wormhole, as

$$\rho_{\mathrm{WH}} = \sum_{\phi_1,\phi_2}$$



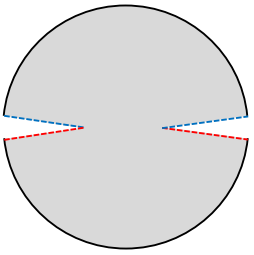
$$=$$



. (2.96)

We get the density matrix of the *thermo-mixed double* (TMD) [208–210] as

$$\rho_{\mathrm{TMD}} =$$



. (2.97)

The ρ_{TMD} is the ironed version of the up or down part of (2.96), and the connected region in the middle can be regarded as the "island". The TMD density is called "janus pacman" to distinguish "pacman" shown in (2.27).

The TMD state can be morally represented in figure 2.12. The density matrix of TMD and TFD are supposed to give out the same density matrix when tracing out the left part L

$$\rho_{BH} = \mathrm{tr}_L(\rho_{\mathrm{TMD}}) = \mathrm{tr}_L(\rho_{\mathrm{TFD}})\,, \tag{2.98}$$

whose von Neumann entropy is S_{BH}. More generally, for n-fold wormholes, we have Herman Verlinde's *replica ansatz* [210]

$$\mathrm{tr}(\rho^n_{\mathrm{TMD}}) = \frac{Z_n}{Z_1^n}\,, \tag{2.99}$$

with n-fold wormhole partition function Z_n.

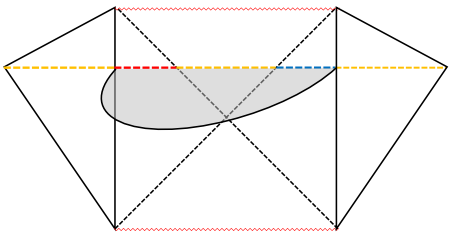
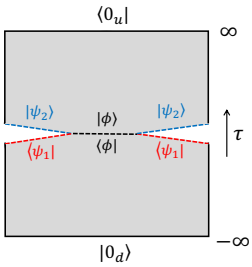


Figure 2.12: TMD state and island region for an old black hole on Penrose diagram

Equipped with the TMD interpretation, let us look at what the replica wormhole means. As discussed in section 2.1, the TFD state can be defined as a state evolved from the Minkowski vacuum. Here, by the same logic, we can also define the TMD as a state evolved from the Minkowski vacuum. The vacuum evolved from $\tau = -\infty$ can be written as $|0_d\rangle$ and the one from $\tau = \infty$ as $|0_u\rangle$. The subscripts u and d mark different possible vacua. Those different vacua can be identified by trace operation. Diagrammatically, when the boundaries representing $|0_u\rangle$ and $|0_d\rangle$ are connected with each other, we have a unified vacuum. When the boundaries form their own circles, we may have different vacua labelled by different subscripts.

Thus, the TMD density matrix can be represented as

$$\langle\psi_1|\,\rho_{\rm TMD}\,|\psi_2\rangle = \sum_{\phi}\langle\phi|\,\langle\psi_1|0_d\rangle\,\langle 0_u|\psi_2\rangle\,|\phi\rangle = \sum_{\phi}$$



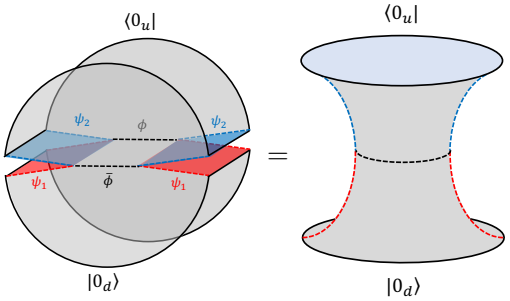
(2.100)

The replica wormhole can be further expressed as

$$\begin{aligned}\mathrm{tr}[\rho^2_{\rm TMD}] &= \sum_{\phi,\bar\phi}\sum_{\psi_1,\psi_2}\langle\psi_1|\,\langle\phi|0_d\rangle\,\langle 0_u|\phi\rangle\,|\psi_2\rangle\left[\langle\psi_1|\,\langle\bar\phi|0_d\rangle\langle\bar\phi|0_u\rangle\,|\psi_2\rangle\right]^* \\ &= \langle 0_u|^*\,\langle 0_u|0_d\rangle\,|0_d\rangle^*\,. \end{aligned} \tag{2.101}$$

Diagrammatically, the above expression is just the inverse process of (2.96)

$$\mathrm{tr}[\rho^2_{\rm TMD}] = \sum_{\psi_1,\psi_2}\sum_{\phi,\bar\phi}$$



. (2.102)

Now we can see from equations (2.101) and (2.102) that the replica wormhole actually counts the transition amplitude from different vacua evolved from $\tau = -\infty$ to $\tau = \infty$. The state at $\tau = -\infty$ and $\tau = \infty$ are not necessarily the same because of the independent circles. This is different from the Hawking saddle, where the up and down circles are connected with each other because of trace operation, as shown in figure 2.11. The final result of the replica wormhole saddle should be determined by the vacuum degeneracy at infrared. The related natural questions to ask are what is the infrared degeneracy in a black

hole system, and is this degeneracy coming from some symmetry breaking pattern...

The reason why we have spent so much time on TMD and vacua degeneracy is that we are trying to build a bridge between the recent progress on the BHIP and infrared physics. The recent progresses emphasise the importance of the non-local effects in the BHIP. And we have shown the close relationship between the replica wormhole and infrared vacuum degeneracy in this subsection. In the following chapters, we will treat the infrared physics of the black hole system seriously. For example, we will put the black hole into a box with boundary conditions on the surface of the box and calculate the entropy of the whole system with special attention to the degrees of freedom related to the nontrivial boundary conditions in chapter 3. In chapter 4, we will give a close look at the would-be-gauge degrees of freedom on the infrared boundary and possible connections between the soft hair and the BHIP. We will review more about the infrared structure of gauge and gravity theory [211] there, and further develop some relevant ideas in chapter 4.

We have to admit that the relationship between the BHIP and infrared physics is still suggestive here. More further studies are needed to understand gravitational non-local effects and infrared physics better.

3 Phase Transitions of Black Holes in Boxes

As discussed in the previous chapter, there are two basic logics to derive the Bekenstein-Hawking entropy from different "quantum" gravity theories. The first one is to add extra microscopic structures near the horizon. Under the assumption that each microscopic degree of freedom approximately occupies one Planck area, the entropy is always proportional to the Bekenstein-Hawking entropy Area/G_N. What exactly the micro-states are usually depends on the details of the theory, such as entanglement or spin network across the horizon. Despite different micro-structures in different approaches, they can always get the right magnitude of Bekenstein-Hawking entropy. The second one is related to symmetries, which claims that there are extra physical symmetries that acts in a well-defined way on the Hilbert space, conformal symmetry as an example. The Hilbert space is extended because of the symmetries, and the extra degrees of freedom give out the black hole entropy. The hidden symmetries can usually be found in extremal black holes. Those two logics both have merits and demerits. The attempts to derive the universal coefficient 1/4 in the Bekentein-Hawking entropy in the first case and generate the symmetry argument to finite temperature black holes are two important aspects of the research. This chapter is trying to understand this problem from a different perspective; maybe the tension between those two logics is a solution to each other.

In this chapter, we evaluate the partition function of U(1) gauge theory in three different backgrounds: flat case, finite-temperature black hole and extremal black hole, using Euclidean path integral. Various behaviour in different temperature limits are studied in detail. We get the standard blackbody radiation results for the flat case at high temperature. As for the black hole case, there is

an additional contribution proportional to the horizon area. The zero modes along the radius direction and the Wilson lines stretched between different boundaries play the dominant role for lower temperature. In this temperature limit, the entropy of the system starts to behave as the area of the boundaries multiplied by temperature squared. All the fluctuation modes are supposed to die off at very low temperature, and we are only left with constant modes and topological modes in the flat case. However, at very low temperature, we still get a Bekenstein-Hawking-like entropy contribution on the black hole background, which also persists in the extremal black hole case. There are two phase transitions and they provide a way to study the low-temperature Bekenstein-Hawking-like entropy from a symmetry breaking viewpoint.

3.1 Introduction

Gauge theories on a fixed background can be regarded as good toy models of gravitational fluctuations around saddle points of gravity theory. Let us consider a general Euclidean path integral over an arbitrary collection of fields Ψ

$$Z = \int \mathcal{D}\Psi \; e^{-\frac{1}{\hbar} S_E[\Psi]} \,, \tag{3.1}$$

where we have written $\hbar$ explicitly. In gravitational theories, We would also like to write the Newton's constant G_N explicitly, and the coefficient can be written as $1/(\hbar G_N)$. Suppose Ψ_{cl} are the classical solutions of action S_E, and $\delta\Psi$ are the fluctuations around the classical solutions $\delta\Psi = \Psi - \Psi_{cl}$. Expanding the action around the vicinity of Ψ_{cl}, we have

$$S_E[\Psi] = S_E[\Psi_{cl}] + \frac{1}{2\hbar} \int dx_1 dx_2 \frac{\delta^2 S_E}{\delta\Psi(x_1)\delta\Psi(x_2)}\Big|_{\Psi=\Psi_{cl}} \delta\Psi(x_1)\delta\Psi(x_2) + \dots \tag{3.2}$$

The linear terms of $\delta\Psi$ vanish because Ψ_{cl} are the stationary points of the action. Putting the above expansion into the path integral, we have Gaussian integrals for the quadratic terms in (3.2), and the path integral can be written as

$$Z = \sum_{\Psi_{cl}} e^{-\frac{1}{\hbar} S_E[\Psi_{cl}]} \times \det \left[\frac{1}{\hbar} \frac{\delta^2 S_E}{\delta\Psi(x_1)\delta\Psi(x_2)}\Big|_{\Psi=\Psi_{cl}} \right]^{-1/2} \times \dots \,. \tag{3.3}$$

The first part is the classical contribution, and the second term can be regarded as the one-loop correction of the theory. We only keep the quadratic terms in (3.2).

Now we can consider a gravitational theory by replacing the action with Einstein-Hilbert action. Let us denote the classical solutions of the theory as $\bar{g}^{(i)}_{\mu\nu}$, and the fluctuations around classical solutions as $h^{(i)}_{\mu\nu}$. Here the superscript (i) represent that we may have different saddle points labelled by i. The gravitational path integral can be written as

$$Z = \sum_i e^{-\frac{1}{\hbar G_N} S_E[\bar{g}^{(i)}_{\mu\nu}]} \times \det \left[\frac{1}{\hbar G_N} \frac{\delta^2 S_E}{\delta h_{\mu\nu} \delta h_{\mu\nu}} \Big|_{g^{(i)}_{\mu\nu} = \bar{g}^{(i)}_{\mu\nu}} \right]^{-1/2} . \tag{3.4}$$

Now, we can see that the one-loop corrections are captured by the linearised Einstein-Hilbert action, which is a massless quadratic Fierz-Pauli action. The massless Fierz-Pauli theory is a gauge theory with two physical polarisations; thus, we can use gauge theories on classical solutions of Einstein's theory, such as black hole, as toy models of metric perturbations. Gauge theory has a much simpler structure to deal with and can properly capture the gauge subtleties of the massless Fierz-Pauli theory. Furthermore, U(1) gauge theory is the simplest one to play with among all other gauge theories. So, we mainly deal with U(1) gauge theory on black hole background using Euclidean path integral in this chapter.

Let us summarise and highlight the main results of this chapter here. We analyse the partition function of U(1) gauge field living between two parallel boundaries in this chapter. The system was also put onto black hole backgrounds, where the U(1) gauge field can be regarded as a toy model of gravitational fluctuations. The basic setup is shown in figure 3.1, where the radius coordinate is labelled by r and the transverse coordinate are x^a. The boundary conditions on different boundaries are shown in equation (3.15), where A_a are set to be Dirichlet, and A_r can fluctuate and have arbitrary dependence on the boundary coordinates x^a. One specific set of boundary conditions defines a Hilbert space for us, and we are going to sum over the physical degrees of freedom within the given Hilbert space in the path integral. This means that we need to work with a fixed $A_a|_{\partial\mathcal{M}}$ configuration, and sum over different $A_r|_{\partial\mathcal{M}}$ configurations. The Euclidean path integral contains four different parts: bulk fluctuation modes $\hat{A}_\mu$, zero modes of A_r along radius direction ϕ, boundary stretched Wilson lines W, and constant modes and winding modes.

There is not too much surprise in the flat case. We expect the bulk fluctuation modes dominate at high temperature, and the modes arising because of the boundary conditions become more important as the temperature becomes lower. After evaluating the path integral, we get the following results. The

bulk fluctuation modes always play the dominant role at very high temperature, whose entropy scales as the volume of the bulk multiplied by temperature curbed

$$\mathcal{S}_{\hat{A}} \propto \text{Volume} \times T^3 \,, \tag{3.5}$$

where we use $\mathcal{S}$ to denote the entropy to avoid confusion with the action S, and T to denote the temperature. As the temperature becomes low, i.e. the ratio between the distance of the two boundaries L and inverse temperature β becomes small, the zero modes ϕ and the Wilson lines W that behave like boundary scalar fields give out the dominant contributions. The entropy of those modes scales as

$$\mathcal{S}_{\phi,W} \propto \text{Area} \times T^2 \,. \tag{3.6}$$

As the temperature becomes super-low, no fluctuation mode can be seen, and we are left with some constant modes and winding modes contributions. The entropy of those modes is approximately the logarithm of the coupling constant and temperature. This is shown in figures 3.3 and 3.4.

The most interesting part lies in the black hole case, where we have one boundary on the stretched horizon and the other at a finite distance L away from the horizon. The setup here is very similar to 't Hooft's "brick wall" model [2], and the difference with the flat case already shows up for bulk fluctuation modes $\hat{A}_\mu$. The corresponding entropy of bulk modes $\hat{A}_\mu$ is shown in equation (3.116), which can be more or less expressed as

$$\mathcal{S}_{\hat{A}} \propto \left(\frac{r_s^2}{\delta^2} + \ln \frac{L r_s}{\delta^2} \right) r_s^3 \times T^3 + \text{Volume} \times T^3 \,, \tag{3.7}$$

where r_s is the radius of the black hole horizon, and δ is defined as the proper distance between the "real" horizon and "stretched" horizon

$$\delta = \int_{r_s}^{r_s+\varepsilon} \sqrt{g_{rr}}\, dr \,. \tag{3.8}$$

The entropy involved with δ can be interpreted as follows [117]. Because of the redshift, any finite frequency modes near the horizon can be regarded as zero-frequency modes by a coordinate observer sitting at infinity. So we have an infinite number of states with very small energy, and the ultraviolet (UV) cutoff is cutting off the summation over states. So besides the volume contribution we have an extra contribution from the modes living very close to the horizon, as shown in equation (3.7). Those modes are used to explain the Bekenstein-Hawking entropy by some authors [2, 3]. Note that this contribution can not

be seen in the extremal black hole case, where the inverse temperature β is much larger than the length scale of the horizon r_s.

As the temperature of the system becomes lower, similar to the flat case, the radius zero modes ϕ and the boundary stretched Wilson lines W behave like two scalar fields living in a lower-dimensional spacetime, whose entropy scales as (3.6). The area contribution (3.6) is the dominate one over all different contributions since the high-frequency modes along the radius direction are gapped.

The biggest surprise appears at the very low temperature. The coefficients in front of $(\partial_\tau \phi)^2$ and $(\partial_\tau W)^2$ in the effective actions diverge, introducing localisation in the space of zero-energy modes $\partial_\tau \phi = \partial_\tau W = 0$ in the Euclidean path integral. The effective actions are shown in equations (3.136) and (3.139). The corresponding entropy of ϕ and W can be expressed as

$$\mathcal{S} \propto \frac{r_s^2}{l_p^2}\,, \tag{3.9}$$

which gives out the right magnitude of Bekenstein-Hawking entropy. The overall entropy of the finite black hole case is shown in figure 3.7. We also show that the localisation and contribution (3.9) persists in the extremal black hole case.

The question is how to understand the large entropy (3.9) we have gotten for the low-temperature black hole case and the extremal black hole case. Let us take ϕ as an example to show where this large entropy comes from. The finite temperature partition function for a massless bosonic field can be written as

$$\ln Z_\phi \propto \beta \text{Area} \cdot \Lambda^3 + \text{Area} \int d^2p \ \ln(1 - e^{-\beta p})\,, \tag{3.10}$$

where Λ is the UV cutoff and $p = \sqrt{\vec{p}^{\,2}}$. The first part proportional to the volume of the whole spacetime divided by the smallest volume unit is the zero-point energy of the field theory, which is a constant piece in the free energy. Constant free energy does not contribute to the entropy, thus can be ignored. The second part is finite and gives out the Area $\times\, T^2$ contribution in entropy. However, it can not be seen at very low temperature.

Now because of the localisation, the zero-point energy of the field is not a constant free energy anymore. The logarithm of the partition function only depends on the area of the boundary because ϕ only depends on the spatial coordinates on the boundary. So the partition function of the zero-point energy

can be written as

$$\ln Z_\phi \propto \text{Area} \cdot \Lambda^2 = \frac{\text{Area}}{l_p^2}\,, \tag{3.11}$$

if we suppose the UV cutoff is at Planck scale. The corresponding entropy is of Bekenstein-Hawking entropy magnitude shown in equation (3.9). There are also constant modes and winding modes contributions at very low temperature, whose entropy scales as logarithm of the coupling constant and other length scales. It is very interesting to notice that the area shown in equation (3.11) only contains the area of the horizon not the area of the other boundary.

Now, we have seen that the Bekenstein-Hawking-like entropies and logarithm corrections come from two different places, shown in (3.7) and (3.9) separately. The first place corresponds to the degrees of freedom living very close to the horizon, and the second place corresponds to the vacuum degeneracy near the horizon. The first contribution only appears in the finite temperature black hole, and the second one only appears in a very low-temperature limit or the extremal black hole case. We thus infer that the attempts in different theories that are trying to add quantum structure near the horizon are counting the same thing as the contributions shown in equation (3.7). Furthermore, the attempts that are trying to find hidden symmetries for extremal black holes and explaining the black hole entropy from a symmetry breaking viewpoint are counting the same thing as shown in (3.9). The Bekenstein-Hawking entropy for the finite-temperature black hole and the extremal black hole might be counting completely different things.

In this chapter, we study the partition function of the U(1) gauge theory in three different situations: flat case, finite temperature black hole, and extremal black hole. The chapter is organised as follows. In section 3.2, we study the symplectic structure of the gauge theory and evaluate the partition function of the gauge theory in different temperature limits. And then, we repeat the same calculations on a black hole background in section 3.3. We compare the difference between the flat and black hole cases. Next, we check if the similar behaviour persists to be true in the extremal black hole case in section 3.4. Finally, we summarise the whole chapter and provide some further discussions in section 3.5. More details of the calculations are exiled to the appendixes.

3.2 Gauge Fields between Flat Parallel Plates

In this section, we study the U(1) gauge theory living between flat parallel plates. The canonical analysis shown in subsection 3.2.1 helps us to see the physical degrees of freedom in the phase space by looking at the Poisson bracket. And we would only include physical degrees of freedom in the Euclidean path integral. It turns out we have four different parts in the path integral, which are bulk fluctuation modes $\hat{A}_\mu$, zero modes of A_r along radius direction ϕ, boundary stretched Wilson lines W, as well as constant modes and winding modes. The zero modes ϕ and the Wilson lines W arising because of the boundary conditions that we will see late, are very similar to Barnich's non-proper gauge degrees of freedom [212, 213]. The basic behaviour of the corresponding entropy is shown in figure 3.4, which scales as the volume of the bulk, area of the boundary, and logarithm of the coupling constant and temperature as the system's temperature cool down. The most crucial difference with the normal black body radiation case is the boundary degrees of freedom Area $\times T^2$ coming from the zero modes and the Wilson lines.

The flat case is mainly used to illustrate the basic concepts, physical modes in the phase space, and basic behaviour of the entropy, which can be regarded as the "background noise" of curved spacetime calculations. The deviations from the flat case in later sections can be regarded as special properties of the black hole system. The reader who already familiar with the physics of gauge theory with boundaries can skip this section, but do keep the figure 3.4 and the behaviour of the zero modes ϕ and the Wilson lines W in mind.

The situation we are mainly interested in is shown in figure 3.1, where we have a Maxwell field theory living between two parallel boundaries on the left- and right-hand side. The original Maxwell theory has the action

$$S = -\frac{1}{4e^2}\int_{\mathcal{M}} d^4x \; F^{\mu\nu}F_{\mu\nu}\,, \tag{3.12}$$

where $\mathcal{M}$ is the 4-dimensional manifold, and e^2 is a dimensionless coupling constant. The coupling constant should be replaced by G_N in the linearised gravity theory. The 4-dimensional box has coordinate system $x^\mu = (x^a, r) = (t, x^2, x^3, r)$, where r is the radius direction, and x^a are the directions along the boundaries. In order to do finite temperature field theory calculations, we Wick rotate the time direction $t \to -i\tau$ such that τ becomes the Euclidean time with a periodicity β which is also known as the inverse temperature of the spacetime. If all the gauge fields die off near the boundary, this is just

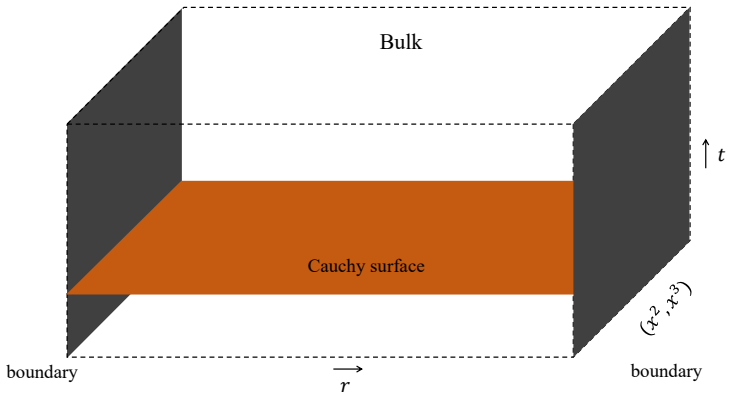


Figure 3.1: U(1) gauge theory living between two parallel boundaries. The orange surface is a Cauchy surface with constant time.

the blackbody radiation with two polarisation degrees of freedom after gauge fixing. However, interesting phenomena start to show up when we release the boundary conditions, meaning that the boundary degrees of freedom are added into the system. Let us suppose the boundaries shown in figure 3.1 are labelled by $r = r_\alpha$. Then, in order to have a well-defined Hilbert space and variation principle, there would be some constraints for the boundary conditions. To see those constraints, let us first look at the variation of the action (3.12), which can be written as

$$\delta S = \frac{1}{e^2}\int_{\mathcal{M}} d\tau d^3x\ \partial_\mu F^{\mu\nu}\delta A_\nu - \frac{1}{e^2}\int_{\partial\mathcal{M}} d^3x\ n_\mu F^{\mu\nu}\delta A_\nu\,. \tag{3.13}$$

For the boundaries with normal vector $n^\mu\partial_\mu = \partial_r$, as shown in figure 3.1, the on-shell variation of the action can be written as

$$\delta S = -\frac{1}{e^2}\int_{\partial\mathcal{M}} d^3x\ F^{ra}\delta A_a\,. \tag{3.14}$$

Here, to have a well-defined variation principle without adding any Gibbons-Hawking-like terms, we have two obvious choices: $\delta A_a\big|_{\partial\mathcal{M}} = 0$ or $F^{ra}\big|_{\partial\mathcal{M}} = 0$. For the first choice $\delta A_a\big|_{\partial\mathcal{M}} = 0$, $A_a\big|_{\partial\mathcal{M}}$ are the configurations that fixed on the boundaries. Meanwhile, there is no constraint for A_r at the boundary, and $F^{ra}\big|_{\partial\mathcal{M}}$ can be arbitrary value. The Hilbert space is well-defined with the fixed boundary configurations $A_a\big|_{\partial\mathcal{M}}$, and we need to sum over different boundary configurations of A_r in the path integral. For the second choice, we need to sum over all the boundary configurations that respect $F^{ra}\big|_{\partial\mathcal{M}} = 0$.

Here we are mainly going to focus on the first choice, and leave the other choices for future studies.

So the boundary conditions we are going to set on $r = r_\alpha$ are

$$A_a\Big|_{r=r_\alpha} = f_a^\alpha(x^a)\,, \tag{3.15}$$

$$A_r\Big|_{r=r_\alpha} = \text{Arbitrary}\,. \tag{3.16}$$

where the superscript α labels different boundaries as shown in figure 3.1. $f_a(x^a)$ are fixed boundary configurations, $3x^4$ as an example. As a contrast, $A_r|_{\partial\mathcal{M}}$ can have arbitrary dependence of x^a and can fluctuate. As a consequence of the above boundary conditions, we have $\delta A_a = 0$ and do not need to add any Gibbons-Hawking-like boundary term.

For the boundary configurations shown in (3.15), there are bulk modes corresponding to them. To see the bulk modes that corresponding to the boundary configurations, we can separate A_a into two parts, the part $\hat{A}_a$ that vanish on the boundaries and the on-shell part B_a that satisfy the boundary conditions (3.15). Thus, the gauge fields A_a can be written as

$$A_a = \hat{A}_a + B_a\,. \tag{3.17}$$

For a given set of boundary configurations f_a^α, we can find solutions of bulk equations of motion that satisfy the above boundary conditions in principle. However, the solutions can be hard to solve. Here we just take a subset of f_a^α: the flat boundary configurations $f_a^\alpha(x^a) = C_a^\alpha + \partial_a\lambda^\alpha$ to solve the bulk equations of motion, where C_a^α are constants, and $\lambda^\alpha(x^a)$ are the gauge parameters on boundaries. One obvious solution can be written as

$$B_a = \left[(C_a^{(r)} - C_a^{(l)}) + \partial_a(\lambda^{(r)} - \lambda^{(l)})\right]\frac{r}{L} + (C_a^{(l)} + \partial_a\lambda^{(l)})\,, \tag{3.18}$$

$$B_r = \frac{1}{L}[x^a\cdot(C_a^{(r)} - C_a^{(l)}) + (\lambda^{(r)} - \lambda^{(l)})]\,. \tag{3.19}$$

Again, L is the distance between the two boundaries as shown in figure 3.1. The superscripts (r) and (l) are the values of α that label the right and left boundaries. (3.18) and (3.19) are the bulk on-shell configurations that corresponding to the boundary conditions (3.15). Note that the on-shell configurations B_r are needed because of the following reasons. A_a can take different values on different boundaries, thus the bulk on-shell configurations B_a must depend on the radius coordinate r. Then B_r is needed such that the bulk equations of motion are satisfied.

Special attentions are needed for the A_r component, which is where the "meat" lies on. $A_r\big|_{\partial\mathcal{M}}$ can fluctuate and have arbitrary x^a dependence; thus we need to separate the bulk and boundary configurations carefully. A_r can be separated as follows

$$A_r(x^\mu) = \hat{A}_r(x^\mu) + \frac{\phi(x^a)}{L} + B_r\,. \tag{3.20}$$

where B_r are the bulk on-shell modes discussed above. $\phi(x^a)$ are the configurations that makes $\hat{A}_r\big|_{r=0} = 0$, i.e.

$$A_r(r, x^a)\Big|_{r=0} = \frac{\phi(x^a)}{L} + B_r\,. \tag{3.21}$$

$\hat{A}_r$ does not equal zero at the boundary $r = L$, and we can also further decompose $\hat{A}_r$ into two parts. The part that satisfies $\int_0^L dr\ \hat{A}_r = 0$ will not be our main concern here, and the part that satisfies $\int_0^L dr\ \hat{A}_r \neq 0$ will be important ingredient in our later calculations. Note that one can also decompose $\hat{A}_r$ into the part that vanishes on both side and the part that captures difference between the two boundaries. Integrating the part that vanishes on both boundaries from one boundary to the other gives out zero, and $\int_0^L dr\ \hat{A}_r$ more or less captures the difference of A_r between the two boundaries.

We will eventually functional integrate over all the physical degrees of freedom in the Euclidean path integral to get the partition function of the system with the given boundary conditions. However, before we actually do the calculation, let us first analyse the theory's canonical phase space to get a feeling of which degrees of freedom are supposed to be integrated over in the path integral.

3.2.1 Canonical formulation

In the Hamiltonian formulation, the phase space can be represented by Γ, which is an even-dimensional manifold with coordinates $x^I = \{q^i, p_j\}$, where q^i and p_j are the canonical coordinates and momenta. For field theories, we have infinite-dimensional phase spaces Γ.

For the current theory at hand, we can decompose the gauge fields into temporal and spatial directions. The notations of different surfaces are shown in figure 3.2. Rewriting the gauge fields A_μ as

$$A_\mu = (-V, A_i), \qquad A^\mu = (V, A^i)\,, \tag{3.22}$$

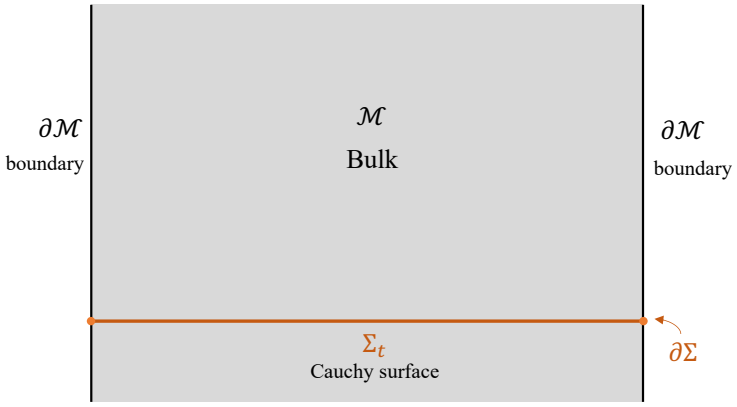


Figure 3.2: The notations that labels the bulk manifold $\mathcal{M}$, the boundaries $\partial\mathcal{M}$, Cauchy surface Σ_t, and the boundaries of the Cauchy surface $\partial\Sigma$ are illustrated in this figure.

we have the Lagrangian density written in terms of V and A_i as

$$\mathcal{L} = \frac{1}{2e^2}(\dot{A}^i + \partial^i V)(\dot{A}_i + \partial_i V) - \frac{1}{2e^2}F^{ij}\partial_i A_j \,. \tag{3.23}$$

The corresponding conjugate momenta of fields V and A_i can be written as

$$\Pi_V = \frac{\partial\mathcal{L}}{\partial\dot{V}} \qquad \Pi^i = \frac{\partial\mathcal{L}}{\partial\dot{A}_i} = \frac{1}{e^2}(\dot{A}^i + \partial^i V)\,, \tag{3.24}$$

where we denoted Π_V as the momentum for V and Π^i as the momenta for A_i.

So, for trivial boundary condition, the phase space we start with is $\Gamma = \{V, A_i, \Pi_V, \Pi^i\}$. Gauge fixing conditions helps us further getting rid of the unphysical degrees of freedom in the phase space. However there are some boundary subtleties in the phase space when we have nontrivial boundary conditions. The degrees of freedom related to the boundary condition need to be added into the phase space. What are those boundary subtleties? Let us turn to the symplectic form of the theory and work out the Poisson bracket between the fields to make those boundary subtleties more explicit. The phase space is equipped with a closed, non-degenerate symplectic two-form Ω, which is defined as

$$\Omega = \frac{1}{2}\Omega_{IJ}\,\mathrm{d}x^I \wedge \mathrm{d}x^J \,. \tag{3.25}$$

Ω_{IJ} is invertible, and the inverse Ω^{IJ} is defined by $\Omega^{IK}\Omega_{KJ} = \delta^I_J$. Now

equipped with the symplectic form, the classical Poisson bracket between functionals F and G can be defined as

$$\{F, G\} = \Omega^{IJ} \frac{\delta F}{\delta x^I} \frac{\delta G}{\delta x^J} \,. \tag{3.26}$$

Quantum commutators can be obtained by adding a factor i in canonical quantisation. The symplectic form can be obtained by the covariant phase space formalism [144] on a chosen Cauchy surface.

Let us consider a field theory with a Lagrange density $\mathcal{L}[\Psi]$, where Ψ denotes an arbitrary collection of fields. Taking the variation of $\mathcal{L}$, we have

$$\delta\mathcal{L} = E \cdot \delta\Psi + \mathrm{d}\Theta \,. \tag{3.27}$$

The equation of motion $E = 0$ kills the first term in the above expression. The (pre)-symplectic potential $\Theta[\Psi, \delta\Psi]$ is a $D-1$ form and can be integrated over a chosen $(D-1)$-dimensional Cauchy surface. The symplectic current ω can be defined as

$$\omega[\Psi, \delta_1\Psi, \delta_2\Psi] = \delta_1\Theta[\Psi, \delta_2\Psi] - \delta_2\Theta[\Psi, \delta_1\Psi] \,, \tag{3.28}$$

where δ_1 and δ_2 can be regarded as variations with respect to two different transformations. Integrating the symplectic current ω over the Cauchy surface Σ, we finally get the symplectic form Ω written as

$$\Omega[\Psi, \delta_1\Psi, \delta_2\Psi] = \int_\Sigma \omega[\Psi, \delta_1\Psi, \delta_2\Psi] \,. \tag{3.29}$$

Note that the choice of the $(D-1)$-form $\omega[\Psi, \delta_1\Psi, \delta_2\Psi]$ also depends on the Cauchy surface. Specifying to the temporal surface Σ_t with normal vector n^t, we can work out each component of ω.

For the U(1) gauge theory at hand, the action was shown in the equation (3.13) with the boundary $\partial\mathcal{M}$ replaced by Cauchy surface Σ. Thus, the symplectic form Ω_Σ can be written as

$$\Omega_\Sigma = -\frac{1}{e^2} \int_\Sigma d^3x \; n^\mu \delta F_{\mu\nu} \wedge \delta A^\nu \,. \tag{3.30}$$

Specifying on the chosen Cauchy surface Σ_t with normal vector $n^\mu\partial_\mu = \partial_t$, the above expression can be written in components as

$$\Omega_{\Sigma_t} = -\frac{1}{e^2} \int_{\Sigma_t} d^3x \; \delta F_{ti} \wedge \delta A^i \,. \tag{3.31}$$

The symplectic form is essential in defining the Hamiltonian dynamics; only the fields equipped with a nontrivial symplectic form with their conjugate momentum can be regarded as dynamical variables in the phase space. We have decomposed the gauge fields into several parts as we discussed before

$$A_a = \hat{A}_a + B_a \,, \tag{3.32}$$

$$A_r = \hat{A}_r + \frac{\phi}{L} + B_r \,. \tag{3.33}$$

We are going to see what is the symplectic partners of all the above configurations in the symplectic form.

Now putting everything back into the symplectic form (3.31), we have

$$\Omega_{\Sigma_t} = -\frac{1}{e^2}\int_{\Sigma_t} d^3x \left[\delta(\hat{F}_{tr} + \frac{\dot{\phi}}{L}) \wedge \delta(\hat{A}^r + \frac{\phi}{L}) + \delta\hat{F}_{t2} \wedge \delta\hat{A}^2 + \delta\hat{F}_{t3} \wedge \delta\hat{A}^3\right] . \tag{3.34}$$

The variation of the on-shell configuration B_i in the bulk is zero because the variations of the boundary configurations are all zero. We can separate the $\hat{A}_i$ part with other parts, and the symplectic form reads as

$$\begin{aligned}\Omega_{\Sigma_t} = & -\frac{1}{e^2}\int_{\Sigma_t} d^3x \; \delta\hat{F}_{ti} \wedge \delta\hat{A}^i \\ & -\frac{1}{e^2}\int_{\Sigma_t} d^2x dr \left[\frac{\delta\dot{\phi}}{L} \wedge \delta\hat{A}^r + \frac{\delta\dot{\phi}}{L} \wedge \frac{\delta\phi}{L} + \delta\hat{F}_{tr} \wedge \frac{\delta\phi}{L}\right] .\end{aligned} \tag{3.35}$$

The first term in (3.35) gives us the usual Poisson bracket of Maxwell's theory. Integrating over r in the second term gives out

$$-\frac{1}{L\cdot e^2}\int d^2x \left[\delta\dot{\phi} \wedge \delta(\int_0^L dr\hat{A}^r) + \delta\dot{\phi} \wedge \delta\phi + \delta(\int_0^L dr\dot{\hat{A}}_r) \wedge \delta\phi\right] , \tag{3.36}$$

where we have used $\delta\hat{A}_\tau\big|_{\partial\Sigma} = 0$. The above symplectic form tells us what are extra physical degrees of freedom in the system. Let us define the quantity W as

$$W = i\int_0^L dr \; \hat{A}_r \,, \tag{3.37}$$

which will be called Wilson lines in this chapter[1]. Note that field W captures

[1]The actual Wilson lines stretched between the two boundaries can be written as

$$\mathcal{W} \propto \mathcal{P}\exp[i(\int_0^L dr \; \hat{A}_r) + i\phi] \,, \tag{3.38}$$

the difference of A_r on the two boundaries. From now on, we will extract W modes out of $\hat{A}_r$ such that $\hat{A}_r$ equal zero at both boundaries.

The symplectic form can be written as

$$\Omega_{\Sigma_t} = -\frac{1}{e^2}\int_{\Sigma_t} d^3x\ \delta\hat{F}_{ti}\wedge\delta\hat{A}^i - \frac{1}{e^2 L}\int d^2x\ \left[-i\delta\dot{\phi}\wedge\delta W - i\delta\dot{W}\wedge\delta\phi + \delta\dot{\phi}\wedge\delta\phi\right]. \tag{3.39}$$

The cross term between W and ϕ can be cancelled by refining the fields. For example, shifting $\phi \to \phi + iW$, the above symplectic form can be written as

$$\Omega_{\Sigma_t} = -\frac{1}{e^2}\int_{\Sigma_t} d^3x\ \delta\hat{F}_{ti}\wedge\delta\hat{A}^i - \frac{1}{e^2 L}\int d^2x\ \left[\delta\dot{\phi}\wedge\delta\phi + \delta\dot{W}\wedge\delta W\right]. \tag{3.40}$$

Now we have the conjugate momentum of W and ϕ be written as

$$\Pi_W = \dot{W}, \quad \Pi_\phi = \dot{\phi}. \tag{3.41}$$

The symplectic form can be written as

$$\Omega_{\Sigma_t} = -\frac{1}{e^2}\int_{\Sigma_t} d^3x\ \delta\hat{\Pi}_i\wedge\delta\hat{A}^i - \frac{1}{e^2 L}\int d^2x\ \left[\delta\Pi_W\wedge\delta W + \delta\Pi_\phi\wedge\delta\phi\right]. \tag{3.42}$$

$\hat{\Pi}^i$ denotes the conjugate momentum of $\hat{A}_i$. The brackets can be derived as

$$\frac{1}{e^2}[\ \hat{\Pi}^i(r,x^2,x^3), \hat{A}_j(r',x'^2,x'^3)\ ] = i\delta^i_j\ \delta(r-r')\ \delta^2(x-x'), \tag{3.43}$$

$$\frac{1}{e^2 L}[\ \Pi_W(x^2,x^3), W(x'^2,x'^3)\ ] = i\delta^2(x-x'), \tag{3.44}$$

$$\frac{1}{e^2 L}[\ \Pi_\phi(x^2,x^3), \phi(x'^2,x'^3)\ ] = i\delta^2(x-x'). \tag{3.45}$$

Note that it is natural to use temporal gauge $A_t = 0$ in the canonical formulation.

As we discussed at the beginning of this subsection, we need to add back the degrees of freedom related to the boundary subtleties, i.e. degrees of freedom related to the boundary configurations of A_r. By the symplectic form analysis, we have found the degrees of freedom that have nontrivial symplectic partner and Poisson bracket. Those are the degrees of freedom needed to add back. So the actual phase space should be

$$\Gamma = \left\{\ \hat{\Pi}^i, \hat{A}_i, \Pi_\phi, \phi, \Pi_W, W\ \right\}. \tag{3.46}$$

with $\mathcal{P}$ denoting the path ordering, which actually is a combination of W and the zero modes ϕ.

Further gauge fixing conditions would help us to get rid of gauge redundancy of $\{\hat{\Pi}^i, \hat{A}_i\}$ such that we are only left with two bulk polarisations. Here ϕ is the zero modes along r direction of A_r, and W are the Wilson lines stretched between those two boundaries. Those degrees of freedom is very similar to the soft photon degrees of freedom [211, 214, 215].

Canonical formula and Euclidean path integral

Let us use a 4 dimensional scalar field Ψ with Lagrangian density

$$\mathcal{L} = -\frac{1}{2}\partial^\mu \Psi \partial_\mu \Psi - \frac{1}{2}m^2\Psi^2 - V(\Psi) \tag{3.47}$$

to demonstrate the relationship between the canonical formula and the Euclidean path integral. The transition amplitude between $|\Psi_1\rangle$ and $|\Psi_2\rangle$ can be represented as

$$\langle \Psi_2 | \, e^{-iHt_0} \, |\Psi_1\rangle = \int \mathcal{D}\Pi \int_{\Psi_1}^{\Psi_2} \mathcal{D}\Psi \times \exp\left[i \int_0^{t_0} dt \int d^3x (\Pi\dot{\Psi} - \mathcal{H}[\Pi, \Psi]) \right] , \tag{3.48}$$

where Π is the conjugate momentum of Ψ, H is the Hamiltonian of the theory and $\mathcal{H}$ is the Hamiltonian density. The partition function can be written as

$$Z = \text{tr } e^{-\beta H} = \int \mathcal{D}\Pi \int \mathcal{D}\Psi \times \exp\left[\int_0^\beta d\tau \int d^3x (i\Pi\dot{\Psi} - \mathcal{H}[\Pi, \Psi]) \right] , \tag{3.49}$$

with the Euclidean time $\tau = it$. The Ψ path integral is over $\Psi(0) = \Psi(\beta)$. Let us now discretise the path integral by defining $\Delta\tau = \beta/N$ with large number

N. Then the discretised partition function can be written as

$$\begin{aligned} Z &= \mathcal{N}\cdot\prod_j^N\int d\Pi_j\int d\Psi_j\times\exp\Delta\tau\int d^3x \\ &\quad\times\left[i\Pi_j\frac{\Psi_{n+1}-\Psi_j}{\Delta\tau}-\frac{1}{2}\Pi_j^2-\frac{1}{2}(\nabla\Psi_j)^2-\mu^2\Psi_j^2-V\right] \\ &= \mathcal{N}'\cdot\prod_j^N\int d\Psi_j\times\exp\Delta\tau\int d^3x \\ &\quad\times\left[-\frac{1}{2}\left(\frac{\Psi_{n+1}-\Psi_j}{\Delta\tau}\right)^2-\frac{1}{2}(\nabla\Psi_j)^2-\mu^2\Psi_j^2-V\right] \\ &= \mathcal{N}'\cdot\int\mathcal{D}\Psi\times\exp\left[-\int_0^\beta d\tau\int d^3x[\frac{1}{2}\dot{\Psi}^2+\frac{1}{2}(\nabla\Psi)^2+\mu^2\Psi^2+V]\right] \\ &= \mathcal{N}'\cdot\int\mathcal{D}\Psi\times e^{-S_E}\,, \end{aligned} \tag{3.50}$$

where $\mathcal{N}$ and $\mathcal{N}'$ are constants. The integrals over conjugate momentum Π_j in the second line are Gaussian integrals, and thus can be easily worked out. Now we have shown the relation between $\mathrm{tr}e^{-\beta H}$ and Euclidean path integral $\int\mathcal{D}\Psi\ e^{-S_E}$.

Now, we are going to include the physical degrees of freedom, analysed in the current section, into the Euclidean path integral to evaluate their contribution to the entropy, which will be our main task for the next subsection. Note that the canonical analysis gives us some hints about what degrees of freedom are physical, but we are not going to use the canonical results as an input. We will work out the Euclidean path integral from scratch, and may use different gauge fixing conditions, if they are more convenient, in the next subsection.

3.2.2 Euclidean path integral

Now, after the canonical analysis, we know the dynamical modes are the bulk fluctuation modes $\hat{A}_\mu$, zero modes along r direction ϕ, and the Wilson lines stretched between the two boundaries W. Those fields are the ingredients that need to be added in the Euclidean path integral. Note that there should only be two physical polarisations for the bulk fields $\hat{A}_\mu$ after gauge fixing. We will deal with bulk gauge fixing conditions after the physics are clear to avoid gauging too much or too little. Note that there are also some modes B_μ shown in equations (3.18-3.19) corresponding to the boundary conditions. For a chosen set of boundary configurations f_a^α, a unitary Hilbert space is fixed

with field contents $\hat{A}_\mu$, ϕ and W. So we will not sum over different boundary conditions B_μ in the path integral here. We will discuss the issues related to summing over boundary conditions at the end of this chapter.

As discussed in the previous subsection, the partition function can be written as a Euclidean path integral

$$Z = \int \mathcal{D}A_\mu \; e^{-S_E} \,, \tag{3.51}$$

where S_E is the Euclidean action

$$S_E = \frac{1}{4e^2} \int_{\mathcal{M}} d\tau d^3x \; F^{\mu\nu} F_{\mu\nu} \,. \tag{3.52}$$

The above is the original formula, and we are going to do some massage according to the hints from the canonical analysis. We can separate the x^a directions with r in the action

$$S_E \;=\; \frac{1}{4e^2} \int_{\mathcal{M}} d\tau d^3x \; F^{ab} F_{ab} + \frac{1}{2e^2} \int_{\mathcal{M}} d\tau d^3x \; F^{ra} F_{ra} \,. \tag{3.53}$$

Again, we are going to separate the gauge fields into different parts

$$A_a \;=\; \hat{A}_a + B_a \,, \tag{3.54}$$

$$A_r \;=\; \hat{A}_r + \frac{\phi(x^a)}{L} + B_r \,. \tag{3.55}$$

The Euclidean action can be written in terms of those modes as

$$\begin{aligned} S_E &= \frac{1}{4e^2} \int_{\mathcal{M}} d\tau d^3x \; F^{ab} F_{ab} + \frac{1}{2e^2} \int_{\mathcal{M}} d\tau d^3x \; (\hat{F}^{ra} - \frac{\partial^a \phi}{L})(\hat{F}_{ra} - \frac{\partial_a \phi}{L}) \\ &= \frac{1}{4e^2} \int_{\mathcal{M}} d\tau d^3x \; \hat{F}^{\mu\nu} \hat{F}_{\mu\nu} + \frac{1}{2e^2} \int_{\mathcal{M}} d\tau d^3x \; \left[-\frac{2}{L} \; \hat{F}^{ra} \partial_a \phi + \frac{\partial^a \phi \partial_a \phi}{L^2} \right] . \end{aligned}$$

Let's denote the second part in the above expression as

$$S = \frac{1}{2e^2} \int_{\mathcal{M}} d\tau d^2x dr \; \left[\frac{\partial^a \phi \partial_a \phi}{L^2} + \frac{2}{L} \; \hat{F}^{ar} \partial_a \phi \right] . \tag{3.56}$$

$\phi(x^a)$ is not a function of radius direction r, so the integral over r can just pass through ϕ, and gives out an extra L in the first term. Noticing that $\hat{A}_a$ equal zero at the boundaries, integrating over r in the above effective action gives out

$$\begin{aligned} S &= \frac{1}{2e^2} \int d\tau d^2x \; \left[\frac{\partial^a \phi \partial_a \phi}{L} - \frac{2i}{L} \; \partial^a (i \int dr \hat{A}^r) \partial_a \phi - \frac{2}{L} \; \hat{A}_a \big|_{r_\alpha} \partial_a \phi \right] \\ &= \frac{1}{2e^2 L} \int d\tau d^2x \; \left[\partial^a \phi \partial_a \phi - 2i \; \partial^a (i \int dr \hat{A}^r) \partial_a \phi \right] . \end{aligned} \tag{3.57}$$

Denoting $W(x^a) = i\int dr\hat{A}_r$, the original action can be rewritten as

$$S_E = \frac{1}{4e^2}\int_{\mathcal{M}} d\tau d^3x\ \hat{F}^{\mu\nu}\hat{F}_{\mu\nu} + \frac{1}{2e^2L}\int d\tau d^2x\ \left[\partial^a\phi\partial_a\phi - 2i\ \partial^a W\partial_a\phi\right]\ . \quad (3.58)$$

The cross term between fields W and ϕ are cancelled by shifting $\phi \to \phi + iW$. An effective action for W can be getting from this procedure, which is the analogy of canonical analysis. We can also get the effective action for field W by integrating over ϕ in the Euclidean path integral. The path integral over ϕ is Gaussian, which gives out $\det(\partial^2)^{-1/2}$ in the partition function, and the effective action for W is the action for 3-dimensional massless scalar field. The determinant $\det(\partial^2)^{-1/2}$ can be rewritten as integral over bosonic field ϕ. With all the above arguments, the effective action can be expressed as

$$S_E = \frac{1}{4e^2}\int_{\mathcal{M}} d\tau d^3x\ \hat{F}^{\mu\nu}\hat{F}_{\mu\nu} + \frac{1}{2e^2L}\int d\tau d^2x\ \left[\partial^a\phi\partial_a\phi + \partial^a W\partial_a W\right]\ . \quad (3.59)$$

All the above calculations are similar to the symplectic form calculation shown in the previous subsection.

Now we can put the effective action back into the path integral (3.51). Note that the integral over A_r component is divided into several different pieces. B_r is the fixed configuration which does not need to be integrated over. The zero modes ϕ and the Wilson lines W are capturing the boundary configurations of A_r. Finally, we are left with the bulk modes $\hat{A}_r$ that satisfy the condition $\int_0^L dr\hat{A}_r = 0$.

The first term is the action for a Maxwell theory with vanishing boundary conditions for gauge fields. Let us denote the partition function of this part as $Z_{\hat{A}}$. The field ϕ and W are 3-dimensional scalar fields living on a surface with coordinate x^a. Now, we can separate the path integral as bulk and boundary parts

$$Z = Z_{\hat{A}} \times \int \mathcal{D}\phi\mathcal{D}W\ \exp\left[-\frac{1}{2e^2L}\int d\tau d^2x(\partial^a\phi\partial_a\phi + \partial^a W\partial_a W)\right]\ , \quad (3.60)$$

and then evaluate the bulk and boundary partition functions separately.

Note that we have not used any gauge fixing condition here, and kept all the components of A_μ untouched. Whereas, we will eventually gauge away the gauge redundancy later. We already analysed the physics that we are interested in, so we can use gauge fixing conditions on the bulk fields $\hat{A}_\mu$ such that the zero modes of A_r and the Wilson lines are not gauged away. One still needs to be careful that we did not kill any other interesting physical degrees

of freedom. Let us suppose we are dealing with a compact U(1) gauge theory. In the Euclidean background, the map between the background time circle $\tau \sim \tau + \beta$ and compact gauge parameter allow us to include some topological modes for component A_τ. The fundamental group of S_1 is $\mathbf{Z}$. So the modes can be expressed as

$$A_\tau \ni \frac{2\pi n}{\beta}\,, \qquad n \in \mathbf{Z}\,. \tag{3.61}$$

Those modes corresponding to the large gauge transformation are physical, and should be add back to the system. We will come back to those modes when we discuss summing over different boundary configurations B_μ. For now, let us only focus on $\hat{A}_\mu$ vanishing on the boundaries and do not sum over any configurations that are non-vanishing.

Bulk fluctuation modes

First of all, let us evaluate the partition function for bulk fluctuation modes $Z_{\hat{A}}$. We will use the Faddeev-Popov method [216] to evaluate the partition function, by inserting the following identity

$$1 = \int \mathcal{D}\lambda \det\left(\frac{\partial G}{\partial \lambda}\right) \delta(G-0)\,, \tag{3.62}$$

with gauge fixing condition $G = \partial_\mu \hat{A}^\mu - c(x)$. Following the standard gauge fixing procedure, in Feynman gauge, we eventually get

$$\begin{aligned} Z_{\hat{A}} &= \int \mathcal{D}\hat{A}_\mu \mathcal{D}C\mathcal{D}\bar{C}\; e^{-\frac{1}{2e^2}\int_{\mathcal{M}} d\tau d^3x\; [\hat{A}^\mu(\partial^2)\hat{A}_\mu + \bar{C}(\partial^2)C]} \\ &= \det(\partial^2)^{-1}\,, \end{aligned} \tag{3.63}$$

where C and $\bar{C}$ are ghost fields. After gauge fixing, the finally result is the partition function for two bosonic polarisations $Z = \det(\partial^2)^{-1}$.

Now if we define the energy and momenta of the gauge fields as (ω, p_r, p_2, p_3), the logarithm of $Z_{\hat{A}}$ can be calculated by working out the determinantal operation

$$\ln Z_{\hat{A}} = -\sum_{\omega} \sum_{p_r, p_2, p_3} \ln\left[\beta^2(\omega^2 + p_r^2 + p_2^2 + p_3^2)\right]\,. \tag{3.64}$$

One can further evaluated the partition function by taking different limits of the length scales in the theory. We are going to evaluate the logarithm of the partition function in the next subsection when we discuss different temperature limits.

Note that we can also use different gauge fixing conditions, like axial gauge. The gauge fixing condition does not make much difference for the fluctuation modes as far as we keep two physical polarisations in the finally result.

Zero modes and Wilson lines

Let us evaluate the partition function of fields ϕ and W in this subsection. The action for ϕ and W can be written as

$$S_{\phi,W} = \frac{1}{2e^2 L}\int d\tau d^2x\ (\partial_a\phi\partial^a\phi + \partial_a W\partial^a W)\,, \tag{3.65}$$

which is the action for two massless scalar fields living on the boundary. Denoting the area of the boundary as "Area", the fluctuation modes of field ϕ and W can be expanded as

$$\phi(x^a) = \sqrt{\frac{2e^2\beta L}{\text{Area}}}\sum_{\omega,p_2,p_3}\tilde{\phi}(\omega,p_2,p_3)e^{i(\omega\tau+p_2x^2+p_3x^3)}\,, \tag{3.66}$$

$$W(x^a) = \sqrt{\frac{2e^2\beta L}{\text{Area}}}\sum_{\omega,p_2,p_3}\tilde{W}(\omega,p_2,p_3)e^{i(\omega\tau+p_2x^2+p_3x^3)}\,. \tag{3.67}$$

The coefficient is chosen such that $\tilde{\phi}$s and $\tilde{W}$s are dimensionless and thus the integrals over $d\tilde{\phi}$ in the path integral give out dimensionless quantities. With this mode expansion, the corresponding partition function can be express as

$$Z_F = \prod_{\omega,p_2,p_3}[\beta^2(\omega^2+p_2^2+p_3^2)]^{-1}\,, \tag{3.68}$$

logarithm of which is read as

$$\ln Z_F = -\sum_{\omega,p_2,p_3}\ln[\beta^2(\omega^2+p_2^2+p_3^2)]\,. \tag{3.69}$$

Z_F is the partition function for two 3-dimensional massless scalar fields. We can then calculate the free energy and entropy of those modes in different temperature limits, and compare it with the bulk fluctuation modes, which will be the task for the next subsection.

There are some other interesting modes of W because of the topology. The Wilson lines stretched between the two boundaries can be denoted as

$$\mathcal{W}_\gamma = \mathcal{P}\exp[i\int_0^L dr\hat{A}_r]\,. \tag{3.70}$$

And because it's always inside of an exponential function, $\int_0^L dr\hat{A}_r$ is compact with periodicity 2π. The requirement that the Wilson lines are single-valued allows us to include the elements of fundamental group S^1. In the Euclidean background, the background time circle $\tau \sim \tau + \beta$ allows the field W to wind around the S^1 circle and have some winding modes $2\pi n\tau/\beta$. Now the field W has compact constant modes and novel winding modes which are interesting to deal with. The constant modes contribution of W can always be written as

$$Z_0 = \int_0^{2\pi\sqrt{\frac{\text{Area}}{e^2L\beta}}} d\tilde{W}_0 = 2\pi\sqrt{\frac{1}{e^2L}} \times \sqrt{\frac{\text{Area}}{\beta}}\,. \tag{3.71}$$

The winding mode contribution can be written as

$$Z_w = \sum_n e^{-\frac{\text{Area}}{2e^2L\beta}(2\pi n)^2}\,. \tag{3.72}$$

which equals 1 when the coefficient $\frac{\text{Area}}{2e^2\beta L}$ is very large, because only $n = 0$ mode dominants. When the coefficient is very small, we can change the sum into a Gaussian integral. The overall partition function of ϕ and W is the product of Z_0, Z_w, and fluctuation modes Z_F.

As a summary of the partition functions for fields W and ϕ calculated in this subsection, we have two collections of fluctuation modes from W and ϕ separately, the partition function of which is denoted as Z_F. Also there are interesting constant modes and winding modes. The logarithm of the partition function can be expressed as

$$\ln Z = -2 \times \frac{1}{2} \sum_{\omega,p_2,p_3} \ln[\beta^2(\omega^2 + p_2^2 + p_3^2)] + \frac{1}{2}\ln[\frac{\text{Area}}{\beta L}] - \ln e + \ln Z_w\,, \tag{3.73}$$

with

$$\ln Z_w = \begin{cases} 0\ ; & \frac{\text{Area}}{2e^2L\beta} \gg 1 \\ -\frac{1}{2}\ln[\frac{\text{Area}}{\beta L}] + \ln e\ . & \frac{\text{Area}}{2e^2L\beta} \ll 1 \end{cases} \tag{3.74}$$

Before we go any further towards comparing bulk and boundary degrees of freedom, let us qualitatively illustrate the basic behaviour of the entropy corresponding to (3.73). At high temperature, the entropy of the fluctuation modes is proportional to $T^2 \times$ Area. While as temperature goes lower and lower, the fluctuation modes play a less and less important role, and the zero modes and winding modes that contribute as the logarithm of temperature and e^2 start to dominate. However, as the temperature becomes very low, the contribution from $\ln Z_w$ cancels the one from zero modes, and the overall entropy goes to a constant. The behaviour of those modes is illustrated in figure 3.3.

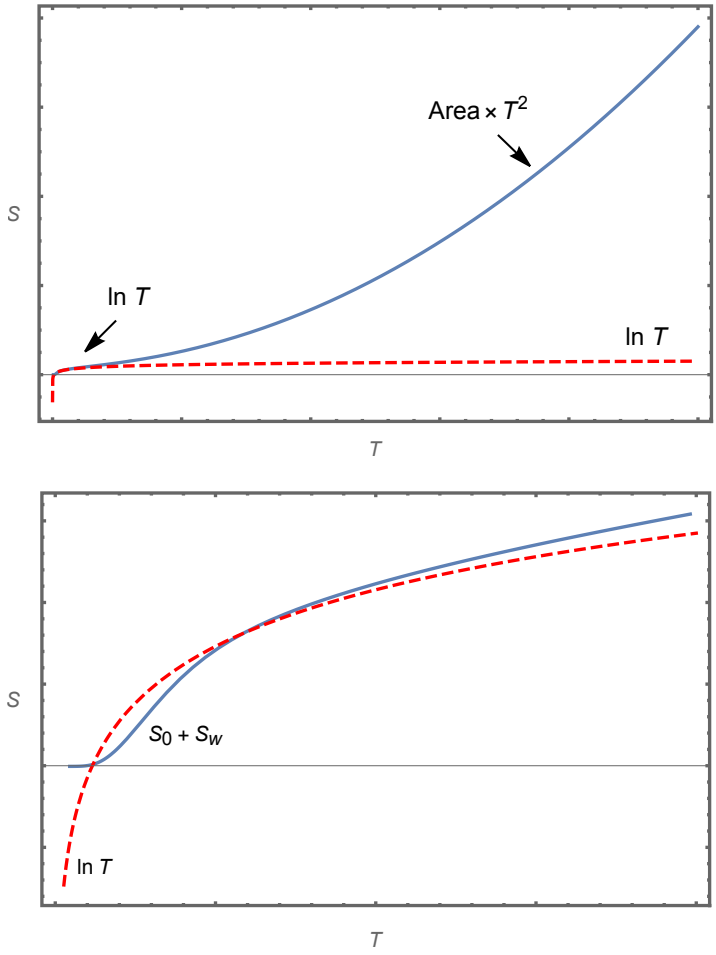


Figure 3.3: A sketch of the entropy of fields ϕ and W as we vary the temperature. The second picture is an enlarged version of the low-temperature region. The red dashed line is an auxiliary line showing $\ln T$. At high temperature, the entropy scales as Area $\times\, T^2$. As temperature goes lower than before, the entropy scales as the logarithm of temperature and coupling constant. Finally, the entropy goes to zero in the extremely low-temperature limit because the contributions coming from zero modes and winding modes cancel with each other.

3.2.3 Different temperature limits

In this section, we will evaluate and compare the partition functions from the bulk fluctuation modes $\hat{A}_\mu$, the zero modes ϕ, and the Wilson lines W in different temperature limits. There are three different dimensional length scales in the theory at hand, the inverse temperature β, the distance between the two boundaries L, and the length scale of the boundary $\sqrt{\text{Area}}$. The coupling constant e^2 is dimensionless. Comparing the inverse temperature β with other length scales in the theory, we can get different behaviour of the partition function.

We are mainly interested in the following three different temperature limits. The so-called high-temperature limit is the limit where we have $\beta \ll L \ll \sqrt{\text{Area}}$. β is the smallest length scale in the system. In this temperature limit, the bulk fluctuation modes $\hat{A}_\mu$ is the most important contribution. The second temperature limit we are interested in is the low-temperature limit, where we have $L \ll \beta \ll \sqrt{\text{Area}}$. The distance between the two boundaries is way smaller than the inverse temperature β, and all the high-frequency modes along the r direction will be gapped. The zero modes ϕ and the Wilson lines W start to play the most important role in this limit. The last case is the super-low temperature limit, where we have $L \ll \sqrt{\text{Area}} \ll \beta$. The temperature is very low, and all the fluctuation partition functions that proportional to the temperature are disappeared. The logarithm contributions of fields W shown in the previous section becomes the most important ones.

Let us discuss those three temperature limits separately. The qualitative behaviour of the entropy is illustrated in figure 3.4. The overall entropy is a summation of different contributions. Figure 3.4 only illustrates the separate contributions, and the solid red line shows the dominant contribution. More details of the calculations can be found in the Appendix A.1.

High temperature limit In the high-temperature limit, nothing is special, and we expect to see the usual result of black body radiation in a box because the bulk fluctuation modes whose entropy is proportional to T^3 are the most important contribution. The partition function $Z_{\hat{A}}$ and entropy $\mathcal{S}_{\hat{A}}$ of bulk modes $\hat{A}_\mu$ are shown below

$$\ln Z_{\hat{A}} = -\frac{1}{8\pi^2}\beta V \times \Lambda^4 + \frac{\pi^2}{45}\frac{V}{\beta^3}\,, \tag{3.75}$$

$$\mathcal{S}_{\hat{A}} = (1-\beta\partial_\beta)\ln Z_{\hat{A}} = \frac{4\pi^2}{45}VT^3\,, \tag{3.76}$$

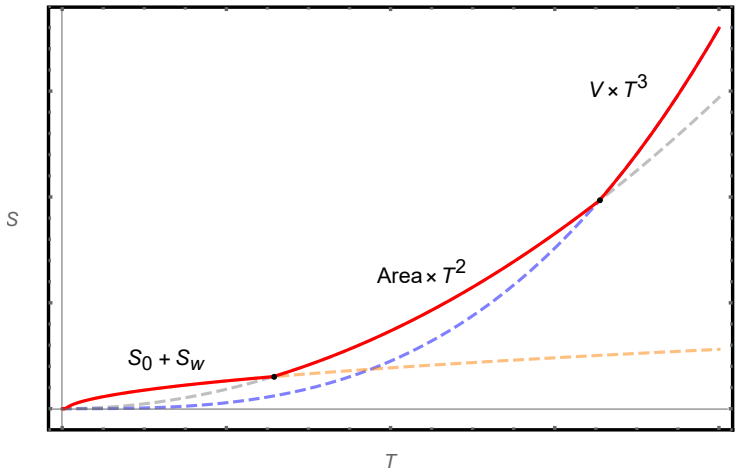


Figure 3.4: A sketch of the entropy of the whole system in different temperature limits. The actual entropy is the sum of different contributions, and the red line demonstrates the dominant contribution. There are two transitions of the dominants shown in the figure. The bulk fluctuation modes always dominate in the high-temperature limit, and the entropy scales as the volume multiplied by temperature cubed. As the temperature becomes lower than before, the area contribution starts to dominate. At very low temperature, the fluctuation contribution is not important anymore, and the only contribution is from the constant modes and winding modes of field W. A more clear curve of the entropy near the origin is shown in the second panel of figure 3.3.

which is exactly the black body radiation result in a flat box. When the temperature is high, the contributions from ϕ and W proportional to the area of the boundary are small compared with the bulk radiation. So in the high-temperature limit, the dominant contribution always comes from the bulk fluctuation modes $\hat{A}_\mu$, which scales as the volume between the two boundaries multiplied by the temperature cubed.

Low temperature limit As the temperature goes lower and lower, when we have $L \ll \beta \ll \sqrt{\text{Area}}$, the situation starts to change. For finite β, $1/L$ is very big, and the energy needed to excite high-frequency modes along the r is very high. In this temperature limit, the modes along the r direction are gapped, and we are only left with zero modes along this direction. All the zero modes of $\hat{A}_a$ is killed by the boundary conditions, and we get no contribution from $\hat{A}_a$ components. Fortunately, there is a survivor the zero modes of A_r, i.e. ϕ. W plays the similar role as ϕ. The partition function and entropy for ϕ and W can be written as

$$\ln Z = \frac{1}{2}\ln\frac{2\pi^2\text{Area}}{e^2\beta L} - \frac{1}{6\pi}\beta\text{Area}\cdot\Lambda^3 + \frac{\zeta(3)}{\pi}\frac{\text{Area}}{\beta^2} + \ln Z_w\,, \tag{3.77}$$

$$\mathcal{S} = (1-\beta\partial_\beta)\ln Z = \frac{3\zeta(3)}{\pi}\text{Area}\times T^2 + \text{logarithm corrections}\,. \tag{3.78}$$

The entropy of the thermal fluctuation modes along the boundary direction is proportional to the area of the plates times temperature squared. There are extra contributions from constant modes and winding modes if e^2 is not very large, which is proportional to the logarithm of temperature and the coupling constant. The logarithm contribution is mainly controlled by the coupling constant e^2, which can surpass the area contribution for suitable e^2. We will discuss the logarithm contribution later because those terms will be the star as the temperature is even lower. So in this temperature limit, the entropy of the system mainly comes from the fluctuation modes of ϕ and W, which is more or less proportional to the area of the plates times the temperature squared.

The situation here is similar to the Kuluza-Klein (K-K) reduction along the radius direction. The energy of the K-K tower is proportional to $1/R$, where R is the length scales of extra dimensions. When R are very small, we can only see the zero modes of the K-K tower, and the effective theory is lower-dimensional.

Super-low temperature limit As the temperature becomes even lower, all the thermal fluctuation contributions proportional to the temperature will not survive. In the so-called super-low temperature, we have $\frac{L\cdot\text{Area}}{\beta^3} \gg 1$. The bulk fluctuation modes are already frozen to death in the previous stage, and now it is the turn for ϕ and W. The logarithm contributions from the constant modes and winding modes can be written as

$$\ln Z = \frac{1}{2}\ln\frac{2\pi^2\text{Area}}{e^2\beta L} + \ln Z_w\,. \tag{3.79}$$

with $\ln Z_w$ shown in the equation (3.74). The coupling constant $1/e^2$ is the control of constant modes. For weak coupling limit, where we have $1/e^2 \ll 1$ such that

$$\frac{1}{e^2}\frac{\text{Area}}{\beta L} \ll 1\,, \tag{3.80}$$

the constant modes contribution of ϕ is cancelled by the winding modes. However, for the strong coupling limit where

$$\frac{1}{e^2}\frac{\text{Area}}{\beta L} \gg 1\,, \tag{3.81}$$

we can always see the contribution of the constant modes, which always scales as the logarithm of the coupling constant and temperature. The basic behaviour of the entropy of constant modes and winding modes is shown in the second picture of figure 3.3.

The entropy of gauge theory in the flat case is summarised in the figure 3.4. We would like to see if a similar phenomenon also shows up in the black hole background and what are the similarities and differences. This question will be answered in the next section.

3.3 Schwarzschild Black Hole at Finite Temperature

Black holes are systems associated with temperatures and entropies. The Euclidean method for finite temperature system also works for black hole. The black hole system is special compare to the flat case because of the Bekenstein-Hawking entropy is always proportional to the area of horizon in the unit of G_N. As we have already seen the phase transitions and contributions from zero modes ϕ and the Wilson lines W in section 3.2.3, one might speculate that the similar phenomenon would also happens in the black hole system,

which might explains something deep for black hole mechanism. We analyse the finite temperature black hole in this section, and leave the extremal case for the next section.

We follow a similar procedure as the flat case, so most of the result is very similar to the previous section. One noticeable but not vital difference is that we are using spherical coordinates rather than Cartesian here. The important difference is we find two extra contributions both scales as the area of the horizon. The first contribution comes from the bulk fluctuation modes, which agrees with the entropy found in the brick wall model [2]. This is shown in section 3.3.2. The second contribution is because the zero modes ϕ and the boundary stretched Wilson lines W are localised in the phase space $\partial_\tau \phi = \partial_\tau W = 0$, whose entropy is coming from the zero point energy and proportional to the area of the horizon in the unit of Planck area. The effective action and corresponding analysis can be found in subsections 3.3.3 and 3.3.4.

The background we are interested in is a static Euclidean black hole with metric

$$ds^2 = (1 - \frac{r_s}{r})d\tau^2 + (1 - \frac{r_s}{r})^{-1}dr^2 + r^2 d\Omega^2 \,, \tag{3.82}$$

where the Schwarzschild radius is $r_s = 2G_N M$. To avoid conical singularity at the horizon, the Euclidean time direction must be compact with a periodicity β, which can be regarded as the inverse temperature of the system. With the periodicity of τ, the geometry of the Euclidean Schwarzschild black hole can be illustrated as the so-call "cigar" geometry, as shown in figure 3.5. We put the black hole in a box with length scale $r_s + L$, as shown in figure 3.6. Thus, we have two boundaries on this background. The first one is a very small distance ε away from the horizon, known as the "stretched horizon" [168]. The other boundary is the surface of the box similar to the last section, while here L is the distance away from the horizon.

The stretched horizon might look strange at first sight. Here the stretched horizon can be regarded as an auxiliary timelike surface outside the event horizon, where we can put boundary conditions on. We will eventually take the real horizon limit. There are different opinions towards the stretched horizon; for example, it is believed that only the physics outside of the stretched horizon is well-described by local QFT. In this case, the stretched horizon is a physical surface [217], and ε can be regarded as the UV regulator of the QFT outside of the horizon and can be taken to but not exceed the Planck scale l_p. Whether the stretched horizon is a physical surface or an auxiliary surface does not make much difference for our calculations here, and ε will be made very small

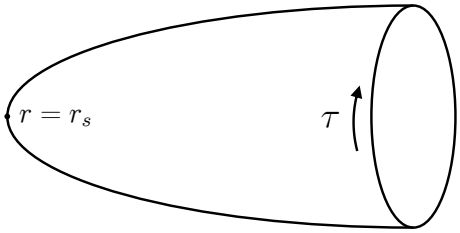


Figure 3.5: Geometry of a Euclidean black hole, where every point on the cigar is an S^2. We put two boundaries on this background, at the stretched horizon ($r = r_s + \varepsilon$) and distance L away from the horizon ($r = r_s + L$) separately.

either way. However, one does need to pay attention to the order of taking different limits of length scales. As we will see, in the low-temperature limit $\beta \gg L$, taking the $\varepsilon \to 0$ limit would introduce localisations on the spaces of some specific modes, which is essential to get an area entropy contribution. One may not see this phenomenon if the order of taking the above limits are messed up.

As in the flat case, we put a U(1) gauge theory between the boundaries, with Euclidean action

$$S_E = \frac{1}{4e^2} \int_{\mathcal{M}} d^4x \sqrt{g}\ F^{\mu\nu} F_{\mu\nu}\,. \tag{3.83}$$

The physics is very similar to the previous section, and we are trying to mimic the calculation in the flat case. We will discuss two different limits of the black hole system, as shown in figure 3.6. In the high-temperature limit $L \gg r_s$, shown in the first panel, the situation should be similar to the flat case because the existence of the small black merely gives a periodicity to the Euclidean time direction to avoid conical singularity at the horizon. On the other hand, in the low-temperature limit $L \ll r_s$, we expect the results to be qualitatively different from the high-temperature limit and share some similarity with the extremal black hole case, which will be discussed in the next section. This means that there will be phase transitions, which should be explained by some symmetry breaking pattern.

So firstly, let us redefine $\rho = r - r_s$ and write the metric as

$$ds^2 = \frac{\rho}{\rho + r_s} d\tau^2 + \frac{\rho + r_s}{\rho} d\rho^2 + (\rho + r_s)^2 d\Omega^2\,. \tag{3.84}$$

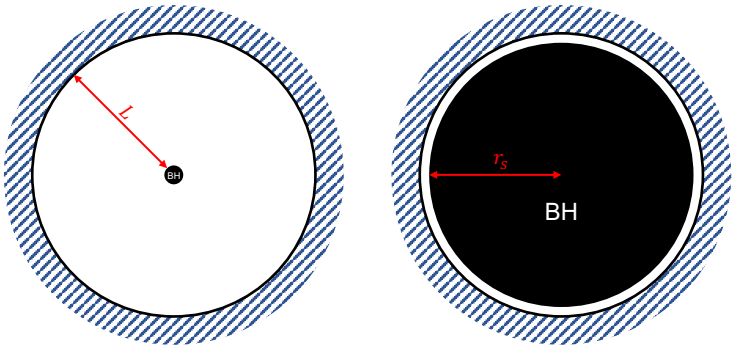


Figure 3.6: Different temperature limits of black holes in a box. **First panel**: high-temperature limit $L \gg r_s$, i.e. the small black hole case. The temperature of the system is $T \propto 1/r_s$. The role that the black hole plays is to give a very high temperature T to the system. We expect the final results to be similar to the flat case. **Second panel**: low-temperature limit of the system $L \ll r_s$, i.e. large black hole case. We expect the results to share some similarity with the extremal black hole case.

Now the two boundaries are at $\rho = \varepsilon$ and $\rho = L$ separately. We will dimensional reduce the gauge theory action to a lower-dimensional spacetime, so it is necessary to define the proper distance y along the radius direction

$$dy = \sqrt{\frac{\rho + r_s}{\rho}} d\rho \,. \tag{3.85}$$

The proper distance y can be integrated out and read as

$$y = \sqrt{\rho\ (\rho + r_s)} + r_s\ \text{arcsinh}\sqrt{\frac{\rho}{r_s}} \,. \tag{3.86}$$

Inverting the above equation, one can represent ρ as a function of y and rewrite the black hole metric as

$$ds^2 = \frac{\rho(y)}{\rho(y) + r_s} d\tau^2 + dy^2 + [\rho(y) + r_s]^2 d\Omega^2 \,, \tag{3.87}$$

which will be the metric we mainly work with. Note that y takes value from y_1 to y_2, with

$$y_1 = \sqrt{\varepsilon(\varepsilon + r_s)} + r_s\ \text{arcsinh}\sqrt{\frac{\varepsilon}{r_s}} \,, \qquad y_2 = \sqrt{L(L + r_s)} + r_s\ \text{arcsinh}\sqrt{\frac{L}{r_s}} \,. \tag{3.88}$$

We are going to focus on the same boundary conditions for gauge field A_μ as the flat case.

On the black hole background, the boundary condition we are interested in are

$$A_a\Big|_{r=r_\alpha} = f_a^\alpha(x^a)\,, \tag{3.89}$$

$$A_y\Big|_{r=r_\alpha} = \text{Arbitrary}\,. \tag{3.90}$$

α labels different boundaries, and $f_a^\alpha(x^a)$ are fixed boundary configurations. Again, there are no constraint on the boundary configurations of A_y, and we need to integrate over the boundary configurations of A_y in the path integral.

Before we continue, let us take a step back and think about whether it is physically reasonable to put boundary conditions on the stretched horizon and what those boundary conditions mean. Mathematically, one can put boundaries at any location. The original physics can be recovered by glueing the two sides along the edge, which means one needs to integrate over the boundary conditions along the boundary. So it is not strange to put a boundary near the horizon. However, is it reasonable to put a physical boundary near the horizon is unknown because any object near the horizon would fall into the black hole. Although we are trying to mimic the flat case calculation, one may need to keep in mind that the black hole situation might not be the same as the case of the parallel plates.

We will not go through all the canonical analysis in this case. Naively one would expect we still have bulk modes $\hat{A}_\mu(x^\mu)$, zero momentum modes along y-direction $\phi(x^a)$, and Wilson lines stretched between the two boundaries $W(x^a)$ in this case. We will justify that they are physical by working out their effective actions. Following the same logic as the flat case, the gauge fields A_μ can be split into several pieces. Denoting B_μ as the bulk on-shell configurations that capture the fixed boundary configuration f_a^α, we can write A_y as

$$A_y = \hat{A}_y(x^\mu) + \frac{\phi(x^a)}{|y|} + B_y\,, \tag{3.91}$$

with $|y| = y_2 - y_1$. Similar argument goes as the flat case, $\phi(x^a)$ captures the boundary modes on the left boundary, and there are extra field captures the difference between the two boundaries. And we can further decompose A_y into two parts. The difference between the two boundaries is captured by the integral $\int_{y_1}^{y_2} dy\ \hat{A}_y$, and what left is trivial on both boundary. We also separate

A_a into

$$A_a = \hat{A}_a + B_a \,. \tag{3.92}$$

$\hat{A}_a$ are the bulk modes that vanish on the boundary. For a given set of configurations f_a^α, we can find bulk on-shell configurations B_μ corresponding to them. Again we will chose flat boundary configurations $f_a^\alpha(x^a) = C_a^\alpha + \partial_a \lambda^\alpha$ to simplify the calculations. This problem is solved in the appendix A.2, and the solutions can be written as

$$\begin{aligned} B_a &= -\frac{r_s(r_s+L)}{(\rho+r_s)L}[C_a^{(r)} - C_a^{(l)} + \partial_a(\lambda^{(r)} - \lambda^{(l)})] \\ &\quad + \frac{r_s+L}{L}(C_a^{(r)} + \partial_a\lambda^{(r)}) - \frac{r_s}{L}(C_a^{(l)} + \partial_a\lambda^{(l)})\,, \end{aligned} \tag{3.93}$$

$$B_y = \frac{\sqrt{\rho}}{(\rho+r_s)^{5/2}} \frac{r_s(r_s+L)}{L} \times [x^a \cdot (C_a^{(r)} - C_a^{(l)}) + (\lambda^{(r)} - \lambda^{(l)})]\,. \tag{3.94}$$

One can check that B_μ indeed satisfy the bulk equations of motion and the boundary conditions (3.89). For example one can simply check that $B_a\big|_{\rho=\varepsilon} = C_a^{(l)} - \partial_a\lambda^{(l)}$ and $B_a\big|_{\rho=L} = C_a^{(r)} - \partial_a\lambda^{(r)}$.

Note that those B_μ configurations are the bulk modes corresponding to a specific boundary condition. If one wants to sum over different boundary conditions, which might be equivalent to sum over different superselection sectors, then functional integral over those B_μ modes are needed. We will discuss those issues at the end of this chapter. For now, we would not include the B_μ modes in the path integral.

3.3.1 Effective action and the Kaluza-Klein reduction

Now the partition function can be written as

$$Z = \int \mathcal{D}\hat{A}_\mu \mathcal{D}\phi \mathcal{D}W \; e^{-S_E}\,. \tag{3.95}$$

Again the boundary stretched Wilson lines $W(x^a)$ is defined as

$$W = i\int_{y_1}^{y_2} dy \; \hat{A}_y\,. \tag{3.96}$$

In this case, the effective action is worked out in the Appendix A.3, which reads as

$$\begin{aligned} S_E &= \frac{1}{4e^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\ \hat{F}^{\mu\nu}\hat{F}_{\mu\nu} + \frac{1}{2e^2|y|^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\left[g^{ab}\partial_a\phi\partial_b\phi\right] \\ &\quad -\frac{1}{e^2|y|}\int d\tau d^2x\left[\sqrt{g}\ g^{ab}\right]_{y=y_1}^{y=y_2} i\partial_a W\partial_b\phi\,. \end{aligned} \tag{3.97}$$

The first part can be directly put into the path integral, so we can write the partition function for the bulk fluctuation modes as

$$Z_{\hat{A}} = \int \mathcal{D}\hat{A}_\mu \exp\left[-\frac{1}{4e^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\ \hat{F}^{\mu\nu}\hat{F}_{\mu\nu}\right]. \tag{3.98}$$

We will discuss bulk gauge fixing condition and evaluate the above partition function in the next subsection. The rest of the action is

$$\begin{aligned} S[\phi,W] &= \frac{1}{2e^2|y|^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\left[g^{ab}\partial_a\phi\partial_b\phi\right] \\ &\quad -\frac{1}{e^2|y|}\int d\tau d^2x\left[\sqrt{g}\ g^{ab}\right]_{y=y_1}^{y=y_2} i\partial_a W\partial_b\phi\,. \end{aligned} \tag{3.99}$$

In order to perform path integral, we are going to rewrite $S[\phi,W]$ into a three dimensional action

$$\begin{aligned} S^{(3)}[\phi,W] &= \frac{1}{2e'^2}\int d\tau d^2x\sqrt{h}\left[h^{ab}\partial_a\phi\partial_b\phi\right] - \frac{i}{2e'^2}\int d\tau d^2x\sqrt{h} \\ &\quad \times\left[\gamma_1\ h^{\tau\tau}\partial_\tau W\partial_\tau\phi + \gamma_2\ h^{\theta\theta}\partial_\theta W\partial_\theta\phi + \gamma_3\ h^{\varphi\varphi}\partial_\varphi W\partial_\varphi\phi\right]. \end{aligned} \tag{3.100}$$

This is more or less like the K-K reduction of a higher-dimensional action to a lower-dimensional one.

After dimensional reduced the action for ϕ and W, the Gaussian path integral for field ϕ can be easily worked out, and we get an effective action for field W in the mean time. The path integral over ϕ gives out $\det(\partial^2)^{-1/2}$. Then, the corresponding action for W can be expressed as

$$S_W = -\frac{1}{2e'^2}\int d\tau d^2x\sqrt{h}\,(\gamma_1^2\ h^{\tau\tau}\partial_\tau W\partial_\tau W + \gamma_2^2\ h^{\theta\theta}\partial_\theta W\partial_\theta W + \gamma_3^2\ h^{\varphi\varphi}\partial_\varphi W\partial_\varphi W)\,. \tag{3.101}$$

Rewriting $\det(\partial^2)^{-1/2}$ as a path integral over field ϕ, the overall partition function can be expressed as

$$Z_{\phi,W} = \int \mathcal{D}\phi\mathcal{D}W\ e^{-S_{\phi,W}}\,. \tag{3.102}$$

with effective action

$$\begin{aligned} S_{\phi,W} &= \frac{1}{2e'^2}\int d\tau d^2x\sqrt{h} \\ &\times \left(h^{ab}\partial_a\phi\partial_b\phi + \gamma_1^2\ h^{\tau\tau}(\partial_\tau W)^2 + \gamma_2^2\ h^{\theta\theta}(\partial_\theta W)^2 + \gamma_3^2\ h^{\varphi\varphi}(\partial_\varphi W)^2\right). \end{aligned} \tag{3.103}$$

This is the partition function and the effective action we will mainly focus on. So the task left is to follow the logic described above and work out the low dimensional effective action for ϕ and W. Then, we can evaluated the partition functions of fields $\hat{A}_\mu$, ϕ and W by the path integral. The first task is to dimensional reduce action (3.99) to get the low dimensional metric h_{ab} and the coupling constant e'^2, and then work out the other parameters in action (3.103).

Dimensional reduction

Now we are going to work out all the parameters in action (3.103) reduced from the higher-dimensional one (3.99). More details of the calculations can be found in appendix A.4. Before actually doing that, let us firstly assume that the 3-dimensional theory uses the same time coordinate τ as the original one, and the 3-dimensional metric takes the following form

$$h_{ab} = \mathrm{diag}(h_{\tau\tau},\ R^2,\ R^2\sin^2\theta)\,. \tag{3.104}$$

R^2 is the undermined parameter in the metric that we are going to fix by dimensional reduction. With this assumption, we have $\sqrt{h} = \sqrt{h_{\tau\tau}}\ R^2\sin\theta$. To match the first parts in actions (3.100) and (3.99), we need to work out the following problem

$$\frac{1}{2e^2|y|^2}\int d\tau d^2x\ (\int_{y_1}^{y_2} dy\sqrt{g}g^{ab})\ \partial_a\phi\partial_b\phi \quad = \quad \frac{1}{2e'^2}\int d\tau d^2x\sqrt{h}\left[h^{ab}\partial_a\phi\partial_b\phi\right] \tag{3.105}$$

The solution can be easily obtained and can be written as

$$h_{\tau\tau} = \frac{LR^2}{F}\,, \tag{3.106}$$

$$\frac{1}{e'^2} = \frac{L}{e^2|y|^2}\sqrt{h^{\tau\tau}} = \frac{1}{e^2|y|^2}\frac{\sqrt{L\ F}}{R}\,, \tag{3.107}$$

with

$$F \approx\ 3r_s^2L + \frac{3}{2}r_sL^2 + \frac{L^3}{3} + r_s^3\ln L/\varepsilon\,. \tag{3.108}$$

One can find more details of the calculations in appendix A.4. Now we have obtained the 3 dimensional metric h_{ab} and the effective coupling constant e'^2. The second step is to calculate the parameters γ_1, γ_2, and γ_3 in the action (3.100). We need to match the rest of action in (3.99) and (3.100), i.e.

$$\begin{aligned}
\frac{1}{e^2|y|}\int d\tau d^2x\,[\sqrt{g}\;g^{\tau\tau}]_{y=y_1}^{y=y_2}\,\partial_\tau W\partial_\tau\phi &= \frac{\gamma_1}{2e'^2}\int d\tau d^2x\sqrt{h}\;[h^{\tau\tau}\partial_\tau W\partial_\tau\phi]\,,\\
\frac{1}{e^2|y|}\int d\tau d^2x\left[\sqrt{g}\;g^{\theta\theta}\right]_{y=y_1}^{y=y_2}\partial_\theta W\partial_\theta\phi &= \frac{\gamma_2}{2e'^2}\int d\tau d^2x\sqrt{h}\;\left[h^{\theta\theta}\partial_\theta W\partial_\theta\phi\right]\,,\\
\frac{1}{e^2|y|}\int d\tau d^2x\,[\sqrt{g}\;g^{\varphi\varphi}]_{y=y_1}^{y=y_2}\,\partial_\varphi W\partial_\varphi\phi &= \frac{\gamma_3}{2e'^2}\int d\tau d^2x\sqrt{h}\;[h^{\varphi\varphi}\partial_\varphi W\partial_\varphi\phi]\,.
\end{aligned} \tag{3.109}$$

The solution of the equations (3.109) can be obtained as

$$\gamma_1 = -2|y|\sqrt{\frac{r_s}{\varepsilon}}\frac{r_s^2}{F}\,,\qquad \gamma_2=\gamma_3=\frac{2|y|}{L}\sqrt{\frac{L}{L+r_s}}\,. \tag{3.110}$$

With all the low-dimensional parameters at hand, the effective action (3.103) can be expressed as

$$S_{\phi,W} = S_\phi + S_W \tag{3.111}$$

with

$$\begin{aligned}
S_\phi &= \frac{L}{2e^2|y|^2}\int d\tau d^2x\;R^2\sin\theta\\
&\quad\Big(\frac{F}{LR^2}\partial_\tau\phi\partial_\tau\phi+\frac{1}{R^2}\partial_\theta\phi\partial_\theta\phi+\frac{1}{R^2\sin^2\theta}\partial_\varphi\phi\partial_\varphi\phi\Big)\\
S_W &= \frac{2}{e^2(L+r_s)}\int d\tau d^2x\;r_s^2\sin\theta\\
&\quad\times\Big(\frac{r_s}{\varepsilon}\frac{Lr_s^2+r_s^3}{F}\partial_\tau W\partial_\tau W+\frac{1}{r_s^2}\;\partial_\theta W\partial_\theta W+\frac{1}{r_s^2\sin^2\theta}\partial_\varphi W\partial_\varphi W\Big)\,.
\end{aligned}$$

Again, the two symbols F and $|y|$ in the above actions are

$$F = 3r_s^2L+\frac{3}{2}r_sL^2+\frac{L^3}{3}+r_s^3\ln L/\varepsilon\,, \tag{3.112}$$

$$|y| = \sqrt{L(L+r_s)}+r_s\;\mathrm{arcsinh}\sqrt{\frac{L}{r_s}}\,. \tag{3.113}$$

Note that it is important to keep the original Euclidean time coordinate τ as the time coordinate for the 3-dimensional theory because the periodicity of

coordinate time τ is the physical inverse temperature for the observer who uses the Schwarzschild metric. If one uses different time coordinates, the physics would be hard to discuss. For example, the low-temperature limit for the coordinate observer can be high-temperature for an observer using a different coordinate system. So we would always keep τ as our time coordinate.

3.3.2 Bulk fluctuation modes

In the next several subsections we are going to evaluate the partition function of the theory in different temperature limits. The partition function for W and ϕ will be treated separately, and we will see different behaviour in different temperature limits. The overall partition function can be written as three parts: bulk contribution $Z_{\hat{A}}$ multiplied by Z_ϕ and Z_W

$$\begin{aligned} Z &:= Z_{\hat{A}} \times Z_\phi \times Z_W \\ &= \int \mathcal{D}\hat{A}_\mu \; e^{-\frac{1}{4e^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\; \hat{F}^{\mu\nu}\hat{F}_{\mu\nu}} \times \int \mathcal{D}\phi \; e^{-S_\phi} \times \int \mathcal{D}W \; e^{-S_W} \, . \end{aligned} \tag{3.114}$$

Firstly, we are going to calculate the partition function for the bulk fluctuation modes $Z_{\hat{A}}$ in the current section. After that we will evaluate Z_ϕ and Z_W in the high and low temperature limits in section 3.3.3 and section 3.3.4.

Now, let us first evaluate the entropy of bulk fluctuation modes. The details of the calculation are exiled to appendix A.5, to avoid being distracted from the main text. This is the standard black body calculation on curved spacetime. For electromagnetism, we have two physical polarisations in the bulk, which are both bosonic and massless. The bulk partition function reads as

$$\ln Z_{\hat{A}} = -\frac{4\pi^3}{45}\frac{1}{\beta^3}\frac{r_s^4}{\varepsilon} - \frac{16\pi^3}{45}\frac{r_s^3}{\beta^3}\ln\frac{L}{\varepsilon} - \frac{4\pi^3}{45}\frac{1}{\beta^3}\left(-\frac{r_s^4}{L} + 6\, r_s^2 L + 2\, r_s L^2 + \frac{L^3}{3}\right) . \tag{3.115}$$

and the corresponding entropy can be written as

$$\begin{aligned} \mathcal{S}_{\hat{A}} &= \frac{16\pi^3}{45}\frac{1}{\beta^3}\frac{r_s^4}{\varepsilon} + \frac{64\pi^3}{45}\frac{r_s^3}{\beta^3}\ln\frac{L}{\varepsilon} \\ &\quad + \frac{16\pi^3}{45}\frac{1}{\beta^3}\left(-\frac{r_s^4}{L} + 6\, r_s^2 L + 2\, r_s L^2 + \frac{L^3}{3}\right) . \end{aligned} \tag{3.116}$$

Let us look at the first two terms

$$\mathcal{S}_0 = \frac{16\pi^3}{45}\frac{1}{\beta^3}\frac{r_s^4}{\varepsilon} + \frac{64\pi^3}{45}\frac{r_s^3}{\beta^3}\ln\frac{L}{\varepsilon} . \tag{3.117}$$

First of all, those contributions can not be seen in the extremal black hole case where we have $\beta \to \infty$, while the radius of the black hole is still finite. For a finite temperature black hole, defining the proper distance from the real horizon to the stretched horizon as δ, we have

$$\delta = \int_{r_s}^{r_s+\varepsilon} \sqrt{g_{rr}}\, dr \approx 2\sqrt{\varepsilon\, r_s}\,. \tag{3.118}$$

Thus, we have $\delta^2 \approx 4\varepsilon\, r_s$. For the finite temperature black hole, with the inverse temperature

$$\beta = 8\pi G_N M\,, \tag{3.119}$$

the entropy (3.117) can be written as

$$\mathcal{S}_0 \propto \frac{r_s^2}{\delta^2} + \ln \frac{L r_s}{\delta^2}\,. \tag{3.120}$$

Those contributions arise because of the redshift between the horizon and the coordinate observer sitting at infinity. Any finite frequency modes near the horizon would have zero frequency as seen by a coordinate observer. We have an infinite number of states with zero energy, and summing over all of those states with a UV cutoff would give us the above result. Those very dense ground state mainly comes from the near-horizon region; thus we have an area contribution [117, 217]. This result was also used to understand the Bekenstein-Hawking entropy by some authors, for example [2], and it can also be interpreted as the entanglement entropy across the stretched horizon [3]. However interpreting those contribution as the origin of the Bekenstein-Hawking entropy is not widely accepted, for example Susskind showed that those contribution should be absorbed by the renormalisation of the Newton's constant G_N due to the loop contribution [117].

The other terms

$$\mathcal{S}_{\hat{A}} = \frac{16\pi^3}{45}\frac{1}{\beta^3}\left(-\frac{r_s^4}{L} + 6\, r_s^2 L + 2\, r_s L^2 + \frac{L^3}{3}\right) \propto \text{Volume} \times T^3 \tag{3.121}$$

are entropy of the thermal fluctuation modes of gauge fields living in the bulk, which can be regarded as the curved spacetime analogue of blackbody result (3.76). $S_{\hat{A}}$ is finite and more or less proportional to the volume between the two boundaries multiplied by T^3.

3.3.3 Zero modes of A_r

In this section, we evaluate the partition function for ϕ, which is the zero modes of A_r. The partition function Z_ϕ can always be written as

$$Z_\phi = \int \mathcal{D}\phi \; e^{-S_\phi} \,, \tag{3.122}$$

with the action

$$S_\phi = \frac{L}{2e^2|y|^2} \int d\tau d^2x \; R^2 \sin\theta \left(\frac{F}{LR^2} \partial_\tau \phi \partial_\tau \phi + \frac{1}{R^2} \partial_\theta \phi \partial_\theta \phi + \frac{1}{R^2 \sin^2\theta} \partial_\varphi \phi \partial_\varphi \phi \right) . \tag{3.123}$$

The action (3.123) has different behaviour in different temperature limits, and we will see how the entropy contribution from ϕ changes as we vary the temperature of the system. One obvious thing to notice is that there is a transition of dominance in function F shown in equation (3.112) at different temperatures. The L^3 term is the dominant contribution for high temperature $L \gg r_s$, and the UV cutoff ε is absent in S_ϕ. While at low temperature $L \ll r_s$, the most important term is $r_s^3 \ln L/\varepsilon$, and we will see different behaviour of the corresponding entropy. We always assume the short distance cutoff ε is always much smaller than L and r_s.

High temperature limit

Let us first discuss the high-temperature limit $r_s \ll L$, as shown in the first panel of figure 3.6. If

$$r_s^3 \ln \frac{L}{\varepsilon} \ll L^3 \tag{3.124}$$

is satisfied, then L^3 is the most important term in F. In this case F can be approximated as

$$F \approx \frac{L^3}{3} \,. \tag{3.125}$$

Function $\operatorname{arcsinh}(x)$ is always much smaller than x for large value of x, so $|y|$ can be written as $|y| \approx L$. Then, the effective action in high-temperature limit can be expressed as

$$S_\phi \;=\; \frac{1}{2e^2 L} \int d\tau d^2x \; L^2 \sin\theta \left(\frac{1}{3} \partial_\tau \phi \partial_\tau \phi + \frac{1}{L^2} \partial_\theta \phi \partial_\theta \phi + \frac{1}{L^2 \sin^2\theta} \partial_\varphi \phi \partial_\varphi \phi \right) . \tag{3.126}$$

The effective action for ϕ is just a scalar field living on a $S^1 \times S^2$, where the length scale of S^1 is β, and the length scale of S^2 is L. The partition function

can be directly evaluated using the standard method discussed in the previous section.

Equipped with the effective action (3.126), we can directly calculate the partition function Z_ϕ shown in (3.122). We can absorb the finite constant $1/3$ in front of $\partial_\tau\phi\partial_\tau\phi$ term in the action into the redefinition of τ to τ'. Let us suppose the fluctuation modes of ϕ are

$$\phi(x^a) \quad = \quad \mathcal{N}_\phi \cdot \sum_\omega \sum_{l,m} e^{i\omega\tau'} Y_{lm}(\theta,\varphi)\tilde{\phi}(\omega,l,m)\,, \tag{3.127}$$

where $\mathcal{N}_\phi$ is a normalisation constant. The partition function for ϕ in canonical ensemble can be written as

$$\ln Z_\phi = -\sum_\omega \ln(1-e^{-\beta'\omega})\,. \tag{3.128}$$

More details of the calculation was demonstrated in a similar calculation for bulk fluctuation modes in appendix A.5. We can change the summation of ω to integration by introducing density of state $g(\omega)$, which gives out

$$\ln Z_\phi \quad = \quad -\int_0^\infty g(\omega)\ln(1-e^{-\beta'\omega})d\omega = \beta'\int_0^\infty \frac{\Gamma(\omega)}{e^{\beta'\omega}-1}d\omega\,. \tag{3.129}$$

$\Gamma(\omega)$ is defined by $d\Gamma = g(\omega)d\omega$. According to the dispersion relation in this background, $\Gamma(\omega)$ can be counted as

$$\Gamma \quad = \quad \sum_{l,m}\frac{\beta'}{2\pi}\sqrt{\frac{l(l+1)}{L^2}} = \frac{\beta'}{2\pi}\sum_l (2l+1)\sqrt{\frac{l(l+1)}{L^2}}\,. \tag{3.130}$$

The summation is from $l=0$ to the level with energy ω. Putting everything back into equation (3.129) and changing the summation of l into integral, we can write the partition function as

$$\ln Z_\phi \quad = \quad \frac{\beta'^2}{2\pi}L^2\int_0^\infty \frac{d\omega}{e^{\beta'\omega}-1}\int_l d\left(\frac{l(l+1)}{L^2}\right)\sqrt{\frac{l(l+1)}{L^2}}\,. \tag{3.131}$$

$\sqrt{l(l+1)/L^2}$ is exactly the energy carried by a particle with angular momentum $l(l+1)$. Thus we can change a variable and get

$$\ln Z_\phi = \frac{\beta'^2L^2}{2\pi}\int_0^\infty \frac{d\omega}{e^{\beta'\omega}-1}\int_0^{\omega^2} dx\ \sqrt{x} = \frac{\beta'^2L^2}{3\pi}\int_0^\infty \frac{\omega^3 d\omega}{e^{\beta'\omega}-1}\,. \tag{3.132}$$

Then the above expression can be worked out, and the logarithm of partition function can be expressed as

$$\ln Z_\phi = \frac{\pi^3}{45}\frac{L^2}{\beta'^2} = \frac{\pi^3}{135}\frac{L^2}{\beta^2}\,. \tag{3.133}$$

Then, the corresponding entropy can be written as

$$\mathcal{S}_\phi = \frac{\pi^3}{45}\frac{L^2}{\beta^2}\,. \tag{3.134}$$

Comparing this result with the entropy of ϕ in the flat case (A.16), we can conclude that the zero modes ϕ shares similar properties with the flat case in the high temperature limit $r_s \ll L$.

So the high-temperature behaviour of field ϕ is just like a 3-dimensional scalar field with inverse temperature β, whose entropy is standard, as shown in equation (3.134). There is nothing different with the flat case.

Low temperature limit

Now, as the temperature cools down, r_s can be larger than L, we arrive at the low-temperature limit of the system $r_s \gg L$, as shown in the second panel of figure 3.6. Let us assume both L and r_s are finite in this case and leave the discussion of the (near)-extremal black hole case for the next section. As we can see, there is a transition of dominance between different terms in F, and in the low-temperature limit F can be approximated as

$$F \approx r_s^3 \ln\frac{L}{\varepsilon}\,. \tag{3.135}$$

We also have $|y|^2 \approx L(L+r_s)$. Then, the effective action for ϕ in the low-temperature limit can be written as

$$\begin{aligned} S_\phi &= \frac{1}{2e^2(L+r_s)}\int d\tau d^2x\; r_s^2\sin\theta \\ &\quad\times\Big(\frac{r_s}{L}\ln\frac{L}{\varepsilon}\partial_\tau\phi\partial_\tau\phi + \frac{1}{r_s^2}\partial_\theta\phi\partial_\theta\phi + \frac{1}{r_s^2\sin^2\theta}\partial_\varphi\phi\partial_\varphi\phi\Big)\,. \end{aligned} \tag{3.136}$$

As discussed before, the time of the 3-dimensional action should always be chosen as the coordinate time τ. Doing so, the coefficient in front of $(\partial_\tau\phi)^2$ is fixed and always much larger than 1. Especially when we take the real event horizon limit, i.e. the $\varepsilon \to 0$ limit, the path integral localises on the space of zero energy modes $\partial_\tau\phi = 0$. Then the partition function of field ϕ only depend on the radius of S^2.

Because of the localisation, the logarithm of the partition function for the zero-point energy is no longer linear in β but only depend on the radius of S^2 and UV cutoff, which can be expressed as

$$\ln Z_\phi \propto r_s^2 \cdot \Lambda^2 \,. \tag{3.137}$$

Let us assume the UV cutoff is the Planck scale, then the corresponding entropy is of Bekenstein-Hawking entropy magnitude

$$\mathcal{S} \propto \frac{r_s^2}{l_p^2} \,. \tag{3.138}$$

In the low-temperature limit, the entropy of ϕ can be written as the area of the horizon divided by the Planck area l_p^2.

Now, Let us summarise what we have gotton for the entropy of zero modes ϕ. The entropy of the system is illustrated in figure 3.7. At high temperature $L \gg r_s$, the entropy of ϕ is more or less like what we have in the flat case, namely the area of the box times temperature squared $L^2 \times T^2$. Here, the presence of a small black hole merely gives a high temperature to the system. This amount of entropy then has a competition with the bulk fluctuation modes. At very high temperature, the volume times temperature curbed wins, and the area times temperature squared wins at lower temperature. This is exactly the same story told in the last section. The difference starts to show up at very low temperature; one can compare figure 3.7 and 3.4 to get a feeling. In the black hole case, the entropy for ϕ is shown in equation (3.138) at very low temperature, which is reminiscent of the Bekenstein-Hawking entropy. In comparison, this phenomenon can not be seen in the flat case, where the entropy has a logarithm behaviour.

3.3.4 Boundary stretched Wilson lines

Now, let us also discuss different behaviour of field W in different temperature limits in this subsection. The effective action for W is

$$\begin{aligned} S_W \quad = \quad & \frac{2}{e^2(L+r_s)} \int d\tau d^2x \; r_s^2 \sin\theta \\ & \times \Big(\frac{r_s}{\varepsilon} \frac{L r_s^2 + r_s^3}{F} \partial_\tau W \partial_\tau W + \frac{1}{r_s^2} \; \partial_\theta W \partial_\theta W + \frac{1}{r_s^2 \sin^2\theta} \partial_\varphi W \partial_\varphi W \Big) \,. \end{aligned} \tag{3.139}$$

with

$$F = 3 r_s^2 L + \frac{3}{2} r_s L^2 + \frac{L^3}{3} + r_s^3 \ln L/\varepsilon \,. \tag{3.140}$$

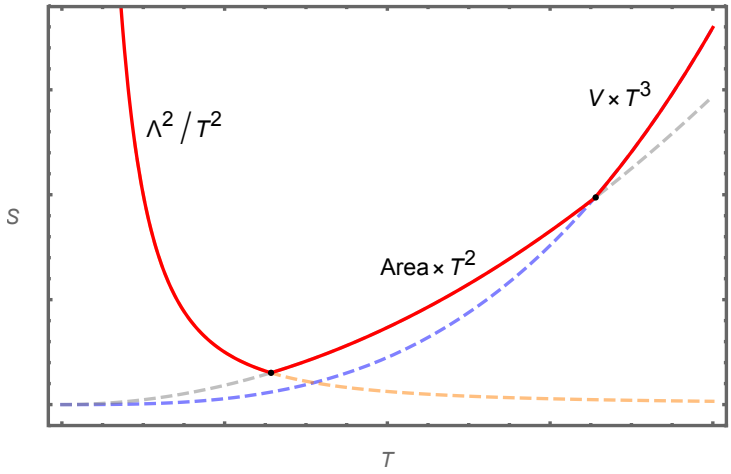


Figure 3.7: A sketch of the entropy of black hole system as we vary the temperature. The high-temperature behaviour is more or less like the flat case in the previous section shown in figure 3.4. At very low temperature, the entropy of the zero modes ϕ is proportional to the area of the horizon, which is proportional to β^2, multiplied by the UV cutoff Λ squared.

The infinite coefficient introduces a localisation on the zero energy modes $\partial_\tau W = 0$, and the entropy of the boundary stretched Wilson lines W is more or less

$$\mathcal{S} \propto \frac{r_s^2}{l_p^2}\,, \tag{3.141}$$

for the same reason discussed in the previous subsection 3.3.3.

Only when the size of the box L is extremely large compare to r_s, namely the high temperature limit, some power of r_s/L may overcome r_s/ε, and we get a finite coefficient. The corresponding entropy can be expressed as $\mathcal{S} \propto r_s^2 \times T^2$, which is a finite constant. In this limit the entropy from W can always be ignored, because the entropy of fluctuation modes always plays the dominant role at high temperature.

As for the constant modes and winding modes of W, their entropy contributions are never comparable with (3.141). We have

$$Z_0 = \sqrt{\frac{e^2\beta(L+r_s)}{\text{Area}}}\,, \qquad Z_w = \sum_n e^0 = \zeta(0)\,, \tag{3.142}$$

which contribute as logarithm of the temperature T, the coupling constant e, and other length scales.

Summary of the main results of the current section Let us summarise this section. Although the entropy of W also contributes as (3.141), the qualitative behaviour of entropy shown in figure 3.7 is not changed. Basically, we have three different phases. At high temperature, the bulk fluctuation modes shown in section 3.3.2 is the most important contribution. The entropy is proportional to the volume of the box multiplied by the temperature cubed, as shown by the right part of the curve in figure 3.7. We also have a contribution (3.120) scales as the area of the horizon plus logarithm correction from the bulk fluctuation modes. As the temperature is lower, the area of the box contribution dominates over other contribution. The corresponding entropy is shown in equation (3.134). At very low temperature, we have an entropy of Bekenstein-Hawking magnitude

$$\mathcal{S} \propto \frac{r_s^2}{l_p^2}\,, \tag{3.143}$$

coming from both the zero modes ϕ and the Wilson lines W. This contribution will not die off at zero temperature; we will see if the similar contribution also appears in the extremal black hole case or not in the next section.

3.4 Extremal Black Hole

In this section, we will see what happens if the temperature of the black hole is approximately zero. We are going to look at a (near)-extremal black hole background and repeat the path integral calculation for U(1) gauge theory on it. To get an extremal black hole, we need to start with the Euclidean Reissner-Nordström (RN) metric, which can be written as

$$ds^2 = f(r)d\tau^2 + f(r)^{-1}dr^2 + r^2 d\Omega_2^2\,, \tag{3.144}$$

with

$$f(r) = 1 - \frac{2G_N M}{r} + \frac{G_N Q^2}{r^2}\,. \tag{3.145}$$

The horizon is a null surface, which can be obtained by solving $g_{rr} = 0$. The inner and outer horizon can be written as

$$r_\pm = (G_N M \pm \sqrt{(G_N M)^2 - G_N Q^2})\,. \tag{3.146}$$

The extremal limit is the limit where we have a double zero, which means

$$r_+ = r_- = G_N M \,. \tag{3.147}$$

Thus, the metric of an extremal black hole can be written as

$$ds^2 = \left(1 - \frac{r_H}{r}\right)^2 dt^2 + \left(1 - \frac{r_H}{r}\right)^{-2} dr^2 + r^2 d\Omega_2^2 \,, \tag{3.148}$$

with

$$r_H = G_N M \,. \tag{3.149}$$

The above metric is the extremal black hole. We can give this black hole system a very tiny temperature by replacing the double zero at the horizon with two single zeros $r_\pm = r_H \pm \varepsilon$ with the short distance cutoff ε. The temperature of the near extremal black hole is calculated in appendix A.6, which is read as

$$\beta = \frac{2\pi r_H^2}{\varepsilon} \,. \tag{3.150}$$

In this limit, the bulk fluctuation modes shown in equation (3.121) will not survive, because of the finite r_H and the small temperature $T = 1/\beta$. So we will mainly focus on fields ϕ and W to see the behaviour of those modes in this section. Actually, all the fluctuation modes contributions whose entropy proportional to the temperature will not play any role in the final result, and we expect to see entropy contributions that look like (3.141).

3.4.1 Effective action

Following exactly the same routine as the finite temperature case, we first shift the radius direction by r_H as

$$\rho = r - r_H \,, \tag{3.151}$$

and then define the proper distance y along the radius direction as

$$dy = \frac{\rho + r_H}{\rho} d\rho \,. \tag{3.152}$$

Solving the above equation, the proper distance y can be expressed as

$$y = \rho + r_H \ln \frac{\rho}{r_H} \,. \tag{3.153}$$

y takes value from y_1 to y_2, with

$$y_1 = \varepsilon + r_H \ln \frac{\varepsilon}{r_H} \,, \qquad y_2 = L + r_H \ln \frac{L}{r_H} \,. \tag{3.154}$$

Unlike $\text{arcsinh}(0) = 0$ in the finite temperature case, we have $\ln(0) = -\infty$ and thus $y_1 \to -\infty$ in this case. Thus, $|y|$ can be approximated as

$$|y| = y_2 - y_1 \approx L + r_H \ln \frac{L}{\varepsilon} \,. \tag{3.155}$$

The Wilson lines stretched between the two boundaries at $y = y_1$ and $y = y_2$ are defined as

$$W = i \int_{y_1}^{y_2} dy \; \hat{A}_y \,. \tag{3.156}$$

Again the action can be separated into two parts

$$S_E = S_{\hat{A}} + S[\phi, W] \,, \tag{3.157}$$

where

$$S_{\hat{A}} = \frac{1}{4e^2} \int_{\mathcal{M}} d\tau d^3x \sqrt{g} \; \hat{F}^{\mu\nu} \hat{F}_{\mu\nu} \,. \tag{3.158}$$

Again, the bulk fluctuation modes will not give out any important contribution in the super-low temperature limit and can be ignored. The part we are going to focus on is

$$\begin{aligned} S[\phi, W] \;\; = \;\; & \frac{1}{2e^2|y|^2} \int_{\mathcal{M}} d\tau d^3x \sqrt{g} \left[g^{ab} \partial_a \phi \partial_b \phi \right] \\ & - \frac{i}{e^2|y|} \int d\tau d^2x \left[\sqrt{g} \; g^{ab} \right]_{y=y_1}^{y=y_2} \partial_a W \partial_b \phi \,. \end{aligned} \tag{3.159}$$

Following the same dimensional reduction procedure as shown in section 3.3.1, we can obtain the 3-dimensional effective action for fields ϕ and W as

$$S_{\phi,W} = S_\phi + S_W \,, \tag{3.160}$$

with

$$\begin{aligned} S_\phi \;\; = \;\; & \frac{L}{2e^2|y|^2} \int d\tau d^2x \; r_H^2 \sin\theta \\ & \times \left(\frac{r_H^2}{\varepsilon L} \partial_\tau \phi \partial_\tau \phi + \frac{1}{r_H^2} \partial_\theta \phi \partial_\theta \phi + \frac{1}{r_H^2 \sin^2\theta} \partial_\varphi \phi \partial_\varphi \phi \right) \,, \\ S_W \;\; = \;\; & \frac{2L}{e^2 (r_H + L)^2} \int d\tau d^2x \; r_H^2 \sin\theta \\ & \times \left(\frac{(r_H + L)^2}{\varepsilon L} \partial_\tau W \partial_\tau W + \frac{1}{r_H^2} \partial_\theta W \partial_\theta W + \frac{1}{r_H^2 \sin^2\theta} \partial_\varphi W \partial_\varphi W \right) \,. \end{aligned}$$

Note that we have kept the coordinate time τ as the time coordinate for the 3-dimensional effective actions. As we can see from the above effective actions,

for finite r_s and L, the coefficients in front of terms $(\partial_\tau \phi)^2$ and $(\partial_\tau W)^2$ are proportional to $1/\varepsilon$. Taking the $\varepsilon \to 0$ limit, the path integral is located on the space of $\partial_\tau \phi = \partial_\tau W = 0$, same as what we discussed in the low-temperature limit of the finite temperature black hole case. Then the entropy of those modes can be calculated as

$$\mathcal{S} \propto \frac{r_H^2}{l_p^2}\,, \tag{3.161}$$

which is exactly matches the low-temperature entropy (3.141) of the finite temperature black hole. The extremal black hole calculation here can be regarded as a consistency check. Also, the partition functions of the constant modes and winding modes can be calculated, similar to equation (3.142). One thing worth noticing is that the entropy of constant modes proportional to

$$\mathcal{S}_0 \propto \ln |y|\,, \tag{3.162}$$

can be a large contribution.

In the (near)-extremal black hole case, the localisation of $\partial_\tau \phi = \partial_\tau W = 0$ space in the path integral is extremely straightforward as shown in the action (3.160). This is different from the finite temperature black hole case where we have to take different limits. So the localisation phenomenon and the entropy (3.161) are intrinsic for the extremal black hole, which should be used as a general mechanism to explain the entropy of the extremal black hole putting in a finite size box.

Note that from the action (3.160), the zero modes ϕ and the Wilson lines W are all effectively living near the horizon rather than the boundary of the box. The result is from direct mathematical calculations, and we do not have an explanation of why this is the case. There might be redshift related arguments, which needs further understanding.

3.5 Conclusion and Discussion

This chapter mainly evaluates the partition functions and entropy for U(1) gauge theory living on different backgrounds with boundaries, using the Euclidean path integral method. The calculation is done in three different situations: the flat capacitor with a temperature, the finite temperature black hole in a box, and the extremal black hole in a box. According to the canonical analysis discussed in section 3.2, the physical degrees of freedom are the bulk fluctuation modes, the zero modes of A_r component, and the boundary

stretched Wilson lines. The above fields have a well-defined symplectic form and Poisson bracket with their conjugate momenta. There is a competition between those modes in different temperature limits, and the partition function goes through some phase transitions as the temperature varies.

Firstly, the flat case is a good place to get familiar with the general method and physics. The results are mainly shown in figure 3.4. At high temperature, the partition function and entropy is the standard one for black-body radiation. The entropy contains two copies of physical polarisations, and is proportional to the volume between the two boundaries multiplied by T^3. As the temperature becomes cooler than before, the only survived zero modes start to be the most important contribution. Also, the boundary stretched Wilson lines W have the same behaviour as ϕ. ϕ and W just behave like two lower-dimensional massless scalar fields living on the boundary. At relatively high temperature, the fluctuation modes of ϕ and W are dominant whose entropies scale as the area of the boundary times T^2 as shown in equation (4.1). At the even lower temperature, the fluctuation contributions die off; and we are left with the constant and winding modes of W, whose entropies more or less scale as the logarithm of coupling constant e^2 and other length scales of the theory as shown in figure 3.3.

For the black hole case, the partition function of the system has a huge difference with the flat case at the low-temperature limit. The main difference between the flat case and finite temperature black hole case can be seen by comparing figure 3.4 and 3.7. The high-temperature limit of the black hole is fairly similar to the flat case, where the presence of the black hole merely gives a high temperature to the system. The dominant contribution comes from the bulk fluctuation modes, whose entropy is shown in equation (3.116). There is also a contribution from the modes living very close to the horizon, whose entropy proportional to the area of the horizon. Interesting things start to happen at low temperature; as we gradually lower the temperature, the entropy of the zero modes ϕ and Wilson lines W firstly behave as the area of the box multiplied by the temperature squared and then scale as the area of the horizon divided by UV cutoff squared. The entropy proportional to the horizon area comes from the modes $\partial_\tau \phi = \partial_\tau W = 0$. This behaviour is depicted in figure 3.7. We also checked that the same behaviour of ϕ and W remains to be true in the extremal black hole limit. The above argument suggests that the low-temperature Bekenstein-Hawking entropy might comes from the zero modes ϕ and boundary stretched Wilson lines W. There are phase transitions

between the low-temperature black hole and high-temperature ones.

As for the phase transitions for the finite temperature black hole, especially the phase transition at the low temperature, there should be some symmetry breaking pattern to explain them. The low-temperature entropy of the flat case scales as logarithm because the degeneracy manifold is a circle in the symmetry-breaking phase of U(1) global symmetry. This can be seen from the bottom of "Mexican hat" potential. In such a sense, the low-temperature Bekenstein-Hawking-like entropy

$$\mathcal{S} \propto \frac{\text{Area}}{l_p^2} ,$$

can also be explained from a symmetry breaking viewpoint. We are sitting in a global symmetry breaking phase at low temperature with degeneracy $\exp \mathcal{S}$. Note that those amount of entropy comes from the modes in the limit

$$\lim_{\omega \to 0} \ \tilde{\phi}(\omega, x^2, x^3) e^{i\omega\tau} \tag{3.163}$$

because of the localisation on $\partial_\tau \phi = 0$ space. This might can be thought of as calculating the entropy of the soft hair of the black hole system. There are some interesting relations between soft hair from symmetry breaking, Barnich's non-proper gauge degrees of freedom, and others, needed to be further understood.

Summing over boundary conditions

In this chapter, we have worked in a frame of a fixed boundary condition for A_a, and the boundary condition fixes a specific Hilbert space for us. Here we have chosen the boundary condition (3.15), such that the variation of the action

$$\delta S = -\frac{1}{e^2} \int_{\partial \mathcal{M}} d^3x \ F^{ra} \delta A_a \, , \tag{3.164}$$

vanishes. All the boundary configurations of A_a are fixed. But the A_r component can be arbitrary on the boundaries, and we need to sum over those configurations of A_r in the path integral. Those boundary configurations of A_r are the origin of ϕ and W, which is where the interesting physics lies in this chapter.

The Hilbert space is fixed by a given boundary configuration f_a^α, so summing over f_a^α amounts to adding different superselection sectors together. In principle, this is not allowed, and we should restrict ourselves in a fixed Hilbert space. However, there can be exceptions if one wants to consider other physical

effects through some physical arguments. As we mentioned in section 3.2.2, the modes that related to the large gauge transformations are

$$A_\tau \ni \frac{2\pi n}{\beta}, \qquad n \in \mathbf{Z}, \tag{3.165}$$

with integer $\mathbf{Z}$. Those modes do not correspond to a fixed boundary configuration, but they can be physical, thus can be counted in the functional integral. So if one wants to count the effect of those modes, one may need to add those modes in the path integral. Furthermore, for a fixed boundary configuration f_a^α, there are some configurations that are related with f_a^α by some gauge transformations. This is a subset of all the general boundary configurations and does not form superselection sectors.

So one can sum over all the boundary configurations that have the same field strength as our fixed f_a^α in the path integral. By doing this, we allow boundary gauge transform and treat the would-be-gauge boundary degrees of freedom as physical ones. The boundary configurations can be denoted as

$$A_a\Big|_{\partial\mathcal{M}} = f_a^\alpha + C_a^\alpha + \partial_a \lambda^\alpha, \tag{3.166}$$

where f_a^α are the fixed boundary configurations, C_a are constants, and λ is the gauge parameter. Fixed boundary configurations do not have a large contribution in the path integral, and we only need to sum over different $(C_a^\alpha, \lambda^\alpha)$ configurations.

Here is the argument why those boundary configurations that have same field strength do not form different superselection sectors. The reason why we do not sum over arbitrary f_a^α in the main context is that the boundary conditions form superselection sectors, and there is no finite energy operator that can take us from one superselection sector to another. Nevertheless, this is not the case for those boundary configurations shown in (3.166), and it does not cost energy to move from different sectors with different $(C_a^\alpha, \lambda^\alpha)$. Thus no superselection sector is formed, and we can sum over those flat boundary configurations. There is similar argument for summing over soft hair or would-be-gauge degrees of freedom [211]. And different flat boundary configurations $(C_a^\alpha, \lambda^\alpha)$ might correspond to different soft hair dressing.

As argued above, if no superselection sector is formed for fixed f_a^α with different $(C_a^\alpha, \lambda^\alpha)$, one might consider summing over those configurations in the Euclidean path integral. The boundary configurations f_a^α correspond to the bulk configurations B_a, thus we need to sum over the B_a configurations corresponding to different $(C_a^\alpha, \lambda^\alpha)$. Let us take the flat parallel plates case shown

in section 3.2 as an example; the corresponding bulk B_a configuration can be expressed as

$$\begin{aligned} B_a &= \left[(C_a^{(r)} - C_a^{(l)}) + \partial_a(\lambda^{(r)} - \lambda^{(l)})\right] \frac{r}{L} + (C_a^{(l)} + \partial_a \lambda^{(l)}) \,, \\ B_r &= \frac{1}{L}[x^a \cdot (C_a^{(r)} - C_a^{(l)}) + (\lambda^{(r)} - \lambda^{(l)})] \,. \end{aligned}$$

where (l) and (r) label the left and the right boundaries. We also worked out the corresponding bulk solutions for black hole case in appendix A.2. In the previous sections, we choose fixed $(C_a^\alpha, \lambda^\alpha)$ configurations and do not sum over those configurations, whereas here we sum over all those different configurations. Let us write the partition function we calculated in the main context with fixed f_a^α and fixed $(C_a^\alpha, \lambda^\alpha)$ configurations as $Z_{C,\lambda}$. The corresponding path integral for summing over different flat boundary configurations $(C_a^\alpha, \lambda^\alpha)$ can be written as

$$Z = \int \mathcal{D}B_a \times \mathcal{D}\hat{A}_\mu \mathcal{D}\phi \mathcal{D}W \; e^{-S_E} = \int dC^{(l)} dC^{(r)} \mathcal{D}\lambda(x^a) \; Z_{C,\lambda} \,. \tag{3.167}$$

Note that each boundary configuration $(C_a^\alpha, \lambda^\alpha)$ is energetically equivalent with each other; thus, $Z_{C,\lambda}$ usually does not depend on what specifically $(C_a^\alpha, \lambda^\alpha)$ is. Then, we can extract $Z_{C,\lambda}$ from the integral in calculating the partition function and work out the functional integral over configurations $(C_a^\alpha, \lambda^\alpha)$. Whereas, if there is any operator that depends on the boundary condition, we can not extract the operator out of the integral (3.167) when calculating correlation functions. Then, we may have concepts of averaging over boundary configurations $(C_a^\alpha, \lambda^\alpha)$ in this case.

Furthermore, one can choose different boundary conditions, for example,

$$F^{ra}\Big|_{\partial\mathcal{M}} = 0 \,. \tag{3.168}$$

We need to sum over all the configurations that respect the above condition in the functional integral. We expect the bulk fluctuation modes should be more or less the same, but the boundary related degrees of freedom might be different. For example, the boundary configurations of A_r can not be arbitrary, but only some flat configurations are allowed. Boundary configurations of A_a should also be added into the path integral because $\delta A_a \neq 0$ and $F_{ab} \neq 0$ in this case. The boundary configurations are harder to control, and we leave this case for further studies.

Two Bekenstein-Hawking-like entropies

In this chapter, the main point we want to make is that the zero modes ϕ and the Wilson lines W have an entropy proportional to the horizon area divided by the Planck area for super-low-temperature black holes, which can be used to understand the Bekenstein-Hawking entropy because of the right magnitude. Note that there is also part of the entropy (3.120) in the bulk fluctuation modes for the finite temperature black hole system, which is also proportional to the area divided by the UV cutoff squared. This entropy was used to understand the microscopic degrees of freedom of finite-temperature black hole systems [2], and was interpreted as the entanglement entropy across the stretched horizon [3, 117]. If we accept the above argument, we might tend to interpret that the finite-temperature Bekenstein-Hawking entropy comes from some extra microscopic structure near the horizon like entanglement across the horizon, but the (near)-extremal black hole entropy comes from different places as the finite-temperature black hole. The (near)-extremal black hole entropy only appears in low temperature and comes from the breaking of global symmetries. Thus, we have two different types of Bekenstein-Hawking-like entropies for finite-temperature and (near)-extremal black holes, and they both behave like the area of the horizon divided by the UV cutoff squared. We leave the symmetry breaking explanation of the phase transitions and other related topics for further study.

Evaporating black hole and alopecia paradox

As a digression, we briefly discuss a possible application of the phase transition of the entropy shown in figure 3.7. In order to better understand the AMPS firewall [169], Yoshida proves the decoupling theorem in paper [173] (also discussed in [175]) to reconstruct black hole interior operators. The basic idea of Yoshida's decoupling theorem can be summarised as follows.

Including an extra system $\mathcal{O}$ with a large Hilbert space $\mathcal{H}_\mathcal{O}$ that has correlations with the black hole system B, the late time radiation H and early radiation R are decoupled under a scrambling unitary evolution of the black hole. The Hawking partner P can be constructed only using the black hole Hilbert space $\mathcal{H}_B$ and the extra system $\mathcal{H}_\mathcal{O}$, and no information of the early radiation R is needed. The construction can be realised with an error using the Petz map.

The theorem implies that including a large extra system in the black hole can avoid AMPS firewall and construct the Hawking partner P. However, there is an important condition, which is the dimension of the Hilbert space $\mathcal{H}_{\mathcal{O}}$ must be very large comparing to the late time Hawking radiation $\mathcal{H}_H$, and the error mentioned above is proportional to

$$\text{Error} \sim \left(\frac{\dim \mathcal{H}_H}{\dim \mathcal{H}_{\mathcal{O}}}\right)^2 \ll 1\,. \tag{3.169}$$

So to properly complete the task, $\dim \mathcal{H}_{\mathcal{O}}$ must be much larger than $\dim \mathcal{H}_H$. This condition will bring us another paradox in the process of black hole evaporation.

Because of the correlation between $\mathcal{O}$ and part of black hole system$B$, the black hole system also encodes the information of $\mathcal{O}$. In paper [175], it was shown that the operator $\mathcal{O}$ can correlate with the black hole system by gravitational dressing. Thus we have

$$\dim \mathcal{H}_B > \dim \mathcal{H}_O\,. \tag{3.170}$$

The combination of equations (3.169) and (3.170) implies

$$\dim \mathcal{H}_B \gg \dim \mathcal{H}_H \tag{3.171}$$

must be true. As the black hole evaporates, $\dim \mathcal{H}_B$ is decreasing. Once the above condition can not be fulfilled, the decoupling theorem fails to hold, and the firewall comes back again. In this sense, we have not completely wiped out the firewall at the Page time but postponed it until the end stage of the evaporation.

If we adopt the idea that the correlations between the extra system $\mathcal{O}$ and the black hole B are through soft hair [175], then we are keep losing the soft hair degrees of freedom as the black hole evaporates because the entropy of the soft hair must be smaller than the coarse-grained entropy of the black hole. When we do not have enough soft hair, the decoupling theorem fails. So we call it

Alopecia paradox: *At the end stage of the black hole evaporation, the black hole system does not have large enough coarse-grained entropy to support large $\mathcal{H}_{\mathcal{O}}$ to construct the Hawking partner.*

The above alopecia paradox is a paradox in a more general setting rather than a paradox that only appears in Yoshida's decoupling theorem. Most of the attempts trying to reconstruct the interior operators using quantum

information protocol rely on a large phase space of the black hole system. If the black hole system only left with a few qubits, even still exponentially hard, the Harlow-Hayden decoding task [171, 172] is not that hard. Thus, we always face the AMPS firewall if the black hole system is only left with several qubits. If the firewall shows up once on the horizon, there would be a firewall all over the horizon because of the general covariance of the general relativity. This paradox can be seen more explicitly by the coarse-grained entropy of an evaporating black hole [160, 161]. The Stephan-Boltzmann law of blackbody radiation can be written as

$$\frac{dM}{dAdt} \propto T^4, \tag{3.172}$$

with the area of the black hole A. Then, the power of the evaporation can be written as

$$\frac{dM}{dt} \propto \frac{1}{M^2}. \tag{3.173}$$

Working out the above integral, we obtain

$$M = M_0 \left(1 - \frac{t}{t_0}\right)^{1/3}, \tag{3.174}$$

where M_0 is the original mass of the black hole, and t_0 is the total time for the evaporation. Similarly, for the entropy of the system, we have

$$S = S_0 \left(1 - \frac{t}{t_0}\right)^{2/3}, \tag{3.175}$$

with the original entropy of the black hole S_0. The coarse-grained entropy of the black hole system is depicted in figure 3.8. As we can see from the figure, at the end of the evaporation the coarse-grained entropy of the system goes to zero, thus can not support large amount of soft hair degrees of freedom for the purpose of decoupling between the late and early time radiation and constructing the Hawking partner.

A possible way to avoid this problem is provided in the calculations of this chapter. We can transfer the entropy shown in figure 3.7 to an evaporating black hole. The actual coarse-grained entropy of the black hole system is illustrated in figure 3.9. As we can see, the entropy follows the same curve as an evaporation black hole shown in figure 3.8 at the early stage of the evaporation. Then the temperature of the system rises, we start to see the phase transition of the field ϕ. After the phase transition, the entropy of ϕ scales as $L^2 \times T^2$, which is large enough at high temperature. Thus, Yoshida's

decoupling theorem still hold at the end of the evaporation, and we can still reconstruct the interior operators and stay away from the firewall.

We have to admit that the above argument about the alopecia paradox is very suggestive, but we think it is interesting to present and discuss it here. At the end of the evaporation, the very tiny black hole might already beyond the scope of general relativity, so the problem might not be very strange, and we do not really need to worry about it.

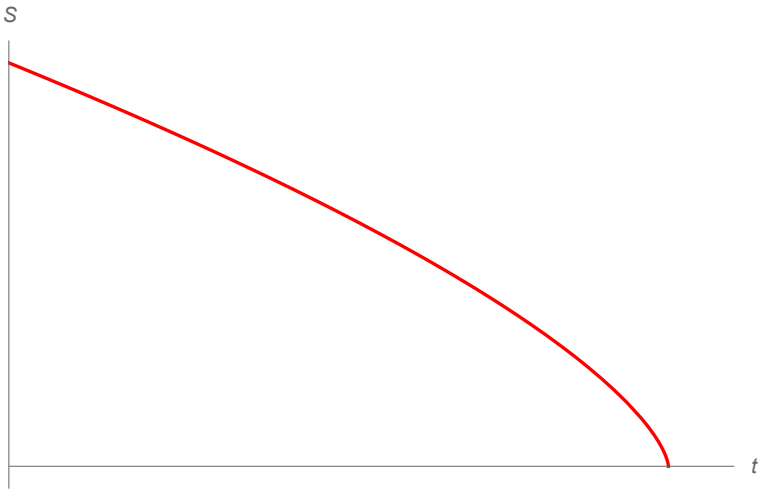


Figure 3.8: The coarse-grained entropy of black hole system decrease to zero as the alopecia paradox for evaporating black hole.

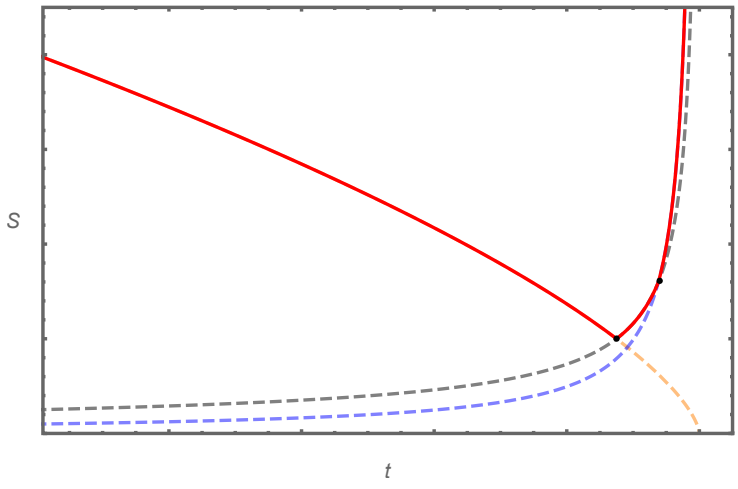


Figure 3.9: The entropy of an evaporating black hole system is illustrated here. Firstly, the entropy decreases, following the same curve as shown in figure 3.8. Then, as the system's temperature becomes high, there is a phase transition of the black hole system, and after that, the entropy scales as Area$\times T^2$ and Volume$\times T^3$.

4 Soft Black Hole Information Paradox

One of the most important black hole crises is the black hole information paradox (BHIP), which is sharpened by the AMPS firewall argument. To avoid the firewall near the horizon, one may need to ask the early and late radiation to be decoupled. Recent works [174, 175] considered soft hair as the core ingredient in the decoupling theorem to avoid firewall and reconstruct Hawking partner. With the help of the seven key assertions of soft hair, Pasterski and Verlinde argued that the infalling observer would never knock into a firewall, and the interior operator can be reconstructed via the Petz map. However, the derivation of the Page curve of a black hole collapsed from a pure state, which is essential in the paradox, is not addressed. Therefore, this chapter aims to further understand the role of soft hair in the BHIP and derives a Page curve using Maxwell soft hair.

Treating Maxwell soft hair as a transition function that relates U(1) gauge fields living in the asymptotic region and the near-horizon region, the U(1) gauge parameter $\lambda(x^a)$ naturally becomes a good label of those Maxwell soft hair degrees of freedom. This interpretation also builds the connection between Maxwell soft hair and U(1) edge modes living in the intermediate region, which admits a well-defined effective action description. We study the statistical properties by Euclidean path integral, which concludes that the soft hair density of state increases with the black hole temperature. Hawking radiation increases black hole entropy by creating entanglement, while the measurement of soft mode projects the black hole onto a lower entropy state. The competition between phase spaces of Hawking radiation and soft hair measurement gives rise to one version of the Page curve consistent with the unitary evolution of the black hole.

4.1 Introduction

The BHIP was put forward by Hawking in 1976 [218], which ironically seems to suggest Hawking's own model of radiation [7, 12] should be modified largely at the late time of evaporation. Even with the recent exciting progress of the "Island prescription" [1, 178–180, 190, 196, 219], there are no general agreements regarding what has been missed by Hawking's calculation. Was Hawking missed quantum effects of gravity? Then we might need a while to understand this problem completely. Or optimistically, if the answer is classical global effects, the problem might be more handleable.

The paradox is that the von Neumann entropy of black holes cannot be larger than the coarse-grained entropy that is proportional to the area of the horizon. Hawking radiation creates entanglement between the black hole and radiation. However, in principle, after a certain time of evaporation t_{page}, no more entanglement can be created, because there is simply no room in the black hole to store the information of the Hawking partner P anymore. We meet the conundrum that fine-grained entropy becomes larger than coarse-grained entropy if the black hole continues to evaporate after the Page time. Naively, there are several ways out. The first one is that the Hawking partner P stored in the black hole can go out of the same horizon as Hawking radiation H to purify the radiation, which is forbidden because of the Almheiri-Marolf-Polchinski-Sully (AMPS) firewall argument [169]. The second one is that the information of P is encoded somewhere else, which seems not possible because of the no-hair theorem [220, 221]. The third one is that the black hole interior is already included in the entanglement wedge of radiation and can be reconstructed from the radiation [222, 223]. The recently proposed island prescription also belongs to this category.

The recent paper by Pasterski and Verlinde (PV) [175] seems to have unified the above arguments, claiming that the global soft hair beyond the no-hair theorem is the handle to reconstruct interior operators and is firewall-free. They treated the supertranslation soft hair as a transition function that connects the asymptotic region and the near-horizon region, which provided a way to understand the soft hair degrees of freedom. A dressed infalling observer would perform a measurement of the classical value of the transition function $f(z^A)$ when crossing the intermediate region. The measurement can project the black hole onto a specific soft hair state that enables the reconstruction of the black hole interior and decreases the black hole entropy. They also gave

seven key assertions that related to the soft hair and checked them in the paper. However, the supertranslation story did not provide a way to evaluate the phase space of supertranslation soft hair and thus cannot answer if the soft hair is powerful enough to give rise to the Page curve [160] consistent with the unitary evolution of the black hole or not.

Maxwell's theory, on the other hand, has a much simpler structure, and we are able to analyse the phase space quantitatively and get an effective action description of Maxwell soft hair. By studying the Maxwell soft hair in the current chapter, the U(1) gauge transition function is identified with the edge modes living in the intermediate region studied in literature [119, 136–138, 159]. The basic idea is the following. Divide the spacetime into three regions: the near-horizon region $\mathcal{M}$, the asymptotic region $\bar{\mathcal{M}}$, and the intermediate boundary $\partial\mathcal{M}$; the gauge fields living in $\mathcal{M}$ and $\bar{\mathcal{M}}$ need to match each other on the boundary. We introduce Lagrange multiplier currents J^μ along the boundary to match the above gauge fields. The path integral over J^μ gives out an effective action for λ which is a massless scalar living on the boundary

$$S[\lambda] = -\frac{1}{2f(L)} \int_{\partial\mathcal{M}} d^3x\sqrt{-h}\ \partial^a\lambda\partial_a\lambda\,. \tag{4.1}$$

Now the would-be-gauge parameter $\lambda(x^a)$ becomes physical degrees of freedom living on the boundary, and the massless bosonic field can be regarded as the Goldstone modes of the large gauge symmetry on the boundary.

The Hawking radiation increases the entropy of the black hole, and the measurement of Maxwell soft hair projects the black hole onto lower entropy states. Fermi's golden rule compares the rates of different physical processes by comparing their phase space, which can also be adopted here to compare the Hawking radiation and soft hair measurement. Equipped with the effective action (4.1), we can directly use the Euclidean path integral to calculate the partition function and evaluate the size of the phase space for Maxwell soft hair. The competition between the phase spaces of the soft hair and black hole, i.e., the rates of measurement and Hawking radiation, gives out one version of the Page curve consistent with unitarity.

Note that soft hair as a potential solution to BHIP is not new and was already hinted at in many papers, such as [107, 109–113, 211, 224–230]. The new ingredients here are the Page curve and the analysis of the phase space of soft hair by studying edge modes effective action.

Another motivation of this chapter is that people are searching for free param-

eters [205, 207, 208] that label the superselection sectors inspired by island prescription and the ensemble average proposal [190, 199, 200]. Swampland program opposes such kind of free parameters for $d > 3$ in quantum gravity [203, 231–233]. However, approximate global symmetry [234] rather than exact ones might survive from swampland and can provide such free parameters.

The remainder of the chapter is organised as follows. In section 4.2, we provide a brief review of the gravitational story, including properties of supertranslation soft hair, gravitational dressing of operators, and quantum information protocol to reconstruct interior operators. In section 4.3, we present an analysis of the relation between U(1) edge modes and Maxwell soft hair and get an effective action for those Goldstone modes. The effective action enables us to calculate the statistical properties of those soft hair degrees of freedom. In section 4.4, we introduce Fermi's golden rule to analyse the rates of two processes that increase and decrease the entropy of the black hole separately. A version of the Page curve can be gotten from the competition of those two processes. We end with some conclusions and further comments in the last section.

4.2 Review of Gravitational Story

Here we provide a brief review of the supertranslation story. The transition function related to the supertranslation between different regions on an asymptotic flat black hole is analysed to make the Maxwell story conceptually easier to accept. We also provide some basic idea of what role soft hair can play in the BHIP by some quantum information protocols. One can consult [107, 175, 211] for more details.

The origin of the gravitational soft hair is the ambiguity in defining the metric in the asymptotic region. For an asymptotically flat Schwarzschild black hole, the asymptotic region is Minkowski spacetime Mink_4 as we take $r \to \infty$ limit. An isometry is a coordinate transformation that leaves the metric invariant, which is a sign of the spacetime symmetry and corresponds to conserved charges. The asymptotic symmetries are the diffeomorphisms that preserve the asymptotic metric because the change of the metric coming from the transformation can also die off as $r \to \infty$. It is an extended version of isometry. At the linearized level, the asymptotic group is generated by the transformations called supertranslations. For a standard Schwarzschild metric in the advanced

Bondi coordinates, the diffeomorphisms

$$g_{\mu\nu} \to g'_{\mu\nu} = g_{\mu\nu} + \mathcal{L}_\zeta g_{\mu\nu} \tag{4.2}$$

that preserve Bondi gauge and the standard falloff conditions at the asymptotic region are generated by vector [107]

$$\zeta_f = f\ \partial_v - \frac{1}{2}D^2 f\ \partial_r + \frac{1}{r}D^A f\ \partial_A\,. \tag{4.3}$$

Note that f is a function of the celestial sphere coordinates.

One of the crucial insights of [175] is that one can change perspective and characterise that diffeomorphism in terms of a transition function that connects the the asymptotic and the near-horizon coordinates. We choose light cone coordinates (u, v) in the asymptotic region such that the metric is

$$ds^2\big|_{\text{asymp}} = \Lambda du dv + r^2\gamma_{AB}dz^A dz^B\,,\ \Lambda \equiv 1 - \frac{2M}{r}\,, \tag{4.4}$$

with γ_{AB} the metric on S^2. In the near-horizon region, we use Kruskal-Szekeres coordinates (U, V) and write the metric as

$$ds^2\big|_{\text{hor}} = -F dU dV + r^2\gamma_{AB}dZ^A dZ^B\,,\ F \equiv \frac{2M}{r}e^{-r/2M}\,. \tag{4.5}$$

Those two coordinates can be patched together via coordinate transformation

$$\frac{U}{4M} = -e^{-u/4M}\,,\quad \frac{V}{4M} = -e^{v/4M}\,,\quad Z^A = z^A\,. \tag{4.6}$$

There is ambiguity in matching those two coordinates, and thus the transformation can be modified by adding linearized soft hair into

$$\frac{U}{4M} = -e^{-u/4M} + \zeta_f^U\,, \tag{4.7}$$

$$\frac{V}{4M} = -e^{v/4M} + \zeta_f^V\,, \tag{4.8}$$

$$Z^A = z^A + \zeta_f^A\,. \tag{4.9}$$

Here ζ_f is exactly the components of the vector field associated with the supertranslation in equation (4.3). The above coordinate transition between the two coordinate systems (4.4) and (4.5) can also be written as [175]

$$v = 4M\ln\frac{V}{4M} - f\,, \tag{4.10}$$

$$u = -4M\ln(-\frac{U}{4M}) - f - \frac{1}{F}e^{(u-v)/4M}D^2 f\,, \tag{4.11}$$

$$z^A = Z^A - \frac{1}{r}D^A f\,, \tag{4.12}$$

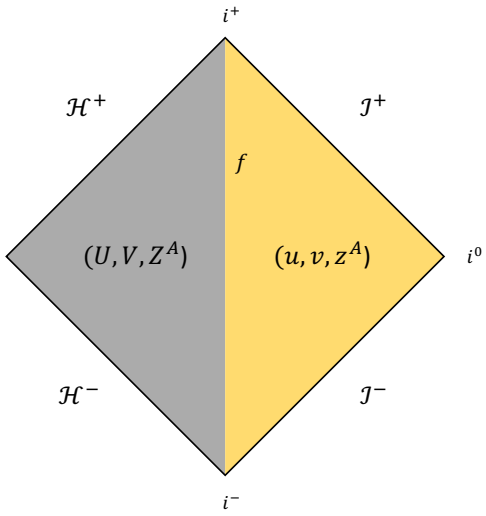


Figure 4.1: Supertranslation soft hair degrees of freedom are encoded in the transition function f that relates different coordinate systems. The relations between those two coordinates are shown in equations (4.7)-(4.9).

where D_A is the covariant derivative with respect to γ_{AB}. The transition function f that relates different coordinate systems is schematically depicted in figure 4.1. Now, as shown in figure 4.2, there are extra degrees of freedom in defining the asymptotic structure for the observers sitting on the horizon, which are labelled by the supertranslation parameter $f(z^A)$. Similarly, the asymptotic observers also have problems in deciding what soft hair state the black hole is. Those extra degrees of freedom are the gravitational soft hair.

The gravitational soft hair discussed above is physical degrees of freedom rather than redundancy in the description. There are conserved charges associated with the supertranslation and canonical conjugate momentum dual to the parameter $f(z^A)$ through standard symplectic form analysis [211]. Promoting f to a quantum operator, $\hat{f}(z^A)$ describes the soft hair Goldstone modes. The soft charge $Q_S(z^A)$ can be written as

$$Q_S(z^A) = \int dv\ \hat{q}_S(v, Z^A)\,, \tag{4.13}$$

and conjugate variable to $\hat{f}$ can be seen from the commutator

$$[\hat{q}_S(v, z^A), f(v', z'^A)] = i\delta(v - v')\delta^{(2)}(z - z')\,, \tag{4.14}$$

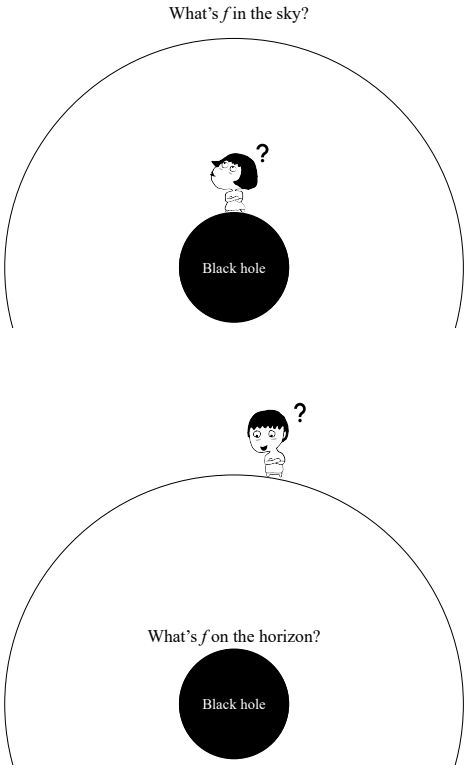


Figure 4.2: Schematic depiction of asymptotic symmetry group and soft hair. **First panel**: for an observer sitting on the horizon, there is an ambiguity in the sky generated by the diffeomorphisms that preserve the asymptotic structure. This ambiguity is labelled by a parameter $f(z^A)$ and is a global effect. **Second panel**: an asymptotic observer meets a similar problem. The black hole is a superposition state of different $f(z^A)$s unless the observer jumps into the black hole to measure it. By jumping across the boundary, the black hole is projected to a state with more specific f, which has lower entropy.

with

$$\hat{q}_S(v, z^A) = \frac{1}{16\pi} D^2(D^2+2)\ \partial_v \hat{f}\,. \tag{4.15}$$

Also, one can implant soft hair by sending in shock waves at $v = v_0$ with some specific stress tensors [107]

$$\hat{T}_{vv} = \frac{1}{16\pi M^2}\left[m - \frac{1}{4}D^2(D^2-1)\hat{f}\right]\ \delta(v-v_0)\,, \tag{4.16}$$

$$\hat{T}_{vA} = \frac{3}{32\pi M} D_A \hat{f}\ \delta(v-v_0)\,. \tag{4.17}$$

where m is the mass of the shock wave. The gravitational soft hair can be changed by the shock wave stress tensor

$$g_{\mu\nu} \to g'_{\mu\nu} = g_{\mu\nu} + \Theta(v-v_0)\left(\mathcal{L}_f g_{\mu\nu} + \frac{m}{r}\delta^v_a \delta^v_b\right)\,, \tag{4.18}$$

which means that $\hat{f}$ is changed to $\hat{f} + f$ by the shock wave. Implanting supertranslation soft hair is another physical argument showing that soft hair is not a description redundancy.

In a gravitational theory, diffeomorphism invariant (physical) operators must commute with the total supertranslation charge $Q_f = Q_S + Q_H$

$$[Q_f, \mathcal{O}_{phys}] = 0\,, \tag{4.19}$$

which requires the physical operators factorise into a product of matter operator $\mathcal{O}$ times a gravitational Wilson line $\mathcal{W}$, i.e.,

$$\mathcal{O}_{phys} = \mathcal{O} \times \mathcal{W}\,. \tag{4.20}$$

The Wilson line $\mathcal{W}$ takes the following form [230]:

$$\mathcal{W}(k, z^A) = e^{-ik\hat{f}(z^A)}\,. \tag{4.21}$$

For an infalling operator $\mathcal{O}(v, z^A)$, the gravitational dressing is just adding an extra phase on the momentum eigenstate e^{ikv}. So the dressed infalling operator can be simply expressed by replacing the original v coordinate by

$$\hat{v} = v - \hat{f}\,. \tag{4.22}$$

Under the action of Q_S, $\hat{f}$ is shifted by an amount f,

$$\hat{f} \to \hat{f} + f\,. \tag{4.23}$$

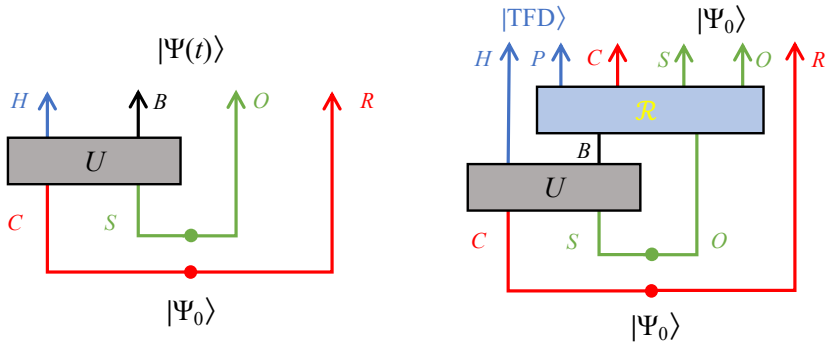


Figure 4.3: Diagrammatic representation of the black hole evolution. **First panel**: a system including code subspace C, early radiation R, soft hair S, and observer O evolve for a period of Hawking evaporation. **Second panel**: including the information of observer Hilbert space, the recovery operator $\mathcal{R}$ can reverse the evolution U, and reconstruct Hawking partner P.

And the v coordinate of a black hole with gravitational soft hair implanted is shifted with the opposite amount $-f$,

$$v \to v - f \tag{4.24}$$

as shown in (4.10). Those two shifted phases cancel each other, and we can conclude that the dressed infalling external operators and the black hole soft hair know the phases of each other. In this sense, the outside observers and the black hole soft hair are entangled with each other, which is essential for the reconstruction of interior operators.

Now with the soft hair degrees of freedom included in the system, we can reconstruct the Hawking partner of the late time radiation using those soft hair degrees of freedom as far as the Hilbert space of soft hair is large. Let us first clarify the notations, and then we can explain the basic idea. We denote the black hole Hilbert space as $\mathcal{H}_B$ with its size $d_B = \dim \mathcal{H}_B$, code subspace $\mathcal{H}_C$ with $d_C = \dim \mathcal{H}_C$, the Hilbert space associated with the observers $\mathcal{H}_O$ with $d_O = \dim \mathcal{H}_O$, and similar pattern for early radiation R, late time radiation H, and mirror degrees of freedom P playing the role of the Hawking partner. As discussed in the previous paragraph, S and O purify each other. The basis states of code subspace C and radiation R can be represented as $|i\rangle_C$ and $|i\rangle_R$, the ones for soft hair S and observer O are $|f\rangle_S$ and $|f\rangle_O$, and the ones for H and P are $|n\rangle_H$ and $|m\rangle_P$.

The state we are starting with is

$$|\Psi_0\rangle = \frac{1}{\sqrt{d_C d_S}} \sum_{i,f} |i\rangle_C |f\rangle_S |f\rangle_O |i\rangle_R \tag{4.25}$$

which is diagrammatically shown in figure 4.3. After embedding the state $|\Psi\rangle_0$ into black hole Hilbert space and time evolving the black hole system to emit one late time Hawking radiation H, the wave function of the system can be written as

$$|\Psi(t)\rangle = \sum_{i,f,n} C_n T_f |i\rangle_C \otimes |n\rangle_H |i\rangle_R |f\rangle_O \,, \tag{4.26}$$

where T_f is the embedding tensor mapping the code subspace to the Hilbert space $\mathcal{H}_f$ with a fixed soft hair eigenvalue f, and C_n is the Kraus operator. The evolution is illustrated in the first panel of figure 4.3.

Inspired by the quantum error correction protocol and Petz map, the recovery map $\mathcal{R}$, which reverses the unitary evolution U and reconstructs the code subspace, and the Hawking partner can be worked out. The role of the recovery map is illustrated in the second panel of figure 4.3, which can be represented as

$$\mathcal{R}\, U |\Psi_0\rangle \simeq |\Psi_0\rangle |\mathrm{TFD}\rangle_{HP} \,, \tag{4.27}$$

where $|\mathrm{TFD}\rangle_{HP}$ is the thermo-field double (TFD) made by the Hawking radiation and its partner. The error of the reconstruct $d_c d_H / d_B \ll 1$ is indicated by the sign $\simeq$. The recovery map within a fixed soft hair sector $\mathcal{R}_f$ can be expressed as [175]

$$\mathcal{R}_f |\Phi\rangle_B = \sum_{f,n} \mathcal{R}_{f,n} \otimes |\Phi\rangle_B |f\rangle_S |n\rangle_P \tag{4.28}$$

where $|\Phi\rangle_B$ is the black hole state. $\mathcal{R}_{f,n}$ can be expressed in terms of the density matrix of black hole and observer in a fixed soft hair sector σ_{Bf}

$$\mathcal{R}_{f,n} = \frac{1}{\sqrt{d_C}} T_f^\dagger C_n^\dagger (\sigma_{Bf})^{-1/2} \,, \qquad \sigma_{Bf} = \frac{1}{d_C} \sum_n C_n T_f T_f^\dagger C_n^\dagger \,. \tag{4.29}$$

It can be shown that

$$\mathcal{R}_{f,n} C_m T_f |i\rangle_C \simeq \sqrt{p_n} \delta_{nm} |i\rangle_C |f\rangle_S \,, \tag{4.30}$$

where p_n are the Boltzmann weights. Then we can accomplish our task by the

following procedures

$$\begin{aligned}\mathcal{R}\,U\,|\Psi_0\rangle = \quad & \frac{1}{\sqrt{d_C d_S}} \times \sum_{n,m,i,f} \mathcal{R}_{f,n} C_m T_f \,|i\rangle_C \otimes |f\rangle_O \,|i\rangle_R \,|n\rangle_P \,|m\rangle_H \\ \simeq \quad & \frac{1}{\sqrt{d_C d_S}} \times \sum_{i,f} |i\rangle_C \,|f\rangle_S \,|f\rangle_O \,|i\rangle_R \otimes \sum_n \sqrt{p_n}\,|n\rangle_H \,|n\rangle_P \\ = \quad & |\Psi_0\rangle \otimes |\mathrm{TFD}\rangle_{HP} \ , \end{aligned} \tag{4.31}$$

as shown in figure 4.3.

The fact that the above procedure can be done without AMPS firewall can be reasoned by Yoshida's decoupling theorem [173], which can be expressed as follows: if U is scrambling and the dimension of $\mathcal{H}_S$ is much larger than $\mathcal{H}_H$, i.e., $d_S \gg d_H$, early and late Hawking radiation are decoupled, and one can reconstruct the Hawking partner P without using the early radiation R. The Hawking-Perry-Strominger soft hair introduces an observer-dependent firewall [175], and the infalling observer will never knock into any AMPS firewall before reaching the singularity.

Several key assertions [175] related to the soft hair degrees of freedom are on the horizon with all the above arguments. We summarise those assertions below. The soft hair degrees of freedom are encoded in the transition function (or diffeomorphism) $f(z^A)$; thus, they are classical and measurable properties of black holes, as can be seen from figure 4.1. The information carried by the function f should be regarded as part of black hole entropy. The observers who are sitting at the horizon or asymptotic region cannot determine what f is; thus as shown in figure 4.2 the soft hair is invisible to them. Only by jumping across the intermediate region and comparing the coordinate on both sides, the dressed infalling observer adopts the so-called sharp focus perspective. As a result, the black hole is projected to a state with a more specific value of $f(z_0^A)$ on location z_0^A. The measurement of soft hair would project the black hole onto a soft hair eigenstate with less entropy and reduce the total entropy of the black hole. The measurement enables the reconstruction of the Hawking partner using a set of quantum information protocols as discussed above. Now the entropy of the black hole is lower than before because of this measurement. However, the soft hair is not necessarily completely projected out by one measurement because $f(z^A)$ can be a lot of configurations rather than a global parameter. So we can gradually reduce the total entropy by repeatedly throwing gravitationally dressed operators across the boundary.

Now we have encoded the gravitational soft hair in terms of transition function

shown in equations (4.7)-(4.9), which makes the definition of the Maxwell soft hair very transparent. Although one can largely mimic the gravitational falloff analysis by imposing some falloff boundary conditions on gauge fields near the infinity and claiming that all the gauge transformations consistent with the falloff boundary conditions can be regarded as physical symmetries. Here we just use the gauge field to replace the metric and use gauge transformation to replace diffeomorphism. However, there is a conceptual difference between the isometry that preserves some metric structure (even asymptotic metric) and other diffeomorphisms. This difference is not obvious in gauge theories. Gauge parameters can have arbitrary dependence on spacetime coordinates x^μ, unless the canonical analysis tells us which of them are physical and which are not. Understanding the whole story by looking at the transition function and getting an effective action for those would-be-gauge degrees of freedom seems more straightforward. Moreover, the transition function interpretation has a great potential to be generalised to a finite distance away from the horizon. So the strategy for Maxwell's theory is to regard the Maxwell soft hair as the transition function that compares gauge fields living in different regions. We will explain more details in the next section.

4.3 Maxwell Soft Hair

In this section, we adopt the language reviewed in the previous section, where we treated gravitational soft hair as a transition function between the asymptotic region and the near-horizon region of the global black hole spacetime. Whereas here, we mainly focus on U(1) gauge theory living on a black hole background and characterise the Maxwell soft hair in terms of a transition function of gauge fields between those different regions. The main difference is that the show's leading role is the U(1) gauge transformation $\lambda(x^a)$ on the boundary x^a, rather than supertranslation. The virtue of Maxwell's theory, in addition to being easier to be handled both conceptually and computationally, is that we can relate it with U(1) edge modes and have an effective Lagrangian description of λ, which can help us to estimate the size of phase space of Maxwell soft hair. Once we know the size of phase space, we are enabled to do more analysis on what soft hair can do to help us understand BHIP as an example, which will be the central subject of the next section.

4.3.1 Maxwell soft hair as a transition function

Similar to the gravity story, we divide the asymptotically flat Schwarzschild black hole into two regions, namely the near-horizon region $\mathcal{M}$ and the asymptotic region $\bar{\mathcal{M}}$. At the place where those two regions meet each other, we interpret the Maxwell soft hair as a transition function that uniformises the gauge fields A_μ and $\bar{A}_\mu$ living in $\mathcal{M}$ and $\bar{\mathcal{M}}$ respectively, as shown in figure 4.4. In this sense, one can think there is a boundary, denoted as $\partial\mathcal{M}$, between the two regions, and the Maxwell soft hair can be regarded as the edge states living on the boundary [136, 137]. One can imagine this boundary as some kind of "Dyson sphere" outside the black hole, and we will mainly adopt the perspective of the observer living inside the Dyson sphere and regard the soft hair degrees of freedom as vacuum degeneracy in the sky. This perspective is shown in the first panel of figure 4.2; as the observer looks up into the sky, the U(1) soft hair is labelled by gauge parameter λ on the boundary.

Before taking any further step, let us first set up our physical background. We are mainly interested in the U(1) gauge theory with action

$$S = -\frac{1}{4}\int_{\mathcal{M}} d^4x\sqrt{-g}\ F^{\mu\nu}F_{\mu\nu}\,. \tag{4.32}$$

The theory is put on an asymptotically flat black hole spacetime with metric

$$ds^2 = -(1-\frac{2GM}{r})dt^2 + (1-\frac{2GM}{r})^{-1}dr^2 + r^2 d\Omega_2^2\,. \tag{4.33}$$

The spacetime is divided into the near-horizon region $\mathcal{M}$ and the asymptotic region $\bar{\mathcal{M}}$. The path integral on this manifold can be written as

$$\int [DA_\mu][D\bar{A}_\mu] \prod_{x\in\partial\mathcal{M}} \delta(A_\mu - \bar{A}_\mu)\ e^{i(S_A + S_{\bar{A}})}\,. \tag{4.34}$$

The gauge fields living on the sides are supposed to match each other. That is the reason why we introduce a boundary delta function in the path integral. The boundary delta function can be dealt by introducing Lagrange multiplier fields on the boundary

$$\prod_{x\in\partial\mathcal{M}} \delta(A_\mu - \bar{A}_\mu) = \int [DJ^\mu] e^{i\int_{\partial\mathcal{M}} J^\mu (A_\mu - \bar{A}_\mu)}\,. \tag{4.35}$$

Lagrange multipliers J^μ can be arbitrary functions on the boundary, which will be integrated over eventually. This part can be regarded as the duplication of coordinate transformation in equation (4.6). For the observer living in the

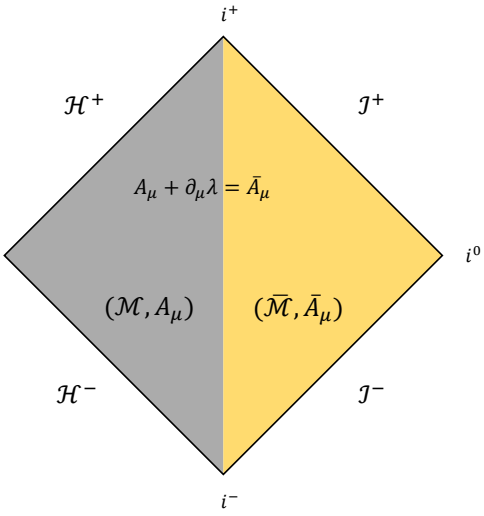


Figure 4.4: Maxwell soft hair degrees of freedom are encoded in the gauge transformation parameter λ between the gauge fields living in two different regions.

near-horizon region, the original gauge theory can be gotten by getting rid of the contributions from $\bar{A}_\mu$. A_μ should be fully gauge fixed in bulk $\mathcal{M}$ for a physical observer, and we can write the gauge field as $A_\mu - i\partial_\mu\lambda$ on the boundary to reserve the ambiguity to match the fields between two sides; hence λ should be interpreted as the transition function that connects the gauge fields living in those two regions. This gauge parameter λ along the boundary has also been interpreted as U(1) edge modes [136, 137]. One can also perform symplectic form analysis to show those modes λ are indeed physical [211].

Following the same logic as the supertranslation case, we interpreted the gauge parameter as the transition function that relates the Maxwell gauge fields between two regions. Thus we can say that λ at the boundary provides a good label of Maxwell soft hair just as supertranslation parameter f did for gravitational soft hair. Moreover, we also provide a concrete relation between the Maxwell soft hair and the would-be-gauge edge states living on the boundary. This relation was also mentioned in [137].

The Maxwell soft hair should have the same seven assertions as supertranslation soft hair discussed by PV. Those properties are essential for soft hair to shed light on the BHIP. The Maxwell soft hair degrees of freedom should

also be classical and physical effects of the global black hole, which are invisible for observers restricted in the asymptotic region or the near-horizon region. Maxwell soft hair can be measured and is exponentially sensitive for an infalling observer who crosses the intermediate region. Those soft degrees of freedom carry a large amount of entropy and can be projected to a lower entropy state by measurement done by the dressed in-falling particles. We will use those assertions for Maxwell's theory as the key ingredients to argue the Page curve in the next section.

4.3.2 Effective field theory of Maxwell soft hair

Now we have interpreted the Maxwell soft hair as a transition function that relates the U(1) gauge fields living in different regions. Moreover, we have built the connection of Maxwell soft hair and the U(1) edge modes living on the intermediate boundary. The virtue of Maxwell soft hair is that we have an effective action description of those edge modes. With the help of the action, we can do some quantitative analysis of the phase space of Maxwell soft hair $\lambda(x^a)$.

Restricted to the near-horizon region $\mathcal{M}$, the effective action for A_μ can be written as

$$\begin{aligned} S_A &= -\frac{1}{4}\int_{\mathcal{M}} d^4x\sqrt{-g}\; F^{\mu\nu}F_{\mu\nu} \\ &\quad + \int_{\partial\mathcal{M}} d^3x\sqrt{-h}\; J^\mu(A_\mu - i\partial_\mu\lambda)\,, \end{aligned} \tag{4.36}$$

Here J^μ are the Lagrange multipliers introduced in equation (4.35) and will be integrated over later. In order to separate the effects from edge modes and bulk modes, we are going to separate A_μ into two parts

$$A_\mu = \tilde{A}_\mu + B_\mu\,, \tag{4.37}$$

where $\tilde{A}_\mu$ vanishes at boundary, and B_μ is on shell in the bulk and takes the same boundary value as A_μ. Then the effective action can be separated into two parts

$$\begin{aligned} S_A &= S_{\tilde{A}} + S_B \\ &= -\frac{1}{4}\int_{\mathcal{M}} d^4x\tilde{F}^{\mu\nu}\tilde{F}_{\mu\nu} \\ &\quad + \int_{\partial\mathcal{M}} d^3x\sqrt{-h}\left[-\frac{1}{2}n_\nu F^{\mu\nu}_{(B)}B_\mu + J^\mu(B_\mu - i\partial_\mu\lambda)\right], \end{aligned} \tag{4.38}$$

where n^μ is the normal vector orthogonal to the boundary and $F^{\mu\nu}_{(B)}$ is the field strength calculated from B_μ. The first part $S_{\tilde{A}}$ that captures the bulk fluctuation contribution is not of interest here. The second part

$$S_B = \int_{\partial\mathcal{M}} d^3x\sqrt{-h}\left[-\frac{1}{2}n_\nu F^{\mu\nu}_{(B)}B_\mu + J^\mu(B_\mu - i\partial_\mu\lambda)\right] \tag{4.39}$$

captures the interesting physical effects and will give out the effective action for Goldstone modes λ. B_μ is on shell in the bulk $\mathcal{M}$, and thus we can solve B_μ in terms of boundary current J^μ. Putting back the solution $B_\mu[J^\mu]$ means that the effective action S_B can be written as a functional of J^μ and λ. To do that, one needs to solve the following bulk problem

$$\nabla_\mu F^{\mu\nu}_{(B)} = 0\,, \qquad \text{with} \qquad n_\nu F^{\mu\nu}_{(B)}\big|_{\text{bdy}} = J^\mu\,. \tag{4.40}$$

Until now, we have not picked any specific gauge fixing condition. No matter what gauge fixing condition we pick for B_μ in bulk, $\tilde{A}_\mu$ should use the same gauge fixing condition. Here we are going to let $B_r = 0$ (i.e. $A_r = 0$) as our bulk gauge fixing condition. Choosing $r = L$ as our the boundary, the boundary condition in equation (4.40) is

$$\partial_r B_a\big|_{\text{bdy}} = J_a\,, \qquad 0 = J_r\,, \tag{4.41}$$

where x^a is the coordinates along the boundary. Working with our gauge fixing condition, one can immediately see that after variable separation, B_a always take the form of

$$B_a = f(r)\cdot J_a\,. \tag{4.42}$$

$f(r)$ can be some very complicated function of r determined by the bulk equation of motion, with $\partial_r f(r)\big|_{r=L} = 1$. This is already enough information for us to get the effective action for Goldstone $\lambda(x^a)$. Putting (4.42) back into the action (4.39), we get

$$S_B = \int_{\partial\mathcal{M}} d^3x\sqrt{-h}\left(\frac{f(L)}{2}J^aJ_a - iJ^a\partial_a\lambda(x^a)\right)\,. \tag{4.43}$$

Functional integrating out $J^a(x^b)$ in the path integral, one gets an effective action for $\lambda(x^a)$, which is read as

$$S[\lambda] = -\frac{1}{2f(L)}\int_{\partial\mathcal{M}} d^3x\sqrt{-h}\;\partial^a\lambda\partial_a\lambda\,. \tag{4.44}$$

The action describes a three-dimensional massless scalar field living on the boundary $r = L$ with a coupling constant that varies with the location of the boundary. This is the effective action description of Goldstone modes $\lambda(x^a)$.

Now we are good to analyse how large the phase space of Maxwell soft hair is. We analyse the statistical properties of $\lambda(x^a)$ by performing a Euclidean path integral at finite temperature. The partition function for λ can be written as

$$Z_\lambda = \int [D\lambda]\ e^{-S_E[\lambda]} \tag{4.45}$$

where the Euclidean action is

$$S_E[\lambda] = \frac{1}{2f(L)} \int_0^\beta d\tau \int d^2x\sqrt{h}\ \partial^a\lambda\partial_a\lambda\,, \tag{4.46}$$

with inverse temperature β as the periodicity of τ direction. This is nothing strange, and the result is well known. The partition function for fluctuation modes of Maxwell soft hair at large L is

$$Z_F = \prod_{n,\mathbf{k}} \left[\frac{\pi}{\beta^2(\omega_n^2+\mathbf{k}^2)}\right]^{1/2}\,, \tag{4.47}$$

where ω_n and $\mathbf{k}$ are Fourier modes along τ and spatial directions. The free energy and entropy from this part are read as

$$\begin{aligned} F &= -T\ln Z_F = -\frac{\zeta(3)}{6\pi}T^3L^2\,, \\ S_F &= -\frac{\partial F}{\partial T} = \frac{\zeta(3)}{2\pi}\frac{L^2}{\beta^2}\,. \end{aligned} \tag{4.48}$$

Note that L is taken to be large compared to the horizon's scale or inverse temperature here, mainly for two reasons. The first one is that, when we are doing asymptotic analysis in the gravity case, we have taken $1/r$ to be small, and the diffeomorphisms we care about are the ones that preserve the asymptotic metric. The extra degrees of freedom are living on the so-called celestial sphere. Here the gauge theory is supposed to have a similar property, and L is taken to be the radius of the celestial sphere. Second, in order to calculate the partition function of λ, we put those modes in a (2+1)-dimensional box with finite temperature $1/\beta$. The background topology $S^1\times S^2$ is taken care of by periodic boundary conditions. Then this is the standard thermal field theory for a scalar field,

$$\ln Z = -\sum_k [\frac{1}{2}\beta k + \ln(1-e^{-\beta k})]\,, \tag{4.49}$$

with $k=\sqrt{\mathbf{k}^2}$. In order to change the sum into an integral, we need to take a large volume limit. $L\gg\beta$ also justifies that we put the scalar field in a large box.

There are also constant modes or topological modes contribution to the entropy. The scalar field λ is compact; thus, the path integral over the constant modes gives out a partition function proportional to the perimeter of the circle up to a normalisation factor. The corresponding entropy proportional to the logarithm of the partition function may dominate when the power-law contribution from thermal modes becomes less important at low temperature. Also, because we are working on a Euclidean background with compact time direction, the fundamental group element of S^1 can be added into the path integral as winding modes. Those topological modes give out theta function contribution in the partition function. Again, both contributions from constant modes and topological modes can become important contributions in the low-temperature limit. They may contribute to the logarithm correction of black hole entropy in the near extremal black hole case. Nevertheless, those contributions are not essential for our discussion here. The size of the Maxwell soft hair phase space is more or less proportional to

$$k \propto e^{S_F} = \exp\left(\frac{\zeta(3)}{2\pi}\frac{L^2}{\beta^2}\right) . \tag{4.50}$$

Note that k changes as the temperature of the spacetime varies. As the temperature of the spacetime becomes higher and higher, k can be very large. In other words, the phase space of soft hair increases as more and more Hawking radiation happens. The variation of k with temperature is the crucial ingredient for our argument to understand the BHIP, which we will see in the next section.

We do not have an excellent angle to argue what L is in this case. According to asymptotic analysis, L should be a large distance cutoff, and the boundary should be regarded as the celestial sphere. Now the transition function prescription in gauge theory does not really care about the falloff conditions, and what we need is just a boundary between these two regions. In this sense, the boundary can be moved to a finite distance away from the horizon. This is the situation that we did not go into too many details and restricted ourselves to large L. But it is interesting to consider more about this case. The scale of L is essential in determining the phase space of soft hair. The situation is not like the extremal black hole case where we have a natural length scale to characterise the boundary between $\mathrm{AdS}_2 \times S^2$ throat and asymptotic region. In principle, L should be determined by length scales G_N and M, or even the history of the evaporation, which means L might vary with time. The point we want to make here is, at the end of the evaporation, where the temperature

for the tiny black hole is very high, k should always be large enough to play a significant role.

4.4 Page Curve from Soft Hair

Having an effective action of Maxwell soft hair at hand, we estimated the size of phase space in the previous section. In this section, we address the problem of how soft hair of the black hole gives rise to one version of "Page curve", which might not be the same as the one originally proposed by Page [160], but certainly does not violate the bound for von Neumann entropy during evaporation. Note that here the Page curve is gotten from a more microscopic perspective. The basic idea is that we should at least allow for two kinds of processes on the black hole background, namely, Hawking radiation and measurement. The measurement is done by a global infalling observer across the boundary. The competition between those two processes, the rate of which is proportional to the size of phase space, will eventually give out the Page curve.

4.4.1 Two types of processes

In general, we should consider two types of physical processes on a black hole background. The Hawking radiation process is the well-understood one, which increases black hole entropy by creating the entanglement between the Hawking radiation and interior partner. There should also be measurement processes, which can be done by U(1) dressed infalling observers. As argued by PV, the measurement of the soft hair f or λ enables the reconstruction of Hawking partners, and they also demonstrated the possible quantum information protocol of the reconstruction using measurement of soft hair. There are still subtleties related to the exchange between code subspace and soft hair subspace, which is not completely clear. However, naively, one can always say that the comparison of f (or λ in the Maxwell case) between two regions projects the black hole onto a given soft hair configuration, a lower entropy state. The entropy of the black hole is reduced once the outside observer in figure 4.2 knows more information about what is f or λ. The soft hair space of the black hole is projected onto a lower entropy state by repeated measurements. From now on, we will say that the measurement can be done by any infalling dressed particle that crosses the intermediate region, and such measurement reduces the entropy of the black hole.

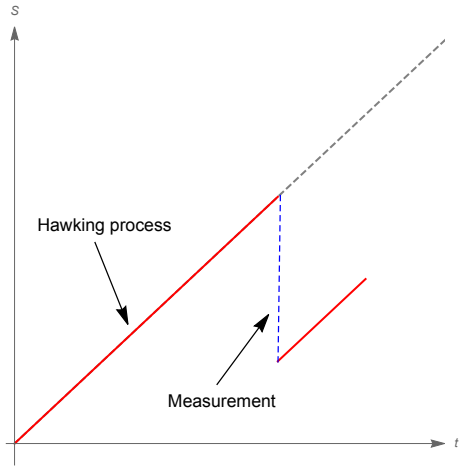


Figure 4.5: The entropy of black hole is increased by Hawking radiation and decreased by measurement that project the black hole onto states with given soft hair configurations.

The entropy of the black hole is depicted in figure 4.5. As usual, we always assume we are starting with a pure state black hole. The entropy of the black hole is increased by Hawking radiation, as shown by the red lines in the figure. The more entanglement between the black hole and Hawking radiation is created, the more entropy the black hole has. This is precisely Hawking's paradox, which confuses people by saying that the process seems to continue forever, and the entanglement entropy can be even larger than the black hole thermodynamics entropy. The measurement that reduces the entropy of the black hole are shown by the blue dashed line, which might be invisible to Hawking's argument, but will be the most important new ingredient for the BHIP. The argument in this section is still very qualitative. A natural question is how much do the measurement reduce the black hole entropy, which requires a competition between the rate of Hawking radiation and soft hair measurement.

4.4.2 Fermi's golden rule

A more accurate version of counting the rate of particle scattering should rely on black hole S-matrix [235–237]. There is some recent progress on black hole S-matrix related to soft particle dressing [228–230]. However, for our purpose here, it is already enough to use a very modest version for two-level systems,

i.e., "Fermi's golden rule" [238, 239]. The Fermi's Golden rule means that we can approximately treat Hawking radiation and infalling particles across the boundary at a given energy as a two-level system. By virtue of Fermi's golden rule, the rate of transition w can be written as

$$w = \langle \mathcal{T} \rangle^2 \rho(E) \,, \tag{4.51}$$

with $\mathcal{T}$ as the scattering matrix of the process, and $\rho(E)$ as the density of state of energy E. We can use the formula (4.51) to calculate the rate of Hawking radiation and the measurement process. Then the whole problem is reduced to a density of state counting, assuming that $\langle \mathcal{T} \rangle$ is more or less of the same order.

The intuitive way of understanding Fermi's golden rule is following. For the black hole with mass M emitting a Hawking quantum with energy δM and then going to a lower energy state, we can naturally portray it as a two-level system. Of course, the rate for this process depends on the density of state of the given black hole state. The more states we have at the given energy of the black hole, the bigger chance Hawking radiation can happen. Similarly, for the infalling particles waiting to cross the boundary, one should think this process is a domain-wall-crossing process, which can also be portrayed as a two-level system. More degeneracy at a given temperature means one has more choices from to choose, and it is easier to make the deal.

Let us first look at the Hawking radiation process. The density of state can be written as an exponent of the Bekenstein-Hawking entropy $e^{S_{BH}}$. Then the rate of a black hole with mass M emitting particles at the horizon can be written as [2]

$$w_H = \langle \mathcal{T}_H \rangle^2 \, e^{S_{BH}(M)} \,. \tag{4.52}$$

where w_H is the rate of our first physical process Hawking radiation. For the second process, the phase space of degeneracy at a given temperature was calculated in the previous section, where we denoted it as

$$k \propto \exp\left(\frac{\zeta(3)}{2\pi} \frac{L^2}{\beta^2} \right) . \tag{4.53}$$

Then the rate of measurement can be written as

$$w_M = \langle \mathcal{T}_M \rangle^2 \exp\left(\frac{\zeta(3)}{2\pi} \frac{L^2}{\beta^2} \right) . \tag{4.54}$$

We need more details of the interacting Hamiltonian to calculate the scattering matrix. For now, let us assume the overall scattering amplitude of those two

processes is comparable with each other. Then one can compare the rate of those processes just based on the size of phase space.

Those two densities of states are both exponents of entropy. At the early stage of Hawking radiation where we have a low-temperature black hole, the rate of measurement can be minimal. In this period, Hawking radiation plays a dominant role. whereas as the temperature gets higher and higher, it is possible to have

$$k > e^{S_{BH}} . \tag{4.55}$$

Then the measurement process, which decreases the entropy of the black hole, plays a significant role. The measurement of the soft hair does not mean the possibility of Hawking radiation being reduced; the entropy of the radiation still goes as Hawking's calculation. That only means we can have a large amount of "invisible" processes that may reduce the entropy of the black hole by measuring the soft hair degrees of freedom.

4.4.3 Page curve

In this subsection, we draw the Page curve of black hole entropy by including the physical effects of measurement of soft hair.

Hawking's calculation of entanglement pair creations should not be changed too much and go until the end of the evaporation. The new ingredient is the measurement process.

The central insight we have from previous subsections is that the phase space of this process increases with time. Without an expression for L at hand, let us assume linear growth of soft hair phase space with time for simplicity [1]. Again for simplicity, let us assume the Hawking radiation is emitted at a constant rate, and one infalling particle crossing the intermediate region does not change the energy of the black hole but reduces the entropy of the black hole by one unit. Now the measurement rate also has a linear growth in time. The Page curve of linear-growth measurement phase space is illustrated in figure 4.7. The red line is continuously growing with time because of Hawking

[1] The explicit phase space always depends on the choice of ensemble. Nevertheless, we can work out the time dependence of density of states by ignoring the time dependence of L and only considering the temperature change because of Hawking radiation. Then the density of state of soft hair is proportional to $\exp(T^2(t))$. If the Hawking radiation $dM/dt = -\alpha/M^2$ with constant α, the shape of measurement rate can be depicted in figure 4.6, which means that the modification from the global effects is relatively low at the beginning of the radiation, then becomes greatly enhanced at late times.

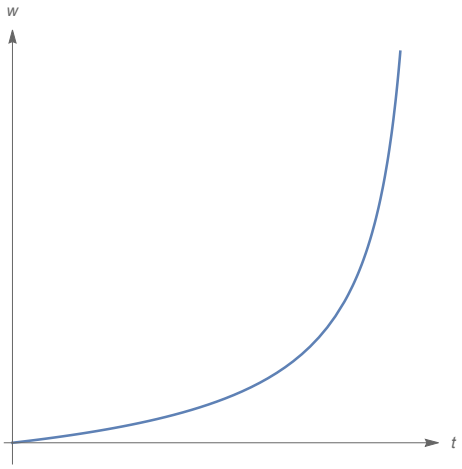


Figure 4.6: A sketch of the rate of measurement w under the assumption of constant L during the evaporation.

radiation creating entanglement. The measurement of soft hair decreases the entropy as shown in figure 4.5; that is the reason for discontinuity in the first panel of figure 4.7. In the second panel of figure 4.7, we take smaller steps in numerical approximation, which gives a more "smooth" curve.

One needs to be aware that those curves are not perfect because of the simplicity assumption we have taken, and if the time dependence of measurement looks like the curve shown figure 4.6, the Page curve will look different. The modification of Hawking's result is relatively small at the early time, so the curve more or less follows a straight line. At the late time, the entropy of the black hole gains a more significant modification from this global measurement, then the entropy quickly drops to zero. Then the curve is more or less the Page curve one expected, and those curves are all consistent with the unitary evolution of the black hole.

Note that our derivation of the Page curve has no contradiction with Hawking's calculation. All the red lines in figure 4.7 are precisely coming from Hawking's calculation. We are just saying that there are physical effects invisible to Hawking and missed by Hawking's calculation. By adding those global effects back to the black hole, the entropy of the black hole can be extensively reduced, thus there is no violation of unitarity during the evaporation. That means that, to get a convincing Page curve, one might not need a full quantum theory of

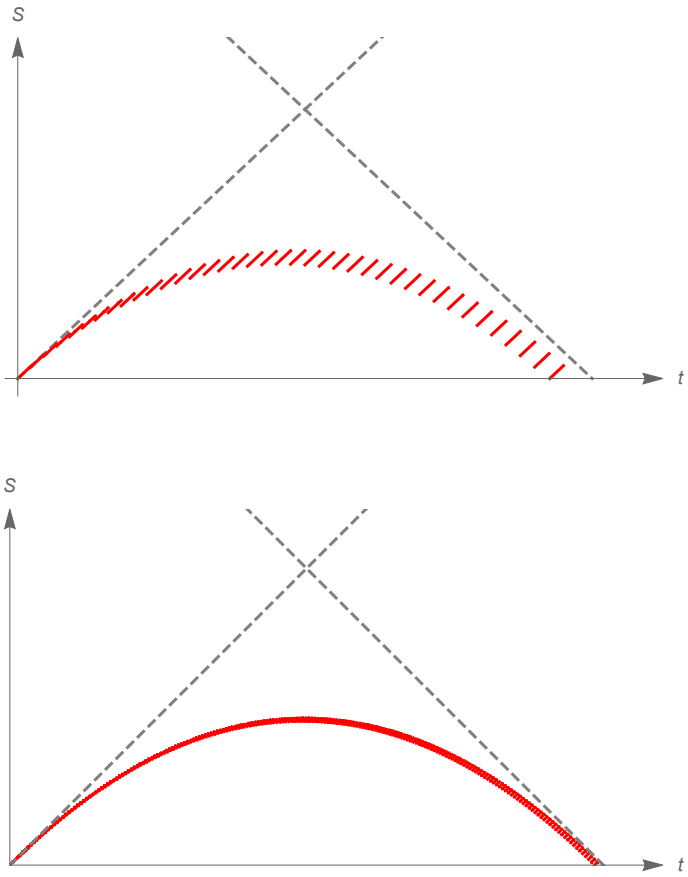


Figure 4.7: Page curve of black hole entropy with the effect of measurement of soft hair being included. The phase space of soft hair is assumed to have a linear growth with time. The first figure is a coarse-grained version, whereas in the second figure the Page curve is "smoothed" by taking smaller steps in time.

gravity, but global effects are indeed needed. The right way to derive a Page curve for radiation, in addition to just saying the whole system is pure, needs further study.

The Page curve we have shown in the second panel of figure 4.7 looks like a smooth curve, rather than the usually expected phase transition at Page time t_{page}. Actually, this curve is not a smooth but a wavy curve, which means that if one enlarges each point there is the same pattern as shown in figure 4.5. The wavy curve is because there are many phase transitions at each measurement, as demonstrated in figure 4.5. It is worth mentioning that this enormous amount of phase transitions is not strange, but getting more and more popular recently because of the island prescription. It has believed that in the calculation of the entropy of black hole or radiation using Euclidean path integral, there are also a large number of other saddles between the fully connected and disconnected geometries [190]. So there should not be only one phase transition, but many relatively small phase transitions between different saddles. Here we show the similar effects in our Page curve shown in figure 4.7.

4.5 Conclusion and Discussion

Following the same strategy as [175], we treated the Maxwell soft hair degrees of freedom as a transition function that relates the Maxwell gauge fields in the asymptotic region and the near-horizon region. We introduced U(1) gauge field A_μ living in the near-horizon region and $\bar{A}_\mu$ in the asymptotic flat region, which are related by a gauge transformation $\bar{A} = A + \mathrm{d}\lambda$ on the boundary between those two regions. The transition function λ, i.e. would-be-gauge degrees of freedom, is regarded as the Goldstone modes that characterise the Maxwell soft hair. The Maxwell edge modes and soft hair are naturally connected.

The advantage of the U(1) gauge case is that an effective action description of those soft hair degrees of freedom can be easily obtained. By separating the bulk fluctuation modes that vanish on the boundary and boundary soft hair modes that are on shell in bulk, we get the effective action for soft hair which is just the one for massless scalar field

$$S[\lambda] = -\frac{1}{2f(L)} \int_{\partial\mathcal{M}} d^3x\sqrt{-h}\ \partial^a\lambda\partial_a\lambda\,. \tag{4.56}$$

Thus one can use the Euclidean path integral in terms of the effective action to study the statistical properties of the Maxwell soft hair. The entropy gotten

from the path integral is roughly $S_{\text{soft}} \propto L^2/\beta^2$. The phase space of soft hair increases as the temperature of the black hole becomes higher and higher.

There should be two basic processes on the black hole background: Hawking radiation and measurement of soft hair. The measurement of soft hair can be done by the infalling observer who measures the U(1) phase difference between two sides. Those global effects are invisible to local observers. It was shown in the paper by PV that the measurement implements a code subspace projection that enables the reconstruction of interior operators. Fermi's golden rule states that, for two-level systems, the rate of changing energy level is proportional to the density of states at each level. At a given energy, we can regard the Hawking radiation and boundary-crossing process as two-level systems. The competition of the size of phase spaces would finally give rise to the Page curve. At the early stage, the Hawking radiation has a large phase space, and thus they dominate. While at the late stage, the boundary-crossing measurement becoming important and thus reduce the entropy of the black hole. One version of the Page curve is shown in figure 4.7.

Our results suggest that soft hair as a global effect might play an essential role in the unitary evolution of black hole and be powerful enough to give the Page curve of black hole evolution. Is Hawking's calculation of radiation valid until the last breath of the black hole? The answer might be yes, and even so, there is no contradiction with the Page curve.

Understanding the exact location of the boundary between the near-horizon region and asymptotic region is essential in determining the phase space of soft hair. The critical question for the brave infalling observer jumping into the black hole is when will the phase difference between the asymptotic region and the near-horizon region start to be felt. We tend to believe L is an infrared cutoff at a large distance because of falloff analysis of asymptotic symmetry [211]. One might argue why crossing a boundary at a large distance can do anything to the black hole. This question might be answered by ER=EPR or emphasising that those are global effects of the black hole.

All the arguments above are based on the assumption that measurement of the soft hair enables the reconstruction of Hawking partners. Although we have seen some evidences of such reconstruction [173, 175], there is no clear argument about how the exchange between the code subspace and soft hair subspace could happen. In other words, if the information of the Hawking partner is stored in the soft hair degrees of freedom, which is invisible for local

Hawking calculation, what is the physical process of information exchange between Hawking radiation and soft hair? Does this happen when the partner crosses the horizon or when the Hawking radiation crosses the boundary $r = L$? Such questions are always the critical ingredients of understanding the problem, which might be very hard to answer within the current framework.

It would be very interesting to understand the connection between the soft hair story and island prescription. One clue of the relation is how to understand the replica wormholes arising from Euclidean path integral derivation of island rule [190, 196]. Should those replica wormholes be understood as some kind of domain-wall-crossing process, it might be easier to build the connection. They indeed share a lot of similarity in terms of phase space counting to decide which process dominates, reconstruction of the interior partner of Hawking radiation, and so on. Another interesting question to understand is that, if the soft hair provides extra parameters labelling superselection sectors, does swampland have something to say about those "soft symmetries"?

5 Emergent Gravity

One of the reasons that make the black hole such a seductive object is its thermodynamics. It is important to study the relationship between gravity, thermodynamics and quantum information. It follows that the whole concept of gravity might be not fundamental but emergent from thermodynamics or quantum information. In this chapter, we are going to explore those ideas further.

We clarify the problem in which occasions can gravitational force be regarded as emergent from thermodynamics, by proposing an entropic mechanism that can extract the entropic gradient existing in spacetime, due to the variation of the Casini-Bekenstein bound in specific quasi-static processes with the heat flux δQ into the whole casual wedge. We explicitly formulate the derivation of inertial force as the emergent gravitational attraction from the entanglement first law. The current entropic mechanism reproduces Newton's second law in Rindler space and the gravitational force (together with derivation of the Einstein equation) beyond the near-horizon region, and can be adapted to AdS/CFT and other generic situations.

5.1 Introduction

Gravitational force is special, whose origin may be approached in a completely different way from other kinds of fundamental forces which have been quantised and unified. Spacetime and gravity have been regarded as emergent phenomenon from microscopic degrees of freedom of quantum field theory (QFT), an insight from the developments of string theory and loop quantum gravity, the two potential candidates of quantum gravity. Question is aroused if the gravitational attraction reflects a fundamental tendency of information or not.

From the AdS/CFT correspondence [240], early attempts [241–243] show entanglement builds spacetime geometry, in the sense that the connection and continuity of spacetime geometry are closely related to the entanglement structure of QFT states. The idea of entanglement generating spacetime [243] then leads to the conjectures of $A = R_B$ [244], and then ER=EPR [245]. They were proposed to save the principle of equivalence against the firewall paradox in AMPS [246] argued from the monogamy of entanglement. While a general rule holds for any quantum system, the entanglement first law is then applied to gravity, and leads to breakthrough results, the derivation of Einstein's equation from AdS/CFT to linearised level [247, 248] as well as to non-linear level [249], and Jacobson's new derivation [250] of Einstein's equation based on *maximal vacuum entanglement hypothesis*. However, since those developments are based on vacuum entanglement, they are not equivalent to explain the tendency of gravitational attraction. One should apply this entanglement first law for perturbing excited states to reconsider old questions put up in Verlinde's emergent gravity theory.

One decade ago, Verlinde remarkably attached information meaning to gravitational attraction through the entropic force conjecture [251]. The basic idea of Verlinde's emergent gravity theory [251, 252] is that the gravitational force is possibly an entropic force $F = T\nabla S$ that usually occurs in macroscopic systems such as colloid and polymer molecules, with the entropy gradient from a variation of "holographic screen" generally existing in spacetime. In this way, the theory attempts to explain the falling tendency of Newton's apples as an entropy-increasing tendency of the thermodynamic second law.

However, this theory requires either subtle improvement or modification because the existence of the entropic gradient remains unclarified. Furthermore, we still do not know if such entropic mechanism exists since the possibility of $dW = dE$ with no entropy varying $dS = 0$, such as the querying in

[253]. Through calculating the back-reaction to the geometry, the conjecture is tested only to hold in near-horizon regions in [254] for area variation of the holographic screen providing too much entropy. So far, it could not explain the gravitational attraction in generic situations, for the original holographic screen approach fails to be generalised beyond the near-horizon region, and it cannot be applied to the sun nor planets. How to interpret gravitational force from an entropic mechanism remains a mystery.

The argument continued recently, after Susskind proposed a new alternative description from *complexity tendency*, together with a query of the explanation ability of the entropic mechanism for the oscillating movement in pure AdS. Whether the gravitational force can be interpreted from an entropic mechanism even becomes a question.

The resolution of the snags could be simple. To match the local gravitational force $F^g_\mu = -\frac{G_N M m}{r^2(1-\frac{2G_N M}{r})}\delta^r_\mu$ derived for Schwazchild solution in general relativity (GR) (see textbook [255, 256]), former research [257] happened to work out in a simple single-mode thermal harmonic oscillator model. It suggests that the entropic change is exactly the variation of Casini-Bekenstein bound [258]. That work was inspired by the observation that the process of a static observer lifting/lowing a box through a long string in Bekenstein's famous thought experiment for the generalised second law [259] is indeed a quasi-static process. It is like the processes that a heat engine endures during the Carnot cycle, after considering the thermalisation of the box by the local Hawking temperature. Therefore this non-unitary process changes the entanglement entropy within the casual wedge, then causes heat flow $\delta Q = T\delta S$ into the exterior region of the black hole, through external influence by the long string.

What makes a difference is to include the thermalisation by the local Hawking temperature $T = T_H/V$ to replace the usual rule of the Unruh temperature in entropic gravity theories to the box regarded as an excited state confined in the subsystem. Since Hawking/Unruh effect happens to different static observers related by the redshift factor V, the entropic gradient comes out along with the temperature gradient when the string slowly moves the box. Then to calculate inertial force, one should adopt the entanglement first law involving excited states modular hamiltonian if we consider the entanglement entropy during this process.

This chapter will illustrate this entropic mechanism and show that it can be developed to explain gravitational attraction in generic situations.

Main Results The primary problem to solve is to find what causes the entropy variation; then, we may be able to calculate how much should such variation be in general. More specifically, we should find which well-defined entropy is necessary for the gravitational force and which thermodynamic process.

Based on the positivity of relative entropy, Casini proved a more concise version of the Bekenstein bound [258] for any relativistic QFT

$$\Delta S \leq \Delta \langle K \rangle \ . \tag{5.1}$$

which is related to the modular Hamiltonian K and the entanglement entropy. The proof is for the Rindler space of Minkowski spacetime but also applies to eternal Schwarzschild black hole that has Hartle-Hawking states as its vacuum. Our main derivation in this chapter is also on these two cases.

This entropy bound is indeed saturated generally in some occasions for infinitesimal perturbation of vacuum, as later tested in [260] the saturation of the bound (5.1) to the first-order variation in the AdS/CFT framework. Also, Dvali recently showed that the saturation of universal entropy bounds is also related to the unitarity of scattering amplitudes [261].

In this chapter, under the semi-static process to extract gravitational force by fixing local measurement of $\Delta \langle H \rangle \to m$ for nearby static observers, we show the saturation of the entropy bound

$$\Delta S = \frac{\Delta \langle H \rangle}{T} \,, \tag{5.2}$$

leads to an entropic gradient generally

$$\nabla_\mu S = \frac{m}{T_0} \nabla_\mu V \,, \tag{5.3}$$

where m is the mass of the test particle and $V = e^\phi$ is the redshift factor with respect to the general gravitational potential ϕ, while T standing alternatively for the local measurement of the Unruh temperature or the Hawking temperature for static observers along with $T = T_0$ for $V = 1$. It proves the necessity of external force for an entropic mechanism.

The covariant version of the external force (necessary to balance the gravitational force) is emerged directly, from the entropic force formula

$$F_\mu = T \nabla_\mu S \,, \tag{5.4}$$

rather than the gravitational force as the inertial force in Verlinde's original theory.

While this entropic force formula is no longer a macroscopic effect after involving fine-grained entropy bound, it is just an approximation of a more general modular Hamiltonian approach we develop. Indeed, the true derivation of the inertial force as emergent gravity actually comes from utilising the entanglement first law to get a work term

$$dW_g = -d\langle O\rangle_1 \ , \tag{5.5}$$

where $O = K_1 - K_0$ is the difference between the modular Hamiltonian of excited states and vacuum states. Here, we further prove that when the bound is saturated, the resulted inertial force does not dependent on the detail of O. The variation of Casini-Bekenstein bound in such quasi-static processes will naturally reproduce Newton's second law in Rindler space and local gravitational force for Schwarzschild black hole.

After this, we reach the core topic, which is to find a holographic interpretation for the gravitational attraction since the saturation of this bound is a condition of holography. Noticing the connection between the saturation of the Casini-Bekenstein bound and the first law of black hole thermodynamics, we interpret the entropic gradient holographically as

$$\nabla_\mu S = \nabla_\mu \left(\frac{\delta A(\Sigma_{r_s})}{4G_N} \right) \ , \tag{5.6}$$

corresponding to the variation of horizon area $\delta A(\Sigma_{r_s})$ as extremal surface, rather than the variation $\delta A(\Sigma_r)$ of the holographic screen at r, which would otherwise provide too much holographic entropy. This holographic interpretation is covariant, and corresponds to the Bousso bound in [262, 263] (reviewed in [264]).

Structure of the Content The structure of the chapter is as follows.

In Section 5.2, we set our stage by reviewing Casini-Bekenstein bound for bipartite systems. Then we show how the entropic gradient raises and reproduces results matching GR.

In Section 5.3, we further develop the techniques to derive inertial force utilising the entanglement first law, which is more rigorous and compare it with the derivation from the entropic force formula. Then we introduce our new holographic interpretation for the entropy change to explain gravitational force,

noticing the connection between the upper entropy bound and the first law of black hole thermodynamics.

In section 5.4, we summarise the whole chapter and provide some further discussions.

5.2 Entanglement and Thermodynamics

In this section, we set our stage on cases of bipartite systems, whose Hilbert space admits a tensor factorization $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$. We consider relativistic QFTs on a stationary geometry background with metric $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$. For them, such a decomposition is not arbitrary, according to the Reeh-Schlieder theorem. Then, we review the Casini-Bekenstein bound, a general result for any relativistic QFT that respects such decomposition. The modification of such bound requires changing the modular flow, which is supposed to be conserved during unitary transformation. We will show no change of local quantity ΔH is the specific condition that leads to an entropic gradient which can reproduce Newton's second law and gravitational force in GR, as in the two cases we are familiar with the definition of this bound, Rindler space and static black hole.

5.2.1 Casini-Bekenstein bound in global causal wedges

For any global state with density matrix $\rho = |\Psi\rangle\langle\Psi|$ in a general quantum system, the state confined in the subsystem A (whose complement is $\bar{A}$) can be described by the reduced density matrix $\rho_A = \mathrm{Tr}_{\bar{A}}\,\rho$. We can always write the reduced density matrix as

$$\rho_A = \frac{e^{-K}}{\mathrm{Tr}\, e^{-K}}\,, \tag{5.7}$$

because it is positive defined and hermitian. K is known as the modular Hamiltonian [265] of ρ_A. The entanglement entropy is defined as the von Neumann Entropy

$$S(\rho_A) = -\,\mathrm{Tr}\,\rho_A \log \rho_A\,. \tag{5.8}$$

Let us consider the special cases in relativistic QFTs whose Hilbert space can be decomposed as a tensor product $\mathcal{H} = \mathcal{H}_R \bigotimes \mathcal{H}_{R^c}$, associated to spatial region R (which has an algebra $\mathcal{A}(R)$ of local operators) and its complementary

set R^c lying on a Cauchy slice. By tracing over $\mathcal{H}_{R^c}$, we get the reduced density matrix

$$\rho_R = \mathrm{Tr}_{R^c}\rho \tag{5.9}$$

Generally, such tensor decomposition in relativistic QFTs is not possible if $|\Psi\rangle$ is cyclic and separating [266], according to the Reeh-Schlieder theorem. Special global causal wedges, such as the Rindler wedge in Minkowski spacetime, are where the decomposition can take place. Therefore, we would rather set the thermodynamics in a global causal wedge than forming a local entropic mechanism by thermodynamics on the "local Rindler horizon" as in [267].

We denote the causal domain of R as $D(R)$. While choosing another spatial region V' which shared the same causal domain $D(R') = D(R)$, the entanglement entropy stay the same

$$S(\rho_{R'}) = S(\rho_R) \tag{5.10}$$

and it does not change under unitary transformations U

$$\rho_{R'} = U^\dagger \rho_R U\,, \tag{5.11}$$

Also, during time evolution, the unitary transformation does not change the entanglement entropy inside the causal domain.

Take the half-space $R = \{t = 0, x \geq 0\}$ in Minkowski spacetime for example first, its causal domain is the Rindler space, called the right Rindler wedge. According to [268], the Minkowski vacuum state confined in the right Rindler wedge is a Gibbs state

$$\rho_R^0 = \frac{e^{-H/T}}{Z}\,, \tag{5.12}$$

and the modular Hamiltonian of the vacuum state is the boost generator $K = \frac{H_\eta}{T_U}$, which is a local operator and generates a conserved modular flow. We can see this from the conserved charge $\int_\Sigma T_{\mu\nu}\chi^\mu d\Sigma^\nu$ associated with the Killing vector χ^μ; thus the expectation value of modular Hamiltonian generates the conserved flow from the local operator

$$H = \int_\Sigma T_{\mu\nu}\chi^\mu d\Sigma^\nu\,. \tag{5.13}$$

We use the expectation value

$$\langle H \rangle_{\rho_R} = \mathrm{Tr}\,\rho_R H \tag{5.14}$$

to replace the role of energy, for the state labeled by its density matrix ρ_R. Since χ^μ is dependent on the trajectory labeled by $\xi = const$, this expectation value are also related by redshift factor V to different observers.

Generally, the vacuum fluctuation causes UV-divergence in $S(\rho_R^0)$. Energy and entropy subtracting the vacuum fluctuation defined in [269, 270] are

$$\Delta \langle H \rangle = \mathrm{Tr}\, \rho_R^1 H - \mathrm{Tr}\, \rho_R^0 H \,, \tag{5.15}$$

and

$$\Delta S = S(\rho_R^1) - S(\rho_R^0) \,. \tag{5.16}$$

Now, let us review Casini's proof. The relative entropy is defined as

$$S(\rho|\sigma) = \mathrm{Tr}\, \rho \log \rho - \mathrm{Tr}\, \rho \log \sigma \ . \tag{5.17}$$

and from the positivity of the relative entropy

$$S(\rho_R^1|\rho_R^0) = \Delta \langle K \rangle - \Delta S \geq 0 \,. \tag{5.18}$$

Casini simply proved

$$\Delta S \leq \Delta \langle K \rangle \ , \tag{5.19}$$

which is

$$\Delta S \leq \Delta \langle H \rangle / T \,. \tag{5.20}$$

when including thermalisation.

In the whole context, we always take the saturation of the entropy bound

$$\Delta S = \Delta \langle K \rangle \ , \tag{5.21}$$

or

$$\Delta S = \Delta \langle H \rangle / T \,. \tag{5.22}$$

Now we have set up the stage and the definition of quantities.

5.2.2 Where does the entropic gradient come from?

This question directly links to the interesting query how one can realise gravitational force as a thermodynamic force.

Between two static observers with different trajectories label by λ' and λ, the local measurement of the conserved quantity $\Delta\langle H\rangle$ and temperature T, both depend on the redshift factor accordingly

$$H'/H = T'/T = V(\lambda)/V(\lambda')\,, \tag{5.23}$$

where the second equality is for the Tolman relation [271], while it is $K = H/T$ that stays the same.

But do remember the entanglement entropy is always the one in this Cauchy slice, so we write

$$\Delta S' = \Delta S\,, \tag{5.24}$$

even if the entropy bound is not saturated.

Let us define

$$\delta\Delta\langle H\rangle = \Delta\langle H'\rangle - \Delta\langle H\rangle \tag{5.25}$$

and

$$\delta\Delta S = \frac{\Delta\langle H'\rangle}{T'} - \frac{\Delta\langle H\rangle}{T} \tag{5.26}$$

for infinitesimal variation.

The thermodynamics comes when one tries to extract gravity, in processes under a special condition

$$\delta\Delta\langle H\rangle = 0\,, \tag{5.27}$$

which will cause the entropy bound change

$$\delta\Delta S = \delta\frac{\Delta\langle H\rangle}{T} = \frac{\Delta\langle H\rangle}{T_0}\delta V\,. \tag{5.28}$$

This condition reveals the origin where the entropic gradient comes into the story.

Or we could consider what happens in the view of the same observer with H. Then, the temperature is fixed $T' = T$ but after the influence, the condition (5.27) is equivalent to $\Delta\langle H\rangle' = \Delta\langle H\rangle V$, so we will still get (5.28).

The expectation value of H is the integration

$$\langle H \rangle_{\rho_R} = \int_\Sigma \langle T_{\mu\nu} \rangle_{\rho_R} \chi^\mu d\Sigma^\nu \tag{5.29}$$

of the expectation value of local operator $T_{\mu\nu}$. Thus, one test particle (we call it box) as excited state localised at the position of one local observer can be made by centralising/massing $\langle T_{\mu\nu} \rangle_{\rho_R}$ into a small region.

The external influence to overcome the redshift effect brings in thermodynamics to form an entropic mechanism for gravity. The external influence then causes the heat flow $\delta Q = T\delta\Delta S$ into the causal wedge. It is easy to ignore that this process is not unitary if the progress changes the fine-grained entropy in the whole casual wedge.

5.2.3 Emergence of Newton's second law in Rindler space

In the coordinate $\{\eta, \xi\}$, the metric of Rindler space is

$$ds^2 = e^{2a\xi}(-d\eta^2 + d\xi^2)\,, \tag{5.30}$$

for the right Rindler wedge of the Minkowski spacetime.

Every orbit $\xi \equiv const$ corresponds to one of the different accelerating observers following a boost killing vector δ_η. Those accelerating orbits share the same Rindler horizon $H^\pm$ as well as the same causal development, which is the right Rindler wedge.

The redshift factor is

$$V(\xi) = \sqrt{-\chi_\mu \chi^\mu} = e^{a\xi} \tag{5.31}$$

where χ_μ is the killing vector.

The surface gravity of the Killing horizon of the wedge is just $\kappa = a$, so the Unruh temperature [272] is

$$T = T_U = \frac{a}{2\pi}\,, \tag{5.32}$$

where the parameter a is also the acceleration of the observer following the orbit $\xi \equiv 0$.

From the proposed entropic gradient expression (5.3), we will get

$$\nabla_\mu S = \frac{m}{T_U}\delta^\xi_\mu \delta_\xi V(\xi) = \delta^\xi_\mu 2\pi m e^{a\xi}\,, \tag{5.33}$$

and the entropic force formula (5.4) produces

$$F_\mu = T_U \nabla_\mu S = \delta^\xi_\mu m a e^{a\xi}\,, \tag{5.34}$$

where the covariant δ^ξ_μ shows the force is in the direction to switch the orbit towards the one with higher acceleration. So the external force $F = \sqrt{F_\mu F^\mu}$ is

$$F = ma\,, \tag{5.35}$$

which exactly agrees with Newton's second Law.

5.2.4 Emergence of gravitational force

We set a stationary background of an asymptotic flat Schwarzschild black hole with the metric

$$ds^2 = -(1 - \frac{2G_N M}{r})dt^2 + (1 - \frac{2G_N M}{r})^{-1}dr^2 + r^2 d\Omega^2\,, \tag{5.36}$$

in the global coordinate. We ignore the back-reaction from our test particle to the geometry.

The redshift factor is

$$V(r) = \sqrt{-\chi^\mu \chi_\mu} = \sqrt{-g_{00}} = \sqrt{1 - \frac{2G_N M}{r}}\,, \tag{5.37}$$

the entropic gradient is

$$\nabla_\mu S \quad = \quad \frac{1}{T_H}\frac{G_N M}{r^2\sqrt{1 - \frac{2G_N M}{r}}}\delta^r_\mu\,, \tag{5.38}$$

and the local measure the Hawking temperature for the static observer with $r \equiv const$ is

$$T = \frac{T_H}{V(r)}\,. \tag{5.39}$$

So the entropic force formula reproduces

$$F_\mu = T\nabla_\mu S = \frac{G_N M m}{r^2(1 - \frac{2G_N M}{r})}\delta^r_\mu\,. \tag{5.40}$$

For the observer at infinity, the force amounts $F = V(r)\sqrt{F_\mu F^\mu} = \frac{G_N M m}{r^2}$.

Notice that it is directly covariant results calculated in GR; see textbooks [255, 256]. These results agree with the local external force

$$\mathbf{F}_{ex} = m a_\mu \tag{5.41}$$

with $a^\mu = U^\nu \nabla_\nu U^\mu$, for the static observer whose four-velocity U^μ is proportional to the time-translation Killing vector δ_t.

Near-Horizon limit is not generalisable We note that, to form a general entropic mechanism, the local Hawking temperature $T(r) = \frac{T_H}{V(r)}$ plays the ordinary role of the Unruh temperature T_U in entropic gravity theories. And our results are directly consistent with the gravitational force, not just in the near-horizon region.

In the near-horizon limit, the black hole geometry approximates the Rindler space while the local Hawking temperature approximates the Unruh Temperature; that is why an entropic mechanism works directly in generic situations can be applied to the near-horizon region, not the other way around.

5.3 The Emergence of Inertial Force

This section develops the entropic mechanism in detail to derive the inertial force from the entanglement first law. It is a specific technique to extract gravitational attraction through thermodynamics. Then we give a holographic interpretation after confirming that the saturation of the Casini-Bekenstein bound is closely related to the first law of black hole thermodynamics, providing the exact amount of entanglement entropy necessary for generic situations.

In Newton's mechanics, to maintain any object of mass m relatively static to the accelerating/inertial frame with acceleration a, we need to add on one external force

$$F = ma\,, \tag{5.42}$$

which is a reframed statement of Newton's second law. While, from the point of view of one accelerating observer, the balance condition

$$F_i + F = 0 \tag{5.43}$$

should be satisfied for one effective force, F_i, which is the inertial force.

However, in GR, the free-falling trajectory is indeed geodesic with no acceleration. We choose the accelerating frame to be static, with the velocity U^μ proportional to the timelike Killing vector χ^μ. The acceleration $a^\mu = U^\nu \nabla_\nu U^\mu$ is for the static observer following a timelike killing vector, and then gravitational attraction becomes the inertial force

$$\mathbf{F}_g = mg^\mu \tag{5.44}$$

where $g^\mu = -a^\mu$ is the gravitational acceleration for the geodesic relative to that static observer.

To calculate the inertial force from thermodynamics, let us form a quasi-static process to move the object a little bit to the nearby trajectory, with the existence of external force satisfying the balance condition $\mathbf{F}_{ex} + \mathbf{F}_g = 0$. Noted that this process will not change the momentum

$$\frac{dp}{d\lambda} = 0\,, \tag{5.45}$$

which is the major divergence from Susskind's situation for complexity tendency.

Modular Hamiltonian We use the expectation value of the modular Hamiltonian as "energy" in the spacetime thermodynamics. We already know the Killing vector χ^μ is associated with a conserved charge

$$E_T = \int_\Sigma T_{\mu\nu}\chi^\mu d\Sigma^\nu\,, \tag{5.46}$$

In Rindler space, this leads to the boost generator

$$H_\eta = a\int_{x>0} d^{d-1}x\, x\, T_{00}\,, \tag{5.47}$$

for the Killing vector

$$\delta_\eta = a(x\delta_t + t\delta_x)\,, \tag{5.48}$$

to the observer of acceleration a. And $K = H_\eta/T_U$ is the modular Hamiltonian of the vacuum state ρ_R^0. For example, the vacuum state for eternal black hole without radiation is Hartle-Hawking state [273]

$$\rho_{HH} \sim e^{-H/T_H} \tag{5.49}$$

where H is the time-translation symmetry operator for the static geometry as (5.13) associated with the Killing vector δ_t for the observer at infinity.

Now we would also clarify that thermodynamics for spacetime is always associated with the quantum expectation value $\langle H\rangle$ along with the temperature T_H, neither classical Komar mass nor ADM mass. Macroscopic thermal temperature is probably irrelevant here. However, the conserved quantum quantity $\Delta\langle H\rangle$ will promisingly approximate to Komar mass or ADM mass in the classical limit.

5.3.1 External work term from the entanglement first law

Previous work [257] derived certain thermodynamic equations to calculate the inertial force, noticing the difference between the thermodynamics first law and entanglement first law. Let us illustrate it in this subsection and then further develop it in the next subsection.

In the last section, we have set our stage on the causal wedge $D(R)$ associated with a special separation of Hilbert space $\mathcal{H} = \mathcal{H}_R \otimes \mathcal{H}_{R^c}$. The spatial region R can be the half-plane $x > 0$ in Minkowski spacetime or the exterior region $r > r_s$ of the two-sided Schwarzschild black hole. This stage allows us to form certain equations for thermodynamic quantities using the entanglement first law in the whole wedge.

The entanglement first law states that if $\rho_R(\lambda)$ of a state in the subsystem varying with one parameter λ, to the first order perturbation $d\lambda$ at $\lambda = \lambda_0$, we always have the following equation

$$\frac{dS(\rho_R)}{d\lambda} = \mathrm{Tr}\left(\frac{d\rho_R}{d\lambda} K_R\right) \tag{5.50}$$

or we can rewrite it as

$$dS = d\langle K_R \rangle \tag{5.51}$$

where $K_R = -\log \rho_R(\lambda = \lambda_0)$ is the modular Hamiltonian of the initial state. A detailed proof can be find in [252]. As a consequence of (5.50), we could take the parameters such as temperature T in $K = H/T$ out of the derivative

$$TdS = d\langle H \rangle \ . \tag{5.52}$$

We note here that there were some relevant papers about the first-law-like relation for entanglement entropy. In [274], the entanglement temperature was defined, and the generalized entanglement first law relation in Gauss-Bonet gravity and Love-Lock gravity was studied. And in [275, 276], entanglement entropy and a first-law-like relation was introduced to explain gravitational force from information erasing. While the modular Hamiltonian was not involved in those papers.

The work term Now we write the entanglement first law for the vacuum state $\rho_R^0 = e^{-H/T} / \mathrm{Tr}\, e^{-H/T}$ as

$$TdS_0 = d\langle H \rangle_0 \ , \tag{5.53}$$

and for the excited state $\rho_R^1 = e^{-K_1}/\operatorname{Tr} e^{-K_1}$ as

$$TdS_1 = d\langle H\rangle_1 + Td\langle O\rangle_1 \,, \tag{5.54}$$

where we take the modular Hamiltonian K_1 of the following form

$$K_1 = H/T + O \,, \tag{5.55}$$

where the operator

$$O = K_1 - K_0 \tag{5.56}$$

is the difference between the modular Hamiltonians of ρ_R^1 and ρ_R^0 .

Subtract (5.53) from (5.54), we get

$$Td\Delta S = d\Delta\langle H\rangle + Td\langle O\rangle_1 \tag{5.57}$$

Compare with the thermodynamic first law $dW + dQ = dE$, one can easily make the hypothesis that the work term is related to

$$dW_g = -Td\langle O\rangle_1 \,. \tag{5.58}$$

By considering the variation of the state in the existence of the external influence, we can extract the work term dW_g accounts for the external work. We noted that the detailed form of modular Hamiltonian for excited states are given in [277]. It supports our hypothesis of $K_1 = K_0 + O$, and O only involves local operators in the visible causal wedge. In another work, the modular Hamiltonian for holographic excited states is also discussed in [278].

While so far, we haven't apply the condition (5.2) $\Delta S = \frac{\Delta\langle H\rangle}{T}$ yet. We will prove that the external work will not depend on the detailed form of O after applying this condition.

As a good example, paper [257] provided a simple scalar model with a single-frequency mode to explicitly show what each term involved in is and how they vary during the process. Accidentally, after applying the entropy bound's saturation, during the quasi-static thermodynamic process below, the dW term turns into the correct expression for the inertial force as gravitational attraction successfully.

Local isoenergic process vs global isothermal process Let us now explain the thermodynamic progress first proposed in [257] in detail. We will

see it is either an isoenergic process or an isothermal process in the eyes of different observers.

In the *Bekenstein thought experiment* (see a review such as [264]), Bekenstein considered a quasi-static progress to classical level (historically it was called Geroch progress), to lower a box towards the black hole with a long string very slowly till Planck-scale-near the horizon and finally to drop it into the black hole.

While, beyond this near-horizon region, it could be still a thermodynamic process. Semi-classically, we consider the Hawking/Unruh effect that thermalises the "box" (we take as an excited particle state). Once the gravitational force is balanced by the external force, to form a thermodynamic process that changes the states, Alice varies the static trajectory $X(\lambda_0)$ a little bit to the nearby trajectory $X(\lambda)$. So the infinitesimal variation $d\lambda$ of the states is to the temperature

$$\frac{d}{d\lambda} = \frac{dT}{d\lambda}\delta_T \,. \tag{5.59}$$

In this quasi-static process, it is the external force that maintains the local measurement of frequency ω of the box not varying

$$\omega = \omega_A \tag{5.60}$$

to the local observer (let us call her the proper observer Alice) moving along with the box, so the local measurement of the energy $E = 2\pi\omega$ also stays the same.

The proper observer Alice who follows the box, will endure a temperature field with the parameter λ

$$T_A = \frac{T(\lambda_0)}{V(\lambda)}V(\lambda_0) = \frac{T_0}{V(\lambda)} \tag{5.61}$$

$$\omega_A \equiv \omega \tag{5.62}$$

where $T_0 = T(\lambda_0)V(\lambda_0)$ is the reference temperature to $V(\lambda) = 1$.

Or the process is equivalent to the fixed observer Bob, who will see the temperature fixed, but frequency changed when Alice moving with the box

$$T_B \equiv T(\lambda_0) \tag{5.63}$$

$$\omega_B = \frac{\omega_A}{V(\lambda_0)}V(\lambda) \tag{5.64}$$

The derivation is with respect to the frequency ω

$$\frac{d}{d\lambda} = \frac{d\omega}{d\lambda}\delta_\omega \tag{5.65}$$

For Alice and Bob, the distribution varies in the same way during the process, since the distribution factor varies as

$$e^{-\frac{\omega N}{T}} \to e^{-\frac{\omega_0 N V(\lambda)}{T_0}} \tag{5.66}$$

for both observers, where N is the particle number of this frequency mode. This agrees with the statement that any time slice in the same Cauchy slice is the same.

However, Alice forgets to include the redshift factor to measure global energy. If she insists on using the local Hamiltonian to measure energy

$$H_A = H|_{\lambda=\lambda_A} \tag{5.67}$$

as if Alice think she is in flat spacetime (to use her measurement of frequency for energy $E = \omega$). In Alice's eyes, objects following geodesic will get gravitational redshift, while the frequency of the box keeps the same.

However, the expectation value for the fixed frequency has changed since the temperature increases for Alice. Thus, the energy changes with the temperature of the state

$$\frac{d\langle H_A\rangle}{d\lambda} = \frac{dT}{d\lambda}\operatorname{Tr}\left(H_A\delta_T\rho_R\right) = \frac{dT}{d\lambda}\delta_T\left(\operatorname{Tr} H_A\rho_R\right) \tag{5.68}$$

since the frequency in the distribution and Hamiltonian operator is fixed.

Emergence of the inertial force In this part, we will combine the saturation condition (5.2) during the temperature-changing process, to see if the inertial force emerges the same as the entropic force formula as we used the entropic gradient. The derivation is independent of the detailed form of ρ_R^1. Let us rewrite the entanglement first law of the vacuum state (labeled by 0) and the excited state (labeled by 1) as

$$TdS_0 = d\langle H_A\rangle_0 \ , \tag{5.69}$$

$$dW_g + TdS_1 = d\langle H_A\rangle_1 \ . \tag{5.70}$$

Subtracting the vacuum fluctuation will simply lead to

$$\begin{aligned} dW_g &= -Td\Delta S + d\Delta\langle H_A\rangle \\ &= -Td\frac{\Delta(\operatorname{Tr} H_A\rho_R)}{T} + d\Delta\langle H_A\rangle \ , \end{aligned} \tag{5.71}$$

where we apply the saturation of the bound (5.2) for the second equality

$$-Td\Delta S = -Td\frac{\Delta(\mathrm{Tr}\, H_A\rho_R)}{T}\,. \tag{5.72}$$

We should be cautious that

$$\begin{aligned} d\Delta(\mathrm{Tr}\, H_A\rho_R) &= \Delta\,\mathrm{Tr}\{(dH_A)\rho_R\} + \Delta\,\mathrm{Tr}\{H_A d\rho_R\}\,, & (5.73)\\ d\Delta\,\langle H_A\rangle &= \Delta\,\mathrm{Tr}\{H_A d\rho_R\}\,, & (5.74)\end{aligned}$$

where $dH_A = 0$ vanishes since the frequency does not change during the process. An example for this is in the single-mode scalar model in [257], where we have

$$\begin{aligned} H_A &= \omega\mathcal{N} & (5.75)\\ O = K_1 - K &= \log\mathcal{N} & (5.76)\end{aligned}$$

where the number operator $\mathcal{N}$ counting the particle number of the single frequency ω mode, so we would say $dH_A = 0$ during this frequency-fixed process.

We end up with the work term simplified to

$$dW_g = -T\times\Delta\,\langle H_A\rangle d\frac{1}{T}\,, \tag{5.77}$$

which does not depend on the detailed form of the operator O. Then local temperature field $T = \frac{T_0}{V}$ for the temperature-changing process leads to the inertial force

$$\mathbf{F}_g = -T\times\frac{\Delta\,\langle H_A\rangle}{T_0}\nabla_\mu V\,. \tag{5.78}$$

This formula is exactly opposite to the external force formula (5.4) with the entropic gradient (5.3). Moreover, for Bob at a fixed position with fixed temperature, the result will be the same, but $\nabla_\mu V$ comes from $\mathrm{Tr}\, dH\rho_R$, since the *isoenergy process* for Alice is a *isothermal process* with frequency varying according to (5.64) for Bob.

Noticing the minus sigh in (5.77) and (5.78), the approach using the entanglement first law will reproduce the inertial force, while the entropic force formula together with the entropic gradient will reproduce the external force, as we expect.

Compare with the entropic force formula When T_A is very low such that the distribution factor $e^{-\omega/T_A} \ll 1$, $\Delta \langle H_A \rangle$ stays almost the same

$$\Delta \langle H_A \rangle' \approx \Delta \langle H_A \rangle \tag{5.79}$$

during the frequency-fixed process. And the entropy bound varies almost the same way as (5.28) in the fixed-energy process

$$d\Delta S \approx \frac{\Delta \langle H \rangle}{T_0} dV \tag{5.80}$$

Thus this process approximates to the energy-fixed process in Section 5.2.2 in low-temperature limit. So we will still get

$$dW_g \approx -T d\Delta S \tag{5.81}$$

which is in the opposite direction to the change of the entropy bound ΔS.

5.3.2 Connection to the first law of black hole thermodynamics

The saturation of Casini-Bekenstein bound is the maximal entanglement entropy in the causal domain associated with the definite amount of energy within. Here we show it is closely related to the first law of black hole thermodynamics: the upper bound for a box outside of a black hole is also the increase of the holographic entropy when the box is merging into the black hole.

For a static observer at r, the modular Hamiltonian H associated with the Killing vector δ_t at r and local measurement of the Hawking temperature comes from the Tolman relation

$$T = \frac{T_H}{V(r)} \,. \tag{5.82}$$

If we introduce the following replacement to the entropy bound (5.2)

$$T \quad \to \quad \frac{T_H}{V(r)} \tag{5.83}$$

$$\Delta \langle H \rangle \quad \to \quad m \tag{5.84}$$

where $T_H = \frac{\kappa}{2\pi}$ is the Hawking temperature with the surface gravity $\kappa = \frac{1}{4G_N M}$ for the Schwarzschild black hole, the entropy bound (5.2) becomes

$$\Delta S = \frac{\Delta \langle H \rangle}{T} \to \frac{m V(r)}{T_H} \,, \tag{5.85}$$

where we can import the detailed form of T_H to get

$$\frac{mV(r)}{T_H} = 2\pi \times 4G_N M m V(r) = \frac{4 \times 2\pi (2G_N M)(2G_N m V(r))}{4G_N} \tag{5.86}$$

Since we know the Schwarzschild radius is $r_s = 2G_N M$, we can write

$$\frac{mV(r)}{T_H} = \frac{8\pi r_s (2G_N m V(r))}{4G_N} \,. \tag{5.87}$$

This result by introducing the T_H reminds us to compare with the 1st law of black hole thermodynamic.

We can also rewrite the bound in a first-law-like form

$$T_H \Delta S = mV(r) \,, \tag{5.88}$$

while the first law of black hole thermodynamics [279] states

$$T_H \delta S_{BH} = \delta M \tag{5.89}$$

if the change of black hole mass δM relates to the change of Bekenstein-Hawking entropy

$$S_{BH} = \frac{A}{4G_N} \,, \tag{5.90}$$

where the area of event horizon is $A = 4\pi r_s^2$, with the Schwarzschild radius $r_s = 2G_N M$. Thus we know $\delta r_s = 2G_N \delta M$ and

$$\delta S_{BH} = \frac{\delta A}{4G_N} = \frac{8\pi r_s \delta r_s}{4G_N} \,. \tag{5.91}$$

By comparing (5.87) and (5.91), we can relate the change of the Bekenstein-Hawking entropy and change of black hole mass as follows

$$\delta S_{BH} = \Delta S \tag{5.92}$$

$$\delta M = mV(r) \tag{5.93}$$

to the entropy bound in the causal domain and local measurement of mass by red-shifting to infinity.

At the same time, we know the perturbation of the conserved energy in asymptotic flat Schwarzschild spacetime is equal to the amount of local measurement of mass m by red-shifting to infinity: $\delta M = mV(r)$. Geometrically, the

Schwarzschild radius will increase by $\delta r_s = 2GmV(r)$, when the black hole absorbs the box completely with the local mass m measured by a static observer at r.

In summary, the introduction of the local Hawking temperature made the entropy bound in the casual wedge equal to the change of the Bekenstein-Hawking entropy when black hole mass increases by $mV(r)$. The connection

$$\Delta S = \frac{\delta A(r_s)}{4G_N} \tag{5.94}$$

and

$$\Delta \langle H \rangle = m \tag{5.95}$$

is the foundation to build the new holographic interpretation for our entropic mechanism.

In [254], the entropic force formula and the entropic gradient that originates from the variation of the horizon area are tested by calculating the back-reaction to the geometry. They confirm the entropic force proposal works in the near-horizon region for a large Schwarzschild black hole, a large electrically charged black hole and a slowly rotating Kerr black hole. However, they find that the original holographic screen proposal does not work in generic situations.

Next, we show our discovery of (5.94) here is the key to a new holographic interpretation beyond the near-horizon region.

5.3.3 Holographic interpretation

We have found that the upper bound of entropy to the mass m of the box in a black hole background is equal to the variation of the new black hole if merged with the mass m. Moreover, it corresponds to the radius variation of the event horizon by $\delta r_s = 2GMmV(r)$. The saturation of the Casini-Bekenstein bound along with the vanishing relative entropy is equivalent to a more general condition of holography for the exterior matter of the black hole horizon.

We can rewrite the Bekenstein-Hawking entropy

$$S_{BH} = \frac{A(\Sigma_{\text{hor}})}{4G_N} \tag{5.96}$$

and the event horizon can be regarded as the minimal surface Σ_{hor} for two-sided AdS black holes.

Quasi-Static to covariant Once we withdraw external influence by setting $\mathbf{F}_{ex} = 0$, the heat flow stops: $\delta Q = T\delta S = 0$. If the quasi-static process stops at r and the mass m starting to free-fall towards the black hole, the entropy change of the new black hole will depend on the final position r

$$d(S'_{bh} - S_{bh}) = d\frac{m}{T(r)} \tag{5.97}$$

The external force measured at infinity in GR exactly matches with the expression

$$\mathbf{F}_{ex} = T_H \nabla_\mu (S'_{bh} - S_{bh}) \tag{5.98}$$

From (5.94), we can write local inertial force in a holographic expression

$$\mathbf{F}_g \approx -\frac{T_H}{V(r)} \nabla_\mu \left(\frac{\delta A(\Sigma_{r_s})}{4G_N}\right). \tag{5.99}$$

We point out that, covariantly this interpretation corresponds to the variation of Bousso bound [262, 263], since this is the same situation to collapse matters to form a new black hole.

The new thing here is that this shows any attempting generalisation will fail, if using the area change δA of the holographic screen at r. Otherwise, the original holographic interpretation from

$$\delta S = \frac{\delta A(\Sigma_r)}{4G_N} \tag{5.100}$$

gives too much entropy that the region interior of the holographic screen is already full of a black hole [254]. Our interpretation is the right answer to generic situations. It simply explains why the original holographic screen approach only works in the near-horizon limit and cannot be generalised directly.

5.3.4 A glimpse to emergent gravity in AdS

Before further developing our theory in the AdS/CFT framework in detail, which remains a future work beyond this thesis, here we can still make prophecies about good properties that our entropic mechanism will have when adapted into this framework, benefiting from its well-established holography.

The major difference from asymptotic flat spacetime comes from that AdS/CFT would provide homologous CFT on the boundary dual to the quantum gravity in the bulk. Thus with the proper decomposition of the entire Hilbert space

of CFT into $\mathcal{H} = \mathcal{H}_B \otimes \mathcal{H}_{\bar{B}}$, the entanglement entropy corresponds to a good geometric object in the bulk, knowing as the extremal surface.

Besides, the vanishing of relative entropy was tested in [260] to the first-order perturbation; we would expect the entanglement entropy is a function of the modular flow (energy) from the saturation of entropy bound.

Therefore, we would expect a better description for our new holographic interpretation, corresponding to the variation of the extremal surface during the process to change the energy in AdS.

Extramal Surfaces When there is matter carrying entropy S_{out} outside of a black hole, the generalized entropy

$$S_{gen} = S_{bh} + S_{out} \tag{5.101}$$

follows the generalised second law (GSL) [259].

In AdS/CFT, it is the geometric subject called *extremal surface* γ_B that corresponds to the S_{gen}

$$S_{gen} = \frac{A(\gamma_B)}{4G_N} + S_{bulk}(\gamma_B) \tag{5.102}$$

for decomposition of boundary into subsystem B and its complement $\bar{B}$. The classical extremal surface for static geometry is the Ryu-Takayanagi surface [96] which minimise the bulk area γ_B, and the bulk contribution can be omitted since it is sub-leading. The HRT formula [95] was proposed as a covariant version at the classical level. While the quantum version was proposed by Faulkner-Lewkowycz-Maldacena (FLM) in [182] and then generalised to the Quantum Extremal Surface (QES) [184] with an extra maximin procedure. For a two-sided AdS black hole, the horizon can be regarded as the extremal surface for the entanglement entropy between two copies of CFT.

During the evaporation of AdS black holes, covariant versions of extremal surfaces do not vary, neither in a classical nor quantum level. This is equivalent to that the entropy bound stays the same in the covariant situation when test particles freely fall toward the black hole as a unitary process.

However, when extracting gravitational force in the bulk, we would expect that the generalised entropy changes, as well as the extremal surface associated with it. So we may again use the entanglement entropy for the decomposition $\mathcal{H} = \mathcal{H}_B \otimes \mathcal{H}_{\bar{B}}$ of the boundary CFTs, to interpret inertial force thermodynamically.

Besides, our entropic mechanism may also work to explain the gravitational force in pure AdS, using $\Delta \langle K_B \rangle$ for the Casini-Bekenstein bound, since there is no temperature.

In all, in the AdS/CFT framework, the role of surface Σ_{r_s} should be taken by the extremal surface, and a similar entropic gradient will reflect on the variation of the extremal surface.

5.4 Summary and Discussion

We build a more concrete entropic mechanism of the emergent gravity theory to explain the gravitational attraction. It somehow differs from the original entropic force conjecture.

The entropic mechanism works under two major conditions:

- This entropic mechanism appears under certain progress. It requires external influence that causes the heat flow $\delta Q = T\delta\Delta S$ into the causal domain, thus varying the entropy bound. Under specific thermodynamics processes, gravity can be extracted, and an entropic gradient occurs.
- The saturation of the entropy bound turns the result matching with that of GR. Fine-grained entropy is thus introduced to explain gravitational attraction. Moreover, this condition leads to implications of spacetime information.

The first condition is consistent with the spirit of the equivalence principle. Only when we are trying to detect gravity through interfering can we feel the existence of gravitational force. Now, it requires information change after considering the thermalisation of the Hawking temperature. Otherwise, we admit that there will be no δQ in unitary processes such as free-falling.

Figuring out the problem when will thermodynamic process arise allows us to distinguish the different occasions between emergent gravity and Susskind's complexity tendency: the latter does not vary the entanglement entropy in the whole causal wedge. So they can co-exist with each other, and we may find a way to connect the entropic gradient we calculated to operator growth through transforming.

For the second condition, the replacement of thermal entropy distinguishes emergent gravity from macroscopic thermal mechanics. This setup allows us to move the stage to quantum systems and utilise the entanglement first law

to form precise thermodynamic equations viewed by local static observers.

The Casini-Bekenstein bound supplies one simple relation between upper entropy bound and energy. After we point out it corresponds to the area variation of the horizon as the extremal surface, it leads to a simple holographic interpretation of the entropic gradient to explain the gravitational attraction.

Extremal Surface The saturation of the bound along with the vanishing relative entropy corresponds to the variation of minimal surface. To covariant meaning, it is the Bousso bound. During dynamic processes, such as the black hole evaporation and matters free-fall towards the black hole, the covariant Bousso bound associated with the extremal surfaces stays the same. In a holographic theory with the AdS/CFT correspondence [240, 280], it is the Ryu-Takayanaki surface (and covariant HRT surface) as well as its quantum versions that correspond to the generalised entropy. We point out that the new holographic interpretation of the entropic gradients will reflect the variation of extremal surfaces in this framework.

Complexity Tendency Recently, Susskind argued that gravitational attraction comes from the complexity tendency [281] by proposing *size-momentum duality* [282], and claimed it is not compatible with an entropic mechanism that may be not able to explain the oscillation of free particles in pure AdS [283].

Again, salvation is natural after our theory: these two kinds of theories are in two considerations of processes, and indeed they can co-exist after distinguishing situation difference. We show that a properly adapted emergent gravity theory to AdS may help understand the gravitational attraction in pure AdS. Moreover, the possibility is to build a connection to transform between the entropic gradient and the operator growth once we know the generic entropic gradient in spacetime and turn it into momentum-change through virtual processes involving intermediate states.

These setups modify the way we think about emergent gravity theories to explain the gravitational force. Afterwards, this could be a rigorous formulation adapted to generic situations. To approve the mechanism in detail, we would expect it also works in AdS/CFT and can be verified in this better holographic frame.

6 Summary and Outlook

If a man keeps cherishing his old knowledge, so as continually to be acquiring new, he can be a master of others.

– Confucius, The Analects (551-479 BC)

This thesis have discussed some topics related to black hole thermodynamics, the black hole information paradox, infrared structure, and emergent gravity. We would like to conclude the whole thesis with a summary of the main results of the thesis and discuss some further research directions in the last chapter.

Firstly, we have reviewed some basic concepts and ideas related to black hole thermodynamics, micro-states, and the black hole information paradox in chapter 2. The black hole is a thermal system with Hawking temperature, whose thermodynamic temperature and entropy can be regarded as our ignorance of the region behind the horizon. The black hole thermodynamics can be derived by looking at the classical saddle in the Euclidean path integral. We have also discussed different approaches in understanding the microscopic degrees of freedom of the black hole, especially hidden conformal symmetry. The black hole thermodynamics naturally lead to the black hole information paradox, which is sharpened by the AMPS firewall argument. We have provided the recent development related to the island prescription. The Page curve of radiation can be obtained by utilising the island rule. The island rule brings us many new ideas about replica wormholes, doubly holography,

thermo-mixed double state, ensemble average, and so on, which have also been introduced in this chapter.

Then, we have turned to understand the physics related to boundary effects of gravitational system, by putting a black hole into a box and looking at gauge theory living on this background. We used Euclidean path integral to calculate the entropy of the system. Firstly, to get an impression of the physics, we have U(1) gauge theory living between two flat parallel plates. There are four different contributions in the path integral because of the boundary condition we are interested in: bulk fluctuation modes $\hat{A}_\mu$, zero modes along radius direction ϕ, Wilson lines W, and other modes. At high temperature, we have the standard blackbody radiation result mainly coming from the Bulk fluctuation modes $\hat{A}_\mu$. As the temperature becomes lower and lower, i.e., two plates move closer and closer, we start to see the zero modes ϕ and boundary stretched Wilson lines W starts to dominant, which give out area multiplied by temperature squared contribution. All the fluctuation modes are not important at every low temperature, and the constant modes and topological modes give out the logarithm of temperature and coupling constant contributions. The black hole system also has two boundaries: stretched horizon and the boundary of the box. The physics are more or less the same as the flat case at high temperature. The bulk fluctuation modes contains a Bekenstein-Hawking-like entropy similar to the brick wall model, along with the standard blackbody contribution. What's more, we have found localisation of zero-energy modes of the above fields ϕ and W in the path integral at very low temperature. The corresponding entropy is proportional to the horizon area in unit of the Planck area. This phenomenon persists in the (near)-extremal black hole case. Thus, we can infer that there are two different Bekenstein-Hawking-like entropies, and the finite-temperature black hole entropy and the (near)-extremal black hole entropy may come from different places. The part contained in the bulk fluctuation modes can be understood as the degrees of freedom living near to the horizon and was used to explain the entropy of finite temperature black hole in the brick wall model. The entropy of the (near)-extremal black hole might come from a different place as the finite temperature case and can be explained from a spontaneous symmetry breaking point of view.

Soft hair was used to resolve the firewall and played an important role in Yoshida's decoupling theorem and reconstruct the black hole interior. Firstly, we have reviewed the relevant ideas and adopted the idea of treating the soft hair as a transition function between the near-horizon region and infinity in

chapter 4. The above treatment enables us to derive an effective action for those soft modes and evaluate the size of phase space of the soft hair. The Hawking radiation increases the entropies of black hole and radiation by creating new entanglement pairs; however, the soft hair measurement decreases the entropies of black hole and radiation by reconstructing the interior of the black hole. The competition between the rate of those two processes gives out a Page curve consistent with the unitary evolution of the black hole system.

Then we have moved to discuss some topics related to emergent gravity in chapter 5, where we were trying to derive the inertial force utilising the entanglement first law. The variation of Casini-Bekenstein bound naturally reproduces Newton's second law in Rindler space and local gravitational force for Schwarzschild black hole, as shown in the main context of the chapter. Then we derived the inertial force from the entanglement first law and discussed a new holographic interpretation for the entropy change to explain the gravitational force.

To conclude, this thesis has tried to cover the topics related to the black hole phase transitions, soft hair of black hole system and its role in solving black hole information paradox, and gravity from entanglement first law. However, there are still many relevant concepts and ideas that the author misses, and we hope to further understand more in the future.

Outlook

Now, we briefly discuss some related topics that are interesting to study further in the future. Most of the topics have been mentioned in the main context, so the current section should be regarded as a brief summary of those points.

Alopecia paradox

As discussed in section 3.5, the alopecia paradox is an important problem in the decoupling theorem, which is always ignored by most of the discussion that trying to avoid firewall by decoupling the late time radiation and early radiation. If one can not avoid the firewall at the end of the evaporation, the firewall is always there and can not be wiped away. Properly describe the paradox and find a proper way to explain the paradox still need more works.

Replica wormholes

As discussed in chapter 2.3, the replica wormholes can be regarded as infrared vacua degeneracy. Properly understand the relation between infrared structure

and replica wormholes is an interesting topic. Also, as discussed by Herman Verlinde in recent papers [210, 284], thermo-mixed double state and wormholes play important roles both in the spectral form factor and the calculation of Renyi entropy. Further understanding the relation of wormholes in those two cases and their relations, i.e., checking the *replica Ansatz*, is an important aspect.

Swampland program and ensemble average
The concepts of ensemble average in holography, vacuum degeneracy, and baby universe are all assuming there are extra parameters in the gravity theory related to global symmetry. However those parameters are forbidden by the swampland program in $d > 3$ dimensions [203]. The question is how to understand those new concepts without any conflict with swampland [233, 285].

Entropy of the extremal black hole
As mentioned in chapter 3, there is a phase transition at the low temperature of the black hole system, and the entropy of the extremal black hole comes with a phase transition. Understanding the difference between finite temperature and (near)-extremal black hole entropy, and explaining the entropy for the (near)-extremal case from some symmetry breaking point of view are topics to explore.

Soft hair and related topics
Also, there are still much more to be understood in terms of soft BHIP. It would be interesting to generalise the U(1) soft hair to non-abelian gauge theories and gravity theory. There are a large number of phase transitions on the Page curve shown in chapter 4; what is the exact relation between those phase transitions and the transitions between different saddles for n-th Renyi entropy is interesting to understand. What is more, soft hair, edge modes, and non-proper gauge degrees of freedom are the concepts related to the would-be-gauge degrees of freedom; however, the relation and subtle difference between those concepts need to be clarified.

A Appendices

A.1 Different Temperature Limits

In this appendix, we will evaluate and compare the partition functions of fields $\hat{A}_\mu$, ϕ, W and other modes in different temperature limits. We will be working on the flat background in this appendix. The conclusions are summarised in the main context of section 3.2.3. Here we would like to provide more details about the calculation. The partition functions we intend to evaluate are the bulk fluctuation modes

$$\ln Z_{\hat{A}} = -\sum_{\omega} \sum_{p_r,p_2,p_3} \ln\left[\beta^2(\omega^2 + p_r^2 + p_2^2 + p_3^2)\right] , \tag{A.1}$$

and the contribution from fields ϕ and W

$$\ln Z = -\sum_{\omega,p_2,p_3} \ln[\beta^2(\omega^2 + p_2^2 + p_3^2)] + \frac{1}{2}\ln[\frac{\text{Area}}{\beta L}] - \ln e + \ln Z_w \,, \tag{A.2}$$

with

$$\ln Z_w = \begin{cases} 0 \ ; & \frac{\text{Area}}{2e^2 L\beta} \gg 1 \\ -\frac{1}{2}\ln[\frac{\text{Area}}{\beta L}] + \ln e \ . & \frac{\text{Area}}{2e^2 L\beta} \ll 1 \end{cases} \tag{A.3}$$

A.1.1 High temperature limit

Firstly, let us take the high temperature limit $\beta \ll L \ll \sqrt{\text{Area}}$. The first task is to evaluate the partition function $Z_{\hat{A}}$ for the bulk fluctuation modes. In this temperature limit, we have

$$\frac{V}{\beta^3} = \frac{L \cdot \text{Area}}{\beta^3} \gg 1 \,, \tag{A.4}$$

which means that we can write

$$\omega = \omega_m = \frac{2\pi m}{\beta} \tag{A.5}$$

$$\sum_{p_r}\sum_{p_2}\sum_{p_3} = \frac{V}{(2\pi)^3}\int dp_r dp_2 dp_3\,. \tag{A.6}$$

One can further write (A.1) as

$$\ln Z_{\hat{A}} = -2V\int \frac{d^3p}{(2\pi)^3}\left[\frac{1}{2}\beta\omega + \ln(1-e^{-\beta\omega})\right], \tag{A.7}$$

where we have $\omega = \sqrt{|p_r^2+p_2^2+p_3^2|}$. This is the result for two copies of bosonic fields. The first part in $\ln Z_{\hat{A}}$ is ultraviolet (UV) divergent and can be evaluated in the presence of a regulator Λ. The integrand of the second part is exponentially small as p goes up; thus, the integral is convergent. After introducing UV cutoff Λ, we have

$$\ln Z_{\hat{A}} = -\frac{1}{8\pi^2}\beta V\times\Lambda^4 + \frac{\pi^2}{45}\frac{V}{\beta^3}\,. \tag{A.8}$$

Note that the first part involving UV cutoff Λ is a constant in the free energy because the logarithm of the partition function is linear in β. Therefore, the entropy of those modes can be written as

$$\mathcal{S}_{\hat{A}} = (1-\beta\partial_\beta)\ln Z_{\hat{A}} = \frac{4\pi^2}{45}\frac{V}{\beta^3}\,. \tag{A.9}$$

For a similar reason, one can show that the partition function for $\phi(x^a)$ and W can be written as

$$\ln Z = -\frac{1}{6\pi}\beta\text{Area}\times\Lambda^3 + \frac{1}{2}\ln\frac{2\pi^2\text{Area}}{e^2\beta L} + \frac{\zeta(3)}{\pi}\frac{\text{Area}}{\beta^2} + \ln Z_w\,. \tag{A.10}$$

The winding modes Z_w shown above depends on the value of coupling constant e^2. When

$$\frac{1}{e^2}\frac{\text{Area}}{\beta L} \gg 1\,, \tag{A.11}$$

we have

$$Z_w = \sum_n e^{-\frac{\text{Area}}{2e^2\beta L}(2\pi n)^2} \approx e^{-\frac{\text{Area}}{2e^2\beta L}(2\pi n)^2}\big|_{n=0} = 1\,, \tag{A.12}$$

thus $\ln Z_w = 0$. However, when e^2 is big enough to make

$$\frac{1}{e^2}\frac{\text{Area}}{\beta L} \ll 1\,, \tag{A.13}$$

the coefficient inside of the exponential function is very small, and we can change the sum into an integral. Thus we have

$$Z_w \approx \int dn\ e^{-\frac{\text{Area}}{2e^2\beta L}(2\pi n)^2} = \left(\frac{2\pi\text{Area}}{e^2\beta L}\right)^{-1/2}, \tag{A.14}$$

the logarithm of which can be written as

$$\ln Z_w = -\frac{1}{2}\ln\frac{2\pi\text{Area}}{e^2\beta L}. \tag{A.15}$$

So when e^2 is large enough, the constant modes contribution $\ln Z_0$ can be cancelled. Nevertheless, this does not matter because, in this temperature limit, the constant modes contribution of ϕ is always much much smaller than the fluctuation modes contributions. The statistical entropy of the fluctuation modes of ϕ and W can be computed as

$$\mathcal{S} \quad = \quad (1-\beta\partial_\beta)\ln Z = \frac{3\zeta(3)}{\pi}\frac{\text{Area}}{\beta^2}. \tag{A.16}$$

Comparing with the volume contribution, all the area and logarithm contributions are not going to be important. As shown in equation (3.76), the most important contribution always comes from the bulk fluctuation modes $\hat{A}_\mu$, which scales as the volume times the temperature cubed.

A.1.2 Low temperature limit

High temperature is boring because we can only see the bulk fluctuation modes. As the temperature goes lower and lower, when we have $L \ll \beta \ll \sqrt{\text{Area}}$, interesting phenomena start to show up.

Firstly, let us look at the bulk fluctuation modes $\ln Z_{\hat{A}}$. In this limit, the distance between the two plates is very small compared to the inverse temperature β. Assuming finite temperature, we have $\omega_m = 2\pi m/\beta$. small L implies the high-frequency modes along the r direction are gapped, and we would only see zero modes along the r direction. $\hat{A}_\mu$ vanish on the boundary, so the zero modes of $\hat{A}_\mu$ along the r direction is killed by the boundary conditions. The only survived zero modes are the zero modes of A_r namely ϕ, which will be discussed separately. In low temperature limit and also in the super-low temperature limit, we will never see any contribution from bulk fluctuation modes anymore. So we can conclude that the entropy from the bulk fluctuation modes is

$$\mathcal{S}_{\hat{A}} = 0. \tag{A.17}$$

However, the zero modes ϕ are survived and the partition functions for ϕ and W are not changed. We still have

$$\ln Z = \frac{1}{2}\ln\frac{2\pi^2\text{Area}}{e^2\beta L} - \frac{1}{6\pi}\beta\text{Area}\cdot\Lambda^3 + \frac{\zeta(3)}{\pi}\frac{\text{Area}}{\beta^2} + \ln Z_w\,. \quad \text{(A.18)}$$

For the case $\frac{1}{e^2}\frac{\text{Area}}{\beta L} \gg 1$, we have $\ln Z_w = 0$, and the overall partition function can be written as

$$\ln Z = \frac{1}{2}\ln\frac{2\pi^2\text{Area}}{e^2\beta L} - \frac{1}{6\pi}\beta\text{Area}\cdot\Lambda^3 + \frac{\zeta(3)}{\pi}\frac{\text{Area}}{\beta^2}\,. \quad \text{(A.19)}$$

The corresponding entropy can be calculated as

$$\mathcal{S} = (1-\beta\partial_\beta)\ln Z \approx \frac{3\zeta(3)}{\pi}\frac{\text{Area}}{\beta^2} + \frac{1}{2}\ln\frac{2\pi^2\text{Area}}{e^2\beta L} + \frac{1}{2}\,. \quad \text{(A.20)}$$

For the case where e^2 is very large $\frac{1}{e^2}\frac{\text{Area}}{\beta L} \ll 1$, the constant modes contribution $\ln Z_0$ is cancelled by the winding modes contribution, we only left with fluctuation modes contribution. The overall entropy is

$$\mathcal{S} = (1-\beta\partial_\beta)\ln Z = \frac{3\zeta(3)}{\pi}\frac{\text{Area}}{\beta^2}\,. \quad \text{(A.21)}$$

The entropy of the system now scales the area times the temperature squared.

A.1.3 Super-low temperature limit

As the temperature becomes lower and lower, we have $L \ll \sqrt{\text{Area}} \ll \beta$, which is the low temperature limit. In this temperature limit, not only the contribution from $\hat{A}_\mu$ can be ignored, the fluctuation modes of ϕ and W are not important at all. As for the constant modes and winding modes, we have

$$\ln Z = \frac{1}{2}\ln\frac{2\pi^2\text{Area}}{e^2\beta L} + \ln Z_w\,. \quad \text{(A.22)}$$

And in the limit $\frac{1}{e^2}\frac{\text{Area}}{\beta L} \gg 1$, we have $\ln Z_w = 0$. The overall entropy can be written as

$$\mathcal{S} = \frac{1}{2}\ln\frac{2\pi^2\text{Area}}{e^2\beta L} + \frac{1}{2}\,. \quad \text{(A.23)}$$

Whereas in the limit $\frac{1}{e^2}\frac{\text{Area}}{\beta L} \ll 1$, we have

$$\ln Z_w = -\frac{1}{2}\ln\frac{2\pi\text{Area}}{e^2\beta L} \quad \text{(A.24)}$$

which cancels the constant modes contribution and the overall entropy tends to a small constant.

A.2 The Solution of B_μ on Euclidean Schwarzschild Background

The main task in this appendix is straightforward, we need to find a solution of the bulk equation of motion in Euclidean Schwarzschild background. To avoid to complicate the story, we will solve the problem using the original metric (3.84)

$$ds^2 = \frac{\rho}{\rho + r_s} d\tau^2 + \frac{\rho + r_s}{\rho} d\rho^2 + (\rho + r_s)^2 d\Omega^2 \,. \tag{A.25}$$

Then we can perform a coordinate transformation from $x^\mu = (\tau, \rho, \theta, \varphi)$ to $x^{\mu'} = (\tau, y, \theta, \varphi)$ to obtain the solutions on the infalling coordinate system

$$A_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} A_\mu \tag{A.26}$$

The problem reads as

$$\begin{aligned} \nabla_\mu F^{\mu\nu}_{(B)} &= 0 \\ B_a\big|_{\rho=\varepsilon} = f_a^{(l)}(x^a) \qquad & B_a\big|_{\rho=L} = f_a^{(r)}(x^a) \,. \end{aligned} \tag{A.27}$$

The bulk equation of motion can be further written as

$$\begin{aligned} \nabla_\mu F^{\mu\nu}_{(B)} &= \partial_\mu F^{\mu\nu}_{(B)} + \Gamma^\mu_{\mu\lambda} F^{\lambda\nu}_{(B)} + \Gamma^\nu_{\mu\lambda} F^{\mu\lambda}_{(B)} = \partial_\mu F^{\mu\nu}_{(B)} + \Gamma^\mu_{\mu\lambda} F^{\lambda\nu}_{(B)} \\ &= \frac{1}{\sqrt{g}} \partial_\mu (\sqrt{g} F^{\mu\nu}_{(B)}) = 0 \end{aligned} \tag{A.28}$$

We may find a solution that satisfies the following equations separately

$$\frac{1}{\sqrt{g}} \partial_\rho (\sqrt{g} F^{\rho\nu}_{(B)}) = 0 \,, \tag{A.29}$$

$$\frac{1}{\sqrt{g}} \partial_a (\sqrt{g} F^{a\nu}_{(B)}) = 0 \,, \tag{A.30}$$

First of all, let's look at the $\nu = 2$ component of (A.29)

$$\frac{1}{\sqrt{g}} \partial_\rho (\sqrt{g} F^{\rho 2}_{(B)}) = 0 \tag{A.31}$$

which is satisfied if

$$F^{(B)}_{\rho 2} = \partial_\rho B_2 - \partial_2 B_\rho = \frac{D_1(x^a)}{(\rho + r_s)^2} \,. \tag{A.32}$$

Supposing B_2 takes the form

$$B_2 = -\frac{D_1(x^a)}{\rho + r_s} + D_2(x^a)\,, \tag{A.33}$$

the boundary condition fix the coefficients D_1 and D_2

$$\begin{aligned} B_2\big|_{\rho=\varepsilon} &= -\frac{D_1}{r_s+\varepsilon} + D_2 = f_2^{(l)}(x^a) \\ B_2\big|_{\rho=L} &= -\frac{D_1}{r_s+L} + D_2 = f_2^{(r)}(x^a)\,. \end{aligned} \tag{A.34}$$

We can get that

$$\begin{aligned} D_1 &= \frac{(r_s+\varepsilon)(r_s+L)}{L-\varepsilon}\left(f_2^{(r)} - f_2^{(l)}\right)\,, \\ D_2 &= \frac{r_s+L}{L-\varepsilon} f_2^{(r)} - \frac{r_s+\varepsilon}{L-\varepsilon} f_2^{(l)}\,. \end{aligned} \tag{A.35}$$

ε is much smaller than r_s and L, so we can write $r_s \approx r_s + \varepsilon$ and $L \approx L - \varepsilon$. Thus the solution for B_2 can be written as

$$B_2 \;=\; -\frac{r_s(r_s+L)}{(\rho+r_s)L}(f_2^{(r)} - f_2^{(l)}) + \frac{r_s+L}{L} f_2^{(r)} - \frac{r_s}{L} f_2^{(l)}\,. \tag{A.36}$$

For the similar reason, we can get the solution for other components of B_a. And the solution of B_a can be written as

$$B_a \;=\; -\frac{r_s(r_s+L)}{(\rho+r_s)L}(f_a^{(r)} - f_a^{(l)}) + \frac{r_s+L}{L} f_a^{(r)} - \frac{r_s}{L} f_a^{(l)}\,. \tag{A.37}$$

With the solution B_a, one can then further fix B_ρ such that the field strength satisfies the bulk equation of motion. we have

$$\begin{aligned} \partial_\rho B_a &= \frac{1}{(\rho+r_s)^2}\frac{r_s(r_s+L)}{L}(f_a^{(r)} - f_a^{(l)}) \\ \partial_a B_\rho &= \frac{1}{(\rho+r_s)^2}\frac{r_s(r_s+L)}{L}(f_a^{(r)} - f_a^{(l)})\,. \end{aligned}$$

Note that the field strength are set to be zero here. Assuming flat boundary configurations

$$f_a^{(r)} = C_a^{(r)} - \partial_a \lambda^{(r)}\,, \qquad f_a^{(l)} = C_a^{(l)} - \partial_a \lambda^{(l)}\,. \tag{A.38}$$

We can get the solution for B_ρ by integrating $\partial_a B_\rho$ over x^a

$$B_\rho = \frac{1}{(\rho+r_s)^2}\frac{r_s(r_s+L)}{L} \times [x^a \cdot (C_a^{(r)} - C_a^{(l)}) + (\lambda^{(r)} - \lambda^{(l)})]\,, \tag{A.39}$$

As a double check, one can put the solutions (A.36), (A.39), and (A.37) into equation (A.27) to check if it is satisfied or not.

Now we can perform the coordinate transformation shown in equation (A.26). we have

$$\frac{d\rho}{dx} = \sqrt{\frac{\rho}{\rho + r_s}}\,, \tag{A.40}$$

thus B_y can be written as

$$B_y = \frac{\sqrt{\rho}}{(\rho + r_s)^{5/2}} \frac{r_s(r_s + L)}{L} \times [x^a \cdot (C_a^{(r)} - C_a^{(l)}) + (\lambda^{(r)} - \lambda^{(l)})]\,. \tag{A.41}$$

So, the solutions can be summarised as

$$\begin{aligned} B_a = &-\frac{r_s(r_s + L)}{(\rho + r_s)L}[C_a^{(r)} - C_a^{(l)} + \partial_a(\lambda^{(r)} - \lambda^{(l)})] \\ &+\frac{r_s + L}{L}(C_a^{(r)} + \partial_a\lambda^{(r)}) - \frac{r_s}{L}(C_a^{(l)} + \partial_a\lambda^{(l)})\,, \end{aligned} \tag{A.42}$$

$$B_y = \frac{\sqrt{\rho}}{(\rho + r_s)^{5/2}} \frac{r_s(r_s + L)}{L} \times [x^a \cdot (C_a^{(r)} - C_a^{(l)}) + (\lambda^{(r)} - \lambda^{(l)})] \tag{A.43}$$

as shown in the main context (3.93) and (3.94).

A.3 Effective Action for Fields ϕ and W

In this appendix, we derive the effective action (3.97) from the original action of the U(1) gauge theory. The Euclidean action for Maxwell theory on a curved background can be written as

$$S_E = \frac{1}{4e^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\; F^{\mu\nu}F_{\mu\nu} \tag{A.44}$$

Now we are going to work out the above Euclidean action in terms of fields $\hat{A}_\mu$, ϕ and W. Working on the Euclidean Schwarzschild black hole background (3.84), the action can be separated into two parts with regards to (3+y) decomposition as

$$S_E = \frac{1}{4e^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\; F^{ab}F_{ab} + \frac{1}{2e^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\; F^{ya}F_{ya}\,. \tag{A.45}$$

Repeating the same calculation as the flat case, we have

$$\begin{aligned} S_E = &\frac{1}{4e^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\; \hat{F}^{ab}\hat{F}_{ab} + \frac{1}{2e^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\; \hat{F}^{ya}\hat{F}_{ya} \\ &+\frac{1}{2e^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g} \times \left(\frac{1}{|y|^2}\partial^a\phi\partial_a\phi - \frac{2}{|y|}\hat{F}^{ya}\partial_a\phi\right)\,, \end{aligned} \tag{A.46}$$

which can be further simplified as

$$\begin{aligned} S_E &= \frac{1}{4e^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\ \hat{F}^{\mu\nu}\hat{F}_{\mu\nu} + \frac{1}{2e^2|y|^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\left[g^{ab}\partial_a\phi\partial_b\phi\right] \\ &\quad -\frac{1}{e^2|y|}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\left[g^{ab}(\partial_y\hat{A}_a - \partial_a\hat{A}_y)\partial_b\phi\right] . \end{aligned} \tag{A.47}$$

Let us denote the first part in the above action as $\hat{S}_0$, the above effective action can be further written as

$$\begin{aligned} S_E &= \hat{S}_0 + \frac{1}{2e^2|y|^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\left[g^{ab}\partial_a\phi\partial_b\phi\right] \\ &\quad -\frac{1}{e^2|y|}\int_{\mathcal{M}} d\tau d^2x dr\ \sqrt{g}\left[g^{ab}(\partial_y\hat{A}_a - \partial_a\hat{A}_y)\partial_b\phi\right] , \\ &= \hat{S}_0 + \frac{1}{2e^2|y|^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\left[g^{ab}\partial_a\phi\partial_b\phi\right] \\ &\quad +\frac{1}{e^2|y|}\int_{\mathcal{M}} d\tau d^2x\ \partial_b\phi(x^a)\cdot(\int_{y_1}^{y_2} dy\sqrt{g}g^{ab}\partial_a\hat{A}_y) \\ &\quad -\frac{1}{e^2|y|}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\left[g^{ab}\partial_y\hat{A}_a\partial_b\phi\right] \\ &= \hat{S}_0 + \frac{1}{2e^2|y|^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\left[g^{ab}\partial_a\phi\partial_b\phi\right] \\ &\quad +\frac{1}{e^2|y|}\int d\tau d^2x\ \left[\sqrt{g}\ g^{ab}\right]_{y=y_1}^{y=y_2}\partial_b\phi\cdot\partial_a(\int_{y_1}^{y_2} dy\hat{A}_y) \\ &\quad -\frac{1}{e^2|y|}\int_{\mathcal{M}} d\tau d^3x\ \partial_y(\sqrt{g}g^{ab})\cdot\partial_a(\int dy\hat{A}_y)\partial_b\phi \\ &\quad +\frac{1}{e^2|y|}\int_{\mathcal{M}} d\tau d^3x\ \partial_y(\sqrt{g}g^{ab})\hat{A}_a\partial_b\phi \\ &= \hat{S}_0 + \frac{1}{2e^2|y|^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\left[g^{ab}\partial_a\phi\partial_b\phi\right] \\ &\quad -\frac{i}{e^2|y|}\int d\tau d^2x\left[\sqrt{g}\ g^{ab}\right]_{y=y_1}^{y=y_2}\partial_a W\partial_b\phi \\ &\quad +\frac{1}{e^2|y|}\int_{\mathcal{M}} d\tau d^3x\ \partial_y(\sqrt{g}g^{ab})\left[\hat{A}_a - \partial_a(\int dy\hat{A}_y)\right]\partial_b\phi\,. \end{aligned} \tag{A.48}$$

Note that the last term in the above expression can be set to zero by some proper gauge choice. So we are going to ignore the last term and only focus

on the following Euclidean action as shown in equation (3.97)

$$\begin{aligned} S_E &= \frac{1}{4e^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\ \hat{F}^{\mu\nu}\hat{F}_{\mu\nu} + \frac{1}{2e^2|y|^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\left[g^{ab}\partial_a\phi\partial_b\phi\right] \\ &\quad -\frac{i}{e^2|y|}\int d\tau d^2x\left[\sqrt{g}\ g^{ab}\right]_{y=y_1}^{y=y_2}\partial_a W\partial_b\phi\,. \end{aligned} \tag{A.49}$$

A.4 Dimensional Reduction

This appendix provides more details about the dimensional reduction shown in section 3.3.1. The exercise is very straightforward, which is to reduce the higher dimensional action (A.50) to 3-dimensional action (A.51) shown below. We have

$$\begin{aligned} S[\phi,W] &= \frac{1}{2e^2|y|^2}\int_{\mathcal{M}} d\tau d^3x\sqrt{g}\left[g^{ab}\partial_a\phi\partial_b\phi\right] \\ &\quad -\frac{i}{e^2|y|}\int d\tau d^2x\left[\sqrt{g}\ g^{ab}\right]_{y=y_1}^{y=y_2}\partial_a W\partial_b\phi\,, \end{aligned} \tag{A.50}$$

and

$$\begin{aligned} &S_{\phi,W} = \frac{1}{2e'^2}\int d\tau d^2x\sqrt{h}\left[h^{ab}\partial_a\phi\partial_b\phi\right] \\ &-\frac{i}{2e'^2}\int d\tau d^2x\sqrt{h}\left[\gamma_1\ h^{\tau\tau}\partial_\tau W\partial_\tau\phi + \gamma_2\ h^{\theta\theta}\partial_\theta W\partial_\theta\phi + \gamma_3\ h^{\varphi\varphi}\partial_\varphi W\partial_\varphi\phi\right]. \end{aligned} \tag{A.51}$$

We are going to divide the problem into two steps. The first is to solve the low dimensional metric h_{ab} and coupling constant e'. Then we are able to solve γ_1, γ_2, and γ_3 with the results from the first step.

A.4.1 Metric and coupling constant e'

As discussed in the main context, we assume the 3-dimensional metric takes the following form

$$h_{ab} = \mathrm{diag}(h_{\tau\tau},\ R^2,\ R^2\sin^2\theta)\,, \tag{A.52}$$

with topology $S^1\times S^2$. The radius of S^1 is β and radius of S^2 is R. We just need to solve the following equations

$$\frac{1}{2e^2|y|^2}\int d\tau d^2x\ (\int_{y_1}^{y_2} dy\sqrt{g}g^{ab})\ \partial_a\phi\partial_b\phi = \frac{1}{2e'^2}\int d\tau d^2x\sqrt{h}\left[h^{ab}\partial_a\phi\partial_b\phi\right].$$

The above equations can be simplified as

$$\frac{1}{2e^2|y|^2}\int_\varepsilon^L d\rho\,\frac{\partial y}{\partial\rho}\sqrt{g}g^{ab}=\frac{1}{2e'^2}\sqrt{h}h^{ab}\,. \tag{A.53}$$

We are left with two independent components

$$\frac{1}{e^2|y|^2}(3r_s^2L+\frac{3}{2}r_sL^2+\frac{L^3}{3}+r_s^3\ln\frac{L}{\varepsilon})=\frac{1}{e'^2}\sqrt{h^{\tau\tau}}R^2\,, \tag{A.54}$$

$$\frac{L}{e^2|y|^2}=\frac{1}{e'^2}\sqrt{h_{\tau\tau}}\,. \tag{A.55}$$

There are three unknown variables and only two independent equations. So we are going to write $h_{\tau\tau}$ and coupling constant e'^2 as a function of radius R. The solution can be written as

$$h_{\tau\tau} = \frac{LR^2}{3r_s^2L+\frac{3}{2}r_sL^2+\frac{L^3}{3}+r_s^3\ln L/\varepsilon}\,, \tag{A.56}$$

$$\frac{1}{e'^2} = \frac{L}{e^2|y|^2}\sqrt{h^{\tau\tau}}=\frac{\sqrt{L}}{e^2|y|^2}\frac{\sqrt{3r_s^2L+\frac{3}{2}r_sL^2+\frac{L^3}{3}+r_s^3\ln L/\varepsilon}}{R}\,. \tag{A.57}$$

A.4.2 γ couplings

The next step is to calculate γ_1, γ_2, and γ_3. To do that we just need to match the rest part of the actions (A.50) and (A.51), which is read as

$$\begin{aligned}&\tfrac{1}{e^2|y|}\int d\tau d^2x\left[\sqrt{g}\;g^{ab}\right]_{y=y_1}^{y=y_2}\partial_aW\partial_b\phi=\\&\tfrac{1}{2e'^2}\int d\tau d^2x\sqrt{h}\left[\gamma_1\;h^{\tau\tau}\partial_\tau W\partial_\tau\phi+\gamma_2\;h^{\theta\theta}\partial_\theta W\partial_\theta\phi+\gamma_3\;h^{\varphi\varphi}\partial_\varphi W\partial_\varphi\phi\right]\,.\end{aligned} \tag{A.58}$$

The useful information from the above equation is

$$\frac{1}{e^2|y|}\left[\sqrt{g}\;g^{\tau\tau}\right]_{y=y_1}^{y=y_2} = \frac{\gamma_1}{2e'^2}\sqrt{h}\;h^{\tau\tau}\,, \tag{A.59}$$

$$\frac{1}{e^2|y|}\left[\sqrt{g}\;g^{\theta\theta}\right]_{y=y_1}^{y=y_2} = \frac{\gamma_2}{2e'^2}\sqrt{h}\;h^{\theta\theta}\,, \tag{A.60}$$

$$\gamma_2 = \gamma_3\,. \tag{A.61}$$

Writing everything explicitly, we have

$$\frac{1}{e^2|y|}\left[\sqrt{\frac{\rho(y)+r_s}{\rho}}(\rho(y)+r_s)^2\right]_{y=y_1}^{y=y_2} = \gamma_1\,\frac{L}{2e^2|y|^2}h^{\tau\tau}R^2\,, \tag{A.62}$$

$$\frac{1}{e^2|y|}\left[\sqrt{\frac{\rho(y)}{\rho(y)+r_s}}\right]_{y=y_1}^{y=y_2} = \gamma_2\,\frac{L}{2e^2|y|^2}\,, \tag{A.63}$$

$$\gamma_2 = \gamma_3\,. \tag{A.64}$$

The above equations can be solved as

$$\begin{aligned}\gamma_1 &= -2|y|\sqrt{\frac{r_s}{\varepsilon}}\frac{r_s^2}{3r_s^2L+\frac{3}{2}r_sL^2+\frac{L^3}{3}+r_s^3\ln L/\varepsilon}\\ \gamma_2 &= \gamma_3=\frac{2|y|}{L}\sqrt{\frac{L}{L+r_s}}\,.\end{aligned} \tag{A.65}$$

A.5 Bulk Partition Function on Schwarzschild Background

In this appendix, we give a detailed analysis of the partition function of the bulk fluctuation modes. The gauge fixing condition can be imposed by inserting identity (3.62) in the path integral. From the flat blackbody radiation calculation shown in section 3.2.2, we already know that after gauge fixing, there are only two polarisation degrees of freedom for Maxwell theory in the bulk, which can be simulated by two massless scalar fields. One can largely mimic the calculation in equation (3.63) by using the standard Faddeev-Popov method. Here we just assume that on black hole background, the gauge fields A_μ also left with two massless bosonic components after gauge fixing.

The metric of Schwarzschild is shown in (3.82). For one free massless particle living on the this background, the motion for geodesics can be expressed as

$$g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}=0\,, \tag{A.66}$$

with λ the parameter along the trajectory. For a Schwarzschild black hole (section 5.4 in [286] for more details), the Killing vector associated with energy can be written as

$$K^\mu=(\partial_\tau)^\mu=(1,0,0,0)\qquad(\text{or } K_\mu=(1-\frac{2G_NM}{r},0,0,0))\,. \tag{A.67}$$

And the Killing vector associated with angular momentum is

$$R^{\mu} = (\partial_{\varphi})^{\mu} = (0,0,0,1) \qquad (\text{or } R_{\mu} = (0,0,0,r^2\sin^2\theta)) \,. \tag{A.68}$$

There are two conserved charges energy and angular momentum on the equatorial plane because of the Killing vectors. The conserved quantities can be expressed as

$$\begin{aligned} \mathbf{E} &= -K_{\mu}\frac{dx^{\mu}}{d\lambda} = -\left(1-\frac{2G_N M}{r}\right)\frac{d\tau}{d\lambda} \\ \mathbf{L} &= R_{\mu}\frac{dx^{\mu}}{d\lambda} = r^2\frac{d\varphi}{d\lambda} \,. \end{aligned} \tag{A.69}$$

On the other hand, we have

$$p_{\mu} = g_{\mu\nu}\frac{dx^{\nu}}{d\lambda} \,. \tag{A.70}$$

Then expression (A.66) can be expressed by the conserved energy and angular momentum as

$$-\left(1-\frac{2G_N M}{r}\right)^{-1}\mathbf{E}^2 + \left(1-\frac{2G_N M}{r}\right)p_r^2 + \frac{\mathbf{L}^2}{r^2} = 0 \,, \tag{A.71}$$

which can be rewritten as

$$p_r^2 = \left(1-\frac{2G_N M}{r}\right)^{-1}\left[\left(1-\frac{2G_N M}{r}\right)^{-1}\mathbf{E}^2 + \frac{\mathbf{L}^2}{r^2}\right] \,. \tag{A.72}$$

For a massless scalar field with mode expansion

$$\bar{A}(\tau,r,\theta,\varphi) = \sum_{\omega}\sum_{l,m} e^{-i\omega\tau}Y_{lm}(\theta,\phi)\tilde{A}(\omega,l,m,r) \,, \tag{A.73}$$

with

$$\partial_r\tilde{A} = ip_r\tilde{A} \,, \tag{A.74}$$

the above expression (A.72) can be written as the dispersion relation for A

$$p_r^2 = \left(1-\frac{2G_N M}{r}\right)^{-1}\left[\left(1-\frac{2G_N M}{r}\right)^{-1}\omega^2 + \frac{l(l+1)}{r^2}\right] \,. \tag{A.75}$$

This expression can also be gotten from different methods, for example from the equation of motion of massless field on curved space time [2].

Let us study the statistical properties of those bulk fluctuation modes. Because of the boundary conditions on $r = r_s + \varepsilon$ and $r = r_s + L$, we have a standing-wave condition along the radius direction, which can be written as

$$n\pi = \int_{r_s+\varepsilon}^{r_s+L} p_r(r, \omega, l)\ dr\,, \tag{A.76}$$

with $n \in Z_n$. The partition function of bulk fluctuation modes can always be written as

$$\ln Z_{\bar{A}} = -2\sum_{\omega} \ln(1 - e^{-\beta\omega})\,, \tag{A.77}$$

where the factor 2 means we have two polarisations, and we ignored the zero point energy in the above expression. Now we can change the summation of ω into integration by introducing density of state $g(\omega)$ and regarding the spectrum to be continuous. We obtain

$$\begin{aligned}
\ln Z_F &= -2\int_0^\infty g(\omega)\ln(1 - e^{-\beta\omega})d\omega \\
&= -2\int_0^\infty \ln(1 - e^{-\beta\omega})d\Gamma(\omega) \\
&= -2\ln(1 - e^{-\beta\omega})\Gamma(\omega)\big|_0^\infty + 2\int_0^\infty \frac{\Gamma(\omega)e^{-\beta\omega}}{1 - e^{-\beta\omega}}\beta d\omega\,,
\end{aligned} \tag{A.78}$$

where $\Gamma(\omega)$ defined by $d\Gamma = g(\omega)d\omega$ is the number of state not exceeding ω. The first part is zero when $\omega \to 0$ and $\omega \to \infty$, so the final result for our partition function can be written as

$$\ln Z_F = 2\beta\int_0^\infty \frac{\Gamma(\omega)}{e^{\beta\omega} - 1}d\omega\,. \tag{A.79}$$

$\Gamma(\omega)$ is the number of states that have energy lower that ω, we have

$$\begin{aligned}
&\Gamma(\omega) = \tfrac{1}{\pi}\textstyle\sum_{l,m} n(\omega, l, m) \\
&= \tfrac{1}{\pi}\textstyle\sum_l (2l+1)\int_{r_s+\varepsilon}^{r_s+L} dr\sqrt{\left(1 - \frac{2G_N M}{r}\right)^{-1}\left[\left(1 - \frac{2G_N M}{r}\right)^{-1}\omega^2 + \frac{l(l+1)}{r^2}\right]} \\
&\approx \tfrac{1}{\pi}\textstyle\int_l (2l+1)dl\int_{r_s+\varepsilon}^{r_s+L} \frac{dr}{1-\frac{2G_N M}{r}}\left[\omega^2 + \left(1 - \frac{2G_N M}{r}\right)\frac{l(l+1)}{r^2}\right]^{1/2}\,.
\end{aligned} \tag{A.80}$$

Note that we have changed the summation of l into integral assuming the area of the boundary is big enough. The summation or integration is from $l = 0$

to the state with energy ω. Now we can put (A.80) into equation (A.79), and the logarithm of the partition function can be written as

$$\begin{aligned}\ln Z_F &= \tfrac{2\beta}{\pi}\int_0^\infty \tfrac{d\omega}{e^{\beta\omega}-1}\int d[l(l+1)]\int \tfrac{dr}{1-\frac{2G_NM}{r}}\left[\omega^2+\left(1-\tfrac{2G_NM}{r}\right)\tfrac{l(l+1)}{r^2}\right]^{1/2}\\ &= \tfrac{2\beta}{\pi}\int_0^\infty \tfrac{d\omega}{e^{\beta\omega}-1}\int d[l(l+1)]\int \tfrac{dr}{1-\frac{2G_NM}{r}}\left[\omega^2+\left(1-\tfrac{2G_NM}{r}\right)\tfrac{l(l+1)}{r^2}\right]^{1/2}.\end{aligned}$$

Now, let us redefine $x=\left(1-\frac{2G_NM}{r}\right)\frac{l(l+1)}{r^2}$. We obtain

$$d[l(l+1)] = \frac{r^2}{1-\frac{2G_NM}{r}}dx\,, \tag{A.81}$$

thus we can rewrite the integral as

$$\begin{aligned}\ln Z_F &= \frac{2\beta}{\pi}\int_0^\infty \frac{d\omega}{e^{\beta\omega}-1}\int \frac{r^2dr}{(1-\frac{2G_NM}{r})^2}\int_0^{\omega^2} dx\left[\omega^2+x\right]^{1/2}\\ &= \frac{2\beta}{\pi}\int_0^\infty \frac{d\omega}{e^{\beta\omega}-1}\int \frac{r^2dr}{(1-\frac{2G_NM}{r})^2}\times(-\frac{2}{3}\omega^3)\\ &= -\frac{4\beta}{3\pi}\int_0^\infty \frac{\omega^3d\omega}{e^{\beta\omega}-1}\int \frac{r^4dr}{(r-2G_NM)^2}\,.\end{aligned} \tag{A.82}$$

Those integrals are straightforward to work out, we have

$$\int_0^\infty \frac{\omega^3d\omega}{e^{\beta\omega}-1} = \frac{\pi^4}{15}\frac{1}{\beta^4} \tag{A.83}$$

and

$$\begin{aligned}\int_{r_s+\varepsilon}^{r_s+l}\frac{r^4dr}{(r-2G_NM)^2} &= \int_\varepsilon^L\frac{(\rho+r_s)^4}{\rho^2}d\rho\\ &= \left[-\frac{r_s^4}{\rho}+6\,r_s^2\rho+2\,r_s\rho^2+\frac{\rho^3}{3}+4\,r_s^3\ln\rho\right]_\varepsilon^L\\ &\approx \frac{r_s^4}{\varepsilon}+4\,r_s^3\ln\frac{L}{\varepsilon}-\frac{r_s^4}{L}+6\,r_s^2L+2\,r_sL^2+\frac{L^3}{3}\,.\end{aligned} \tag{A.84}$$

All in all, the logarithm of partition function can be written as

$$\begin{aligned}\ln Z_F &= -\frac{4\pi^3}{45}\frac{1}{\beta^3}\left(\frac{r_s^4}{\varepsilon}+4\,r_s^3\ln\frac{L}{\varepsilon}-\frac{r_s^4}{L}+6\,r_s^2L+2\,r_sL^2+\frac{L^3}{3}\right)\\ &= -\frac{4\pi^3}{45}\frac{1}{\beta^3}\frac{r_s^4}{\varepsilon}-\frac{16\pi^3}{45}\frac{r_s^3}{\beta^3}\ln\frac{L}{\varepsilon}\\ &\quad -\frac{4\pi^3}{45}\frac{1}{\beta^3}\left(-\frac{r_s^4}{L}+6\,r_s^2L+2\,r_sL^2+\frac{L^3}{3}\right).\end{aligned} \tag{A.85}$$

The corresponding entropy can be calculated as

$$\begin{aligned}\mathcal{S}_F &= (1-\beta\partial_\beta)\ln Z_F \\ &= \tfrac{16\pi^3}{45}\tfrac{1}{\beta^3}\tfrac{r_s^4}{\varepsilon}+\tfrac{64\pi^3}{45}\tfrac{r_s^3}{\beta^3}\ln\tfrac{L}{\varepsilon}+\tfrac{16\pi^3}{45}\tfrac{1}{\beta^3}\left(-\tfrac{r_s^4}{L}+6\,r_s^2L+2\,r_sL^2+\tfrac{L^3}{3}\right).\end{aligned} \tag{A.86}$$

A.6 Temperature of Near-Extremal Black Holes

Let us calculate the temperature of the near-extremal black hole in this appendix. The near extremal metric is

$$ds^2 = f(\rho)d\tau^2 + f(\rho)^{-1}d\rho^2 + (r_H+\rho)^2 d\Omega_2^2\,, \tag{A.87}$$

with

$$f(\rho) = \frac{\rho^2-\varepsilon^2}{r_H^2+\rho^2} \tag{A.88}$$

Let's mainly focus on the outer horizon $\rho=\rho_+$. The temperature of the geometry can be obtained by looking into the near horizon region. In the near horizon region, where $\rho-\rho_+$ is very small, we can expand $f(\rho)$ as

$$\begin{aligned} f(\rho) &= f(\rho_+)+f'(\rho_+)(\rho-\rho_+)+\mathcal{O}(\rho-\rho_+)^2 \\ &\simeq f'(\rho_+)(\rho-\rho_+)\,. \end{aligned} \tag{A.89}$$

$\simeq$ denotes 'equal' in near horizon region. The proper distance along radius direction is

$$dz = \frac{d\rho}{\sqrt{f(\rho)}} \simeq \frac{d\rho}{\sqrt{f'(\rho_+)(\rho-\rho_+)}}\,. \tag{A.90}$$

Integrating out ρ, one can easily get

$$z = \frac{1}{\sqrt{f'(\rho_+)}}\sqrt{\rho-\rho_+}\,, \tag{A.91}$$

and thus $f(\rho)$ can be expressed in terms of z as

$$f(\rho)\simeq f'(\rho_+)(\rho-\rho_+) = [\frac{1}{2}f'(\rho_+)]^2z^2 \equiv \kappa^2z^2\,. \tag{A.92}$$

The Euclidean metric can be written as

$$\begin{aligned} ds^2 &\simeq \kappa z^2d\tau^2+dz^2+r_H^2d\Omega_2^2 \\ &= z^2d\eta^2+dz^2+r_H^2d\Omega_2^2\,, \end{aligned} \tag{A.93}$$

where we defined $\eta \equiv \kappa\tau$. In order to avoid conical singularity, we are forced to set η to be periodic $\eta \sim \eta + 2\pi$. This means that τ also is periodic with periodicity β

$$\tau \sim \tau + \beta = \tau + \frac{2\pi}{\kappa}\,. \tag{A.94}$$

The temperature of the spacetime is then

$$T = \frac{\kappa}{2\pi} = \frac{f'(\rho_+)}{4\pi} = \frac{\varepsilon}{2\pi r_H^2}\,. \tag{A.95}$$

Bibliography

[1] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, *The entropy of hawking radiation*, 2006.06872v1.

[2] G. 't Hooft, *On the quantum structure of a black hole*, *Nuclear Phys. B Proc. Suppl.* **256** (1985) 727.

[3] M. Srednicki, *Entropy and area*, *Phys. Rev. Lett.* **71** (1993) 666.

[4] S.W. Hawking, *Gravitational radiation from colliding black holes*, *Phys. Rev. Lett.* **26** (1971) 1344.

[5] D. Christodoulou, *Reversible and irreversible transformations in black-hole physics*, *Phys. Rev. Lett.* **25** (1970) 1596.

[6] D. Christodoulou and R. Ruffini, *Reversible transformations of a charged black hole*, *Phys. Rev. D* **4** (1971) 3552.

[7] S.W. Hawking, *Black hole explosions?*, *Nature* **248** (1974) 30.

[8] J.D. Bekenstein, *Black holes and entropy*, *Phys. Rev. D* **7** (1973) 2333.

[9] J.M. Bardeen, B. Carter and S.W. Hawking, *The four laws of black hole mechanics*, *Comm. Math. Phys.* **31** (1973) 161.

[10] J.D. Bekenstein, *Black holes and the second law*, *Lettere Al Nuovo Cimento Series 2* **4** (1972) 737.

[11] B. CARTER, *Rigidity of a black hole*, *Nature Physical Science* **238** (1972) 71.

[12] S.W. Hawking, *Particle creation by black holes*, *Communications In Mathematical Physics* **43** (1975) 199.

[13] J.R. Oppenheimer and H. Snyder, *On continued gravitational contraction*, *Phys. Rev.* **56** (1939) 455.

[14] J.I. Kapusta and C. Gale, *Finite-Temperature Field Theory*, Cambridge University Press (2006), 10.1017/cbo9780511535130.

[15] W.G. Unruh, *Notes on black-hole evaporation*, *Phys. Rev. D* **14** (1976) 870.

[16] T. Hartman, *Lectures on quantum gravity and black holes*, *http://www.hartmanhep.net/topics2015/gravity-lectures.pdf* .

[17] J.M. Maldacena, *Eternal black holes in ads*, *J. High Energy Phys.* **2003** (2003) 021 [hep-th/0106112].

[18] G.D. Birkhoff, *Relativity and modern physics*, Cambridge, Harvard University Press (1923).

[19] D. Kubiznak and R.B. Mann, *P-v criticality of charged ads black holes*, *J. High Energy Phys.* **2012** (2012) [1205.0559].

[20] A. Chamblin, R. Emparan, C.V. Johnson and R.C. Myers, *Holography, thermodynamics and fluctuations of charged ads black holes*, hep-th/9904197.

[21] M.M. Caldarelli, G. Cognola and D. Klemm, *Thermodynamics of kerr-newman-ads black holes and conformal field theories*, *Classical Quantum Gravity* **17** (1999) 399 [hep-th/9908022].

[22] S.W. Hawking and D.N. Page, *Thermodynamics of black holes in anti-de sitter space*, *Comm. Math. Phys.* **87** (1983) 577.

[23] D. Kubiznak and R.B. Mann, *Black hole chemistry*, *Can. J. Phys.* **93** (2015) 999.

[24] D. Kubiznak, R.B. Mann and M. Teo, *Black hole chemistry: thermodynamics with lambda*, *Classical Quantum Gravity* **34** (2017) 063001 [1608.06147].

[25] A.M. Frassino, R.B. Mann and J.R. Mureika, *Lower-dimensional black hole chemistry*, *Phys. Rev. D* **92** (2015) .

[26] C. Teitelboim, *The cosmological constant as a thermodynamic black hole parameter*, *Phys. Lett. B* **158** (1985) 293.

[27] J. Brown and C. Teitelboim, *Neutralization of the cosmological constant by membrane creation*, *Nuclear Phys. B Proc. Suppl.* **297** (1988) 787.

[28] J.D.E. Creighton and R.B. Mann, *Quasilocal thermodynamics of dilaton gravity coupled to gauge fields*, *Phys. Rev. D* **52** (1995) 4569.

[29] T. Padmanabhan, *Classical and quantum thermodynamics of horizons in spherically symmetric spacetimes*, *Classical Quantum Gravity* **19** (2002) 5387.

[30] I. Papadimitriou and K. Skenderis, *Thermodynamics of asymptotically locally AdS spacetimes*, *J. High Energy Phys.* **2005** (2005) 004.

[31] B.P. Dolan, D. Kastor, D. Kubiznak, R.B. Mann and J. Traschen, *Thermodynamic volumes and isoperimetric inequalities for de sitter black holes*, *Phys. Rev. D* **87** (2013) .

[32] R. Banerjee and D. Roychowdhury, *Thermodynamics of phase transition in higher dimensional AdS black holes*, *J. High Energy Phys.* **2011** (2011) .

[33] R. Banerjee, S.K. Modak and D. Roychowdhury, *A unified picture of phase transition: from liquid-vapour systems to AdS black holes*, *J. High Energy Phys.* **2012** (2012) .

[34] C. Niu, Y. Tian and X.-N. Wu, *Critical phenomena and thermodynamic geometry of reissner-nordström-anti-de sitter black holes*, *Phys. Rev. D* **85** (2012) .

[35] Y.-D. Tsai, X.N. Wu and Y. Yang, *Phase structure of the kerr-AdS black hole*, *Phys. Rev. D* **85** (2012) .

[36] D. Kastor, S. Ray and J. Traschen, *Enthalpy and the mechanics of AdS black holes*, *Classical Quantum Gravity* **26** (2009) 195011.

[37] B.P. Dolan, *The cosmological constant and black-hole thermodynamic potentials*, *Classical Quantum Gravity* **28** (2011) 125020.

[38] B.P. Dolan, *Pressure and volume in the first law of black hole thermodynamics*, *Classical Quantum Gravity* **28** (2011) 235017.

[39] D. Kubiznak and R.B. Mann, *P - v criticality of charged AdS black holes*, *J. High Energy Phys.* **2012** (2012) .

[40] M. Cvetic, G.W. Gibbons, D. Kubiznak and C.N. Pope, *Black hole enthalpy and an entropy inequality for the thermodynamic volume*, *Phys. Rev. D* **84** (2011) .

[41] L. Smarr, *Mass formula for kerr black holes*, *Phys. Rev. Lett.* **30** (1973) 71.

[42] P. Cheng, S.-W. Wei and Y.-X. Liu, *Critical phenomena in the extended phase space of kerr-newman-ads black holes*, *Phys Rev D* **94** (2016) [1603.08694].

[43] S.H. Hendi and M.H. Vahidinia, *Extended phase space thermodynamics andP-vcriticality of black holes with a nonlinear source*, *Phys. Rev. D* **88** (2013) .

[44] S.-B. Chen, X.-F. Liu and C.-Q. Liu, *P—v criticality of an AdS black hole in f (r) gravity*, *Chinese Phys. Lett.* **30** (2013) 060401.

[45] R. Zhao, H.-H. Zhao, M.-S. Ma and L.-C. Zhang, *On the critical phenomena and thermodynamics of charged topological dilaton AdS black holes*, *Eur. Phys. J. C* **73** (2013) .

[46] R.-G. Cai, L.-M. Cao, L. Li and R.-Q. Yang, *P-v criticality in the extended phase space of gauss-bonnet black holes in AdS space*, *J. High Energy Phys.* **2013** (2013) .

[47] E. Spallucci and A. Smailagic, *Maxwell's equal-area law for charged anti-de sitter black holes*, *Phys. Lett. B* **723** (2013) 436.

[48] E. Spallucci and A. Smailagic, *Maxwell's equal area law and the hawking-page phase transition*, *Journal of Gravity* **2013** (2013) 1.

[49] W. Xu, H. Xu and L. Zhao, *Gauss–bonnet coupling constant as a free thermodynamical variable and the associated criticality*, *Eur. Phys. J. C* **74** (2014) .

[50] D.-C. Zou, S.-J. Zhang and B. Wang, *Critical behavior of born-infeld AdS black holes in the extended phase space thermodynamics*, *Phys. Rev. D* **89** (2014) .

[51] J.-X. Mo and W.-B. Liu, *P-v criticality of topological black holes in lovelock-born-infeld gravity*, *Eur. Phys. J. C* **74** (2014) .

[52] J.-X. Mo and W.-B. Liu, *Ehrenfest scheme forP-vcriticality of higher dimensional charged black holes, rotating black holes, and gauss-bonnet AdS black holes*, *Phys. Rev. D* **89** (2014) .

[53] D.-C. Zou, Y. Liu and B. Wang, *Critical behavior of charged*

gauss-bonnet-AdS black holes in the grand canonical ensemble, *Phys. Rev. D* **90** (2014) .

[54] B. Mirza and Z. Sherkatghanad, *Phase transitions of hairy black holes in massive gravity and thermodynamic behavior of charged AdS black holes in an extended phase space*, *Phys. Rev. D* **90** (2014) .

[55] B.P. Dolan, A. Kostouki, D. Kubiznak and R.B. Mann, *Isolated critical point from lovelock gravity*, *Classical Quantum Gravity* **31** (2014) 242001.

[56] H. Xu, W. Xu and L. Zhao, *Extended phase space thermodynamics for third-order lovelock black holes in diverse dimensions*, *Eur. Phys. J. C* **74** (2014) .

[57] J. Xu, L.-M. Cao and Y.-P. Hu, *P-vcriticality in the extended phase space of black holes in massive gravity*, *Phys. Rev. D* **91** (2015) .

[58] B.P. Dolan, *Black holes and boyle's law –the thermodynamics of the cosmological constant*, *Modern Phys. Lett. A* **30** (2015) 1540002.

[59] J.-L. Zhang, R.-G. Cai and H. Yu, *Phase transition and thermodynamical geometry of reissner-nordström-AdS black holes in extended phase space*, *Phys. Rev. D* **91** (2015) .

[60] M. Azreg-Aïnou, *Black hole thermodynamics: No inconsistency via the inclusion of the missing P-v terms*, *Phys. Rev. D* **91** (2015) .

[61] S.-W. Wei and Y.-X. Liu, *Clapeyron equations and fitting formula of the coexistence curve in the extended phase space of charged AdS black holes*, *Phys. Rev. D* **91** (2015) .

[62] S.-W. Wei, P. Cheng and Y.-X. Liu, *Analytical and exact critical phenomena ofd-dimensional singly spinning kerr-AdS black holes*, *Phys. Rev. D* **93** (2016) [1510.00085].

[63] S. Gunasekaran, D. Kubiznak and R.B. Mann, *Extended phase space thermodynamics for charged and rotating black holes and born-infeld vacuum polarization*, *J. High Energy Phys.* **2012** (2012) .

[64] S.-W. Wei and Y.-X. Liu, *Triple points and phase diagrams in the extended phase space of charged gauss-bonnet black holes in AdS space*, *Phys. Rev. D* **90** (2014) .

[65] A.M. Frassino, D. Kubiznak, R.B. Mann and F. Simovic, *Multiple reentrant phase transitions and triple points in lovelock thermodynamics*, *J. High Energy Phys.* **2014** (2014) .

[66] Z. Sherkatghanad, B. Mirza, Z. Mirzaeyan and S.A.H. Mansoori, *Critical behaviors and phase transitions of black holes in higher order gravities and extended phase spaces*, *Int. J. Mod. Phys. D* **26** (2017) 1750017 [1412.5028].

[67] N. Altamirano, D. Kubiznak, R.B. Mann and Z. Sherkatghanad, *Kerr-AdS analogue of triple point and solid/liquid/gas phase transition*, *Classical Quantum Gravity* **31** (2014) 042001.

[68] N. Altamirano, D. Kubiznak and R.B. Mann, *Reentrant phase transitions in rotating anti-de sitter black holes*, *Phys. Rev. D* **88** (2013) .

[69] N. Altamirano, D. Kubiznak, R. Mann and Z. Sherkatghanad, *Thermodynamics of rotating black holes and black rings: Phase transitions and thermodynamic volume*, *Galaxies* **2** (2014) 89.

[70] B. Liang, S.-W. Wei and Y.-X. Liu, *Quasinormal modes and van der waals like phase transition of charged ads black holes in lorentz symmetry breaking massive gravity*, *Int. J. Mod. Phys. D 28, 1950113 (2019)* (2017) [1712.01545].

[71] S.-W. Wei and Y.-X. Liu, *Photon orbits and thermodynamic phase transition of d -dimensional charged AdS black holes*, *Phys. Rev. D* **97** (2018) 104027.

[72] S.-W. Wei, Y.-X. Liu and Y.-Q. Wang, *Probing the relationship between the null geodesics and thermodynamic phase transition for rotating kerr-ads black holes*, *Phys. Rev. D 99, 044013 (2019)* (2018) [1807.03455].

[73] S.-W. Wei and Y.-X. Liu, *Insight into the microscopic structure of an AdS black hole from a thermodynamical phase transition*, *Phys. Rev. Lett.* **115** (2015) .

[74] E. Witten, *Anti-de sitter space, thermal phase transition, and confinement in gauge theories*, *Adv. Theor. Math. Phys.* **2** (1998) 505.

[75] J. Maldacena, *The large n limit of superconformal field theories and*

supergravity, *Internat. J. Theoret. Phys.* **38** (1999) 1113 [hep-th/9711200].

[76] S. Gubser, I. Klebanov and A. Polyakov, *Gauge theory correlators from non-critical string theory*, *Phys. Lett. B* **428** (1998) 105.

[77] E. Witten, *Anti de sitter space and holography*, *Adv. Theor. Math. Phys.* **2** (1998) 253.

[78] B.P. Dolan, *Bose condensation and branes*, *J. High Energy Phys.* **2014** (2014) .

[79] D. Kastor, S. Ray and J. Traschen, *Chemical potential in the first law for holographic entanglement entropy*, *J. High Energy Phys.* **2014** (2014) .

[80] E. Caceres, P.H. Nguyen and J.F. Pedraza, *Holographic entanglement entropy and the extended phase structure of STU black holes*, *J. High Energy Phys.* **2015** (2015) .

[81] C.V. Johnson, *Holographic heat engines*, *Classical Quantum Gravity* **31** (2014) 205002.

[82] A. Karch and B. Robinson, *Holographic black hole chemistry*, *J. High Energy Phys.* **2015** (2015) 1.

[83] M.R. Visser, *Holographic thermodynamics requires a chemical potential for color*, 2101.04145.

[84] A. Strominger and C. Vafa, *Microscopic origin of the bekenstein-hawking entropy*, *Phys.Lett.B379:99-104,1996* **379** (1996) 99 [hep-th/9601029].

[85] C.G. Callan and J.M. Maldacena, *D-brane approach to black hole quantum mechanics*, *Nuclear Phys. B Proc. Suppl.* **472** (1996) 591 [hep-th/9602043].

[86] A.W. Peet, *Tasi lectures on black holes in string theory*, hep-th/0008241.

[87] S.R. Das and S.D. Mathur, *The quantum physics of black holes: Results from string theory*, gr-qc/0105063.

[88] J. Maldacena and A. Strominger, *Black hole greybody factors and*

d-brane spectroscopy, *Phys.Rev.D55:861-870,1997* **55** (1996) 861 [hep-th/9609026].

[89] S.D. Mathur, *The fuzzball proposal for black holes: an elementary review*, *Fortschr Phys* **53** (2005) 793 [hep-th/0502050].

[90] S.D. Mathur, *The quantum structure of black holes*, *Classical Quantum Gravity* **23** (2006) R115 [hep-th/0510180].

[91] S.D. Mathur, *What exactly is the information paradox?*, *Lect.Notes Phys.769:3-48,2009* (2008) [0803.2030].

[92] J.D. Brown and M. Henneaux, *Central charges in the canonical realization of asymptotic symmetries: an example from three dimensional gravity*, *Comm. Math. Phys.* **104** (1986) 207.

[93] A. Strominger, *Black hole entropy from near-horizon microstates*, *J. High Energy Phys.* **1998** (1998) 009 [hep-th/9712251].

[94] D. Birmingham, I. Sachs and S. Sen, *Entropy of three-dimensional black holes in string theory*, *Phys. Lett. B* **424** (1998) 275 [hep-th/9801019].

[95] V.E. Hubeny, M. Rangamani and T. Takayanagi, *A Covariant holographic entanglement entropy proposal*, *JHEP* **07** (2007) 062 [0705.0016].

[96] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602 [hep-th/0603001].

[97] R. Emparan, *Black hole entropy as entanglement entropy: a holographic derivation*, *JHEP0606:012,2006* (2006) [hep-th/0603081].

[98] G. Immirzi, *Quantum gravity and regge calculus*, *Nuclear Physics B - Proceedings Supplements* **57** (1997) 65.

[99] J.F.B. G., *Real ashtekar variables for lorentzian signature space-times*, *Phys. Rev. D* **51** (1995) 5507.

[100] C. Rovelli, *Quantum gravity*, Cambridge University Press, Cambridge, UK New York (2004).

[101] K.V. Krasnov, *Counting surface states in the loop quantum gravity*, gr-qc/9603025.

[102] C. Rovelli, *Black hole entropy from loop quantum gravity*, *Phys.Rev.Lett.77:3288-3291,1996* **77** (1996) 3288 [gr-qc/9603063].

[103] A. Ashtekar, J. Baez, A. Corichi and K. Krasnov, *Quantum geometry and black hole entropy*, gr-qc/9710007.

[104] M. Domagala and J. Lewandowski, *Black hole entropy from quantum geometry*, gr-qc/0407051.

[105] K.A. Meissner, *Black hole entropy in loop quantum gravity*, gr-qc/0407052.

[106] A. Ghosh and P. Mitra, *Counting black hole microscopic states in loop quantum gravity*, *Phys. Rev. D* **74** (2006) 064026.

[107] S.W. Hawking, M.J. Perry and A. Strominger, *Superrotation charge and supertranslation hair on black holes*, *J High Energy Phys* **2017** (2017) [1611.09175v2].

[108] S.W. Hawking, M.J. Perry and A. Strominger, *Soft hair on black holes*, *Phys Rev Lett* **116** (2016) [1601.00921v1].

[109] M. Mirbabayi and M. Porrati, *Shaving off black hole soft hair*, *Phys Rev Lett* **117** (2016) [1607.03120].

[110] R. Bousso and M. Porrati, *Soft hair as a soft wig*, *Classical Quantum Gravity* **34** (2017) 204001 [1706.00436].

[111] S. Haco, S.W. Hawking, M.J. Perry and A. Strominger, *Black hole entropy and soft hair*, *J High Energy Phys* **2018** (2018) [1810.01847].

[112] L. Donnay, G. Giribet, H.A. Gonzalez and A. Puhm, *Black hole memory effect*, *Phys Rev D* **98** (2018) [1809.07266].

[113] S. Haco, M.J. Perry and A. Strominger, *Kerr-newman black hole entropy and soft hair*, 1902.02247.

[114] V.P. Frolov, D.V. Fursaev and A.I. Zelnikov, *Statistical origin of black hole entropy in induced gravity*, *Nucl.Phys. B486 (1997) 339-352* **486** (1996) 339 [hep-th/9607104].

[115] V.P. Frolov and D.V. Fursaev, *Mechanism of generation of black hole entropy in sakharov's induced gravity*, *Phys.Rev.D56:2212-2225,1997* **56** (1997) 2212 [hep-th/9703178].

[116] V. Frolov, D. Fursaev and A. Zelnikov, *Cft and black hole entropy in induced gravity*, *JHEP 0303 (2003) 038* **2003** (2003) 038 [hep-th/0302207].

[117] L. Susskind and J. Uglum, *Black hole entropy in canonical quantum gravity and superstring theory*, *Phys Rev D* **50** (1994) 2700 [hep-th/9401070].

[118] S.N. Solodukhin, *Entanglement entropy of black holes*, *Living Rev Relativ* **14** (2011) [1104.3712].

[119] W. Donnelly and A.C. Wall, *Geometric entropy and edge modes of the electromagnetic field*, *Phys Rev D* **94** (2016) [1506.05792v1].

[120] L. Bombelli, R.K. Koul, J. Lee and R.D. Sorkin, *Quantum source of entropy for black holes*, *Phys. Rev. D* **34** (1986) 373.

[121] J.L. Cardy, *Operator content of two-dimensional conformally invariant theories*, *Nuclear Phys. B Proc. Suppl.* **270** (1986) 186.

[122] H.W.J. Blöte, J.L. Cardy and M.P. Nightingale, *Conformal invariance, the central charge, and universal finite-size amplitudes at criticality*, *Phys. Rev. Lett.* **56** (1986) 742.

[123] S. Carlip, *Black hole entropy from conformal field theory in any dimension*, *Phys. Rev. Lett.* **82** (1999) 2828 [hep-th/9812013].

[124] S.N. Solodukhin, *Conformal description of horizon's states*, *Phys. Lett. B* **454** (1999) 213 [hep-th/9812056].

[125] M. Guica, T. Hartman, W. Song and A. Strominger, *The kerr/cft correspondence*, *Phys. Rev. D* **80** (2009) 124008.

[126] A. Castro, A. Maloney and A. Strominger, *Hidden conformal symmetry of the kerr black hole*, *Phys. Rev. D* **82** (2010) [1004.0996].

[127] Y.-Q. Wang and Y.-X. Liu, *Hidden conformal symmetry of the kerr-newman black hole*, *J. High Energy Phys.* **2010** (2010) .

[128] C.-M. Chen and J.-R. Sun, *Hidden conformal symmetry of the reissner-nordström black holes*, *J. High Energy Phys.* **2010** (2010) .

[129] G. Compère, *The kerr/CFT correspondence and its extensions*, *Living Rev Relativ* **20** (2017) .

[130] W. Song and A. Strominger, *Warped AdS3/dipole-CFT duality*, *J. High Energy Phys.* **2012** (2012) .

[131] S. Carlip, *Symmetries, horizons, and black hole entropy*, 0705.3024.

[132] S. Carlip, *Effective conformal descriptions of black hole entropy*, 1107.2678.

[133] S. Carlip, *Black hole entropy from bms symmetry at the horizon*, *Phys. Rev. Lett.* **120** (2018) [1702.04439].

[134] S. Carlip, M. Clements, S. DellaPietra and V. DellaaPietra, *Sewing polyakov amplitudes i: Sewing at a fixed conformal structure*, *Comm. Math. Phys.* **127** (1990) 253.

[135] E. Witten, *On holomorphic factorization of WZW and coset models*, *Comm. Math. Phys.* **144** (1992) 189.

[136] A. Blommaert, T.G. Mertens and H. Verschelde, *Edge dynamics from the path integral: Maxwell and yang-mills*, *J High Energy Phys* **2018** (2018) [1804.07585v2].

[137] A. Blommaert, T.G. Mertens, H. Verschelde and V.I. Zakharov, *Edge state quantization: Vector fields in rindler*, *J High Energy Phys* **2018** (2018) [1801.09910v2].

[138] W. Donnelly and L. Freidel, *Local subsystems in gauge theory and gravity*, *J High Energy Phys* **2016** (2016) [1601.04744v2].

[139] K.S. Thorne, R.H. Price and D.A. MacDonald, *Black holes: The membrane paradigm* (1986).

[140] J. Lee and R.M. Wald, *Local symmetries and constraints*, *J. Math. Phys.* **31** (1990) 725.

[141] R.M. Wald, *Black hole entropy is the noether charge*, *Phys. Rev. D* **48** (1993) R3427.

[142] V. Iyer and R.M. Wald, *Some properties of noether charge and a proposal for dynamical black hole entropy*, *Phys.Rev. D50 (1994) 846-864* (1994) [gr-qc/9403028].

[143] V. Iyer and R.M. Wald, *A comparison of noether charge and euclidean methods for computing the entropy of stationary black holes*, *Phys.Rev. D52 (1995) 4430-4439* (1995) [gr-qc/9503052].

[144] R.M. Wald and A. Zoupas, *A general definition of "conserved quantities" in general relativity and other theories of gravity*, gr-qc/9911095.

[145] T. Regge and C. Teitelboim, *Role of surface integrals in the hamiltonian formulation of general relativity*, *Ann. Physics* **88** (1974) 286.

[146] V.O. Soloviev, *Boundary values as hamiltonian variables. i. new poisson brackets*, *J. Math. Phys.* **34** (1993) 5747.

[147] J. ichirou Koga, *Asymptotic symmetries on killing horizons*, *Phys.Rev.D64:124012,2001* (2001) [gr-qc/0107096].

[148] S. Silva, *Black hole entropy and thermodynamics from symmetries*, *Class.Quant.Grav. 19 (2002) 3947-3962* (2002) [hep-th/0204179].

[149] G. Barnich and F. Brandt, *Covariant theory of asymptotic symmetries, conservation laws and central charges*, *Nucl.Phys.B633:3-82,2002* (2001) [hep-th/0111246].

[150] G. Compère, *Symmetries and conservation laws in lagrangian gauge theories with applications to the mechanics of black holes and to gravity in three dimensions*, 0708.3153.

[151] D. Harlow and J. qiang Wu, *Covariant phase space with boundaries*, *J. High Energy Phys.* **2020** (2019) [1906.08616].

[152] I.M. Gel'fand and D.B. Fuks, *The cohomologies of the lie algebra of the vector fields in a circle*, *Funct Anal Appl+* **2** (1969) 342.

[153] J. Bardeen and G.T. Horowitz, *Extreme kerr throat geometry: A vacuum analog ofads2×s2*, *Phys. Rev. D* **60** (1999) 104030.

[154] S. Carlip, *Extremal and nonextremal kerr/CFT correspondences*, *J. High Energy Phys.* **2011** (2011) .

[155] S. Carlip, *Near-horizon bms symmetry, dimensional reduction, and black hole entropy*, 1910.01762.

[156] Y. Nambu, *Quasi-particles and gauge invariance in the theory of superconductivity*, *Phys. Rev.* **117** (1960) 648.

[157] J. Goldstone, A. Salam and S. Weinberg, *Broken symmetries*, *Phys. Rev.* **127** (1962) 965.

[158] J. Goldstone, *Field theories with superconductor solutions*, *Il Nuovo Cimento* **19** (1961) 154.

[159] W. Donnelly and A.C. Wall, *Entanglement entropy of electromagnetic edge modes*, *Phys Rev Lett* **114** (2015) [1412.1895v2].

[160] D.N. Page, *Information in black hole radiation*, *Phys Rev Lett* **71** (1993) 3743 [hep-th/9306083v2].

[161] D.N. Page, *Time dependence of hawking radiation entropy*, *J Cosmol Astropart Phys* **2013** (2013) 028 [1301.4995v3].

[162] P. Hayden and J. Preskill, *Black holes as mirrors: quantum information in random subsystems*, *J. High Energy Phys.* **2007** (2007) 120.

[163] W.K. Wootters and W.H. Zurek, *A single quantum cannot be cloned*, *Nature* **299** (1982) 802.

[164] D. Dieks, *Communication by EPR devices*, *Phys. Lett. A* **92** (1982) 271.

[165] D.A. Lowe, J. Polchinski, L. Susskind, L. Thorlacius and J. Uglum, *Black hole complementarity vs. locality*, *Phys.Rev.D52:6997-7010,1995* (1995) [hep-th/9506138].

[166] Y. Kiem, E. Verlinde and H. Verlinde, *Black hole horizons and complementarity*, *Phys.Rev. D52 (1995) 7053-7065* (1995) [hep-th/9502074].

[167] L. Susskind and L. Thorlacius, *Gedanken experiments involving black holes*, *Phys.Rev.D49:966-974,1994* (1993) [hep-th/9308100].

[168] L. Susskind, L. Thorlacius and J. Uglum, *The stretched horizon and black hole complementarity*, *Physical Review D* **48** (1993) 3743 [hep-th/9306069].

[169] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, *Black holes: Complementarity or firewalls?*, *J High Energy Phys* **2013** (2013) [1207.3123].

[170] B.M. Terhal, *Is entanglement monogamous?*, *IBM Journal of Research and Development* **48** (2004) 71.

[171] D. Harlow and P. Hayden, *Quantum computation vs. firewalls*, *J. High Energy Phys.* **2013** (2013) [1301.4504].

[172] S. Aaronson, *The complexity of quantum states and transformations: From quantum money to black holes*, 1607.05256.

[173] B. Yoshida, *Observer-dependent black hole interior from operator collision*, *Physical Review D* **103** (2019) 046004 [1910.11346].

[174] B. Yoshida, *Soft mode and interior operator in the hayden-preskill thought experiment*, *Phys. Rev. D* **100** (2019) 086001.

[175] S. Pasterski and H. Verlinde, *Hps meets amps: How soft hair dissolves the firewall*, 2012.03850.

[176] D. PETZ, *SUFFICIENCY OF CHANNELS OVER VON NEUMANN ALGEBRAS*, *Q. J. Math.* **39** (1988) 97.

[177] M.O. Denes Petz, *Quantum Entropy and Its Use*, Springer Berlin Heidelberg (Mar., 2004).

[178] G. Penington, *Entanglement wedge reconstruction and the information paradox*, 1905.08255v2.

[179] A. Almheiri, N. Engelhardt, D. Marolf and H. Maxfield, *The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole*, *J High Energy Phys* **2019** (2019) [1905.08762v3].

[180] A. Almheiri, R. Mahajan, J. Maldacena and Y. Zhao, *The page curve of hawking radiation from semiclassical geometry*, *J High Energy Phys* **2020** (2020) [1908.10996v2].

[181] A. Lewkowycz and J. Maldacena, *Generalized gravitational entropy*, 1304.4926.

[182] T. Faulkner, A. Lewkowycz and J. Maldacena, *Quantum corrections to holographic entanglement entropy*, *JHEP* **11** (2013) 074 [1307.2892].

[183] T. Barrella, X. Dong, S.A. Hartnoll and V.L. Martin, *Holographic entanglement beyond classical gravity*, *JHEP 1309:109,2013* **2013** (2013) [1306.4682].

[184] N. Engelhardt and A.C. Wall, *Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime*, *JHEP* **01** (2015) 073 [1408.3203].

[185] A.C. Wall, *Maximin Surfaces, and the Strong Subadditivity of the*

Covariant Holographic Entanglement Entropy, *Class. Quant. Grav.* **31** (2014) 225007 [1211.3494].

[186] P. Hayden and G. Penington, *Learning the alpha-bits of black holes*, 1807.06041.

[187] C. Akers, N. Engelhardt, G. Penington and M. Usatyuk, *Quantum maximin surfaces*, 1912.02799.

[188] J. Maldacena and L. Susskind, *Cool horizons for entangled black holes*, *Fortschr Phys* **61** (2013) 781 [1306.0533v2].

[189] J. Cotler, P. Hayden, G. Penington, G. Salton, B. Swingle and M. Walter, *Entanglement wedge reconstruction via universal recovery channels*, *Phys. Rev. X 9, 031011 (2019)* **X9** (2017) 031011 [1704.05839].

[190] G. Penington, S.H. Shenker, D. Stanford and Z. Yang, *Replica wormholes and the black hole interior*, 1911.11977v2.

[191] H.Z. Chen, R.C. Myers, D. Neuenfeld, I.A. Reyes and J. Sandor, *Quantum extremal islands made easy, part i: Entanglement on the brane*, 2006.04851v3.

[192] H.Z. Chen, R.C. Myers, D. Neuenfeld, I.A. Reyes and J. Sandor, *Quantum extremal islands made easy, part ii: Black holes on the brane*, 2010.00018v1.

[193] J. Hernandez, R.C. Myers and S.-M. Ruan, *Quantum extremal islands made easy, partiii: Complexity on the brane*, 2010.16398.

[194] L. Randall and R. Sundrum, *An alternative to compactification*, *Phys. Rev. Lett.* **83** (1999) 4690.

[195] L. Randall and R. Sundrum, *Large mass hierarchy from a small extra dimension*, *Phys. Rev. Lett.* **83** (1999) 3370.

[196] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, *Replica wormholes and the entropy of hawking radiation*, *J High Energy Phys* **2020** (2020) [1911.12333v2].

[197] R. Jackiw, *Lower dimensional gravity*, *Nucl Phys B* **252** (1985) 343.

[198] C. Teitelboim, *Supergravity and hamiltonian structure in two spacetime dimensions*, *Phys Lett B* **126** (1983) 46.

[199] P. Saad, S.H. Shenker and D. Stanford, *Jt gravity as a matrix integral*, 1903.11115.

[200] R. Bousso and E. Wildenhain, *Gravity/ensemble duality*, 2006.16289v2.

[201] J. Pollack, M. Rozali, J. Sully and D. Wakeham, *Eigenstate thermalization and disorder averaging in gravity*, *Phys Rev Lett* **125** (2020) [2002.02971v1].

[202] B. Freivogel, D. Nikolakopoulou and A.F. Rotundo, *Wormholes from averaging over states*, 2105.12771.

[203] J. McNamara and C. Vafa, *Baby universes, holography, and the swampland*, 2004.06738.

[204] A. Hebecker, T. Mikhail and P. Soler, *Euclidean wormholes, baby universes, and their impact on particle physics and cosmology*, 1807.00824.

[205] D. Marolf and H. Maxfield, *Observations of hawking radiation: the page curve and baby universes*, 2010.06602.

[206] D. Marolf and H. Maxfield, *The page curve and baby universes*, 2105.12211.

[207] D. Marolf and H. Maxfield, *Transcending the ensemble: baby universes, spacetime wormholes, and the order and disorder of black hole information*, *J High Energy Phys* **2020** (2020) [2002.08950].

[208] H. Verlinde, *Er = epr revisited: On the entropy of an einstein-rosen bridge*, 2003.13117.

[209] A. del Campo and T. Takayanagi, *Decoherence in conformal field theory*, *JHEP02(2020)170* (2019) [1911.07861].

[210] H. Verlinde, *Wormholes in quantum mechanics*, 2105.02129.

[211] A. Strominger, *Lectures on the infrared structure of gravity and gauge theory*, 1703.05448v2.

[212] G. Barnich, *Black hole entropy from non-proper gauge degrees of freedom: Ii. the charged vacuum capacitor*, *Phys Rev D* **99** (2019) [1806.00549v2].

[213] G. Barnich and M. Bonte, *Soft degrees of freedom, gibbons-hawking contribution and entropy from casimir effect*, 1912.12698.

[214] T. He, P. Mitra, A.P. Porfyriadis and A. Strominger, *New symmetries of massless qed*, 1407.3789.

[215] D. Kapec, M. Pate and A. Strominger, *New symmetries of qed*, 1506.02906.

[216] L. Faddeev and V. Popov, *Feynman diagrams for the yang-mills field*, *Physics Letters B* **25** (1967) 29.

[217] L. Susskind and J. Lindesay, *The stretched horizon*, in *An Introduction to Black Holes, Information and the String Theory Revolution*, pp. 61–68, WORLD SCIENTIFIC (2004), DOI.

[218] S.W. Hawking, *Breakdown of predictability in gravitational collapse*, *Phys Rev D* **14** (1976) 2460.

[219] A. Almheiri, R. Mahajan and J. Maldacena, *Islands outside the horizon*, 1910.11077v3.

[220] W. Israel, *Event horizons in static vacuum space-times*, *Phys Rev* **164** (1967) 1776.

[221] W. Israel, *Event horizons in static electrovac space-times*, *Comm. Math. Phys.* **8** (1968) 245.

[222] K. Papadodimas and S. Raju, *An infalling observer in ads/cft*, *J High Energy Phys* **2013** (2013) [1211.6767].

[223] E. Verlinde and H. Verlinde, *Black hole entanglement and quantum error correction*, *J High Energy Phys* **2013** (2013) [1211.6913].

[224] S.W. Hawking, *The information paradox for black holes*, 1509.01147v1.

[225] A. Strominger, *Black hole information revisited*, 1706.07143v1.

[226] Y. Nomura, *Reanalyzing an evaporating black hole*, *Phys Rev D* **99** (2019) [1810.09453].

[227] Y. Nomura, *The interior of a unitarily evaporating black hole*, *Phys Rev D* **102** (2020) [1911.13120].

[228] P. Betzios, N. Gaddam and O. Papadoulaki, *Black hole s-matrix for a scalar field*, 2012.09834.

[229] N. Gaddam and N. Groenenboom, *Soft graviton exchange and the information paradox*, 2012.02355.

[230] E. Himwich, S.A. Narayanan, M. Pate, N. Paul and A. Strominger, *The soft s-matrix in gravity*, *J High Energy Phys* **2020** (2020) [2005.13433].

[231] D. Harlow and H. Ooguri, *Symmetries in quantum field theory and quantum gravity*, 1810.05338.

[232] D. Harlow and H. Ooguri, *Constraints on symmetry from holography*, 1810.05337.

[233] D. Harlow and E. Shaghoulian, *Global symmetry, euclidean gravity, and the black hole information problem*, *Journal of High Energy Physics* **2021** (2020) [2010.10539].

[234] S. Fichet and P. Saraswat, *Approximate symmetries and gravity*, 1909.02002.

[235] T. Dray and G. 't Hooft, *The gravitational shock wave of a massless particle*, *Nuclear Phys. B Proc. Suppl.* **253** (1985) 173.

[236] G. 't Hooft, *The scattering matrix approach for the quantum black hole, an overview*, *Int. J. Modern Phys. A* **11** (1996) 4623 [gr-qc/9607022v1].

[237] G. 't Hooft, *Black hole unitarity and antipodal entanglement*, *Found Phys* **46** (2016) 1185 [1601.03447v4].

[238] P.A. Dirac, *The quantum theory of the emission and absorption of radiation*, *Proc. R. Soc. London A.* **114** (1927) 243.

[239] E. Fermi, *Nuclear Physics: A Course Given by Enrico Fermi at the University of Chicago*, Midway reprint, University of Chicago Press (1950).

[240] J.M. Maldacena, *The Large N limit of superconformal field theories and supergravity*, *Int. J. Theor. Phys.* **38** (1999) 1113 [hep-th/9711200].

[241] B. Swingle, *Entanglement Renormalization and Holography*, *Phys. Rev. D* **86** (2012) 065007 [0905.1317].

[242] M. Van Raamsdonk, *Comments on quantum gravity and entanglement*, 0907.2939.

[243] M. Van Raamsdonk, *Building up spacetime with quantum entanglement*, *Gen. Rel. Grav.* **42** (2010) 2323 [1005.3035].

[244] E. Verlinde and H. Verlinde, *Black Hole Entanglement and Quantum Error Correction*, *JHEP* **10** (2013) 107 [1211.6913].

[245] J. Maldacena and L. Susskind, *Cool horizons for entangled black holes*, *Fortsch. Phys.* **61** (2013) 781 [1306.0533].

[246] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, *Black Holes: Complementarity or Firewalls?*, *JHEP* **02** (2013) 062 [1207.3123].

[247] N. Lashkari, M.B. McDermott and M. Van Raamsdonk, *Gravitational dynamics from entanglement 'thermodynamics'*, *JHEP* **04** (2014) 195 [1308.3716].

[248] T. Faulkner, M. Guica, T. Hartman, R.C. Myers and M. Van Raamsdonk, *Gravitation from Entanglement in Holographic CFTs*, *JHEP* **03** (2014) 051 [1312.7856].

[249] T. Faulkner, F.M. Haehl, E. Hijano, O. Parrikar, C. Rabideau and M. Van Raamsdonk, *Nonlinear Gravity from Entanglement in Conformal Field Theories*, *JHEP* **08** (2017) 057 [1705.03026].

[250] T. Jacobson, *Entanglement Equilibrium and the Einstein Equation*, *Phys. Rev. Lett.* **116** (2016) 201101 [1505.04753].

[251] E.P. Verlinde, *On the Origin of Gravity and the Laws of Newton*, *JHEP* **04** (2011) 029 [1001.0785].

[252] M. Van Raamsdonk, *Lectures on Gravity and Entanglement*, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings (TASI 2015): Boulder, CO, USA, June 1-26, 2015*, pp. 297–351, 2017, DOI [1609.00026].

[253] D.-C. Dai and D. Stojkovic, *Inconsistencies in Verlinde's emergent gravity*, *JHEP* **11** (2017) 007 [1710.00946].

[254] S. Bhattacharya, P. Charalambous, T.N. Tomaras and N. Toumbas, *Comments on the entropic gravity proposal*, *Eur. Phys. J.* **C78** (2018) 627 [1803.05822].

[255] S. Carroll, *Spacetime and Geometry: An Introduction to General Relativity.*

[256] R.M.Wald, *General Reltivity*, The University of Chicago Press (1984).

[257] Y. An, *Inertial force, Hawking Temperature and Quantum Statistics*, 1811.03765.

[258] H. Casini, *Relative entropy and the Bekenstein bound*, *Class. Quant. Grav.* **25** (2008) 205021 [0804.2182].

[259] J.D. Bekenstein, *Generalized second law of thermodynamics in black hole physics*, *Phys. Rev.* **D9** (1974) 3292.

[260] D.D. Blanco, H. Casini, L.-Y. Hung and R.C. Myers, *Relative Entropy and Holography*, *JHEP* **08** (2013) 060 [1305.3182].

[261] G. Dvali, *Entropy Bound and Unitarity of Scattering Amplitudes*, 2003.05546.

[262] R. Bousso, *A Covariant entropy conjecture*, *JHEP* **07** (1999) 004 [hep-th/9905177].

[263] R. Bousso, *Holography in general space-times*, *JHEP* **06** (1999) 028 [hep-th/9906022].

[264] R. Bousso, *The Holographic principle*, *Rev. Mod. Phys.* **74** (2002) 825 [hep-th/0203101].

[265] R. Haag, *Local quantum physics: Fields, particles, algebras* (1992).

[266] E. Witten, *APS Medal for Exceptional Achievement in Research: Invited article on entanglement properties of quantum field theory*, *Rev. Mod. Phys.* **90** (2018) 045003 [1803.04993].

[267] T. Jacobson, *Thermodynamics of space-time: The Einstein equation of state*, *Phys. Rev. Lett.* **75** (1995) 1260 [gr-qc/9504004].

[268] J.J. Bisognano and E.H. Wichmann, *On the Duality Condition for Quantum Fields*, *J. Math. Phys.* **17** (1976) 303.

[269] D. Marolf, D. Minic and S.F. Ross, *Notes on space-time thermodynamics and the observer dependence of entropy*, *Phys. Rev.* **D69** (2004) 064006 [hep-th/0310022].

[270] D. Marolf, *A Few words on entropy, thermodynamics, and horizons*, in *General relativity and gravitation. Proceedings, 17th International Conference, GR17, Dublin, Ireland, July 18-23, 2004*, pp. 83–103, 2004, DOI [hep-th/0410168].

[271] R.C. Tolman, *On the weight of heat and thermal equilibrium in general relativity*, *Phys. Rev.* **35** (1930) 904.

[272] W.G. Unruh, *Notes on black hole evaporation*, *Phys. Rev.* **D14** (1976) 870.

[273] J.B. Hartle and S.W. Hawking, *Path Integral Derivation of Black Hole Radiance*, *Phys. Rev.* **D13** (1976) 2188.

[274] W.-z. Guo, S. He and J. Tao, *Note on Entanglement Temperature for Low Thermal Excited States in Higher Derivative Gravity*, *JHEP* **08** (2013) 050 [1305.2682].

[275] J.-W. Lee, H.-C. Kim and J. Lee, *Gravity from Quantum Information*, *J. Korean Phys. Soc.* **63** (2013) 1094 [1001.5445].

[276] J.-W. Lee, H.-C. Kim and J. Lee, *Gravity as a quantum entanglement force*, *J. Korean Phys. Soc.* **66** (2015) 1025 [1002.4568].

[277] S. Balakrishnan and O. Parrikar, *Modular Hamiltonians for Euclidean Path Integral States*, 2002.00018.

[278] R. Arias, M. Botta-Cantcheff, P.J. Martinez and J.F. Zarate, *Modular Hamiltonian for (holographic) excited states*, 2002.04637.

[279] J.M. Bardeen, B. Carter and S.W. Hawking, *The Four laws of black hole mechanics*, *Commun. Math. Phys.* **31** (1973) 161.

[280] E. Witten, *Anti-de Sitter space and holography*, *Adv. Theor. Math. Phys.* **2** (1998) 253 [hep-th/9802150].

[281] L. Susskind, *Computational Complexity and Black Hole Horizons*, *Fortsch. Phys.* **64** (2016) 44 [1403.5695].

[282] L. Susskind, *Why do Things Fall?*, 1802.01198.

[283] L. Susskind, *Complexity and Newton's Laws*, 1904.12819.

[284] H. Verlinde, *Deconstructing the wormhole: Factorization, entanglement and decoherence*, 2105.02142.

[285] P.-S. Hsin, L.V. Iliesiu and Z. Yang, *A violation of global symmetries from replica wormholes and the fate of black hole remnants*, 2011.09444.

[286] S.M. Carroll, *Spacetime and Geometry*, Cambridge University Press (jul, 2019), 10.1017/9781108770385.

Samenvatting

Zwarte gaten zijn fascinerende objecten in het heelal. Het begrijpen van de fysica van zwarte gaten is altijd een belangrijk onderwerp voor kwantumzwaartekracht, astrofysica, kosmologie en onze nieuwsgierigheid. Dit proefschrift bespreekt enkele onderwerpen die verband houden met de thermodynamica van zwarte gaten, de informatieparadox, infraroodstructuur en opkomende zwaartekracht.

Allereerst geven we in hoofdstuk 2 een inleiding op enkele basisconcepten en achtergrondmateriaal die relevant zijn voor dit proefschrift. We proberen vooral een algemeen overzicht te geven van de fysica van zwarte gaten, puzzels en spannende ontwikkelingen. We besteden vooral ook wat tijd aan de recente ontwikkelingen met betrekking tot de informatieparadox van zwarte gaten en nieuwe inzichten in holografische dualiteit. Meer specifiek bekijken we de thermodynamica van het zwarte gat en de relevante ideeën met betrekking tot het holografische principe en bieden we verschillende benaderingen om de microtoestand van het zwarte gat te begrijpen. Daarna bespraken we hoe we het verborgen conforme systeem in het zwarte gatsysteem kunnen zoeken, en hoe de conforme veldentheoriemethode de universaliteit van de Bekenstein-Hawking-entropie verklaart. Ten slotte bespreken we de basisconcepten van de informatieparadox van het zwarte gat en het eilandrecept.

In hoofdstuk 3 evalueren we de partitiefunctie van de U(1) ijktheorie in drie verschillende achtergronden: vlak geval, eindige temperatuur zwart gat en extreem zwart gat, met behulp van Euclidische padintegraal. Verschillende gedragingen in verschillende temperatuurgrenzen worden in detail bestudeerd. We krijgen de standaard blackbody-stralingsresultaten voor de platte behuizing bij hoge temperatuur. Wat het geval van een zwart gat betreft, is er een extra bijdrage die evenredig is met het horizongebied. De nulmodi langs de straalrichting en de Wilson-lijnen die tussen verschillende grenzen zijn uitgerekt, spelen de dominante rol voor lagere temperaturen. In deze temper-

atuurlimiet begint de entropie van het systeem zich te gedragen als het gebied van de grenzen vermenigvuldigd met temperatuur in het kwadraat. Alle fluctuatiemodi worden verondersteld uit te sterven bij zeer lage temperaturen, en we blijven alleen achter met constante modi en topologische modi in het platte geval. Bij zeer lage temperatuur krijgen we echter nog steeds een Bekenstein-Hawking-achtige entropiebijdrage op de achtergrond van het zwarte gat, die ook aanhoudt in het geval van een extreem zwart gat. Er zijn twee faseovergangen en bieden een manier om de Bekenstein-Hawking-achtige entropie bij lage temperatuur te begrijpen vanuit een symmetrie-doorbrekend gezichtspunt.

Vervolgens bespreken we in hoofdstuk 4 hoe het zachte haar van een zwart gat kan helpen om de informatieparadox van het zwarte gat te begrijpen. Door Maxwell zacht haar te behandelen als een overgangsfunctie die U(1)-meetvelden in het asymptotische gebied en het nabij-horizongebied met elkaar in verband brengt, wordt de U(1)-meetparameter $\lambda(x^a)$ natuurlijk een goed label van die vrijheidsgraden van Maxwell zacht haar. Deze interpretatie bouwt ook de verbinding op tussen Maxwell zacht haar en U(1) edge-modi die in het tussenliggende gebied leven, wat een goed gedefinieerde effectieve actiebeschrijving toelaat. We bestuderen de statistische eigenschappen aan de hand van Euclidische padintegraal, die concludeert dat de dichtheid van zacht haar toeneemt met de temperatuur van het zwarte gat. Hawking-straling verhoogt de entropie van zwart gat door verstrengeling te creëren, terwijl de meting van de zachte modus het zwarte gat op een lagere entropietoestand projecteert. De concurrentie tussen faseruimten van Hawking-straling en zacht haarmetingen geeft aanleiding tot één versie van de Page-curve die consistent is met de unitaire evolutie van het zwarte gat.

In hoofdstuk 5 verduidelijken we het probleem waarbij zwaartekracht kan worden beschouwd als voortkomend uit de thermodynamica, door een entropisch mechanisme voor te stellen dat de entropische gradiënt kan extraheren die in de ruimtetijd bestaat, als gevolg van de variatie van de Casini-Bekenstein-binding in specifieke quasi- statische processen met de warmtestroom in de hele losse wig. We formuleren expliciet de afleiding van traagheidskracht als de opkomende zwaartekracht uit de eerste wet van verstrengeling. Het huidige entropische mechanisme reproduceert de tweede wet van Newton in de Rindler-ruimte en de zwaartekracht (samen met de afleiding van de Einstein-vergelijking) buiten het bijna-horizongebied en kan worden aangepast in AdS/CFT en andere algemene situaties.

Hoofdstuk 6 geeft een samenvatting van de belangrijkste onderwerpen en re-

sultaten die in dit proefschrift zijn besproken en biedt enige vooruitzichten.

Die hoofdstukken worden gevolgd door een bibliografie, de Nederlandse versie van de samenvatting en dankwoord.

Acknowledgements

I can't imagine I would have gotten this far without the help and support of the people around me. It is so nice to have you all during the journey in the Netherlands.

Jan, thank you for your help during the last four years, and it has been a great time. You not only taught me physics knowledge and insights but also how to be an honest physicist. I hope I can learn more from you in future. I am grateful that you accepted me to the Amsterdam string theory group four years ago; this beautiful journey can come true from my dream all because of you.

Diego, I am greatly indebted to you. Thank you for sharing so many wonderful physical insights with me. You also inspired me a lot with your critical thinking. I sincerely appreciate all your help during the last four years.

Yu-Xiao, it has been four years since I left Lanzhou, thank you for still concerning me and treating me as part of the group, and thank you for bringing me to the physical world at the beginning. Shao-Wen, thank you for helping me along the way. Also, many thanks to all my friends in Lanzhou, especially, Wen-Di, Yu-Peng, and Hao. I also appreciate the help from Song He, Pujian Mao, Chen-Te Ma and Xin Zhang.

Yang, It's a great gain to meet you in Amsterdam and has you as my friend and collaborator. I enjoy so much the time when we are discussing physics and new articles from arXiv. I would also like to thank other friends, Emanuele Alesci, Yu-Sen An, Bin Guo, Shan-Ming Ruan, and Jia Tian, with whom we had so many great physics discussions together.

Special thanks to Antonio, Ankit, Sagar, Bahman, Horng Sheng, Nava, Ricardo, Greg, Jeremy, Evita, Beatrix, and Carlos. Thank you guys for the time we shared during PhD schools and all the discussions related to physics

and life. Moreover, I would like to thank all people in the Amsterdam string theory group, in particular Jay, Daniel, Alejandra, Miranda, Ben, Jan Pieter, Erik, Marcel, Tarek, Mert, Shira, Jackson, Gui, Claire, Ramesh, Edward, Dora, Gabriele, Stathis, Jorrit, Lars, Manus, Sam, and Gerben. Thank you for your contributions to all the seminars, journal clubs, discussions, and other activities we had together. Many thanks to our secretaries at IoP, especially Klaartje.

A life without a friend is a life without a sun. I would like to take this chance to express my gratitude to my Chinese friends in the Netherlands. Jingchao Wu, Xiaobai Chen, Junhao Cao, Xiang Zhao, Siyun Zhao, Dong-Gang Wang, Zheng Ying, Lifei Yan, Lingyu Lu, Haorui Peng, Tiantian Wang, and Zhendong Yuan, thank you! We should have taken more chances to have dinner together before the pandemic. Life has been made so hard during the last year and a half; I am glad it is coming to an end soon. I hope we can still get together and hang out when we are back in China.

This thesis is dedicated to my grandma. Grandma, I owe you too much and thank you for your endless love. My family is always my solid backing and supports me unconditionally. I'm so proud I was born and raised in such an affectionate and warm family.

Shuangli, I love you! Thank you for your love and company during the time in the Netherlands. We have known each other since we were twelve years old, and I know we will spend the rest of our life together. We have gone through all the ups and downs, and shared happiness and sadness with each other, we are the best match all over the world! I am such a lucky man to meet you and fall in love with you.