Emergent laws of spacetime mechanics

Ilgin, I.

Publication date
2019

Document Version
Final published version

License
Other

Link to publication

Citation for published version (APA):
Ilgin, I. (2019). *Emergent laws of spacetime mechanics*. [Thesis, fully internal, Universiteit van Amsterdam].

Emergent Laws of Spacetime Mechanics

Irfan Ilgin

Emergent Laws

of

Spacetime Mechanics

This work has been accomplished at the Institute for Theoretical Physics (ITFA) of the University of Amsterdam (UvA) and funded by the Spinoza Grant of the Dutch Science Organization (NWO).

UNIVERSITY OF AMSTERDAM

Emergent Laws
of
Spacetime Mechanics

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

op woensdag 30 oktober 2019, te 13.00 uur

door

Irfan Ilgin

geboren te Izmir

Promotiecommissie

Promotor

prof. dr. E.P. Verlinde,	Universiteit van Amsterdam

Co-Promotores

dr. C.N. Cheng,	Universiteit van Amsterdam

Overige leden

prof. dr. D.D. Baumann,	Universiteit van Amsterdam
prof. dr. J. de Boer,	Universiteit van Amsterdam
prof. dr. S. J. G. Vandoren,	Universiteit Utrecht
dr. A. Castro Anich,	Universiteit van Amsterdam
dr. B. W. Freivogel,	Universiteit van Amsterdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Publications

This thesis is based on the following publications:

[1] I. Ilgin and I-S. Yang
Causal Patch Complementarity: The Inside Story for Old Black Holes
Phys. Rev. **D89 no. 4**, 044007 (2014), arXiv:1311.1219 [hep-th].

[2] I. Ilgin
Bekenstein bound in the bulk and AdS/CFT
on submission, arXiv:1809.05770 [hep-th].

[3] I. Ilgin, E. Verlinde and M. Visser
Emergent Laws of Spacetime Mechanics
on preparation.

[4] I. Ilgin and I-S. Yang
Energy carries Information
Int. J. Mod. Phys. **A29 no. 20**, 1450115 (2014), arXiv:1402.0878 [hep-th].

Preface & Thesis Guide

This thesis is an exposition of some of my research that I carried out during my doctoral studies. The main focus of the thesis is understanding emergent laws of gravitation from an information theoretic point of view.

The thesis consists of seven chapters. First chapter 1 is an introduction on the *it from qubit* paradigm and how this research is connected to other works. The second chapter 2 provides reviews for the material that is strongly used in the rest of the thesis.

The chapter 3 focuses on the connection between black hole physics and quantum information. In particular, black hole information paradox and firewall paradox.

The fourth and fifth chapters focus on the same theme. Mainly how to reconcile bulk physics with information theory in the context of AdS/CFT. In chapter 4, the connection between Bekenstein bound in the bulk and corresponding information theoretic relation in the underlying theory is presented. In chapter 5, we study the impact of first law of entanglement entropy on general bulk surfaces using differential entropy in three dimensions. This chapter can be seen as a proof of concept of the last two chapter where we explore emergent first law with its full generality.

The last two chapters (6, 7) can be seen as the core of this thesis. In 6 we generalize the first law of black hole mechanics to codimension two spacelike surfaces for a general theory of gravity. Based on the analogies between gravity and elasticity, we identify the first law of deformations in spacetime. In chapter 7, we build on the entropic gravity proposal and elaborate/generalize it as adiabatic reaction force. We propose that microcanonical action in a general theory of gravity measures the Gibbs volume entropy of the underlying microscopic theory.

Contents

1 Introduction

The Universe is made of stories,
not of atoms.

— Muriel Rukeyser,
The Speed of Darkness.

Gilles Deleuze defines philosophy as an activity of creating concepts that are new [5]. In this sense theoretical physics strongly involves philosophical activity. History of physics contains abundance of examples of concepts created along the development of new theories. Energy, entropy, particles, fields, aether, quanta, strings are few of these concepts. Concepts establish metaphysical foundations of theories. They are the ground level building blocks of the language/story provided by each theory. Although as transcendental and universal as it sounds, theoretical physics is a human activity (at least in this planet) and in many cases it is strongly entangled with the concepts of the era theories are built in and this thesis is an example of such an influence.

Our language transforms together with the developments taking place in our daily life. Technology have always played a significant role in our understanding of world. There are strong correlations between technological advances and the concepts of physical theories developed in the same century. Technology has strong influence on how we perceive and express the world. Perhaps it is this influence that fundamental physics, in particular the quantum theory of gravity, today strives to employ the concept of information as the underlying building block of the Universe more than ever.

Maybe the messenger was Wheeler, in his memorable report [6] *Information, physics, quantum: the search for links*, he visioned the underlying entity of the existence as the information, semantic information in other words meaning. In a very different way this is the beginning of a new paradigm of unification. It is information that unifies all concepts of physics. According to this view, information is primary and underlies all other concepts of physics. Matter, energy, forces, fields, strings, space, time were all constructs of information. Although as speculative as it sounds, today this paradigm echoes in the community of high energy physics. The slight modification of the slogan coined by Wheeler is hailed between physicists. ***It from qubit.*** Information is represented by qubit, quantum version of a classical bit. Physics is far from what Wheeler envisioned, yet considering information as the primary underlying entity already opened new possibilities to understand the missing piece of the puzzle in radically different way.

We organized this introduction in three sections. In the first section 1.1, we provide a brief review of how information find its place in physics. The central theme will be the concept of entropy and how it evolves from thermodynamics to statistical physics followed by quantum mechanics. In addition we will introduce concept of adiabatic principle and its connection to information theory. We will introduce the first law of adiabatic principle as a reminiscent of first law of thermodynamics. In the second section 1.2, of the introduction, we discuss information theory in the context of black hole physics. This part of the introduction is connected to the first part of the thesis where firewall paradox is studied 3. In this section we start by introducing thermodynamics of black holes that leads to information paradox. We give a historical account on how information paradox evolved into modern form as firewall paradox. We lay out connection of recent ideas on firewall paradox to the idea advocated in the thesis. In the third section of the introduction 1.3, we present another dimension of information, namely its connection with spacetime geometry. Without going to back in the history, we present chain of events that leads us to the main research theme of this thesis.

1.1 Brief history of information in physics[1]

Information firstly appeared in physics implicitly through the concept of entropy in thermodynamics introduced by Rudolf Clausius to describe all transformations of a body through heat exchange with its environment. The concept firstly introduced as the mechanical equivalent of heat [7] for the formulations of first and second laws of thermodynamics. In its modern formulation, these laws are

$$\delta E = T\delta S \qquad\qquad \delta S \geq 0. \tag{1.1.1}$$

where S is the entropy and E is the energy of the system. Interestingly initial name proposed for entropy was *equivalence value* which indicates the identical character of energy and entropy at the infinitesimal level as established by the first law (1.1.1). As the definition, *mechanical equivalent* suggests, entropy at the

[1]The author is not an expert on the history of physics hence he will be presenting occasions/developments that has impact on his understanding of the universe from the perspective of information.

macroscopic level usually explained as the order/disorder of a physical body which in essence is an information theoretic notion.

The microscopic explanation behind the concept of entropy first introduced by the Ludwig Boltzmann. According to this definition, entropy provides a measure for the number of possible microscopic configurations a macrostate can be at.

$$S = k_B \log \Omega \tag{1.1.2}$$

where Ω is the number of microstates corresponding to a macrostate and k_B is Boltzmann constant. This definition in essence is an information theoretic one, as it quantifies the lack of knowledge on the microscopic configuration of the system. In other words, observer's lack of knowledge on the arrangement of molecules of a matter in a macrostate is the entropy. In an equivalent way, entropy is the uncertainty on the microscopic configuration of the body. These are all information theoretic statements and equivalent to each other. Unfortunately when Boltzmann formulated entropy in this way, the existence of atoms and molecules were quite controversial hence Boltzmann's discovery didn't get much attention on those days.

Boltzmann's entropy formula assumes that each microscopic configuration associated to a macrostate is equally likely hence it is not the most general form of a statistical entropy. At a given macrostate some configurations can be much more likely than others, such as low energy ones are more likely than high energy states at a given temperature. The generalization of (1.1.2) is known as Gibbs entropy.

$$S = -k_B \sum_i p_i \log p_i \tag{1.1.3}$$

where p_i is the probability of a microstate i to be realized in an ensemble. When all microstates are equally likely $p_i = 1/\Omega$, Gibbs entropy becomes (1.1.2). Once again Gibbs entropy measures observer's lack of knowledge about the underlying details of the system. This lack of knowledge manifest itself as an ensemble average, it is the primary factor on our experience regarding the world. The story is very interesting as it indicates how powerful the role of information on our description of the world. Not only the information itself, but even lack of information as it is also a form of information.

Let us introduce one more type of entropy known as Gibbs volume entropy which will be of interest in chapter: 7. Gibbs volume entropy [8] measures the entropy associated to the volume of the phase space. It provides an entropic measures even there is no degeneracy associated to the macroscopic state or at zero temperature. The volume entropy of Gibbs is defined as,

$$S = k_b \log \Omega(E) \qquad \Omega(E) = \sum_{E_i \leq E} n_{E_i} \tag{1.1.4}$$

where n_{E_i} is the degeneracy of an energy level. While Gibbs volume entropy counts all the states below a certain energy level, Boltzmann's entropy counts the number of microstates only at a certain energy. One can not distinguish between these two different definitions of entropy for a state having large degeneracy. Yet volume of the phase space can account for an entropic invariant even at zero temperature systems.

1.1.1 Adiabatic principle and reaction force

Adiabatic principle is one of the fundamental principles of nature that reflect itself in many different branches of physics. In essence it is about the fact that when two systems which are separated by a gap in the timescale hierarchy then these two systems can not efficiently communicate. In other words, a system adjust itself to an affect that is slow compared to the timescale of the system. The phase space is preserved in an adiabatic affection and usually referred as *adiabatic invariant*. Similar to thermodynamics, there is a first law that follows from adiabatic affection of the system system [9, 10].

The simplest system one can observe adiabatic principle is a harmonic oscillator. Consider a single harmonic oscillator whose frequency depends on a slow variable x, which is interpreted as the position of a probe that can be affected infinitely slowly. Under such deformation one can determine the reaction force on the probe through an information theoretic relation similar to the first law of thermodynamics.

The principle has also an application in quantum mechanics. Using the standard semi-classical correspondence one learns that the Bohr-Sommerfeld integral

$$J = \frac{1}{2\pi} \oint p dq \tag{1.1.5}$$

remains constant during an adiabatic affection of the system. The integral is taken over a closed orbit in the phase space which becomes $J = E/\omega(x)$ for an harmonic oscillator. For a classical harmonic oscillator, such as pendulum, the action J corresponds to average kinetic energy for a period, which is the volume of the phase space, and is invariant under adiabatic deformation of the system. Semi-classically the action satisfies the Bohr-Sommerfeld quantization condition. The behaviour of energy levels with respect to the slow deformation is presented in figure: 1.1 where $J = (n + \frac{1}{2})\hbar$ and serves as an invariant.

These statements hold not just for the harmonic oscillator, but for any dynamical system with one degree of freedom with a slowly varying Hamiltonian $H(p.q; x)$. The invariance of the action integral J can also be derived by pure classical means and follows from the fact that under slow changes of a dynamical system the phase space volume contained inside the classical orbit remains invariant. Indeed, J is precisely equal to that volume. Let us now explain how this can be used to determine the reaction force F. Invariance of the phase space leads to

$$F = -\left(\frac{\partial E}{\partial x}\right)_J \tag{1.1.6}$$

where the quantity in the subscript is fixed. Force can be expressed in alternative ways, by using the state equation of action variable. For this, the following standard relation between the quantities E and J and the angular frequency ω of the closed classical orbit is used,

$$dE = \omega dJ - F dx. \tag{1.1.7}$$

This identity, which follows from the action principle and Hamilton-Jacobi theory played an important role in Bohr's derivation of the quantization rule from the correspondence principle. In fact, it looks very much like the first law of thermodynamics, but the quantities ω and J have clearly a very different meaning than their thermodynamic analogues T and S. We would like to emphasize that, it is not merely an epistemological difference, especially in the framework of gravity. We don't speak of thermodynamical entropy while studying the microscopics of spacetime that is locally vacuum, it only enters when one studies states that are highly excited compared to vacuum and radiate thermally, such as black holes. Let us go back to adiabatic first law (1.1.7), and interpret alternative expressions one can extract

$$F = \omega \left(\frac{\partial J}{\partial x}\right)_E, \qquad \frac{1}{\omega} = \left(\frac{\partial J}{\partial E}\right)_x. \tag{1.1.8}$$

The beauty of the adiabatic principle is that one can extract information about the system without having the detailed knowledge of the dynamics. In the case of a single harmonic degree of freedom, the adiabatic first law contains the information about the spacing of energy levels. According to the correspondence principle the level spacing between the energy levels equals $\hbar\omega$, and that is reflected in the second equation above where the change in the energy with respect to phase space volume is equal to energy spacing between each level. The first equation of (1.1.8) on the other hand is analog of what has been proposed as the entropic force [11]. One should be careful on the interpretation of this equation. For example in the case of a harmonic oscillator there is no thermal character of the system. The systems is integrable and allowed motions follow fixed orbits in the phase space. Moreover even though one employs the force expression through (1.1.8) that does not mean that the action variable will increase in the direction of the reaction force. Rather it will serve as an invariant such that energy will adjust itself and cause a reaction force on the slow variable. Note that the force points in the

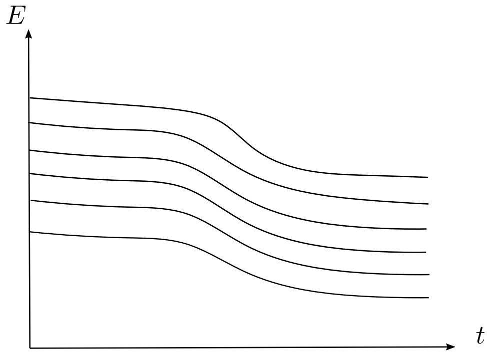

Figure 1.1: *Adiabatic principle in a quantum harmonic oscillator. When the frequency of the oscillator is deformed slowly over time, the energy of each eigenstate changes accordingly yet number of states below a certain energy level stays the same. Bohr-Sommerfeld quantization condition corresponds to adiabatic invariant of a quantum harmonic oscillator* $J = \frac{1}{2\pi} \oint pdq = (n + \frac{1}{2})\hbar$

.

direction in which the action would increase if the energy were to stay constant. But the energy has to adjust itself so that the action stays constant, which means that the energy has to decrease in that same direction.

Although we refer to the adiabatic regime in general, the same interpretation also holds in the more specific entropic regime. Namely an entropic force does not yield change in the entropy. The distinction, in terms of dynamics, is important when it comes to the information theoretic interpretation of the adiabatic invariant in gravity.

1.1.2 A new form of information

The early 20th century gave birth to two cornerstones of modern physics: quantum mechanics, describing physics of microscopic structures, and Einsteins general theory of relativity, marrying space, time, matter and energy in the dance of stars. Although unification, having a single description, a single principle, a fundamental truth is the dream of physicists, how to unify these two descriptions of nature is still an open problem.

Quantum mechanics is strange in many aspects but perhaps main departures of it from classical mechanics are the phenomena of entanglement and process of measurement. Measurement problem will not be discussed here. An entangled quantum state cannot be factored as a product of states of its local constituents, in other words, components of the system are not individual things but are an inseparable whole. There are different measures quantifying entanglement: the entanglement of formation, the entanglement cost, the distillable entanglement, the relative entropic measures, the squashed entanglement, log-negativity, the robustness monotones, entanglement negativity. Apart from log-negativity, the robustness monotones, entanglement negativity all reduce to von Neumann entropy for a pure state [12]. von Neumann entropy is a generalization of Gibbs entropy (1.1.3) to field of quantum mechanics.

$$S = -\mathrm{tr}(\rho \log \rho) \tag{1.1.9}$$

When the definition is applied to a density matrix obtained by tracing out a factor in a pure state, it measures the entanglement of state ρ with its complement. The connection with Gibb's definition can be observed clearly when density matrix is diagonalized in a suitable eigenbasis, $\rho = \sum_i p_i |i\rangle\langle i|$, where p_i denotes the probabilities of a measurement to outcome an eigenvalue E_i corresponding to eigenstate $|i\rangle$ according to Copenhagen interpretation.

$$S = -\sum_i p_i \log p_i \tag{1.1.10}$$

It is clear in this expression why von Neumann entropy is quantum mechanical extension of the Gibbs entropy in statistical physics. The difference between classical correlations and quantum entanglement manifest itself in the measurement process. Measurement of one of the parties in an entangled state determines the outcome of the other party. This fact dazzles the physicists in the early days of quantum revolution. Rather than understanding the origin of the phenomena, if there is, physicists accept it as the intrinsic characteristic of the nature. It is a prime example of concept creation in physics, the language of the classical physics has to be enriched. Also entropy of a classical system is always greater than entropy of the subsystems it composed of.

$$S_c(AB) \geq \max(S_c(A), S_c(B)) \tag{1.1.11}$$

which is definitely not the case for von Neumann entropy, as it vanishes for a pure state that might possesses entanglement between its subsystems. Entanglement entropy of a subsystem can also be considered as lack of knowledge, yet in a very different way than its classical counterpart. It is a lack of knowledge born out of the missing information on how subsytems of a system are connected to each other, which can dissolve when one access to entire state. Also, it is the information on how the measurements on one of the parties are dependent on or affect the measurements on the other party. Entanglement is also the primary ingredient in the phenomena of decoherence [13] that illuminates why we don't observe quantum superposition in classical world. Observers lack of knowledge how things are coupled to environment and on the environment itself result in the classical world[2].

The relation between energy and information in the form of entropy as the first law does not only exist in thermodynamics. It is an emergent law, that shows up in all the branches of physics. It represents the unified character of energy and entropy as information. There is also a first law associated to entanglement entropy. It can be derived from positivity of relative entropy (4.2).

$$\delta S = \delta \langle \hat{H} \rangle \tag{1.1.12}$$

where $\hat{H}_A = \log \rho$ is an hermitian operator, known as modular Hamiltonian (2.4). This relation is extensively studied in this thesis, hence we will not be introducing it further here.

1.2 Black holes and information

The other revolution of 20th century physics was the advent of general theory of relativity. Theory of relativity initiates series of unifications. First it unifies matter and energy then inertial mass with gravitational mass. Finally it connects spacetime with energy.

The connection between information and black holes goes back to the work of Bekenstein [14–16] built upon the works of Penrose [17], Christodoulou and Rufini [18, 19] later elaborated by Hawking [20, 21]. The complete thermodynamical analogy is laid out in a seminal work by Bardeen, Cooper and Hawking [22].

$$S_{BH} = \frac{A}{4G_{\text{N}}} \tag{1.2.1}$$

where G_{N} is gravitational constant which is in fact an information theoretic unit that measures number of microscopic degrees of freedom associated to a black hole. How does gravity knows about short distance physics as it is long distance phenomenon? If Wilsonian arguments were true, would gravity knew about short distance structure? This is a subtle question, subject to much debate.

Black holes like other macroscopic system acquiring some entropy, follows the laws of thermodynamics. The zeroth law states that black hole temperature is constant over the horizon. First and second law is more interesting and follows the same

[2] Decoherence itself is not enough to explain how only one of the possible outcomes realizes in a measurement process. Decoherence illuminates why we do not observed quantum superpositions.

form with ordinary matter that is thermalized.

$$\delta M = \frac{\kappa}{2\pi}\frac{\delta A}{4G_{\mathrm{N}}} \qquad\qquad \delta A \geq 0. \tag{1.2.2}$$

Black hole entropy is universal in the sense that it applies a whole zoo of black holes whether it is rotating or charged etc. Here we prefer to present the simplest version of the laws of black hole mechanics as it is sufficient to convey our message. In the derivation of second law, weak energy condition is assumed. Extension of the first law to general spacelike surfaces will be the main subject of this thesis.

The entropy associated to black hole horizon is based on thermodynamical arguments. Yet it is hard to believe the concept of continuum in a quantum world. Thermodynamical entropy is an emergent concept hence natural question to ask is what underlies black hole entropy. What are the degrees of freedom of a black hole? For a class of five-dimensional extremal black holes in string theory, the number of possible configuration one can wrap D-branes in extra dimensions corresponds to black hole entropy [23]. AdS/CFT, a realization of *holographic principle*, was the underlying mechanism of the success behind this counting. The number of microstates is counted through the dual CFT, using Cardy formula [24]. It provides the entropy of a state in a two-dimensional conformal field theory in a high temperature regime. In a nutshell,

$$S = 2\pi\sqrt{\frac{c}{6}\left(L_0 - \frac{c}{24}\right)} \tag{1.2.3}$$

where c is the central charge that corresponds to number of degrees of freedom in the theory and L_0 is the number of excitation. While counting microscopic states based on the dual CFT is great success of string theory, it doesn't reveal information about how degrees of freedom organized spatiotemporally. Are they localized on the horizon as bits encoded on the unit area of G_{N}, or are they non locally spread around the thermal atmosphere of black holes [25–27]? Maybe degrees of freedom are thread like structures that represent the flow of information or connectivity of spacetime [28]. These thread like structures ends up at the horizon from an outside perspective and the information that is obtained from the cross section of the tread corresponds to the area of black hole. We have shown in section 6, first law of black hole mechanics can also be understood based on a information flow picture. The problem of where the black hole information is subject to much debate, particularly centered around *firewall paradox* [29].

1.2.1 Black hole information paradox

The fact that black holes radiate like ordinary matter provides an answer to the old question: what happens to objects that fall into black hole? If black hole radiates then everything that falls into the black hole in the past will be emitted back as thermal radiation just like a piece of paper thrown into fire. Yet there was a problem in this hopeful logic. Unlike any other form of matter, black hole radiation seem to be perfectly thermal. In other words, it wouldn't matter if you threw a piano or Schrdinger's cat into the black hole, it is not possible to distinguish them in the radiation as it is same thermal distribution[3]. Initially it is proposed that black holes by their nature causes loss of information and there is no S-matrix for process of black hole formation and evaporation. Their dynamics are governed by operators that maps density matrices to density matrices [30]. It is hard for a physicist to give up fundamental principles of physics, hence this proposal is not widely accepted. The expected resolution of the paradox is to find out subtle corrections to the thermal radiation that purifies it.

One possibility that center of mass energy of the collision between ingoing and outgoing shells is Planckian for any radiation emitted after a time scale of $M \log M$ and shouldn't be neglected along the Hawking's computation [31, 32]. In addition it is shown that a naive calculation using CPT invariance and treating black holes as elementary particles is sufficient to derive the density of states of a black hole [32]. This calculation provides further evidence for existence of an S-matrix for black hole. Later 't Hooft could calculate an S-matrix based on the shift of the horizon due to massless shock waves representing infalling and outgoing particles [33]. These attempts were not enough to reach a consensus on the unitary formation and evaporation of a black hole. Susskind and Thorlacius point out that, unitary evolution, clones the information of an infalling shell on a special Cauchy surface as the information of outgoing Hawking radiation and information of the infalling object inside the black hole [34]. To prevent such violations, *black hole complementarity principle* is proposed [35, 36]. Black hole complementarity is a truly information theoretic principle. According to the principle, laws of physics has to be formulated in an observer-participant form. According to black hole complementarity, observers who individually access a copy of the infalling information, can never verify quantum cloning due to causality restrictions. Principle puts strong emphasis on observer-participant reality of physics. If no observer verifies a violation of any principle/law then how can one argue such law/principle is violated. Stronger version of the principle states that information theoretic content of the descriptions of an outside and infalling observer is equivalent [37]. However there is no clear mathematical formulation of the black hole complementarity.

[3] In fact it has correction in the form of greybody factor coefficients but these factors does not depend on the information of what has formed the black hole, hence these corrections do not affect the argument regarding the loss of information.

1.2.2 Firewall formulation of the paradox

Existence of a unitary S-matrix for black hole formation/evaporation was the way going forward for many physicists. However one could still formulate a different version of the information paradox under the assumption that evaporation is unitary. The new formulation is centralized around monogamy of entanglement [29]. In a nutshell, monogamy of entanglement states that, entanglement can not be shared. If two states are maximally entangled, individually, they can not be entangled with a third party. In the case of black hole evaporation, the monogamy of entanglement creates conflict between unitarity of the evaporation and local entanglement structure of the vacuum. Let us start by explaining the formation of entanglement and dissolution of entanglement in a unitary process.

Consider a partition of the Hilbert space as factors of black hole and Hawking radiation, $\mathcal{H} = \mathcal{H}_{BH} \otimes \mathcal{H}_R$. Given that initial state is pure, its entanglement entropy is zero. Moreover entanglement entropy of states in each factor equal to each other due to purity of initial state. During the black hole evaporation the entanglement between radiation and the black hole increases, since Hawking radiation can be considered to originate from entangled states near the horizon. However entanglement between radiation and the black hole can not increase indefinitely since at the end of the evaporation, the radiation states should be a unitary transform of the initial state hence entanglement entropy should vanish for the final state. This is known as the *Page curve* [38, 39] and it constitutes the essence of the firewall paradox. The point where the course grained black hole entropy becomes approximately equal to entanglement entropy of the black hole with its radiation is known as *Page time*. According to Firewall paradox, it is argued that at the Page time, there is no room for the entanglement of the vacuum near the horizon and this entanglement should be broken by formation of a firewall [29]. The firewall paradox manifests itself as violation of strong subadditivity of the entropy. Let A be early radiation, B be outgoing Hawking mode at Page time, and C be its interior partner mode figure: 1.2. In this partition, S_{AB} becomes black hole entropy and after the Page time, its entropy is less than radiation entropy, $S_{AB} < S_A$. Local vacuum structure of spacetime necessitates maximal entanglement and hence $S_{BC} = 0$ which further implies $S_{ABC} = S_A$. Using stong subadditivity one ends up in inconsistency $S_A \geq S_B + S_A$.

Many different counter arguments proposed against firewall. One possibility is the modification of the effective field theory (EFT) by non-local terms in a non violent way due to thermal atmosphere around the black hole [25, 40]. In other words, non-local modifications to the EFT should be comparable to the curvature scale such that local observers are not affected dramatically. Interestingly such theories might be within the observational window in near future [41]. Others

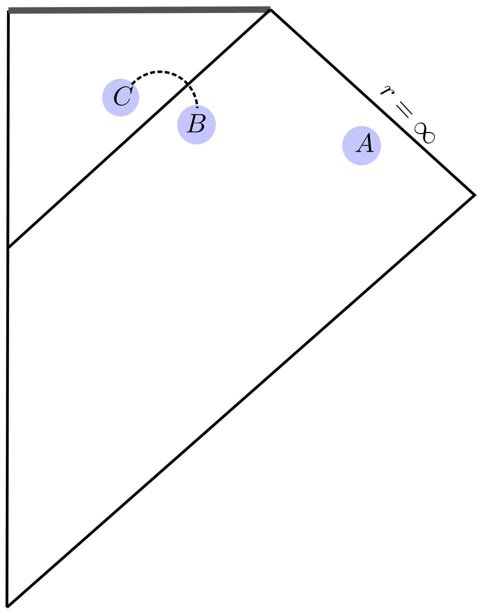

Figure 1.2: Illustration of the firewall paradox. A represents early radiation, B is outgoing Hawking mode at Page time, and C is its interior partner mode. Entanglement between B and C is a necessary condition for vacuum. Unitary evolution creates entanglement between A and the black hole. At the page time this entanglement saturates and becomes maximal hence there is no more space to sustain entanglement between B and C. Breaking entanglement between B and C tears apart the spacetime continuum by creating a firewall.

also claim that information is not localized on the ourgoing Hawking quanta but rather it is encoded on Hawking pairs in a non local way [27,42]. A protocol similar to quantum teleportation swaps entanglement on the Hawking pair to black hole and its radiation. During the entanglement entanglement swap information is also released. However there is no Hamiltonian provided which swaps the entanglement and construct Page curve.

In a similar philosophy, it is argued that subtle violations of locality could lead information to flow out of the black hole [43,44]. According to proposal locality in quantum gravity is an approximate notion and deviations from it can be observed by large number of operations insertion by a super-observer. They pointed out arguments regarding the strong subadditivity requires a clear factorization of the Hilbert space such that information regarding the early and late hawking mode can be associated to each factor. Yet they argue such factorizations and distribution of information is not possible for quantum gravity [45,46]. Moreover they also provide a mapping where one can construct operators for black hole interior, interestingly such operators are microstate dependent. Another interesting observation regarding whether one can manipulate black hole interior having access to early Hawking radiation comes from Susskind and Maldacena [47]. They suggested a thought experiment where one forms another black hole using early hawking radiation. These two black holes are expected to appear as an double sided eternal black hole geometrically connected by a wormhole representing a thermofield double state. Therefore one can throw some material to black hole formed by Hawking radiation and affects the interior of the other. This thought experiment provide some evidence that operators inside the black hole can be constructed by an outside observer. In fact this is what black hole complementarity is about.

Another possibility that prevents monogamy of entanglement is a modification of quantum mechanics by post selection in the form of a final state [48]. Yet it is hard to say such attempts were fruitful.

Many other approaches centralized around whether violations of monogamy is measurable by arbitrary number of observers that are restricted by principles of physics such as causality. These approaches focus on how the firewall paradox is formulated, rather than attempting to solve how information is recovered during the black hole evaporation. All approaches that are presented in this category is a good example of observer-participant nature of physics that is introduced in [6]. In this thesis we also follow such an approach. Let us first briefly introduced other attempts along the same direction.

Central argument in the firewall paradox can be reduced to a thought experiment where observers measure the entanglement of qubits that are supposed to be monogamous and compare their observations. However there are restrictions on whether such experiment can be conducted. Harlow and Hayden claimed that an observer needs to distill the qubit from ntire radiation to be able detect whether it is entangled. However distilling an information from early radiation is highly demanding in quantum computational sense and in average it requires a time that is more than the total evaporation time of the black hole [49]. Moreover it is argued that, Hilbert spaces should take into account observables that are restricted by computational complexity.

In this thesis, in chapter 3, our perspective regarding the firewall paradox is presented. Our approach is also in the class of ideas that suggest a more careful look whether the thought experiment verifying the violation of monogamy of entanglement is consistent with principles of nature within the domain of validity of EFT. One strong principle that needs to be respected by all EFTs is causality. Each observer has an effective description within their causal patch. We propose that a more restricted version of the complementarity constrained by the causal patches of observers should applied to black hole information paradox. Remember that, black hole complementarity is the equivalency of the information content of different observers in the process of black hole evaporation/formation. The problem of quantum cloning in the context of unitary black hole formation/evaporation was resolved via same kind of philosophy, namely observers can not verify such cloning because of causality restrictions. Firewall paradox make it harder to dispute the argument using causality simply moving the qubit that needs to be compared to near horizon regime. However we provide a careful analysis by restricting to physics within causal patches, and show that the experiment can not be conducted for a generic black hole within the domain of validity of EFT [1]. Our approaches provide a solution without introducing any new ingredient into to laws of physics, but rather applying them from an observer-participant point of view.

1.3 Spacetime geometry and information

Black holes are not the only objects of general relativity where information theoretic nature reveals itself. Ted Jacobson, in his remarkable work, showed that Einstein field equations can be expressed locally as the first law of thermodynamics [50]. In this work, Jacobson showed that Einstein equations are equivalent to entropic reaction of local Rindler horizons to thermal flux flowing through it. Since Jacobson used local arguments to derive Einstein equations, the derivation applies all spacetimes that satisfies Einstein equations locally.

$$T\delta S = \delta Q \qquad \Longleftrightarrow \qquad G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G_{\mathrm{N}} T_{\mu\nu} \tag{1.3.1}$$

This was probably the first incident where gravity is suggested to be an emergent phenomena. Emergent gravity paradigm also explains why gravity as a field theory is not renormalizable. On the other hand, the structure and nature of underlying degrees of freedom that gives rise to thermodynamics was a big mystery. It has understood two decades later, Jacobson's work could only be accounting for linearized gravity. This was the tip of an iceberg.

With the advent of AdS/CFT duality [51–53], a new window opened for physicists where they can gather more information regarding the underlying degrees of freedom of gravity. In a nutshell AdS/CFT states that gravitational theory on $d+1$-dimensional AdS background is dual to d dimensional conformal field theory that resides on the boundary of AdS. The gravitational side of the duality is usually referred as *bulk*. Gauge/gravity duality further encouraged emergent gravity paradigm. Many physicists interpret gravitational side of the duality as emergent from boundary CFT [54]. Some argued that duality should be seen different from emergence [55]. According to counter arguments, the direction of

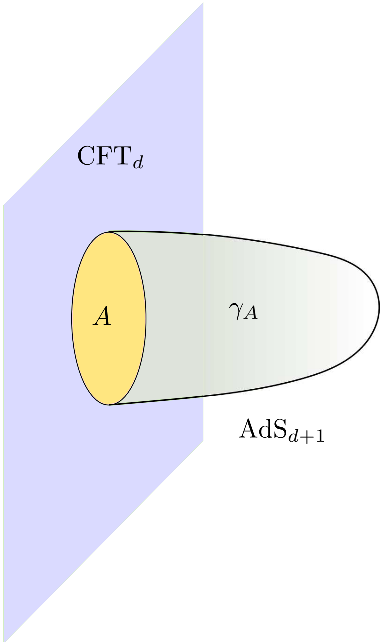

Figure 1.3: *Ryu-Takayanagi formula yields the entanglement entropy of a state in the CFT$_d$ through the area of the minimal surface γ_A in AdS$_{d+1}$.*

a standard emergence phenomenon is that of decreasing fundamentality[4] that is the emergence of less fundamental, high-level entities, out of more fundamental, low-level entities. However in a duality, a more fundamental entity can emerge out of a less fundamental one. In any case, AdS/CFT provides an exceptional framework to draw quantitative, rigorous conclusions. Author believes that final theory of emergent gravity will have a bulk based framework, however physics is probably not at that level of abstraction yet and has to draw lots of conclusions from AdS/CFT to bring about the quantum jump needed for such a description. In this thesis, we have tried to draw information theoretic conclusions in the bulk using lessons from AdS/CFT.

Within the AdS/CFT revolution, a mini revolution took place by the work of Ryu and Takayanagi. Their work revealed the deep connection between geometry and quantum information in the framework of holography [57, 58]. According to Ryu-Takayanagi proposal, von Neumann entropy of states on the CFT correspond to areas of minimal surfaces that are homologous to the boundary regions where

[4]Fundamentality in the sense of reductionism yet, perhaps more is different? [56]

states are defined on.

$$S\text{ent.} = \frac{A}{4G_{\mathrm{N}}} \tag{1.3.2}$$

The conjecture is first proposed for static solutions and later extended to time-dependent geometries [59]. Quantum information meets the geometry of spacetime in a very rigorous way through RT proposal. Soon after, it is shown that connected, continuous spacetime is the result of entanglement in the underlying theory [60]. If entanglement on the underlying theory is tuned to zero, spacetime starts to pinch off and separate into disconnected pieces [60], see figure: 1.4. In other words, *spacetime is entanglement.*

It is well known that Rindler horizons exhibits first law type relations similar to the first law of black hole thermodynamics. The first law of Rindler horizons in AdS translates to CFT as first of entanglement entropy, (1.1.12) [61]. The first law of entanglement entropy simply follows from positivity of relative entropy and satisfy in all quantum mechanical systems. It is shown that first law of entanglement in the CFT implies linearized Einstein equation in the bulk for a general theory of gravity [62]. Contrary to Jacobson's derivation, AdS/CFT version of the derivation was only applicable around the global vacuum. To generalize the derivation to higher order derivative theories of gravity, covariant phase space formulation is used [63–65]. Soon after, source term for linearized equation is introduced as a result of the quantum correction to entanglement entropy in the boundary. The quantum corrections to entanglement in the boundary corresponds to entanglement of quantum fields in the bulk [66]. One lesson learned from derivation of Einstein equation is that the first law of thermodynamics for Rindler horizons is the first law of entanglement entropy in the underlying theory. In the same spirit of what had been proposed by Popescu, Winter and Short on the foundations of thermodynamical entropy [67]. They proposed a decade ago that notion of thermal entropy is derived from entanglement entropy. This conclusion is very essential because Rindler horizons are not special surfaces, any surface on spacetime can be represented as Rindler horizons by accelerated observers. The generalization of the first law of entanglement to arbitrary spacelike surfaces in spherically symmetric asymptotically AdS solution in three dimensions in chapter: 5. In this generalization the notion of differential entropy is used. Differential entropy [68] can be seen as an extension of RT proposal to general surfaces in the bulk. One drawback of differential entropy is that it is only valid in three dimensions and for spherically symmetric situations. Although higher dimensional generalization is proposed [69], it is only applicable to planar symmetric geometries and hence it is trivial extension of 3-dimensional case. Our study on chapter: 5 clearly shows that the information theoretic nature of surface deformations in the bulk is not limited to minimal surfaces or Rindler horizons. It is much more general and applicable

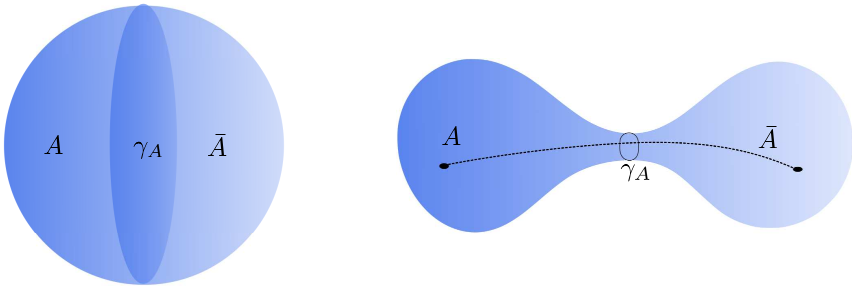

Figure 1.4: *Illustration of entanglement is spacetime idea. The geodesic distance between points on the complementary regions of the boundary of AdS_{d+1} increases when entanglement entropy between boundary regions A and $\bar{A}$ is reduced. The spacetime pinches off into disjoint pieces when entanglement vanishes.*

to any surface, however it is the CFT side of the story that gets complicated when one picks a general surface. In other words one can always map surface deformations to information theoretic relations in the underlying theory. We prove this statement at the linearized level in three dimensions. Along the thesis we tried to pull information theory from CFT to bulk. One attempt along this line is to derive the Bekenstein bound in the bulk [2]. The Bekenstein bound, [70, 71] is a limit on the entropy that can be contained in a physical system or object with a given size and total energy.

$$S \leq 2\pi RE \tag{1.3.3}$$

where R is the size of the box and E is the energy of the excitation. It is introduced in more detail in reviews: 2. Bekenstein bound has been derived for a QFT based on the positivity of relative entropy [72]. In the original work of Bekenstein, the derivation involves gravitational physics and also applicable to system that backreacts under the condition that system is weakly self gravitating. The information theoretic formulation of the bound in the bulk including backreaction is never touched upon in the literature. In 4 the Bekenstein bound in the bulk is derived. We clarify the definition of the size of the box in the presence of a back reaction. In this chapter we also studied distinguishing aspects of a mixed state and pure state excitations. Our motivation for such an analysis comes from the proposal that our universe follows a volume law entanglement due to its thermal nature [73].

1.3.1 The first law of spacetime deformations

The derivation of linearized Einstein equation around the empty AdS in a general theory of gravity displays the value of covariant phase formalism to establish first law like relations. The formalism initially used by Wald to interpret the black hole entropy as the Noether charge [63]. The formalism unifies different theories of gravity on their applications to black hole physics. In a general theory of gravity, the first law can be established based on two quantity, symplectic current and Noether charge independent of the details of the Lagrangian. Although formalism is much more general than its application to black hole horizons, its generality is not fully exploited in the literature. Perhaps it is because, the black holes are considered to be the only objects of general relativity where one can observe the thermodynamical nature. However, work of Jacobson [50, 74], Ryu-Takayanagi formula, entropic gravity proposal [11] and derivation of Einstein equations via AdS/CFT suggested that even empty space enjoys the first law. Natural question follows immediately is: first law of what? Thermodynamics, entanglement..? Along the thesis we have provide attempts to answer these questions by applying covariant phase space formalism to general codimension-2 spacelike surfaces.

The intuitive and mathematical connection between geometric theory of gravity and theory of elasticity has been studied widely in the literature [75–82]. Recently it has been proposed [73] that the underlying theory from which gravity emerges has two phases according to how entanglement entropy of the underlying theory is reflected geometrically. According to this proposal, volume law phase of the theory can be modeled as an elastic phenomena. In other words it is proposed that dark energy has elastic behaviour. To sum up there are deep analogies when gravity and elasticity are considered purely from an information theory perspective. Application of covariant phase space formalism yields the information theoretic analogy between them through the first law of thermodynamics. The connection between gravity and elasticity based on emergent information theoretic first laws broaden our understanding on the emergent gravity paradigm. It can help us to allocate microscopic degrees of freedom of the underlying theory onto spacetime to rebuild it.

By studying covariant phase space formalism, from an information theoretic perspective, on general spacelike surfaces, we have discovered its connection to Brown-York quasilocal charge densities [83, 84]. Although Wald himself also established this connection [65], it has never studied in the context of emergent first laws. On a general surface, change of quasilocal charge densities amounts to change on the entropy that can be measured on the surface. Integral version of the first law is one of the main results of our thesis. This relation is interpreted as the first law of deformations in spacetime based on its resemblance to first law of elastic deformations.

In [73] the entropy of cosmological horizon is distributed over the universe as entropy density per unit volume. Building on that postulate, galactic rotation curves are derived without the hypothesis of dark matter. Although phenomenology of the theory is limited to spherically symmetric systems, the connection presented in chapter 6 is established on more general grounds. Our results provides necessary tools on the gravitational side to go beyond spherical symmetry.

Finally application of covariant phase space formalism in a vector field representation clarifies the identity of information contents of general homologous spacelike surfaces. In chapter 6, we present a connection between bit threads picture of holographic entanglement entropy [28] and covariant phase space formalism in its vector field representation. We conclude that gravitational fields are carriers of information and they act as prefect conductors of information.

1.3.2 Origin of inertia and adiabatic principle

The conceptual underpinnings of emergent gravity paradigm is structured in the work of Verlinde [11]. It was proposed that, gravity should not be treated as fundamental force rather should be derived from the emergent information theoretic notions. Not only Einstein field equations, but also the postulate of geodesic motion and even other forces of the nature follows from the same principle according to the proposal. It is important to distinguish gravitational field equations from postulate of geodesic motion. Geodesic motion enters to theory of relativity as a postulate.

Verlinde provide a sketch of a derivation of Newton's law of inertia based on simple assumptions about emergent aspects of the underlying theory, such as area law behaviour of entanglement entropy and form of equipartition theorem. The idea is criticized based on the relation of entropy and the force, that would require increase of entropy indefinitely even in orbital motions. We clarified these misconceptions originated after the work of Verlinde, in chapter: 7.

The idea that geodesic motion emerges as a result of adiabatic invariants in the underlying theory is used to explain the missing mass problem without the hypothesis of dark matter in [73]. While adiabatic invariant scales with areas of surfaces in the ordinary gravity, it has a volume like component that starts to be dominant in large scales. Based on such motivations we elaborated the connections between entropic gravity proposal and general relativity in chapter: 7.

2 Reviews

2.1 Symplectic mechanics

Symplectic mechanics is central formalism for the covariant phase formulation of gravity. In the Hamiltonian approach to field theories, one specifies a time coordinate and a Cauchy surface, which enforces a decomposition between space and time, therefore breaks the covariant form of general relativity. On the other hand, covariant phase space formalism that is based on the symplectic structure of Hamiltonian mechanics, preserves the covariant structure of the theory. In this short section, we will give a short introduction on the main concepts of the symplectic mechanics.

A *sypmlectic manifold* is a smooth manifold of even dimensionality, $(\mathcal{M}, \Omega)$ equipped with a *symplectic structure*, Ω. Symplectic structure is closed, non-degenerate, 2-form on $\mathcal{M}$. The condition that Ω being non-degenerate means, the only vector field, X on $\mathcal{M}$, satisfying $X \cdot \Omega = 0$ is the one vanishing uniformly. The sypmlectic structure takes the form, $\Omega = dp^i \wedge dq_i$, in a special coordinate system (q_i, p^i), known as Darboux frame. It is always possible to bring symplectic structure into this form locally. The connection between symplectic structur e and Hamiltonian dynamics is made through the Poisson brackets. The inverse of the symplectic structure exists due to its non-degeneracy which is used used to define the Poisson bracket on $\mathcal{M}$.

$$\{f, g\} = \Omega^{ab} \partial_a f \partial_b g \tag{2.1.1}$$

where Ω^{ab} is defined through the relation, $\Omega^{ac}\Omega_{cb} = \delta^a_b$. The relation between Poisson bracket and symplectic structure becomes manifest in the Darboux chart. The definition through sypmlectic structure satisfies the properties of the Poisson bracket. Specifically Jacobi identity is ensured by the closedness of Ω.

Given that, Hamiltonian, H is a smooth function, then unique vector field ξ_H satisfying, $\xi \cdot \Omega = dH$ is called Hamiltonian vector field. In other words Hamiltonian determines a vector field, $\xi^a = \Omega^{ab}\partial_b H$, that generates a congruence, $\gamma(t)$ over the phase space, where t is the parameter a along the congruence.

$$\{f, H\} = \Omega^{ab}\partial_a f \partial_b H = \mathcal{L}_\xi f = \frac{df}{ft} \tag{2.1.2}$$

where $\xi^a = \frac{dx^a}{dt}$.

The symplectic form can be used to define symplectic symmetries over the phase space. A vector field ξ defines the symplectic symmetries through the following equation.

$$\mathcal{L}_\xi \Omega = 0 \tag{2.1.3}$$

which implies through the Cartan's formula,

$$d(\xi \cdot \Omega) = 0, \qquad \xi \cdot \Omega = dH_\xi \tag{2.1.4}$$

hence there exists a function, H_ξ which is the generator of the evolution along the symmetry vector field, ξ through the Poisson bracket. The evaluation of H_ξ over the equations of motion yields the conserved charge.

2.1.1 Symplectic structure of gauge theories

Here we will explain how the sypmlectic structure explained in the previous part has been adapted to the field theories with local gauge symmetries. The discussion is useful to understand the details of the formalism which we will use to study the first law of black hole mechanics on general surfaces. The connection between the perturbation on the Hamiltonian and the presymplectic form is essential for the generalization of first law type relations.

Any field configuration $\Phi(x)$ in the spacetime corresponds to a point in the phase space. Field configurations doe not need to satisfy equations of motion. On-shell field configurations forms a subspace in the phase space. An infinitesimal field perturbation, denoted by $\delta\Phi$ over a configuration Φ corresponds to tangent vector in the phase space at the point $\Phi(x)$. Let us denote the tangent vectors in the phase by $(\delta\Phi)^A$, A runs over the coordinate system covering the phase space. When the entire phase space is considered for gauge theories, it has degeneracy directions as a result of the local gauge symmetry. Reduction over the entire phase space by a symplectic quotient yields physical phase space (Γ, Ω). Let us make the connection with a field theory that is given through a Lagrangian description. We will denote all the field content of the theory by $\Phi(x)$. The presymplectic potential, which is a $(d-1)$-form in spacetime and 1-form in the phase space is defined via the field variation of the Lagrangian, $\boldsymbol{L}$ by,

$$\delta \boldsymbol{L} = \boldsymbol{E}\delta\Phi + d\Theta(\delta\Phi, \Phi) \tag{2.1.5}$$

δ can be view as exterior derivative in the space of field configuration, since $\delta_1\delta_2\Phi - \delta_2\delta_1\Phi = 0$. The definition of the presymplectic structure is ambiguous up to an exact form, which does not change the field variations of the Lagrangian. The presymplectic current, $\omega(\delta_1\Phi, \delta_2\Phi)$, is a $(d-1,2)$ form defined[REF] by,

$$\omega(\delta_1\Phi, \delta_2\Phi, \Phi) = \delta_1\Theta(\delta_2\Phi, \Phi) - \delta_2\Theta(\delta_1\Phi, \Phi) \tag{2.1.6}$$

The covariance of the framework is hidden in the closedness of the presymplectic current when Φ corresponds to a field variation satisfying equation of motion and $\delta_{1,2}\Phi$ satisfies linearized e.o.m.

$$\begin{aligned} d\omega(\delta_1\Phi, \delta_2\Phi, \Phi) &= \delta_1 d\Theta(\delta_2\Phi, \Phi) - (1 \leftrightarrow 2) \\ &= \delta_1(\delta_2\boldsymbol{L}(\Phi) - \boldsymbol{E}(\Phi)\delta_2\Phi) - (1 \leftrightarrow 2) \\ &= \delta_1\boldsymbol{E}(\Phi)\delta_2\Phi - \delta_2\boldsymbol{E}(\Phi)\delta_1\Phi \approx 0 \end{aligned} \tag{2.1.7}$$

Symplectic form in the physical phase space is defined by,

$$\Omega = \Omega_{AB}(\delta_1\Phi)^A(\delta_2\Phi)^B = \int_C \omega(\delta_1\Phi, \delta_2\Phi, \Phi) \tag{2.1.8}$$

The integral is defined over a spacelike surface. When the field configuration on the boundary of the manifold is fixed, the sypmlectic form does not depend on the choice of the hypersurface for on-shell field configurations, which can be shown by Stokes theorem together with closedness of presypmlectic current. Therefore we have concluded that presymplectic form is same for every hypersurface which is a necessary consequence of the covariant phase space approach. Now we are ready to prove the following lemma.

Lemma 2.1.1. *Variation of the generator of local gauge transformation generated by the vector field ξ takes the following form*

$$\delta H_\xi = \int_C \omega(\delta\Phi, \delta_\xi\Phi, \Phi) \tag{2.1.9}$$

*Proof.*Let us show that, the definition (2.1.9) leads to generators along the ξ, in the Poisson algebra. Using the notation introduce at (2.1.8),

$$\begin{aligned} \{F(\Phi), H_\xi\} &= \Omega^{AB}\frac{\delta F}{\delta\Phi^A}\frac{\delta H}{\delta\Phi^A} \\ &= \Omega^{AB}\Omega_{BC}(\delta_\xi\Phi)^C\frac{\delta F}{\delta\Phi^A} \\ &= \delta_\xi F \end{aligned} \tag{2.1.10}$$

hence H_ξ is the generator of the gauge transformation.

2.2 Covariant phase space formulation of gravity

In this section we sketch Wald's derivation of black hole entropy as being the diffeomorphism Noether charge for the horizon generating Killing field, evaluated at the bifurcation surface. Along the thesis, formalism will be used in different settings. In chapter 6 we will not constrain ourselves to black hole horizons or bifurcation surfaces and use the formalism in a more general setting in which the surface can be chosen arbitrarily. In 6, we will show that the first law of black hole thermodynamics is a special case of more general relation between the surface charges and total energy of spacetime. We will use the covariant phase space approach to understand the variational relation between surface charges and the energy of the system.

Wald's derivation applies to any diffeomorphism invariant theory defined by a Lagrangian D-form $\boldsymbol{L}$, where D is the spacetime dimension. The variation of $\boldsymbol{L}$ induced by a field variation is given by,

$$\delta \boldsymbol{L} = \boldsymbol{E}\delta\phi + d\boldsymbol{\Theta}(\phi, \delta\phi) \tag{2.2.1}$$

where dynamical fields denoted collectively by ϕ and $\boldsymbol{E}$ stands for the field equation.

$$\boldsymbol{E}\delta\phi = \boldsymbol{E}_g^{ab}\delta g_{ab} + \boldsymbol{E}_{\psi}\delta\psi = 0$$

where a sum over the matter fields ψ is understood. $\boldsymbol{E}_g$ and $\boldsymbol{E}_\psi$ are locally constructed out of the dynamical fields ϕ and their derivatives. Equations of the motion of the theory are,

$$\boldsymbol{E}_g^{ab} = 0 \text{ and } \boldsymbol{E}_\psi = 0$$

The $D-1$ form, $\boldsymbol{\Theta}$, defined by eq.(2.2.1) is called "symplectic potential". It is locally constructed out of ϕ, $\delta\phi$ and their derivatives and is linear in $\delta\phi$. The anti-symmetrized field variation of $\boldsymbol{\Theta}$ defines another $D-1$ form called "symplectic current".

$$\boldsymbol{\omega}(\phi, \delta_1\phi, \delta_2\phi) = \delta_1\boldsymbol{\Theta}(\phi, \delta_2\phi) - \delta_2\boldsymbol{\Theta}(\phi, \delta_1\phi) \tag{2.2.2}$$

The symplectic form Ω on the phase space of solutions is defined relative to a Cauchy surface, C.

$$\Omega = \int_C \boldsymbol{\omega}(\phi, \delta_1\phi, \delta_2\phi)$$

Keeping these in mind, we now proceed to the demonstration of conserved quantities associated to diffeomorphisms. Consider the variation induced by a diffeomorphism generated by a vector field, ξ,

$$\delta_\xi \phi = \mathcal{L}_\xi \phi$$

Since we will be concerned only with diffeomorphism invariant theories, the Lagrangian will be diffeomorphism covariant in the sense that action induced by diffeomorphism generating vector field on the Lagrangian, is equivalent to action of the generator on the field content of the theory, $f^* \boldsymbol{L}(\phi) = \boldsymbol{L}(f^*(\phi))$, where f^* is the action induced by diffeomorphism, $f^* : M \to M$. Therefore

$$\delta_\xi \boldsymbol{L} = \mathcal{L}_\xi \boldsymbol{L} = d(i_\xi \boldsymbol{L})$$

in the second equality Cartan's formula, $\mathcal{L}_\xi \alpha = \xi \cdot d\alpha + d(\xi \cdot \alpha)$ is used together with the fact that $\boldsymbol{L}$ is a D form and hence its exterior derivative vanishes. Note that variation on the Lagrangian is a total derivative. The vector fields, ξ, generate symmetries of the dynamics and each ξ is associated with an $(D-1)$ form called the Noether current defined using eq.(2.2.1)

$$\boldsymbol{J}_\xi = \boldsymbol{\Theta}(\phi, \mathcal{L}_\xi \phi) - \xi \cdot \boldsymbol{L} \tag{2.2.3}$$

Note that Noether current is conserved on-shell.

$$d\boldsymbol{J}_\xi = -\boldsymbol{E} \mathcal{L}_\xi \phi \tag{2.2.4}$$

This implies that Noether current is closed when the equations of motion are satisfied. Since $\boldsymbol{J}$ is closed for all diffeomorphism generating vector fields, ξ, it follows that there exists an $(D-2)$ form, $\boldsymbol{Q}$- locally constructed out of the fields appearing in $\boldsymbol{L}$ and ξ^μ such that when evaluated on solutions to the equations of motion, we have

$$\boldsymbol{J}_\xi \approx d\boldsymbol{Q}_\xi \tag{2.2.5}$$

where, "$\approx$" symbol denotes the on-shell equality. $\boldsymbol{Q}_\xi$ is $(D-2)$-form constructed locally from the fields and their derivatives. The integral of $\boldsymbol{Q}_\xi$ over a closed $(D-2)$ surface S is called the "Noether charge" of S relative to ξ. The full off-shell expression is given in the appendix.

The formalism is particularly useful on the construction of the "first laws". There exist a $D-2$-form, $\boldsymbol{\chi}$, which is closed when the e.o.m. and linearized one is satisfied. Then constructing the the "first law" is actually equivalent to implementation of the Stokes theorem. The fundamental theorem of covariant phase space formulation of field theories which we have demonstrated at (2.1.7) shows that symplectic current is an exact form on-shell including the linearized solutions of the perturbed field configurations.

$$d\boldsymbol{\chi}_\xi \approx \boldsymbol{\omega}(\phi, \delta\phi, \mathcal{L}_\xi\phi) \tag{2.2.6}$$

The on-shell equality holds for any vector field ξ. On the other hand, the form $\boldsymbol{\chi}$ becomes closed when ξ is a Killing vector field. To obtain the explicit expression for a closed $D-2$-form $\boldsymbol{\chi}$, let's start with the expression for the symplectic form.

$$\Omega = \int_C \boldsymbol{\omega}(\phi, \delta\phi, \mathcal{L}_\xi\phi) \tag{2.2.7}$$

$$= \int_C \delta\boldsymbol{\Theta}(\phi, \mathcal{L}_\xi\phi) - \mathcal{L}_\xi\boldsymbol{\Theta}(\phi, \delta\phi) \tag{2.2.8}$$

$$= \int_C \delta\boldsymbol{J}_\xi + \delta(i_\xi\boldsymbol{L}) - i_\xi d\boldsymbol{\Theta} - d(i_\xi\boldsymbol{\Theta}) \tag{2.2.9}$$

$$= \int_{\partial C} \delta\boldsymbol{Q}_\xi - \xi\cdot\boldsymbol{\Theta} \tag{2.2.10}$$

In the second line we used the definition of symplectic current, (2.2.2) in the third line (2.2.3). Second and forth terms in the third line cancel each other on-shell. In the forth line, we used $\delta\boldsymbol{J} = \delta d\boldsymbol{Q}$, which is true when the linearized equation of motion is satisfied in addition to e.o.m. The vector field is rigid under the variation and is used to compare two manifolds: initial and perturbed one.

If ξ generates symmetry of the field configuration, $\mathcal{L}_\xi\phi = 0$, symplectic current vanishes and yields an identity relating the surface term variations away from that solution. Note that initial surface does not need to be a complete Cauchy surface. The identity, $\int_{\partial C}\chi = 0$, can be obtained for the boundaries of a partial Cauchy surface, as a result of the Stokes' theorem. Namely one can construct a first law, not only for the black hole solutions but for any surfaces in the bulk. The question of what are the physical quantities in the surface term variations for an arbitrary surface are studied in 6.

Moreover in section 6.A, we will describe Wald formalism in the frame field (a la Cartan) formulation of general relativity. It will simplify the expressions for surface term variations.

2.3 Differential entropy

The notion of differential entropy is an attempt to generalize the Ryu-Takayanagi proposal (5.1.2) to more general surfaces in the bulk. Although it is clear that minimal surfaces measure entanglement in the CFT, there is no proof if the area of the minimal surface also measures the entanglement between the degrees of freedoms that one can establish in the bulk. It seems logical to assume that area as a classical notion emerges from the amount of entanglement of degrees of freedom that builds the spacetime. Starting from this postulate, the minimal surface entanglement (5.1.2) immediately generalizes to any surface in spacetime. Moreover, such an assumption leads one to realize that, entanglement of the underlying state is encoded much more naturally in the bulk, and only a special class of bulk surfaces manifest themselves as a spatial entanglement in the dual CFT. This is almost an artifact of the current formulation of quantum gravity, namely AdS/CFT. While extremal surfaces are not special from the bulk point of view they are very special from the CFT point when entanglement entropy is considered.

There are different proposals and constructions in the direction of establishing d.o.f. of the microscopic theory within the bulk [69, 85–87]. The RT proposal provides a benchmark for these constructions. Adopting a similar point one can still ask how a generic codimension two spacelike bulk surface is encoded in the boundary CFT. The question has an answer in $3d$. This boundary quantity is called *differential entropy*. The discrete version is defined as,

$$S_{DE} = \sum_{k=1}^{n} \left[S(I_k) - S(I_k \cap I_{k+1})\right] \tag{2.3.1}$$

where the intervals I_k resides on a time slice in the boundary. It was first observed in [88] that S_{DE}, in the continuum limit yields the area of the bulk curve which is defined through the tangency of each extremal surface homologous to a boundary region I_k. In other words *the hole* is defined by the region that can not be accessed by any of the boundary observers having access to $\{I_k\}$. Although higher dimensional extensions are constructed, [89] they are related to examples AdS_3 through the planar symmetry. Before proving that differential entropy converges to the area of the hole, let us look at the expression in the continuum limit

$$S_{DE} = \frac{1}{2}\int_0^{2\pi} d\theta \, \left(d_\alpha S(\alpha)\right)_{\alpha=\alpha(\theta)} \tag{2.3.2}$$

where $d_\alpha := \frac{d}{d\alpha}$. The $\alpha(\theta)$ represent the half angular width of each geodesic centered at θ on the boundary. Note that constant time slice of the boundary is S^1. Depending on the shape of the hole α varies at each angle θ_0. For a spherical (circular) hole, $\alpha(\theta) = \alpha_0$. We will present expressions in most general case yet explicit calculations will be provided for spherical holes. The differential entropy can also be expressed in terms of the conditional information between I_k/I_{k+1} and $I_k \cap I_{k+1}$ [90]. Now, let us demonstrate differential entropy on a simple exercise, the hole-ographic construction of [68, 88] namely the spherical hole in AdS. We will use a coordinate system in which the entire solution fits into a finite radius.

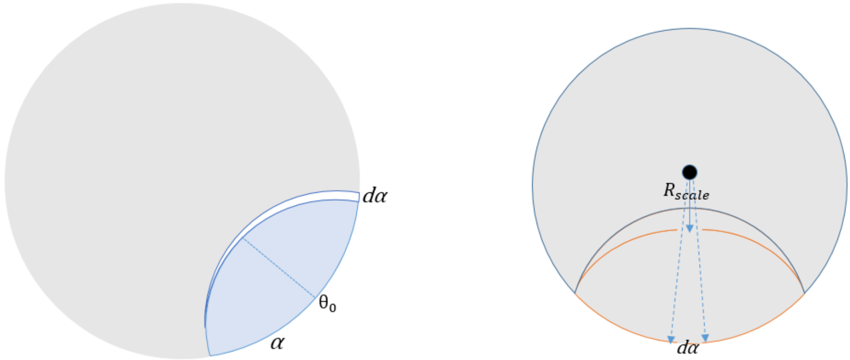

Figure 2.1: *On the left, we see holographic entanglement entropy of two regions of size 2α and $2\alpha + d\alpha$. Figure on the right stands for geometric relation between conditional information, $S(d\alpha|\alpha)$ and radial direction, or radial arc of size $d\alpha$. The relation holds also for excited states. The black point represents the conical defect.*

A consconvenient choice is hyperbolic slicing of AdS_3.

$$ds^2 = L^2 \left[\frac{4(dr^2 + r^2 d\theta^2)}{(1 - r^2)^2} - \left(\frac{1 + r^2}{1 - r^2} \right)^2 dt^2 \right] \tag{2.3.3}$$

one can easily see that for $t = 0$, $\sinh(\rho) = 2r/(1 - r^2)$ and $\theta = \tilde{\theta}$, the metric becomes, $ds^2 = d\rho^2 + \sinh^2 \rho d\tilde{\theta}^2$ which is the metric of the hyperbolic space. The nice property of (2.3.3) is that, in this system spacelike geodesics on a constant time slice are circles intersecting the boundary S^1. Consider the geodesic with endpoints denoted by $\gamma_L(\lambda) = u$ and $\gamma_R(\lambda) = v$. Let's initially consider geodesics with fixed opening angle on the boundary, namely $u - v = 2\alpha_0$, α_0 denotes the half opening angle of the boundary geodesic on the boundary. Then associated entanglement entropy of the state os the subsystem determined by, α and the complement, on the boundary CFT is given by

$$S(u, v) = \frac{c}{3} \log \left(\frac{2L}{\mu} \sin \alpha_0 \right) \tag{2.3.4}$$

The quantity, c is the central charge and is related to gravitational constants by $c = \frac{3L}{2G_N}$ [91]. μ is the UV cutoff set by the radial coordinate of the AdS_3. Removing the cuttoff would cause the entanglement entropy to diverge because of the large number of UV modes that contribute near the surface. Using (2.3.2), and considering geodesics with fixed opening angle α_0 on the boundary,

$$S_{DE} = \frac{1}{4G_N} \int_0^{2\pi} d\theta \, \cot \alpha_0. \tag{2.3.5}$$

Using the fact that geodesics are circles intersecting the boundary, one can easily obtain the location of the tip of the geodesic. This will yield the radius of the hole since for fixed geodesic size. The tip of each geodesic is identical to to point of tangency of the geodesic to the hole. $r_{\text{tip}} = 1/\cos\alpha_0 - \tan\alpha_0$ and using the relation between global coordinates and hyperbolic slicing $R = \frac{2r}{1-r^2}$ one can easily see that $R_{\text{hole}} = \frac{\cos\alpha_0}{\sin\alpha_0}$. Therefore we have shown that,

$$S_{DE} = \frac{1}{4G_{\text{N}}} 2\pi R_{\text{hole}} \tag{2.3.6}$$

which is the area of the hole, hence $S_{DE} = S_{BH}(R_{\text{hole}})$. Explicit computation is carried for a hole of fixed radius but we will prove that differential entropy always reproduces the gravitational entropy in more general settings. We will demonstrate this connection using the integral geometry in section 5.2.

2.4 Modular hamiltonian

Consider a smooth entangling surface Σ which divides a quantum state spatially into two parts, a region A and its complement $\bar{A}$. Upon integrating out the degrees of freedom in $\bar{A}$, one is left with the reduced density matrix ρ describing the remaining degrees of freedom in A. The reduced density matrix is both hermitian and positive semidefinite, hence it can be expressed as,

$$\rho = e^{-H} \tag{2.4.1}$$

where H is some hermitian operator, known as modular Hamiltonian.

The modular or entanglement Hamiltonian plays a central role in computing relative entropies and the first law of entanglement entropy. In most of the cases, entanglement Hamiltonian does not admit a local expression. There are few special cases when the modular Hamiltonian may be expressed as an integral over the local energy-momentum tensor of the field theory. A well known example modular Hamiltonian for the vacuum of a QFT defined on a d-dimensional Minkowski space $\mathbb{R}^{1,d-1}$. When state ρ corresponds to the vacuum density matrix of half-space $x_1 > 0$, modular Hamiltonian becomes,

$$H_A = 2\pi \int_A x_1\, T_{00}(x)\, d^{d-1}x. \tag{2.4.2}$$

Modular Hamiltonian is just the boost generator in the x_1 direction. For a CFT this result can be mapped to the modular Hamiltonian of $(d-2)$ dimensional sphere of radius R in the vacuum. This is possible due to special conformal transformation together with translation which leaves the vacuum invariant yet maps the half space to ball shaped region. In this case the expression is given by,

$$H_A = 2\pi \int_A \frac{R^2 - r^2}{2R} T_{00}(x)\, d^{d-1}x \tag{2.4.3}$$

where A becomes the ball bounded by S^{d-2} centered around the origin in Poincare coordinates. We will use this expression to provide consistency checks of approximate modular Hamiltonian for the states describing conical defect solutions (5.4.6).

In $2d$ CFT, it is even possible to obtain exact expression of modular Hamiltonians for a finite interval in an infinite system at finite temperature. These states are dual to BTZ black holes with a planar symmetry in the context of AdS/CFT. Differential entropy of such states are studied in the appendix 5.A. Let us also provide the derivation of these modular hamiltonian here.

Consider a finite interval of size $2R$ on an infinite system. This setup is conformally equivalent to annulus when a regularization scheme applied on the endpoints of A. The modular Hamiltonian becomes the conserved charge associated to a conformal boost vector ζ on the Euclidean thermal cylinder:

$$H_A = \int_A d\Sigma^\mu\, T_{\mu\nu}\zeta^\nu. \tag{2.4.4}$$

To find ζ explicitly, choose a complex coordinate z on the thermal cylinder with periodic identifications $z \sim z + i\beta$. Region A is the segment $\mathrm{Re}(z) \in [-R, R]$ at $\mathrm{Im}(z) = 0$. The conformal mapping $\alpha = \left(\frac{e^{2\pi z/\beta} - e^{-2\pi R/\beta}}{e^{2\pi R/\beta} - e^{2\pi z/\beta}}\right)$ sends this region to half line on complex α plane. The conformal map which sends half plane to the annulus then becomes, $\omega = \log(\alpha)$. Therefore overall map which sends the finite interval in thermal state to the annulus is given by,

$$\omega = f(z) = \log\left(\frac{e^{2\pi z/\beta} - e^{-2\pi R/\beta}}{e^{2\pi R/\beta} - e^{2\pi z/\beta}}\right) \tag{2.4.5}$$

The entanglement Hamiltonian on the annulus simply the generator of translations around the annulus in the direction $v = \mathrm{Im}(\omega)$.

$$H_A = \int_{v=\text{constant}} T_{vv} du = \int_{f(C)} T(\omega) d\omega + \int_{f(\bar{C})} \bar{T}(\bar{\omega}) d\bar{\omega} \tag{2.4.6}$$

which becomes

$$H_A = \int_C \frac{T(z)}{f'(z)} dz + \int_{\bar{C}} \frac{\bar{T}(\bar{z})}{f'(\bar{z})} d\bar{z} \tag{2.4.7}$$

The factor of $f'(z)^{-1}$, is the product of the $f'(z)^{-2}$ occurring in the transformation rule for T, and the Jacobian factor $f'(z)$. Let us use this expression to check the half interval modular Hamiltonian expression. $\omega = \log(z)$ is the map to annulus for semi-infinite region, which yields the well known expression $H_A = 2\pi \int xT_{tt}(x)dx$, since $f'(z) = \frac{1}{z}$.

If we use (2.4.7) for the coordinate transformation (2.4.5) that maps the finite interval $A : [x_0 - R, x_0 + R]$ in an infinite thermal system to an annulus, one obtains the following modular Hamiltonian

$$H_A = 2\beta \int_A \frac{\sinh\left(\pi(R + x_0 - x)/\beta\right) \sinh\left(\pi(R + x - x_0)/\beta\right)}{\sinh\left(2\pi R/\beta\right)} T_{00}(x)dx. \quad (2.4.8)$$

2.5 Bekenstein bound

The Bekenstein bound, [70, 71] is a limit on the entropy that can be contained in a physical system or object with a given size and total energy. The heuristic yet deep derivation of the bound employes the black hole thermodynamics together with *generalized second law*(GSL). Generalized second law is the extension of the the second law of thermodynamics to the systems involving black holes. The law states that the environment and black hole system together evolves in such a way that total entropy of the combined system does not decrease. Bekenstein derived the bound mainly employing this principle.

Let us have a look at the following *gedanken* experiment. A composite system of radius R with total energy E and entropy S_{box}, falls into a Schwarzschild black hole black hole of mass, M where $M \gg E/c^2$, such that temperature of the black hole stays same in the process. Suppose the system is dropped from a large distance, such that before it becomes part of the black hole, equal amount of entropy is radiated by the black hole. Hence at the end of the process black hole mass stays the same and therefore black hole entropy does not change. Given the process is reversible, the radiation entropy becomes E/T_{BH}.[3] Thus the overall change in the entropy of the universe becomes,

$$\Delta S = E/T_{BH} - S_{\text{box}} \quad (2.5.1)$$

One can choose the Schwarzschild radius larger than then the size of the box R, such that system will fall into black hole without being torn apart, *i.e.* the relation between size of the box and Schwarzschild radius is controlled by an $\mathcal{O}(1)$ parameter λ, $R_{\text{BH}} \sim \lambda R$. Since GSL implies $\Delta S_{\text{universe}} \geq 0$, one puts a bound on the total entropy that can be contained in the box.

$$S_{\text{box}} \leq \lambda RE/c\hbar \quad (2.5.2)$$

[3] Taking curvature into account may alter the amount of entropy emitted by thermal radiation, however these are $\mathcal{O}(1)$ effects that can be absorbed into the coefficients in the bound.

What happens to the bound when the number of species in the system increase was a long standing puzzle. One alternative and more elaborate derivation of the bound has been given in [72] which yields some understanding on how bound preserves its validity when the number of species is increased. The quantum information theoretic derivation of the bound is based on the positivity of relative entropy and has been derived for ball shaped regions in the CFT for the excitations around the vacuum density matrix and for QFT s around the density matrix corresponding the to thermal state in the Rindler space. The derivation is given for few cases where local expression for modular Hamiltonian is known.

Let us briefly review the derivation of the bound in QFT through positivity of relative entropy. The derivation of the bound using the modular Hamiltonian of the ball shaped region in the CFT is more illuminating as it yields a natural definition for the system of size R. Let us consider this system as the ball shaped region itself. The positivity of relative entropy implies, $S(\rho|\rho_{\text{vac.}}) \geq 0 \implies \Delta S_{\text{box}} \leq \Delta\langle H\rangle$ where Δ represents the vacuum subtracted quantities. Inserting the local expression for modular Hamiltonian in general dimensions,

$$\Delta S_{\text{box}} \leq 2\pi \int_0^R dr\, r^{d-1} \int d\Omega_{d-1} \frac{R^2 - r^2}{2R} \Delta\langle \hat{T}_{00}\rangle \tag{2.5.3}$$

For spherically symmetric distributions one can take the integral and turn the local integral over the energy density into a total energy relation. For example for a localized source at the center one obtains $\Delta S_{\text{box}} \leq \pi R\, \Delta E$ and for a uniform energy distribution $\Delta S_{\text{box}} \leq \pi R\, \Delta E/(d+2)$, both of which satisfies the bound up to an order one factor. In the original derivation of the bound, gravity plays a central role, yet the expression is independent of G_{N}. On the other hand, the derivation of the bound based on the positivity of relative entropy is completely quantum mechanical hence explains why the bound is independent of G_{N}. Although Bekenstein bound is independent of G_{N} it exists for gravitational system even when the self gravitation is strong. Saturation of the bound for black holes is a nice example of this situation. In this case size of the box becomes Schwarzschild radius, which is the radial coordinate rather than the geodesic distance. Since the information theoretic derivation of the bound exploits vacuum subtracted quantities, size of the box becomes ambiguous when system has back-reaction on the geometry. One needs to find a reference manifold with respect to which, the size of the box is defined. We will touch upon the ambiguity regarding the back reaction in chapter 4. We will show that size of the system is fixed with respect to the AdS_{d+1} when bound is formulated in the bulk using information theoretic relations in the underlying theory.

2.6 Holographic entanglement entropy in $\mathrm{AdS}_3/\mathbb{Z}_n$

There are different ways of representing the conical defect spacetime in a dual field theory. One approach is to consider the conical defect solution as an excited state of AdS_3. In this picture, one starts with a CFT having central charge of $c = 3L/2G_\mathrm{N}$, where L is the curvature of AdS and G_N is the Newton constant in $3d$. The vacuum of the theory corresponds to empty AdS and conical defect solutions are particular excited states of the theory. The theory should not be considered as a pure gravitational one, since conical defects are nothing but solutions of Einstein equations with Dirac delta sources. This is the standard way of considering conical defect geometries. However we will follow a more elaborate description in which quantum mechanical identification of these geometries is more precise. This type of description for conical defect solutions is also used to study long geodesics having winding numbers from the CFT point of view [92].

The conical defect geometries with integer deficit angle can be described as $\mathrm{AdS}_3/\mathbb{Z}_n$ which can be regarded as angular identification of a covering AdS_3 spacetime. The covering space "ungauges" the $\mathbb{Z}_n$ discrete gauge symmetry. Physical quantities computed in the ungauge theory must be invariant under $\mathbb{Z}_n$. This $\mathbb{Z}_n$ invariance is reflected naturally in the method of images along the uplifting procedure. In this perspective, the space where the field theory dual to conical defect lives is n-times longer circle. Spatial locations x in the CFT_c lift to $\tilde{x}$ in the fundamental domain of the covering space and corresponding $\mathbb{Z}_n$ translates. In other words for each point in x in CFT_c, there are n identified points in the covering theory. Hence an interval A with an opening angle 2α in the CFT_c lifts to n evenly spaced intervals $\tilde{A}_i$ each of angular size $2\tilde{\alpha} = 2\alpha/n$ when the circle of the covering CFT is normalized to 2π. Not all operators of the parent theory descent to conical defect theory. The projection to the defect theory only leaves the operators that are $\mathbb{Z}_n$ symmetric. $\mathbb{Z}_n$ invariance of correlation functions best reflected via method of images, where correlation function between $\mathcal{O}_1(x)$ and $\mathcal{O}_2(y)$ is computed from the geodesics between lifts $(\tilde{x}, \tilde{y})$ of (x, y) to the parent theory and all the $\mathbb{Z}_n$ translations of these locations. The leading contribution comes from the minimal geodesics that connects neighboring intervals. In addition, the uplifted operators should have an explicit $\mathbb{Z}_n$ invariance which is guaranteed through a linear combination in the form $\sum_{i=0}^{n-1} g^i \tilde{\mathcal{O}}$, where g^i is an element of $\mathbb{Z}_n$. The correlation functions of quantum fields in the conical defect and BTZ spacetimes, particularly the ones having a geometric description in the bulk, are in many cases computed by the method of images explained above [93, 94].

AdS_3 with conical defect is not identical to the covering theory in its ground state as emphasized in [92]. We will see that central charges are differ by a factor of

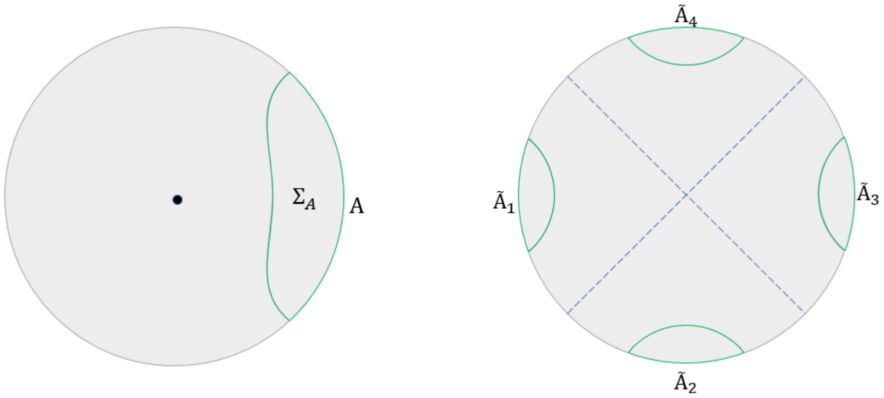

Figure 2.2: *The conical defect solutions as integer quotient AdS/$\mathbb{Z}_n$. The interval A on the defect theory CFT$_c$ and corresponding holographic entanglement entropy (minimal surface) lifts to $n = 4$ evenly spaced interval $\tilde{A}_i$ on the covering theory CFT$_{\tilde{c}}$. Size of the each boundary interval in the covering theory reduces by a factor of n.*

n. In addition the Fourier expansion of the stress energy tensors have different content in terms of the Virasoro generators (5.4.4). Simply because, not all the generators of the covering theory descends to defect theory.

We will find a local expression for the local modular Hamiltonian of the state dual to conical defect by uplifting it to covering theory and using the method of images that imposes $\mathbb{Z}_n$ invariance naturally. Before going through our derivation on the approximate Hamiltonian, we would like to continue to follow [92] on the derivation of holographic entanglement entropy from covering theory, which will be the guide to our derivation.

The essential ingredient in the entanglement entropy of a 2d-CFT is the central charge the theory. What is the central charge $\tilde{c}$ of the covering theory CFT$_{\tilde{c}}$ in terms of the central charge of CFT$_c$? Geometric description of the covering space is ungauged AdS$_3$. Therefore asymptotic symmetry algebra is the Virasora algebra by Brown-Henneaux construction. Central charge $\tilde{c}$ is the central extension of the algebra generated by large diffeomorphisms at asymptotics of the spacetime of the covering theory

$$[\tilde{L}_k, \tilde{L}_s] = (s-k)\tilde{L}_{k+s} + \frac{\tilde{c}}{12}k(k^2-1)\delta_{k+s}. \tag{2.6.1}$$

The diffeos that descend to the defect theory is the subset of diffeomorphisms of the covering space. Only the ones that preserve $\mathbb{Z}_n$ symmetry descends to defect theory. It is argued in [95] that the subalgebra is restricted to the one generated by $\tilde{L}_{nk}$, which also form a Virasoro algebra with the following redefinition of the generators,

$$L_k = \frac{1}{n}\tilde{L}_{nk}, \quad k \neq 0 \qquad\qquad L_0 = \frac{1}{n}\left(\tilde{L}_0 - \frac{\tilde{c}}{24}\right) + \frac{n\tilde{c}}{24}. \tag{2.6.2}$$

This is the sub-algebra that descends to the defect theory, which has a central extension $c = n\tilde{c}$. Using this identification one can calculate the entanglement entropy of a region in the defect theory via its uplift to the covering theory. First let us focus on the case where the interval A in CFT_c is less than the half space *i.e.* $2\alpha < \pi$ which corresponds to the phase of vanishing mutual information at the leading order in the parent theory. Employing Ryu-Takayanagi formula in the defect theory,

$$S(\mathrm{A}) = \frac{1}{4G_{\mathrm{N}}} l(\alpha) = \frac{c}{3}\log\sin\frac{\alpha}{n} + \text{const.} \tag{2.6.3}$$

Constant piece stands for the contribution of the cut-off dependence. If the region A is larger than half the field theory circle $2\alpha > \pi$ then the entanglement entropy becomes,

$$S(\mathrm{A}) = \frac{1}{4G_{\mathrm{N}}} l(\pi - \alpha) = \frac{c}{3}\log\sin\left(\frac{\pi - \alpha}{n}\right) + \text{const.} \tag{2.6.4}$$

This is the phase where region A has mutual information with its uplifted copies in the covering space. Let us now, derive the same expression via the covering theory.

The interval A in CFT_c lifts to n evenly spaced intervals $\tilde{A}_i$ each of which has angular size $2\tilde{\alpha} = 2\alpha/n$ given that the covering space has periodicity 2π. Application of the Ryu-Takayanagi formula in $\mathrm{CFT}_{\tilde{c}}$ tells us that entanglement entropy of the union of intervals $\left(\cup_i^n \tilde{A}_i\right)$ is computed from the length of minimal curve in empty AdS_3 that is homologous to the union of intervals on the boundary. The minimal curve that is homologous to the union of intervals on the boundary consists of n disjoint geodesics each having angle $2\alpha/n$ or $2(\pi - \alpha)/n$ depending on whether $2\alpha \lessgtr \pi$

$$S\left(\cup_i^n \tilde{\mathrm{A}}_i\right) = n\frac{\tilde{c}}{3}\log\sin\tilde{\alpha} + \text{const.} = \frac{c}{3}\log\sin\frac{\alpha}{n} + \text{const.} \tag{2.6.5}$$

which is equal to the entanglement entropy in the defect theory $S(\mathrm{A})$. The expression is the leading order term in $1/N$ expansion. There are also $1/N$ corrections that amount to the mutual information between the regions in the parent theory [96]. The derivation of the other cases $2\alpha > \pi$, follows the same logic except the minimal surfaces become the curves that connects neighboring regions on the boundary.

3 Causal Patch Complementarity

3.1 Introduction and Summary

The black hole information paradox [97] has always been an inspiring topic. It is widely recognized that two well motivated believes, purity of Hawking radiation and in-falling vacuum of the horizon, are at odd with each other [98]. Giving up the former leads to information/unitarity loss [97], while giving up the later leads to an energy curtain [99] or a firewall [29]. Other attempts to reconciliate certain aspects of their coexistence always lead to other problems that requires modifications to either quantum mechanics or gravity at low-energy [25, 47, 100–103]. The reason why we have not been actively modifying low-energy effective theories is the concept of complementarity:

Weak Complementarity:

The low-energy effective quantum theory, semi-classically coupled to weak gravity, only needs to be self-consistent within individual causal patches.

Strong Complementarity:

The boundaries of causal patches are governed by UV theories (of quantum gravity) which can fix the apparent pathologies in global descriptions.

Stated above is the general form of a global-local complementarity, which might be also useful in cosmology [104–107]. However they are most well studied as the black hole complementarity. As shown in Fig.3.1-left, the state of the horizon is only described in the casual patch of in-falling observer, while Hawking radiation is in the patch of a distant observer. Since these two potentially contradicting ingredients belong to two different causal patches, weak complementarity guarantees that our everyday application of low-energy effective theory is justified. It remains an interesting topic to find an appropriate statement in strong complementarity, in terms of descriptions of the horizon (and/or the black hole interior) for a distant observer [35, 36, 108–113]. That however belongs to UV physics instead of

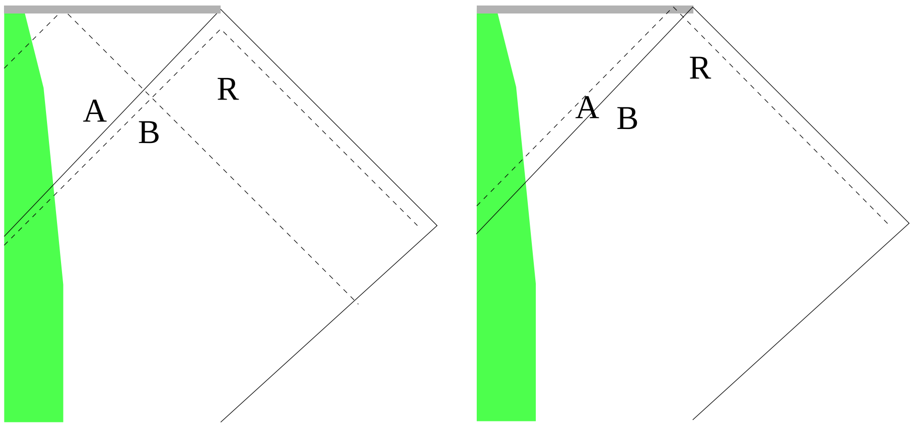

Figure 3.1: *The left figure shows the old story for complementarity: an in-falling causal patch needs to confirm that the interior mode A and the exterior mode B are maximally entangled to ensure a normal horizon; an outside causal patch needs to confirm that B is entangled with the rest of Hawking radiation R to preserve unitarity. The right figure shows that in a late, in-falling causal patch, both entanglements need to be confirmed, and that is a paradox for requiring duplicated information in A and R.*

modifications to low-energy physics.

More than one year earlier, Almheiri, Marolf, Polchinski and Sully (AMPS) formulated a challenge against the status quo [29]. They argued that for a sufficiently old black hole, the information paradox can be revived even abiding the standard of weak complementarity. This new challenge has been paraphrased in many different versions [100–102, 114–117], and many of them involve tricks like using a distillation process and/or a boundary CFT dual. It is very important to clarify the situation being which one of the following:

- Those tricks help to make the paradox clear.
  In this case there should be a more passive and pristine version of the paradox that demonstrates the core of the problem.

- Those tricks are essential to establish the paradox.
  In this case those tricks need more scrutinizations, since the paradox might be an artifact of our naïve idealization of those tricks.

In this paper, we will ask a well-defined question to clarify the above situation. *Without any distillation process, does low-energy theory in individual causal patches run into any pathology?* The naïve answer seems to be *yes*. It comes from an observation that the two causal patches shown in Fig.3.1-left are extreme cases: either one jumps into the black hole right away, or one stays outside forever. As shown in Fig.3.1-right, we can find a generic causal patch between those two. It belongs to an observer who stays outside long enough but eventually falls in. Note that the purity of Hawking radiation has to be verified when the black hole is more than half evaporated, and by that time the black hole (interior) can still be large. Therefore such a late, in-falling causal patch seems to include both ingredients for the paradox.

We will go beyond this naïve answer and examine this late, in-falling causal patch in more details. In particular, we emphasize that merely "fitting into a causal patch" is not good enough. Otherwise, we do not even need the AMPS argument to have a paradox. The very ancient information paradox uses the in-falling matter and Hawking radiation as its two conflicting ingredients, and they do coexist in the outside causal patch as shown in Fig.3.2-left. The reason why such situation cannot qualify as a paradox for weak complementarity is because in the outside causal patch, those two ingredients are never *space-like separated and both low-energy*.

In Sec.3.2, we will explicitly state our standard for two physical quantities to "fit into a causal patch as space-like separated, low-energy quantities". Note that energy is frame dependent, but whether there exists at least one frame satisfying this standard is a frame independent property of the causal patch. We will show that if such standard is not upheld, then no single observer can properly observe them both. Therefore, a potential paradox built from these two quantities is invalidated by the spirit of complementarity. Note that in [118], the integrated boost between two physical quantities was proposed to determine whether they can legitimately form a paradox. That standard can still be defined globally for quantities cannot fit into one causal patch. Our standard is totally within a causal patch and directly related to their simultaneous-observability[1].

In Sec.3.3 we apply the above standard to a Schwarzschild black hole. We explicitly show that within a late, in-falling causal patch, it is impossible for the interior mode and the early quanta to be space-like separated and both low-energy. Either the early Hawking radiation has Planckian wavelength, or the interior region has a size exponentially smaller than the Schwarzschild radius $\sim M$. We then generalize this result in Sec.3.4 to include possible operations on the early Hawking quanta, including confining them into a box and further thermalization. We find that our conclusion remains unchanged. Within a late, in-falling causal patch, either the black hole interior is exponentially small, or the information within early Hawking radiation goes beyond low-energy physics.

A more dynamical picture of our result is that within low-energy physics, the

[1] Despite this major difference, if we first limit ourselves to one causal patch and apply the integrated boost standard, then it qualitatively agrees with our standard in the case of a Schwarzchild black hole.

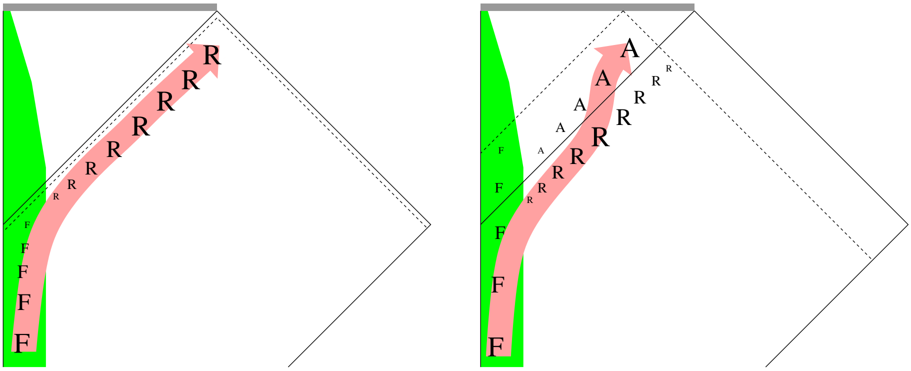

Figure 3.2: *The left panel shows the physical interpretation of the outside causal patch, and the right panel shows that for a late, in-falling causal patch (causal patches are bounded by the dotted lines). The size of the letters (F, A and R) represents the "wavelength" of the corresponding physical quantities (in-falling matters, interior modes and Hawking radiation) according to the description within the causal patch. Large letters represent low-energy quantities and small letters represent UV quantities. The pink arrows represent the information flows, which only need to follow low-energy quantities in the usual way. When the carrier of such information becomes UV quantities, unknown UV processes can direct the information to the appropriate places to avoid paradoxes. For the outside patch, such UV flow of information resolves the information-loss paradox. For the late, inside patch, another such UV flow resolves the AMPS paradox.*

early Hawking radiation only exists in the early times, and the interior mode only exists in the late times. Since they are strictly time-like separated, the double-entanglement (quantum-cloning) problem does not apply. On every time-slice, this also provides a clear distinction between the "recent" Hawking quanta B that live in the low energy theory, and the "earlier" Hawking quanta that form R and hide in the UV. If one wishes to think about an $A = R$ map, it can be unambiguously only applied to the UV quantities R but not to the low energy B, which avoids the frozen vacuum problem [102, 117].

In Fig.3.2, we put the conceptual information flow of this "inside story" side-by-side with the well-known "outside story" of black hole complementarity and observe their similarity.

Outside Story: In the causal patch of an outside observer, the information initially follows the collapsing matter (or anything thrown in later). When the information flows to the boundary, it is transferred to the UV physics. Instead of leaving this causal patch, the UV physics keeps the information in the Planckian neighborhood of the boundary, and later returns it through Hawking radiation. Note that here the boundary of the causal patch is exponentially close to the black hole horizon. Sometimes it is mistaken that the special UV property only occurs for black hole horizon because it is a special place. By the spirit of complementarity, it is the boundary of the causal patch that is a special place for the theory within.

Inside Story: In the causal patch a late, in-falling observer, the first half of the story is the same. Note that the causal patch boundary here is also exponentially close to the black hole horizon, although from the inside. Later, Hawking radiation will approach the outside boundary of this causal patch, which means that it again belongs to UV physics. From there on, we are free to claim that the unknown UV physics guides the information in unexpected ways. That is a good news, because strictly after the early Hawking radiation flows to this UV zone, the interior mode A emerges and demands that information. We can simply claim that through unknown UV physics, the information flows there.

In both stories above, we can see there is no need to modify low-energy physics.

- The information flow is time-like, so there is no need of non-localities [25].
- The information flow is future-directed, so there is no need of final-state quantum mechanics [103].
- There is no duplicate information, so there is no need of firewalls [29].
- Everything happens within the standard black hole geometry, so there is no need of Einstein-Rosen bridges [47].

Thus, we have verified that a pristine, distillation-free version of the AMPS paradox does not exist. The pathology explicitly resides in the unknown and idealized distillation process. In Sec.3.5 we point out a possibility to address the remaining distillation problem with our approach. We also discuss possible implications on cosmological horizons.

3.2 Weak Complementarity

In Sec.3.2.1 we present a generic way to describe physics within a causal patch and provide the standard for a double-entanglement (quantum-cloning) paradox within the low energy theory: "*Within this causal patch, if there are no space-like surfaces on which both copies of information are carried by low energy quantities, then it is illegal to form a paradox with them.*

In Sec.3.2.2 we show that our standard is directly connected to whether a single observer can observe any contradiction. We provide pictorial examples to show that our standard has no unphysical side-effects. It does not invalidate paradoxes

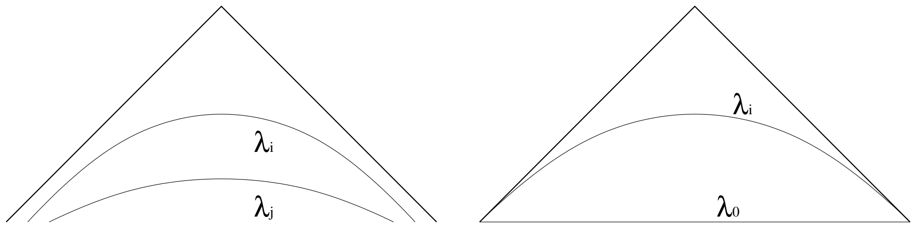

Figure 3.3: *Examples of foliations within causal patches. The left panel shows two slices of the standard, semi-infinite hyperbolic foliation given by Eq. (3.2.1). The right panel shows the cut-off version where the evolution starts on the initial condition given in the first slice $\mathcal{F}(\lambda_0)$.*

in usual situations. It only intervenes when the causal structure of the problem obstructs the practical observability of the paradox.

3.2.1 Theory within a causal patch

Definitions

- A causal patch $\mathcal{C}$: the entire space-time region within the past light-cone of a point.
- A foliation $\mathcal{F}(\lambda)$: one parameter family of 3-dimensional space-like surfaces such that every $\mathcal{F} \subset \mathcal{C}$ and $\bigcup_t \mathcal{F} = \mathcal{C}$. In addition, for all $\lambda_i > \lambda_j$, every future directed path from every point in $\mathcal{F}(\lambda_j)$ passes through $\mathcal{F}(\lambda_i)$, and every past directed path from every point $\mathcal{F}(\lambda_i)$ passes through $\mathcal{F}(\lambda_j)$.

The foliation allows us to describe the dynamical evolution as a closed system. For example, the causal patch of the point $(t = 0, \vec{x} = 0)$ in Minkowski space is naturally foliated by a family of hyperboloids with constant time-like separations from the tip.

$$\mathcal{F}(\lambda) = \{(t, \vec{x}) \mid t < 0,\ \lambda < 0,\ t^2 - |\vec{x}|^2 = \lambda^2\}\ . \tag{3.2.1}$$

Of course, this is not a unique choice. A monotonic map $\lambda \to \lambda'$ leads to a different foliation $\mathcal{F}'$ that is still legal as long as each slice remains space-like. Instead of extending to past infinity, we can also consider a cutoff. Starting from a space-like surface which sets the initial conditions, the foliation only needs to evolve it forward. These example are depicted in Fig.3.3.

Standard

A low-energy effective theory within this causal patch has to be consistent in all possible foliations. In particular, since we have conveniently chosen to view it as

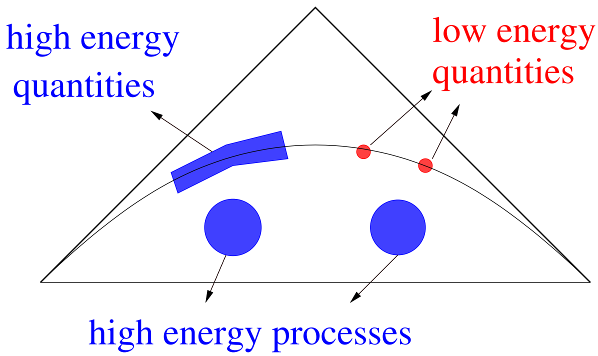

Figure 3.4: *Given low-energy input on the initial slice, a consistent low-energy theory can allow high energy events during the evolution or even in parts of the outcome (blue patches). However if the low-energy parts of the outcome (the red dots) show inconsistency, then there is a paradox and the theory needs to be fixed.*

a closed system, the evolution should be unitary. It looks like a simple job to establish the AMPS paradox. We just need one special example: one causal patch and one foliation in which energy and curvature stay small, but the evolution violates unitarity or has other pathologies.

In fact, we will hold an even lower standard for a paradox. We will allow the condition on being "low-energy" to be only partially satisfied. First of all there can be gaps in the foliation. As long as one finds two slices $\mathcal{F}(\lambda_i)$ and $\mathcal{F}(\lambda_j)$ that energy is low on both of them, then their states need to be related by a unitary transformation, and the propagation of information should be causal. In between them there might be all sorts of high energy activities like a nuclear bomb or a black hole formation-evaporation, during which even the foliation cannot be consistently defined. We will not use those to disqualify the paradox. Furthermore, we will also allow some regions of the slices to contain high energy quantities. As long as the low-energy regions contain sufficient evidence for pathologies, such as violating the monogamy of entanglement, we accept the existence of a paradox and the necessity to modify low-energy physics. This standard is pictorially summarized in Fig.3.4.

3.2.2 Fitting into the causal patch while being low-energy

Note that "energy", or the length scale, is a frame-dependent quantity. Naturally we define it as the inner product between the momentum vector of a particle and the 4-vector normal to the spacelike surface, so it depends on the choice of foliation. As shown in Fig.3.5, a low-energy quantity in one foliation can become high energy on a different one. It is very reasonable to question the physical meaning of such

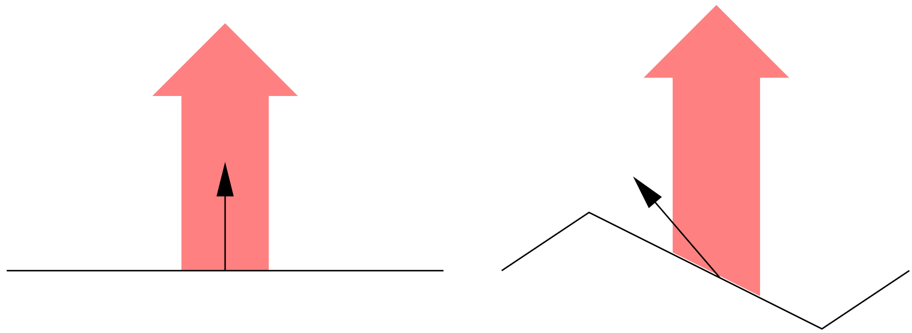

Figure 3.5: *In the left panel shows, the big red arrow going up shows represent a low-energy physical quantity and the time-like vector of its natural frame. The arrow is aligned with the time-like vector specified by the foliation (the small black arrow going up), so such physical quantity is low-energy in the foliation. The right panel shows the same quantity in a different foliation, in which there is a large relative boost between the two arrows. Due to such boost, the same quantity appears to be high energy on this foliation.*

frame-dependent standard. Indeed the properties of one foliation has no reason to have a deep physical meaning. However, it is a frame-independent physical fact if in all possible choices of foliations, something never happens.

In other words, we should try our very best to find foliations that makes things as low energy as possible, and see that sometimes it is just impossible. In this section, we will go over pictorial examples to demonstrate when the requirement of fitting a foliation into a causal patch is very restrictive. The remaining freedom can be insufficient to make relevant physical quantities (for the paradox) appear as low-energy. We will then point out that in those situations, these physical quantities in-principle cannot be observed together. This shows that our standard in Sec.3.2.1 is closely related to observability.

Consider the situation in Fig.3.6 in which the initial slice of a foliation is evolved to a later slice where we identify two subsystems A and R. If A and R contains the same quantum information, then we may try to establish a paradox on the basis of quantum cloning.

First of all, there are some natural frames in which both A and R are low-energy separately. These are most likely their rest frames or the rest frames of the sources if they are radiations. Since A is in the center of the patch and its velocity aligns with the normal 4-vector of the foliation, it remains to be low-energy in this foliation. On the other hand, R is on the edge of the patch and has a large relative boost to the foliation, so it might be high energy. For an observer whose experience is confined within this patch, this situation implies a true limitation. This observer

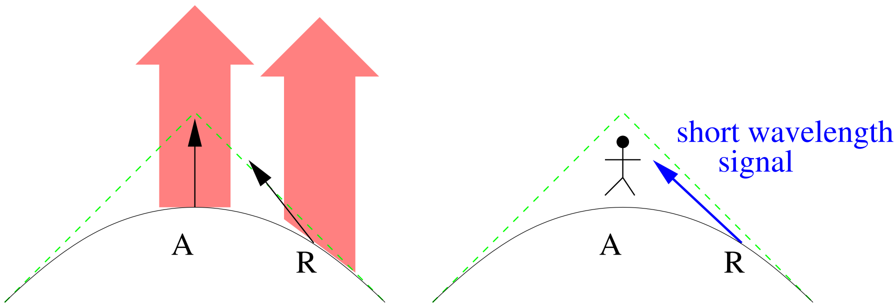

Figure 3.6: *A and R are two subsystems on a slice of the foliation within this causal patch (past light-cone in dashed green lines). It may be a paradox if they contain the same quantum information. However if R appears to be high energy due to a large relative boost, then any observer whose experience is limited to this causal patch cannot verify the paradox. As shown in the right panel, for an observer who sees A, any information from R has to be at ultra-short wavelengths. That is beyond the applicability of low-energy physics.*

has to be at the top of the patch to see A, from where he can only catch a fleeting glimpse into R before it exits the patch. If this glimpse does not last long enough, this observer cannot decipher information from R, at least not when he is still inside this patch.

If this causal patch is from an interior point, then we cannot reject the paradox this way. Since the limitation only applies to a confined observer, one can choose a later observer who is less confined. Equivalently, as shown in Fig.3.7, one can take a later (larger) causal patch, and in this new patch A and R are easily both low-energy in some foliation. This is a consistency check that our standard does not reject paradoxes while it should not.

On the other hand, if the tip of the patch is at a singularity, then the limitation is truly meaningful. Now there is no bigger causal patch. One can consider another patch to the right of this one. It can put R at the center but then A will appear to be high energy. In this case no single observer can collect information from both A and R, therefore the presumed paradox is not observable and should be rejected. This is also shown in Fig.3.7.

In Fig.3.8 we see another consideration that we can evolve both A and R backward in time, and they may appear less boosted with respect to an earlier slice of the foliation. Indeed if A and R together appear on an earlier slice and are both low-energy, that slice suffices to establish the paradox. Later we will see that in the AMPS case, R is low-energy at an early time while A is not, and A is low-energy at a later time while R is not. They are never both low-energy in at the same time. This is a clear sign that we should probably reject the paradox, because the

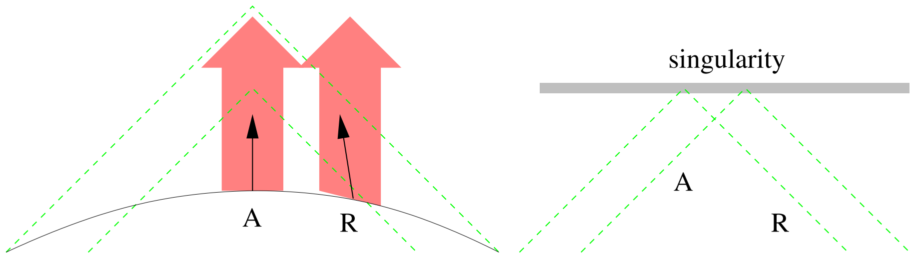

Figure 3.7: *The left panel shows that if the causal patch in Fig.3.6 is from an interior point, then one can take a bigger causal patch which totally covers the previous one, and it can easily accommodate both A and R in low-energy. The right panel shows that if the causal patch is from a boundary point, for example at future singularity, then there is no bigger causal patch. Shifting it to the left or right will not be enough to make A and R both low-energy, or even have trouble to include them both.*

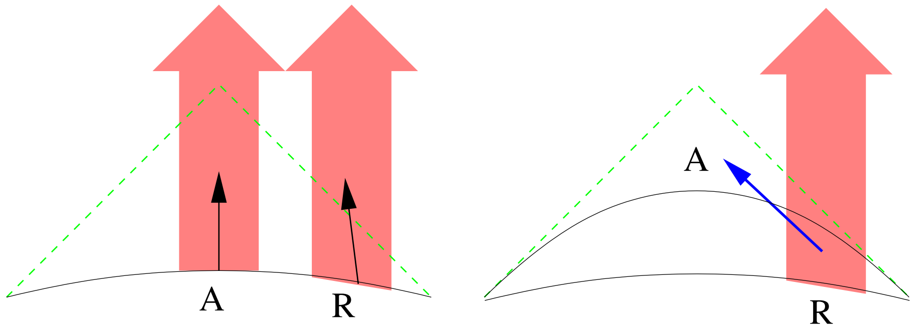

Figure 3.8: *The left panel shows the case that A and R appear to be low-energy on an earlier slice. The right panel shows the case that A does not have a low-energy past, while R only becomes low-energy on an earlier slice when it is in the causal past of A. This allows the interpretation that R gets burned out at the "horizon" and propagates to A (following the blue arrow) through an unknown high energy process.*

cloning problem only sustains when the full Hilbert space contains the product of their individual Hilbert spaces, $\mathcal{H}_A \times \mathcal{H}_R \subset \mathcal{H}$. This picture is natural only when A and R are space-like separated. If R is in the past of A, then it could *be* the past of A, and it is not a paradox for them to have the same quantum information.

A natural interpretation for time-like separated A and R demanding the same information is simple: R gets burned out at the "horizon" (the outside boundary of the patch, not the black hole horizon) and its information content travels to A. Note that this is a process in UV physics, so our lack of expectation of such an information flow, or any low-energy setup to block it, cannot be a sufficient reason to insist on the paradox[2].

[2] We thank Daniel Harlow for pointing this out.

3.3 Causal patch for an old Schwarzschild black hole

Here we will apply the standard in Sec.3.2 to examine causal patches in the geometry of an old Schwarzschild black hole. We will focus on the two ingredients which are used to form the information paradox in [29]. The first ingredient is the near horizon region of size $\sim M$. This region generates Hawking radiation by separating the pair of interior A and exterior B modes which formed the local vacuum state. When the exterior mode B has a wavelength $\sim M$, then it is officially a out-going Hawking quantum. Due to how it was generated, its state is maximally entangled with its interior partner A also with a wavelength $\sim M$ [119].

The second ingredient is the early Hawking radiation R. Because the evaporation process is unitary, an out-going Hawking quantum B should be maximally entangled with R. This double entanglement requires A and R to carry duplicated quantum information. So if they both appear within this causal patch and are low-energy quantities on certain slice of some foliation, then it is indeed a paradox for weak complementarity.

Here we present an explicitly calculation to show that for all possible foliations in all such causal patches, there is no single slice on which A and R coexist as low-energy quantities.

Our calculation has two parallel sessions addressing the inside and the outside. Our inside calculation is in the Kruskal coordinates and includes the black hole interior and some outside region near the horizon. Our outside calculation deals with regions far away from the black hole where the early Hawking radiation has propagated to, and we can simply use the Schwarzschild coordinates. This inside-outside separation is illustrated in Fig.3.9. There are two important anchors in our calculations. The first anchor X is the tip of the causal patch at the singularity, and the second anchor Y is where the slice intersects the horizon. It is useful to draw the past light-cone from both anchor points as we did. Since the slice we are looking for must be bounded between these two light-cones.

These two light-cones extend through both parts of our calculation. The most convenient way to connect the calculations is through Δt, the Schwarzschild time difference when these two light-cones intersects the boundary between the two parts of our calculation. In other words, it represents the distance between the two anchors[3]. This turns out to be the only relevant parameter. We will present

[3]Note that there is a hidden spherical symmetry in the diagram, but the anchor points X and Y are only one particular point even in the angular coordinate. In the inside calculation, we are calculating exactly the distance along this particular angular direction. In the outside it

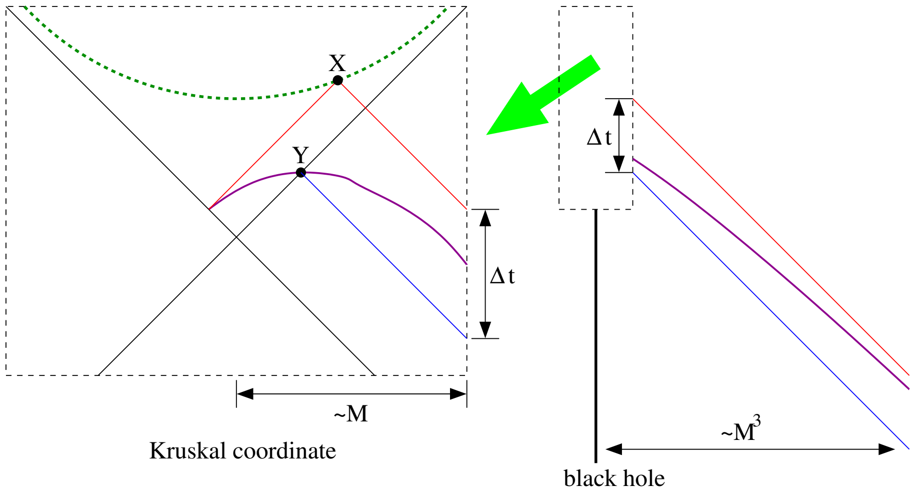

Figure 3.9: *The right side shows the outside calculation where the typical length scale involved is M^3. The near horizon $\sim M$ region of the black hole is effectively a point. We zoom in to that region (the dashed box) in the Kruskal coordinate. The two anchor points: X is the tip of a causal patch (enclosed by the red past light-cone) on the singularity (the green curve); Y is where the space-like slice (yellow curve) intersects with the black hole horizon. The past light-cone (blue line) from Y is convenient for defining the distance Δt between these two anchor points, and the space-like slice must be bounded between these two light-cones.*

the detail calculation in the next two sub-sections. Here we briefly summarize the result.

The interior mode A only exist on the segment of the slice to the left of point Y. We define λ_A to be the proper length of this segment and find it bounded from above by

$$\lambda_A \lesssim M e^{-\Delta t/4M} \ . \tag{3.3.1}$$

Namely, inside this causal patch, we need $\Delta t < M$, otherwise the interior mode is exponentially small.

On the other hand, the wavelength of the early Hawking quanta on this slice (the inverse of its 4-momentun projected onto the normal timelike vector) is bounded by

$$\lambda_R < \sqrt{\frac{\Delta t}{M}} \ . \tag{3.3.2}$$

looks like we are also only calculating the Hawking quanta flowing along this direction. However, the extension of these light-cones to other angular directions will keep the same separation Δt, so our outside calculation is actually valid in all directions.

This means that we need $\Delta t > M$, otherwise these quanta have Planckian wavelengths. Since Δt cannot be both bigger and smaller than M, either we cannot address the information content in R, or we do not describe the near horizon process that generates late Hawking quanta. The two ingredients for the AMPS paradox have failed to coexist in the low-energy theory within a causal patch.

In the situation where we choose $\Delta t \sim M$, the near horizon region of this slice agrees with an in-falling observer free-falling from $\sim M$, so the interior mode A looks normal. However the early Hawking quanta suffer a large boost $\gamma \sim M$ which contracted their wavelengths from $\sim M$ to order one. This shows that our standard qualitatively agrees with the standard of integrated boosted [118].

3.3.1 Inside: the interior mode A

Consider a half-evaporated black hole with current mass M. The metric is usually given in the Schwarzschild form.

$$ds^2 = -\Big(1 - \frac{2M}{r}\Big)dt^2 + \Big(1 - \frac{2M}{r}\Big)^{-1} dr^2 + r^2 d\Omega^2 \ . \tag{3.3.3}$$

However to analyze the geometry across the horizon, it is more convenient to remove the coordinate singularity by going to the Kruskal-Szekeres coordinates.

$$ds^2 = \Big(\frac{32M^3}{r}\Big)e^{-r/2M}(-dV^2 + dU^2) + r^2 d\Omega^2 \tag{3.3.4}$$

This is related to the Schwarzschild coordinates by

$$\begin{aligned} V &= \Big(1 - \frac{r}{2M}\Big)^{1/2} e^{r/4M} \cosh\left(\frac{t}{4M}\right) \ , \\ U &= \Big(1 - \frac{r}{2M}\Big)^{1/2} e^{r/4M} \sinh\left(\frac{t}{4M}\right) \ , \end{aligned} \tag{3.3.5}$$

in the interior ($V^2 > U^2, r < 2M$), with a singularity $V^2 - U^2 = 1$. In the exterior ($V^2 < U^2, r > 2M$), they are related by

$$\begin{aligned} V &= \Big(\frac{r}{2M} - 1\Big)^{1/2} e^{r/4M} \sinh\left(\frac{t}{4M}\right) \ , \\ U &= \Big(\frac{r}{2M} - 1\Big)^{1/2} e^{r/4M} \cosh\left(\frac{t}{4M}\right) \ . \end{aligned} \tag{3.3.6}$$

The coordinates of the two anchor points are given by

$$X = (U_X, V_X) = (b - \frac{1}{4b}, b + \frac{1}{4b}) \ , \tag{3.3.7}$$

$$Y = (U_Y, V_Y) = (a, a) \ . \tag{3.3.8}$$

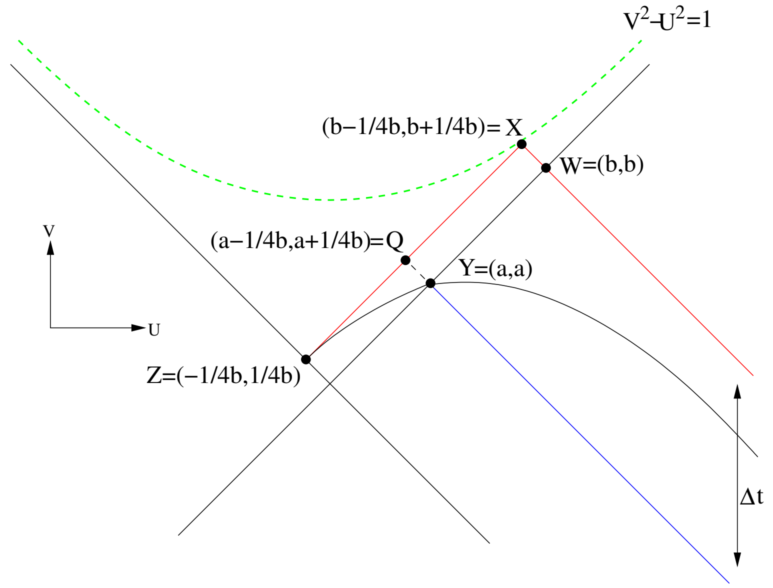

Figure 3.10: *The calculation of the length of the interior mode A in the Kruskal coordinates. Points X and Y are the same as those in Fig.3.9. The (U,V) coordinates of other important points in the calculation are also shown here.*

The variable b is chosen such that the past light-cone from X intersects the horizon at $W = (b,b)$. These points and a few that we will define later are shown in Fig.3.10. The choice of variables a and b simplifies the relation to the important parameter Δt, which is the difference in the Schwarzschild time if we follow the past light-cones from these two points to some $r > 2M$ outside.

$$e^{\Delta t/4M} = \frac{b}{a} \ . \tag{3.3.9}$$

We would like to calculate the length of the interior mode A on a space-like surface which passes through point Y and bounded by the past light-cone from X. This length will be bounded by the space-like geodesic distance between point Y and point $Z = (-\frac{1}{4b}, \frac{1}{4b})$[4].

$$\begin{aligned} \lambda_A &< \int_Z^Y \sqrt{\frac{32M^3}{r}} e^{-r/4M} \sqrt{dU^2 - dV^2} \qquad (3.3.10) \\ &\leq \sqrt{\frac{32M^3}{r_Q}} e^{-r_Q/4M} \int_Z^Y \sqrt{dU^2 - dV^2} \leq \frac{a}{b} \sqrt{\frac{32M^3}{r_Q}} e^{-r_Q/4M} \ . \end{aligned}$$

[4]Point Z is on the other horizon of an eternal black hole, which technically speaking does not exist in the geometry of a black hole formed by collapse. However that means before this slice extends to point Z, it would have run into the collapsing matter and stopped. Thus the distance between Y and Z is a good upper-bound.

The point $Q = (a - \frac{1}{4b}, a + \frac{1}{4b})$ is where the metric factor reaches the maximal value for any space-like path between Y and Z.

$$\left(1 - \frac{r_Q}{2M}\right) e^{r_Q/2M} = \frac{a}{b} = e^{-\Delta t/4M} \ . \tag{3.3.11}$$

After taking out the metric factor, the remaining integral is simply maximized by a straight line.

According to Eq. (3.3.11), r_Q decreases as Δt increases. That means the upper bound in Eq. (3.3.11) decreases as Δt increases. For example, if

$$\frac{\Delta t}{4M} = \ln \frac{b}{a} > \ln 2 \ , \tag{3.3.12}$$

then we have

$$\lambda_A < 4\sqrt{2} M e^{-1/4 - \Delta t/4M} < 2\sqrt{2} M e^{-1/4} \ . \tag{3.3.13}$$

So we would like $\Delta t > M$ for this causal patch to contain any interior mode A with $\lambda_A \gtrsim M$.

3.3.2 Outside: the early Hawking quanta R

In the outside we only need to consider the early Hawking quanta, which are $\sim M^3$ away from the black hole. That means we can even ignore the Schwarzschild metric and treat it as flat space. The relative error we are making is $\sim M^{-2}$, which is negligible in the limit of a large black hole.

Our goal is as shown in Fig.3.11. The early Hawking radiation will pass through some region of this slice, and the slice is bounded between the two past light-cones. We would like to minimize the energy of the early Hawking quanta in the frame determined by this slice. That means minimizing the relative boost between the Schwarzschild time and the time-like direction orthogonal to this slice. Since we do not care about other regions on this slice which do not contain any early Hawking quanta, the maximum is reached by the slice shown in Fig.3.11. This boost is given by

$$\begin{aligned} v &= \frac{M^3}{M^3 + \Delta t} \ , \\ \gamma &= \frac{1}{\sqrt{1 - v^2}} \approx \sqrt{\frac{M^3}{\Delta t}} \ . \end{aligned} \tag{3.3.14}$$

That means the early Hawking quanta on this slice will have wavelength

$$\lambda_R < \frac{M}{\gamma} = \sqrt{\frac{\Delta t}{M}} \ . \tag{3.3.15}$$

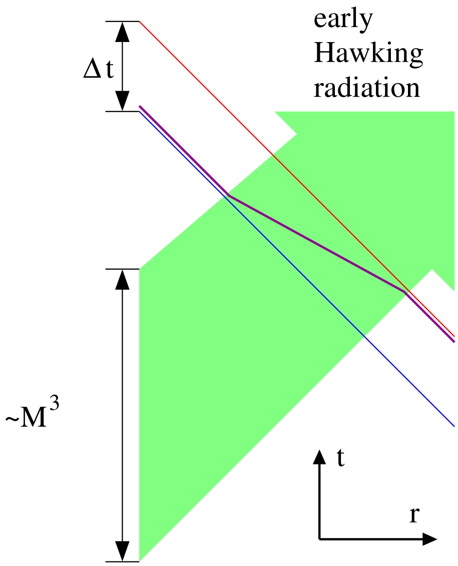

Figure 3.11: *The outside calculation. Early Hawking radiation comes from the black hole during a time duration $\sim M^3$ and flows as the big green arrow. The slice (purple) intersects with this flow and is bounded by the blue and red light-cones with separation Δt. In order to minimize the relative boost between the slice and the Schwarzschild frame (which is equivalent to maximizing the proper length of this segment), it must go through the corners.*

This means that we need $\Delta t > M$ to make $\lambda_R > 1$, so the early Hawking quanta are not Planckian. That is in direct conflict with Eq. (3.3.13) which needs $\Delta t < M$ to make $\lambda_A \gtrsim M$.

3.4 Generalizations

In Sec.3.3 we showed that the problem R might run into in a late, in-falling causal patch is that the wavelength of individual Hawking quantum becomes the Planck scale. That is a combination of their original wavelength M and the large distance M^3 from the black hole. These two things can be modified. For example, Hawking radiation can interact with a cloud of dust and further thermalize into more quanta with longer wavelengths. One can also try to confine the quanta into a box so they are not so far away from the black hole[5]. In this section we will argue that our conclusion still holds under those changes.

[5] Since it is hard to construct perfect reflecting surface (especially for gravitons), the ideal setup is embedding the black hole in an AdS space [120].

The first issue is that those changes, thermalization and confinement, also change the number (density) of quanta. "Wavelength should be smaller than the Planck scale" is the standard for one quantum, and we need to find a generalization of such standard for more quanta. A natural guess is

$$N \ll \lambda L \ . \tag{3.4.1}$$

If there are N quanta of wavelength λ in a region of size L, then it is actually a black hole. In this sense one quantum is the special case that $L = \lambda$.

Eq. (3.4.1) is exactly the standard we will follow. In particular, note that λ and L are frame-dependent quantities. We will argue that whether information is accessible is also frame-dependent. This is similar to the examples in Sec.3.2. If Eq. (3.4.1) cannot be satisfied on all possible foliations with a causal patch, then it means no observer within this causal patch can read such information.

First consider the situation in Fig.3.12-left: on a usual space-like surface, within a shell of radius L, we have N quanta of wavelength λ. We claim that the information content in this group of quanta is illegal if Eq. (3.4.1) is not satisfied. The obvious reason is that this region actually forms a black hole.

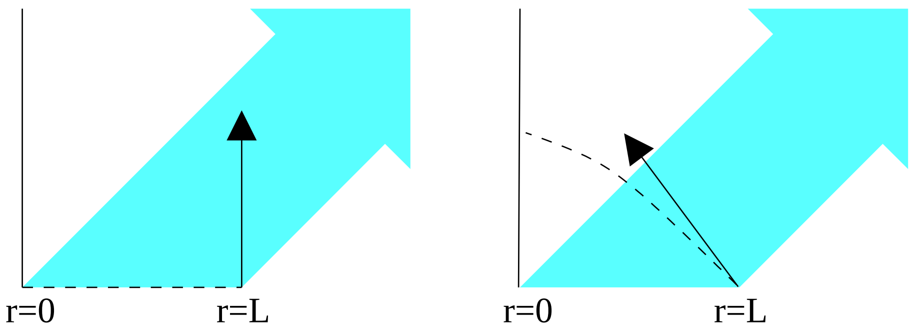

Figure 3.12: *Both figures have a suppressed spherical symmetry around $r = 0$. Whether the information (blue flow) on a space-like surface (the dashed line/curve) qualifies as low-energy is related to whether observers in the rest frame of such surface (arrows) can read such information when they flow out.*

Now take a more observer oriented point of view. Let a group of observers sitting at the shell and read a flow of information from those N quanta. If they find that Eq. (3.4.1) is violated, they should conclude that such information is a lie. Because if there was such information inside the shell, the observers should have collapsed into a black hole with it.

Next we consider the same situation with another space-like surface as in Fig.3.12-right. As we already saw in Sec.3.3, the choice of this kind of surface is enforced by the need to fit into some causal patch and to include another ingredient of the paradox. For simplicity we say that the relevant region (where the information flow through) has a boost γ relative to the surface in Fig.3.12-left.

Our claim is that we should use λ and L in the frame of this slice in Eq. (3.4.1) to determine whether the information within these quanta is legal or not. In this case those two length scales are blue-shifted.

$$N \ll \frac{\lambda}{\gamma}\frac{L}{\gamma} \ . \tag{3.4.2}$$

One might have objections to this standard. Since if Eq. (3.4.1) is satisfied, the information itself is not forming a black hole. However there is always some γ large enough to invalidate Eq. (3.4.2). So what is the physical meaning of this frame-dependent standard?

The answer is similar to the logic we demonstrated in Sec.3.2.2: such frame dependent standard is related to the practical observability. We should ask that whether a group of observers in the frame of this slice (confined to the same causal patch that forces us to choose this slice) can read the relevant information. In Fig.3.12-right, that means a group of observers following a shrinking shell with a boost γ. The quanta they read in their frame will have wavelengths (λ/γ), that means they need to carry N units of such size to keep those information. Now coming back to the rest frame, it means that this group of observers will be carrying N quanta of wavelength (λ/γ^2). They will form a black hole unless

$$N\frac{\gamma^2}{\lambda} \ll L \ , \tag{3.4.3}$$

which is exactly the same standard as Eq. (3.4.2). Therefore, Eq. (3.4.1), applying to quantities in the frame of a slice, is the appropriate standard to determine whether information on such slice is legal within low-energy theories.

Now we have established the standard, it is straightforward to show that neither thermalization nor confinement changes our conclusion in Sec.3.3. Thermalization changes N and λ together proportionally, so Eq. (3.4.1) is not affected at all. Confining Hawking radiation into a region of size L around the black hole can reduce the boost factor in Eq. (3.3.14), but such change is exactly canceled by the explicit presence of L in Eq. (3.4.2).

$$N = M^2 \ll \frac{\lambda}{\gamma}\frac{L}{\gamma} = \frac{M}{\sqrt{L/\Delta t}}\frac{L}{\sqrt{L/\Delta t}} = M\Delta t \ . \tag{3.4.4}$$

This is exactly the same as Eq. (3.3.15), so our conclusion remains unchanged.

3.5 Discussion

3.5.1 Information paradox and the quantum second law

We have verified that a distillation-free version of the AMPS paradox does not exist. It also seems to be clear that if a distillation process is possible, then the paradox does exist. Since such a process allows us to extract the relevant information into a few qubits R_B instead of keeping track of the entire R, this much smaller number of quanta is obviously quite mobile and can likely be further operated to satisfy the standard of weak complementarity.

On the other hand, the standard picture of distillation processes has it own problems. For example, it might take a very long time [49] or result in back-reactions [121], and both concerns can neutralize the paradox. In fact, our discovery here works coherently with the back-reaction argument. We showed that without a distillation process, R is strictly in the past of A. Since a distillation process acts on R, by causality it can of course affect A.

Still, we understand that the arguments involving a distillation process cannot be easily resolved unanimously. The true problem is that such a process is intrinsically unknown. It depends on how the information is encoded in R, which depends on the unknown black hole S-matrix. Our work can also provide a possibility to improve such situation. We can connect the information paradox to another type of distillation process which is not related to the unknown S-matrix.

In order to make such connection, first note that our generalizations in Sec.3.4 involves thermalization: take N Hawking quanta of wavelength λ and split them into αN quanta with wavelength $\alpha\lambda$. We assume that the information has to become hidden in all αN quanta, which is essential to maintain our argument. The paradox can be restored if we discard such assumption.

In other words, one can look for a distillation process that increases the information/energy ratio, namely some form of information density. If there is a general process in quantum thermodynamics that takes energy away without taking information away, then we can increase such information density. This will allow the information carrier to have arbitrarily low energy and avoid the problem we pointed out. Alternatively, one can interpret our work as claiming that if there is no information paradox, then there must be an in-principle obstruction against reducing quantum information density. At least, we should not be able to reduce the energy of a qubit of information from M to an arbitrarily low value within time M^3.

3.5.2 Causal patch complementarity

The two information flows in Fig.3.2 base on the same principle. However, only the situation for an outside observer was widely appreciated before, and it might be misleading in some way. It is easy to believe that the black hole horizon has the mysterious UV property to guide the information. The AMPS paradox and our work together serves as a reminder that from the very beginning, the black hole horizon is not special. It only looks special for outside observers, because it coincides with the boundary of their causal patches. If we keep in mind that all boundaries of causal patches can be similarly special, then the paradoxes can be avoided.

This realization may have a profound consequence. For outside observers, going through the black hole horizon is dropping to the "inside" of something, and it is natural to believe that information eventually comes back. However, in most cases the causal patch boundary leads to "outside", and there are rarely good reasons for the information to come back from there. The late, in-falling patch is a good example showing that if we look hard enough, then we might discover a reason that information must come back instead of flowing "outside".

This makes us wonder whether we should also look harder in other situations, like for cosmological horizons. We just eliminated the naïve distinction between information going "inside" and "outside", so it becomes less crazy to think that some information leaving a cosmological horizon is actually not lost. Up to this moment, all these unexpected information flows have been demanded by the necessity to avoid paradoxes. It will be fascinating if one can come up with a more direct condition to determine whether a causal patch boundary retains information or not.

4

Bekenstein bound in the bulk and AdS/CFT

4.1 Introduction

Bekenstein bound [70] is the universal upper bound on the entropy S or information that can be contained in a physical system or object with given size and total energy. If R is the radius of a sphere that encloses a given system, while E is its total energy including any rest masses, then its entropy is bounded by,

$$S \leq \lambda R E \tag{4.1.1}$$

where λ is a numerical constant of order one. Although the derivation of the bound uses *generalized second law* around black holes [16], the bound seems to be independent of the gravitational physics. This fact manifests itself in G_{N} independence of the bound. Moreover the size of the box R is the geodesic distance in flat space. These observations indicate that the Bekenstein bound is valid in flat space hence can be derived via information theoretic inequalities in QFTs. In [72] the bound is derived by employing positivity of relative entropy for certain class of excited states with respect to vacuum. The derivation exploits the local expressions of modular Hamiltonians of certain spatial sections of vacuum density matrices. On the other hand, the Bekenstein bound manifests itself also on the systems having strong self gravitation. It is a well known fact that the bound is saturated for the Schwarzschild black hole. In other words the Bekenstein bound is saturated when Schwarzschild energy is put into a box of Schwarzschild radius. This is a strong indication that the bound preserves its validity beyond the weak self gravitating systems and hence should have a formulation for systems having back-reactions. One difficulty that is encountered when the system has back-reaction on the spacetime, we don't have at hand a natural definition of the size of the box. For example, in the case of a Schwarzschild black hole the radius of the box is not geodesic distance but the radial coordinate corresponding the black hole horizon. Because of the above observations, natural questions arise. How to define the radius of the box R in the presence of back reaction? What is the energy of the system E for a strong self gravitating system or in a setup that allows backreaction? Let us emphasize that for a black hole, size of the box R is the coordinate radius and energy of the system is the ADM mass of the solution which includes the binding energy of the entire solution. In a way, size of the box is determined with respect to the vacuum solution. Vacuum solution provides a reference grid where excited system can be compared to. Our main goal is to identify the boundary information theoretic observables corresponding the Bekenstein bound in the bulk including systems having non perturbative backreactions on the spacetime metric, such as black holes. We will study the problem using AdS/CFT [51–53] and find the corresponding information theoretic inequality on the CFT that describes the bound in the bulk. We will give the formulation of Bekenstein bound in the bulk for certain class of excitations on asymptotically AdS_{d+1} through the information inequalities in the dual CFT.

Before giving the formulation of the Bekenstein bound in the bulk via the underlying theory, we will clarify issues regarding the first law of entanglement entropy on a simple exercise involving conical defects. This will be a simple demonstration on how quantities involved in the first law type relations in CFT are identified with the quantities in spacetime. In this identification local expressions for modular hamiltonians [122] and Ryu-Takayanagi formula [57, 58, 123] play the central role. The so called first law of entanglement due to presence of conical defects provides us a puzzle which we address in the third section. The solution of this puzzle will also clarify the differences between pure state and thermal (in general mixed state) perturbations when first law of entanglement is considered. The complete knowledge of the pure state puts strong constraints on the expectation value of local stress energy on the dual CFT. Such constraints do not exists for mixed state perturbations as we will demonstrate.

In the forth section we have exercised the initial conical defect setup using the fundamental expression of covariant phase space formulation [64] of the first law of black hole thermodynamics. The formulation is valid for perturbative excitations[1] and conical defect solutions can be studied in this regime. In this formulation, origin of the differences between the relative entropies of the complementary regions will be clear. Due to the pure state nature of the conical defect perturbation, one can reduce the differences between relative entropies of the complementary spherical regions ($\{A, \bar{A}\}$) on the boundary into differences of modular hamiltonians.

[1] Along the chapter, perturbation is considered for two cases. It is either indicating a perturbation on the underlying state, *i.e.* deformation of the state into a nearby one in the Hilbert space or the metric field is perturbed by a δg. As it will be shown these two cases do not have to match at every order of perturbation.

Moreover the differences in modular energies, $\Delta H_{\bar{A}} - \Delta H_A$ can be identified with the bulk modular hamiltonian in the perturbative regime. Using the boundary expression of differences between modular energies of the complementary spatial sections we extend the notion of bulk modular energy of a spherical region around the 'origin' to non perturbative excitations. The result of the calculation of the vacuum subtracted full modular hamiltonian for excited states are very elegant and does not depend on the dimension of the spacetime. Here we introduce the notion of *sphere of ignorance* which is the bulk region that is not accessible from any boundary interval having size below a certain length scale. In other words an observer having access to the region A in the underlying theory can not decode anything in the bulk beyond the bulk sphere $< R_{\text{scale}}$. The essential relation we derive is that the modular energy contribution of the bulk excitation depends linearly to the scale of the system that contains it. This is closely related with UV-IR correspondence in AdS/CFT, and will be explained in more detail throughout the chapter. The full modular hamiltonian—which we also named as *bulk modular energy* inside the sphere of ignorance—for spherical excitation is given by

$$\Delta H_{\bar{A}-A} = 2\pi R_{\text{scale}}\, \Delta M_{\text{ADM}}. \tag{4.1.2}$$

The definition of full modular hamiltonian is $\hat{H}_{\bar{A}-A} \equiv \hat{H}_{\bar{A}} - \hat{H}_A$ and R_{scale} is the radial position of the tip of the minimal surface that is homologous to region A on the boundary.

ΔM_{ADM} is the vacuum subtracted ADM mass. The expression is very elegant as it is valid in any dimension. The above expression is not only valid perturbatively but holds also for excited states that can not be expressed as infinitesimal deformations of the vacuum. In the non perturbative case, R denotes the radial position of the point of ignorance —tip of the geodesic —with respect to the global vacuum.

A similar expression is used in [73] as a change of area in the weak field limit for certain identification of the manifolds and interpreted as the amount of entanglement entropy reduced by bulk excitation from its surrounding. In our interpretation, it quantifies the modular energy contribution of the bulk excitation to the entanglement wedge that contains it. The presence of bulk excitation in the entanglement wedge is the source of the differences between modular energies of complementary regions in the underlying theory.

For the perturbations around AdS_{d+1}, one can relate the boundary modular energy with the bulk modular integral as explained above. However, this is only possible for pure state deformations around the vacuum as it is explicitly using the relation $\Delta S_A = \Delta S_{\bar{A}}$ for pure states. Detailed explanation about this can be found in Section 4. On the other hand one can come up with a mathematically similar, yet physically quite different expression when the underlying state is thermal, or

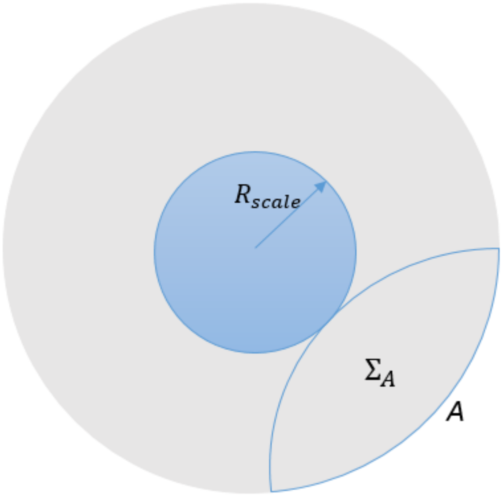

Figure 4.1: *Blue sphere denotes the region that observers who have access to the boundary region of size(A) are blind to. Radius of the sphere, R_{scale}, is the deepest point that can be decoded by accessing A. In a spherically symmetric state, sub-states having support on the region of same size, have the same change in information content, hence the only region that can not be decoded by the observer A is the blue sphere which is referred as the sphere of ignorance along the chapter.*

in general, a mixed state. In a mixed state, one can not constrain the change of entanglement entropy of the complementary regions via their equality. This allows a net change in the total energy of the system [124] when the perturbation is of thermal nature, which we have detailed in section 3 starting from a conical defect exercise. Using the first law of entanglement entropy and freedom to increase total energy of the system in a mixed state perturbation we end up with the following expression for mixed states when the local bulk excitation is confined into the sphere of ignorance

$$\delta S_{\bar{A}-A} = \delta S^{\text{bulk}}_{\mathbb{S}^{d-1}_R} = 2\pi R_{\text{scale}}\, \delta M_{\text{ADM}} \tag{4.1.3}$$

where $\delta S^{\text{bulk}}_{\mathbb{S}^d_R}$ is the bulk entropy that corresponds to $1/N$ corrections to the entanglement entropy in the CFT [66]. We distinguish vacuum subtracted quantities when the state is perturbation around the vacuum $|\psi\rangle = |0\rangle + \epsilon|\phi\rangle$, by lowercase (δ). This entropy resides inside the $d-1$-sphere of radius R. We find it interesting to compare the differences between pure states and thermal states excitations, since these differences are closely related with the volume law [73] and ER-EPR conjectures [47]. It is crucial to realize the two main differences between these two equations. Firstly, while the eq. (4.1.2) can not take place at the linear level of the perturbation on the underlying state, the later, (4.1.3) can be derived only through using the first law of entanglement hence at the linear level. In other words in the first case the constraining equation is equality of change of entanglements for complementary states, in the second case the first law of entanglement entropy, *i.e.* equality of change of modular energy and entanglement entropy for each state. Therefore these two results have different physical principles behind them. Secondly the quantities involved in the relation, that is change of modular energy, $\Delta H_{\bar{A}-A}$ vs change in the entropy $\delta S_{\bar{A}-A}$. In the first case the bulk excitation does not possess any entropy in the form of entanglement with the purifying auxiliary system. The mixed state deformations of the vacuum at the linear level is particularly important since this is the only case where we can directly show that the entropy inside the sphere of ignorance is equal to the difference of change in the entanglement entropies of the underlying theory, $\delta S^{\text{bulk}} = \delta S_{\bar{A}-A}$. Equality holds when all the excitation is localized inside this sphere.

Finally and most importantly, we will extend the result for thermal states to non perturbative level where differences are defined with respect to vacuum. In this case one can not use the first law of entanglement entropy, which was used in (4.1.3) to obtain the equality. In the generalization, we have expressed (4.1.3) in terms of the differences of relative entropies of the complementary states. For a rotational invariant state imposing the positivity together with the monotonicity of relative one can come up with inequality version of equation (4.1.3). The differences of modular hamiltonians can be expressed in terms of the bulk quantities. This is the direct analog of the first equation (4.1.2). We have observed that difference of the change in the entanglement entropies of the complementary regions is bounded by the associated bulk modular energy, which was defined also through (4.1.2). The expression is interesting as it is strikingly reminiscent of the Bekenstein bound [70] in the bulk.

$$\Delta S_{\bar{A}-A} \leq 2\pi R_{\text{scale}}\, \Delta M_{\text{ADM}} \tag{4.1.4}$$

The radius R is interpreted on AdS_{d+1}. It is the radius of the ignorance sphere defined with respect to the boundary region A. We argue that for spherically symmetric state the entropy that resides inside the sphere of ignorance is bounded by the entropic differences of the complementary boundary balls $S^{\text{bulk}}_{\mathbb{S}^{d-1}_R} \leq \Delta S_{\bar{A}-A}$. We will read this expression backwards—in a sense—by asking what the information theoretic extensions of the observables of underlying theory to the bulk physics. It will be explained why the differences in entanglement entropies of the complements in the underlying state manifest itself as the Bekenstein bound in the bulk. Boundary perspective that is presented throughout the chapter have some overlapping content presented in [125] [126] yet the bulk interpretation is completely novel according to our knowledge.

4.2 A simple first law

We would like to start with a simple explicit example for the gravitational counterpart of the first law of entanglement. We will shed light onto interesting set of constraints for the first law entanglement entropy. Let us consider simple case of AdS_3 which will be sufficient for our purposes.

First law of entanglement entropy holds for any state since the relative entropy vanishes at the linear level perturbatively. To see this, consider a reference state ρ_0 and an arbitrary state ρ_1. One can construct a family of interpolating density matrices

$$\rho(\lambda) = (1-\lambda)\rho_0 + \lambda\rho_1 \tag{4.2.1}$$

where λ can be positive or negative, yet the relative entropy $S(\rho(\lambda)|\rho_0)$ is positive for either sign of λ by the positivity of relative entropy. Hence first derivative of the relative entropy with respect to λ vanishes. This is the first law of entanglement since relative entropy can always be expressed as,

$$S(\rho(\lambda)|\rho_0) = \Delta\langle \hat{H}_A\rangle - \Delta S_A \tag{4.2.2}$$

where $\Delta\langle \hat{H}_A\rangle = \mathrm{tr}(\rho(\lambda)\hat{H}_A) - \mathrm{tr}(\rho_0 \hat{H}_A)$ and $\Delta S_A = -\mathrm{tr}(\rho(\lambda)\log\rho(\lambda)) - \mathrm{tr}(\rho_0\log\rho_0)$. In the leading order of λ relative entropy exactly vanishes, which is known as the first law of entanglement entropy,

$$\delta S_A = \delta\langle \hat{H}_A\rangle \tag{4.2.3}$$

where we denoted the linear perturbation by lowercase delta (δ). However, there are only few cases where the local expression for $\hat{H}_A$ is known explicitly [127]. One such case is the ball shaped region in the vacuum state of a CFT in any dimension. In this case modular hamiltonian has a local expression in terms of the local stress energy tensor of the CFT.

$$\hat{H}_A = 2\pi \int_A \zeta^\mu\, \hat{T}_{\mu\nu}\, d\Sigma^\nu \tag{4.2.4}$$

where ζ^μ is the conformal Killing vector that leaves the causal diamond of the ball shaped region invariant. This is an exact operator expression obtained by the conformal transformation from the half of the Minkowski space via employing the invariance of vacuum under global conformal transformations. We will give a straightforward yet illuminating application of this local expression of the modular hamiltonian.

Consider a 2d CFT on $\mathbb{R} \times \mathbb{S}^1$ with a classical gravitational dual. Suppose the vacuum is perturbed to a nearby pure state, $|\psi\rangle = |0\rangle + \epsilon|\phi\rangle$, such that the perturbed state has uniform expectation value for energy density *i.e.* $\delta\langle T_{00}\rangle = \mu$. Let us consider a ball shaped region on the spatial slice *i.e.* a constant time slice of the CFT. For an interval on the boundary, explicit expression for the change of modular energy is given by,

$$\begin{aligned}\delta\langle\hat{H}_A\rangle &= 2\pi\int_A \left(r\frac{\cos(\theta_0-\theta)-\cos(\alpha)}{\sin\alpha}\right)\langle\hat{T}_{00}\rangle\, r d\theta \\ &= 2r(1-\alpha\cot\alpha)\delta E_{\mathrm{CFT}}\end{aligned} \tag{4.2.5}$$

where θ_0 is the center of the region in the boundary, α is the half of the total angular size of the boundary ball and r is the radius of the $\mathbb{S}^1$. δE_{CFT} is the change in the energy of the state, which is equal to the change in the total energy of spacetime, given by $2\pi r\mu$. The combination, $r\delta E_{\mathrm{CFT}}$ is the dimensionless factor that we should expect to identify in the gravitational dual.

The modular hamiltonian side of the calculation does not refer to the gravitational dual. We simply use the local expression for the modular hamiltonian on the CFT which is exact thanks to conformal symmetry. One can check the first law explicitly via holography, by identifying the geometric description of the state in the gravitational description. We have considered the perturbation to be a pure state with a uniform asymptotic energy density. The latter ensures the bulk solution to be a spherically symmetric one. One possible geometric description of the state is the conical defect. Conical defect geometries are 3d solutions of Einstein equation with a Dirac delta type source distribution. Although the expression for the change of modular energy is valid between any state and vacuum, the first law only holds for small perturbations around vacuum. Therefore we would like to look at change of entanglement entropy in the presence of a conical defect with small deficit angle which serves as the perturbation parameter.

The metric of the conical defect is given by,

$$ds^2 = -\left(\gamma^2+\frac{R^2}{L^2}\right)dT^2+\left(\gamma^2+\frac{R^2}{L^2}\right)^{-1}dR^2+R^2d\theta^2 \tag{4.2.6}$$

where $0< \gamma <1$ and related to the deficit angle as $\delta\theta = 2\pi(1-\gamma)$. The minimal surface that is homologous to a boundary region A measures the entanglement entropy of the subsystem that resides on A on the dual CFT.

$$S(\alpha) = \frac{2L}{4G_{\mathrm{N}}}\log\left(\frac{2L}{\gamma\epsilon}\sin(\gamma\alpha)\right) \tag{4.2.7}$$

ϵ corresponds to the UV cutoff in the underlying theory, and serves as a IR cutoff in the bulk that regularizes the infinite area of the boundary sphere in AdS. Although this expression is infinite in the limit that sends the cutoff to zero, the change of entanglement with respect to vacuum is finite once the cutoff is fixed. Looking at the change of entanglement entropy or any other quantity defined on different manifolds requires a comparison scheme. One such physical scheme is to keep number of degrees of freedom fixed since it is the characteristic of the theory describing these two states. From the spatial point of view, fixing degrees of freedom is to keep the ratio of the size of the system to the cutoff fixed.

$$\Delta S(\alpha) = \frac{2L}{4G_{\rm N}} \log\left(\frac{\sin(\gamma\alpha)}{\gamma\sin(\alpha)}\right) \tag{4.2.8}$$

which in the small deficit limit, $\delta\theta/2\pi \ll 1$, perturbatively becomes,

$$\delta S(\alpha) = 2L(1-\alpha\cot\alpha)\delta M_{\rm ADM} \tag{4.2.9}$$

since the change of ADM energy is $\delta M_{\rm ADM} = \delta\theta/8\pi G_{\rm N}$. Thus identification of $L\delta M_{\rm ADM} \equiv r\delta E_{\rm CFT}$ completes the demonstration of the first law. However, in the next section, we will have a closer look at this *so-called* first law. As we will see, what seems like a first law is, in fact, not a first law from the information theoretic point of view. It does not concern the linear level of the parameter that connects density matrices of the underlying theory.

4.3 Puzzles about the first law

In the previous section we have explicitly demonstrated, the first law of entanglement entropy through a gravitational calculation where excited state is considered to be a conical defect. We have matched the change of modular energy on the CFT to change of area of the minimal surface in the bulk due to appearance of a defect. Although it looks like we have computed different quantities in different theories, the RT conjecture maps them. A more careful look into what we have done will reveal a greater understanding.

Initially, we consider AdS_3 solution with a boundary $\mathbb{S}^1\times\mathbb{R}$. Usually when one is restricted to a subspace on the boundary, there is no need to specify the change on the state having support on the complementary region to study the first law. Yet considering the subspace with its complement puts strong constraints on the $\delta\langle T_{\mu\nu}\rangle$. Remember that we considered a uniform perturbation on the vacuum, $\delta T_{00}(\theta) = \mu$ which is identified with conical defect solutions that has Dirac delta type sources in Einstein equation [128]. These solutions can be arbitrarily close to AdS_3 as it is possible to choose $\delta\theta/2\pi \ll 1$. Hence one can consider the bulk dual of the perturbation on the vacuum as a conical defect solution. Of course, this is just one particular solution with the given boundary energy density, one may come up with different semiclassical gravitational descriptions having same energy density. For example the perturbation could also be due to thermal fluctuations around vacuum in which case dual would be thermal AdS_3. We will study the thermal perturbations later on in this chapter. First we will focus to the conical defect. It is crucial that, as explained under eq. (5.3.17), to be able match the change in modular energy to $\delta S(\alpha) = S(\alpha)_{\text{con.}} - S(\alpha)_{AdS_3}$ one needs to rescale the IR cut-off for the conical defect solution. This corresponds to keeping number of degrees of freedom fixed by fixing the proportions of UV cut-offs to the size of the systems where the underlying theory lives. To be explicit, if one considers the conical defect geometry as a angular cut, then one should rescale the UV cut-off on the boundary such that $2\pi r/\epsilon_{UV}$ is fixed. This is one such beauty of AdS/CFT that it provides an unambiguous way to compare observables on nearby solutions.

The question we would like to raise is if there exists a first law on each boundary interval as a result of conical defect type excitations. Before answering this question, let us consider a perturbation on the vacuum, which changes the state into another pure state that is infinitesimally close

$$|\psi\rangle = |0\rangle + \epsilon|\phi\rangle. \tag{4.3.1}$$

We didn't specify how the energy density of the perturbation is organized spatially. Suppose that, such a change in the state causes a localized linear perturbation $\delta\langle T_{\mu\nu}\rangle$. If the perturbation is completely localized inside of a spherical region A in the CFT then change in the modular Hamiltonian of the complement $\bar{A}$ vanishes. The first law ensures that the change of entanglement entropy also vanishes, $\delta S_{\bar{A}} = 0$. As you will realize, we end up with a contradiction. Because the perturbed state was also assumed to be a pure state. In that case we would expect $\delta S_A = \delta S_{\bar{A}}$, yet by the first law of entanglement, $\delta S_A \neq 0$. This would also violate the positivity of relative entropy, because $\delta\langle H_A\rangle = \delta S_A = \delta S_{\bar{A}} > 0$, yet $\delta\langle H_{\bar{A}}\rangle = 0$, which implies, $\delta\langle H_{\bar{A}}\rangle - \delta S_{\bar{A}} < 0$. What is the resolution of this apparent paradox [61]? Indeed we have made a false assumption, by considering pure state perturbation localized only in a region at the linear level of the perturbation theory of the underlying state. One needs to put equal amount of energy to the complement in a pure state. This fact can be demonstrated simply through a field theory argument. Let us start by constructing the following operator,

$$\hat{H} = \hat{H}_A - \hat{H}_{\bar{A}}. \tag{4.3.2}$$

This operator, known as *full modular hamiltonian*, generates conformal transformations that keeps the spherical region fixed, hence annihilates the global vacuum state. The simplest way to understand why this operator annihilates the global vacuum is to consider the half space in QFT, where modular hamiltonian is the generator of rotation in the euclidean plane. Then the combination $H = H_A - H_{\bar{A}}$ generates a boost on the whole state, which leaves the vacuum invariant

$$\hat{H}|0\rangle = (\hat{H}_A - \hat{H}_{\bar{A}})|0\rangle = 0. \tag{4.3.3}$$

Now if the state changes according to (4.3.1), then

$$\begin{aligned}\delta\langle H_A\rangle &= \epsilon\left(\langle\phi|H_A|0\rangle + \langle 0|H_A|\phi\rangle\right)\\ &= \epsilon\left(\langle\phi|H_{\bar{A}}|0\rangle + \langle 0|H_{\bar{A}}|\phi\rangle\right) = \delta\langle H_{\bar{A}}\rangle\end{aligned} \tag{4.3.4}$$

The above equality indicates that whenever one creates some some localized wave packets inside a region, to stay in a pure state, some energy density needs to be introduced outside. It is the purity of the state that enforces such constraint.

Let us further study what kind of constraints we have on the perturbation of the expectation value of stress energy tensor. Assume that constant time slice has the topology of $\mathbb{S}^{d-1}$ with radius r. The modular Hamiltonian for $(d-2)$ dimensional spherical entangling surfaces surrounding a cap of the S^{d-1} specified by the angle α is given by,

$$H_{S_A^{d-2}} = 2\pi \int_0^\alpha r^{d-1}\, d\Omega_{d-2} \sin^{d-2}\theta\, d\theta \left(r\frac{\cos\theta - \cos\alpha}{\sin\alpha}\right) T_{00}(\vec{r}) \tag{4.3.5}$$

This expression is the generalization of what we have used for CFT$_2$. Now we can obtain $H_{\bar{A}}$ by sending the origin of the spherical cap, $\mathbb{S}^{d-1}$ to π and α to $\pi - \alpha$. Then the full modular Hamiltonian $H_{\bar{A}-A} \equiv H_{\mathbb{S}_A^{d-2}} - H_{\mathbb{S}_{\bar{A}}^{d-2}}$ becomes,

$$H_{\bar{A}-A} = 2\pi \int_0^\pi r^{d-1}\, d\Omega_{d-2} \sin^{d-2}\theta \left(r\frac{\cos\theta - \cos\alpha}{\sin\alpha}\right) T_{00}(\vec{r}) \tag{4.3.6}$$

For a pure state, by the first law of entanglement entropy, $\delta\langle H_A\rangle - \delta\langle H_{\bar{A}}\rangle = 0$ at the linear level. Using this equality one can put constraints to the possible $\delta\langle T_{00}\rangle$ for pure state perturbations. To see this explicitly, multiply $\delta\langle H_{\bar{A}-A}\rangle$ by $\tan\alpha$ and take the derivative w.r.t to α. The second term is eliminated and we have an example of such constraint which is $\int d\Omega_{d-1}\cos\theta\,\delta\langle T_{00}\rangle = 0$. Then using this one, we show easily, $\int d\Omega_{d-1}\,\delta\langle T_{00}\rangle = 0$. These are not the only constraints because the first one is obtained by placing the origin of the cap on the z axis. The full modular hamiltonian vanishes at the linear level independent of the choice of origin for boundary balls. It is true for all possible balls. Therefore one obtains the following set of constraints [124]

$$\int d\Omega_{d-1}\,\delta\langle T_{00}(\Omega)\rangle = 0, \qquad \int d\Omega_{d-1}\,\hat{\omega}\,\delta\langle T_{00}(\Omega)\rangle = 0. \tag{4.3.7}$$

$d\Omega_{d-1}$ is the volume form on $\mathbb{S}^{d-1}$ and $\hat{\omega}$ is the unit vector parametrizing the points on $\mathbb{S}^{d-1}$. Note that second constraint is generalization of $\int d\Omega_{d-1} \cos\theta\, \delta\langle T_{00}\rangle = 0$ in which case one focuses on the z-component of $\hat{\omega}$. The first constraint resolves our initial confusion. One can not introduce local excitations on some regions without balancing them with negative energy contributions. Another interesting distribution that is violating the second constraint is Dirac delta sources unless equal amount of positive and negative charges introduced at the same point, which is very constraining. It would be interesting to see examples of distributions satisfying these constraints, yet we leave it for future studies. We are now ready to further puzzle ourselves with the first law in the presence of a conical defect.

If we now go back to our initial construction, where we consider the conical defect solution with a small deficit as the bulk dual of a homogeneous perturbation of the boundary theory around the vacuum. The change of area of the minimal surfaces due to conical defect was equal to $\delta\langle H(\alpha)\rangle$ for each boundary region. When $\delta\langle T_{00}\rangle$ is uniform then conical defect is at the center. Their fusion yields us the generalization of the first(?) law to arbitrary surfaces which we present in chapter 5. At this point we need to reexamine this first law in the light of the puzzles we had uncovered.

First of all, as we have derived above, the first law of entanglement for any pure state perturbation puts some constraints on the total energy of the perturbed state such that it vanishes (4.3.7). On the other hand, we expect the conical defect geometry to represent a pure state in the underlying theory, yet clearly its energy does not vanish. Therefore it violates the constraints above. Furthermore, assuming that for each boundary region, A, there is a first law, then we would expect that the state on $\bar{A}$, also satisfies the first law. However through the Wald formalism, it is clear that when perturbation is sourced by stress tensor then the first law is modified by contribution from stress energy of the perturbation, this will be further explained in the following sections. Hence one does not have a first law for the complement $\bar{A}$ whose entanglement wedge includes the defect. In this case, one has $\delta H_{\bar{A}} - \delta S_{\bar{A}} \sim \delta$, where δ quantifies the deficit angle. Therefore either the perturbed state is not pure or this is not the first law of entanglement entropy. We know that the perturbed state is pure, since it is obtained from the vacuum by adding a localized non thermal mass. Let us be more careful about the order of perturbation. The first law takes place at the linear order in ϵ, where fields perturbed according to $\phi \to \phi + \epsilon\, \phi^{(1)} + O(\epsilon^2)$. However the perturbation of the vacuum, by a classical mass distribution takes place in second order as $T_{\mu\nu}$ is quadratic in fields. Indeed, what we have called as first law was taking place between the second order and first order. Let us look at it in more detail.

From this point on, we will replace our notation for the variation with Δ to denote that difference is beyond the linear order. The perturbation of the solution by addition of classical matter, as explained above, takes place at the second order in perturbation theory. If the perturbed state is a pure state, $|\psi\rangle = |0\rangle + \epsilon|\phi\rangle + O(\epsilon^2)$, then the full modular Hamiltonian at this level does not vanish. Remember that the vanishing of that at the linear level yields us the constraints on the change of $\delta\langle T_{00}\rangle$, which does not exist beyond linear level.

$$\langle\psi|H|\psi\rangle = \epsilon^2\langle\phi|H|\phi\rangle + O(\epsilon^3) \tag{4.3.8}$$

Therefore we are not restricted to the constraints in (4.3.7) at this level. Remember that the first law of entanglement entropy is derived from the relative entropy at the linear level (4.2.1). Next to leading order, due to positivity of the relative entropy, one has,

$$\Delta H_A - \Delta S_A \geq 0 \tag{4.3.9}$$

The interesting point is that, when the perturbation is due to some localized mass distribution outside of the entanglement wedge A one has $\Delta H_A = \Delta S_A$ at the leading order of perturbation by $\delta g_{\mu\nu}$. That was what we observed when the perturbation was due to conical defect. On the other hand, for the complement of the region A, whose size is more than half space, $\alpha > \pi/2$, $\Delta H_A \geq \Delta S_A$ and the difference is proportional to T_{00}^{bulk} of the localized source.

4.3.1 Thermal perturbation

In this part, we would like to point out the differences when the perturbation is a mixed state. Although Einstein equations are agnostic whether the source is thermal or pure, in the microscopic description these two cases are substantially different. For example, in the presence of a BH, when the subsystem size reaches a critical value, the difference, $S_{\bar{A}} - S_A$ saturates the Araki-Lieb bound namely, $S_{\bar{A}} - S_A = S_{A\cup\bar{A}}$ due to homology constraint. In general for a mixed state, one expects separation of the minimal surfaces of a subsystem and its complement which are same in a pure state. Hence the more thermal the system is greater the ignorance becomes. In the thermal case, the amount of information (better to say amount of uncertainty) of the complementary regions of a quantum state does not match due to thermal entropy. Therefore it is not possible to extend the constraints of the previous section to the thermal cases. Let us look at the simple illustration of this fact for a thermal state perturbation on the vacuum,

$$\rho = \frac{|0\rangle\langle 0| + \sum_i e^{-\beta E_i}|i\rangle\langle i| + ...}{1 + \sum_i e^{-\beta E_i} + ...}. \tag{4.3.10}$$

Here we consider the low tempreture expansion of a thermal state. The CFT side of the story had been studied in [126, 129].

The knowledge that the state is thermal itself is not enough to determine the ontological character of the entropy, as it can be seen as either being entanglement or thermal entropy [67]. Surely it is thermal entropy but at the same time one can consider it as the entanglement entropy with its purification[2]. Therefore we will not distinguish these two cases, but when we refer the concept as entanglement entropy, it is the entanglement with respect to the purification, not between the complementary regions of the underlying theory. In other words, we do not refer to the entanglement between subsystem of the mixed state which is very difficult to quantify for an arbitrary mixed state.

It is a well known fact that entanglement entropy of a system and of its complement do not match for a thermal state. This is also true for the change of entanglement in the leading order. Hence there is no such constraints eq. (4.3.7) when the perturbation is a mixed state, which can be shown through non vanishing of full modular hamiltonian. In this case, expectation value of the full modular hamiltonian around the vacuum up to first order does not vanish anymore.

$$\delta\langle H_A\rangle = \sum_i \mathrm{tr}[(\mathrm{tr}_{\bar{A}}|i\rangle\langle i| - \mathrm{tr}_{\bar{A}}|0\rangle\langle 0|)H]e^{-\beta E_i} \tag{4.3.11}$$

the differences of modular hamiltonians for complementary regions become,

$$\delta\langle H_{A-\bar{A}}\rangle = \sum_i \langle i|H_{A-\bar{A}}|i\rangle e^{-\beta E_i} \tag{4.3.12}$$

where $H_{A-\bar{A}} \equiv H_A - H_{\bar{A}}$, each of which acts trivially outside its domain. Since $H_{\bar{A}-A}$ does not annihilate excited states, there is no equality between $\delta\langle H_A\rangle$ and $\delta\langle H_{\bar{A}}\rangle$. Although this relation is simply reflecting the fact that entanglement entropies of complementary regions in a mixed state are different, the first law always satisfied at the linear order, without having any further constraint.

4.4 Relative Entropy through Energy and Scale

At this point, it is clear that the first law-like relations for pure states due to a localized excitation, are actually occurring at the nonlinear level in field variations. In this part we will have a closer look to the origin of the mismatch between $\Delta\langle H_A\rangle$ and $\Delta\langle H_{\bar{A}}\rangle$ using covariant phase space approach. Although the former equals the

[2]Purification of a mixed state is not unique, yet the mixed state having different purifications have same von Neumann etropy.

change of entanglement entropy in the case of a conical defect perturbation, the latter is always greater than that. So we can begin by asking what the difference $\Delta\langle H_{\bar{A}}\rangle - \Delta S_{\bar{A}}$ corresponds to in the bulk. Firstly, we will study the difference between $\Delta\langle H_A\rangle$ and ΔS_A for a state near the vacuum via perturbation theory and later we will generalize the result to any spherically symmetric excitation.

4.4.1 Perturbation theory at the non linear level

To see the origin of the difference between modular hamiltonian and entanglement entropy, let us study the first law via covariant phase space formulation. We will not give the review of the formulation here rather we will use the fundamental theorem of the formalism which ties the linearized equation of motion in the bulk to change of surface charges. The change of surface charges around the vacuum associated to Rindler horizon generating vector fields match with the information theoretic quantities in the microscopic theory. This identification has been used to derive linearized equation in the bulk through the first law of entanglement entropy in CFT [62]. The fundamental theorem of covariant phase space approach states that (5.3.10),

$$d\chi_\xi = \omega(\Phi, \delta\Phi, \mathcal{L}_\xi\Phi) - 2\xi^a \delta E(\Phi)_{ab}\epsilon^b \tag{4.4.1}$$

Φ stands for the whole field content of the theory including gravitational fields. ω is presymplectic potential and ϵ^a is d dimensional volume form. E_{ab} is the equation motion derived from full lagrangian including gravitational part.

The equation above is valid when the equations of motion satisfied for the unperturbed state. χ_ξ is a $(d-1)$ form, whose integral on the boundary region homologous to a minimal surface, yields the change of modular energy and the integral on the minimal surface itself gives the change of the area when the perturbation is considered around the vacuum.

$$\int_\Sigma d\chi_\xi = \oint_{\partial\Sigma} \chi_\xi = \Delta\langle H_A\rangle - \Delta S_A \tag{4.4.2}$$

where Σ denotes the d dimensional timelike hypersurface bounded by minimal surface and infinity. Although the theorem is valid for any solution and the perturbations around it. The correspondence between integral of χ_ξ and the information theoretic quantities in the underlying theory is constructed around the vacuum. When ξ is a Killing field, the presymplectic potential vanishes identically, as $\mathcal{L}_\xi g = 0$.

The derivation of linearized Einstein equations via the first law of entanglement entropy can be demonstrated simply by the fundamental theorem of the covariant phase space approach, where vanishing of $\Delta H_A - \Delta S_A$ implies the vanishing of $\int \xi^a \delta E_{ab} \epsilon^b$. When the relation is satisfied for every boundary ball, it can be turned into a local equality that is equivalent to linearized Einstein equations. However in our case, we have shown explicitly that the first law of entanglement entropy does not hold. Therefore the difference between modular hamiltonian and entanglement entropy is equal to the linearized equation of motion sourced by the bulk matter stress tensor. If one only considers the gravitational part of the field content, any addition of matter stress can be included at the level of linearized equation as the right hand side of the equality. Hence, the linearized equations are not source free for the perturbed state, which is the bulk point of view on the mismatch between modular energy and entanglement entropy. We should emphasize that we are not deriving linearized Einstein equation with classical source around the vacuum. The derivation of linearized Einstein equation with source had been proposed in [130]. However the source term $\langle T_{\mu\nu} \rangle$ in that derivation was semiclassical by nature hence it can appear as the leading term in the perturbation theory around the vacuum, which corresponds to quantum ($1/N$) corrections on the CFT. Here we assume that the geometry is perturbed by a classical stress energy tensor which appears at the quadratic order in the perturbation theory. Yet one can study the backreaction of the gravitational field on the matter fields through linearized Einstein equations as we will do. It does not seem possible to us that one can derive linearized Einstein equation with a classical source around the vacuum via a perturbation theory on the microscopic state since the source term and gravitational part appear at different orders. To summarize, the difference between the modular Hamiltonian and the entanglement entropy (relative entropy) is equal to modular integral of the bulk stress energy. The exact expression can be obtained simply by inserting the source term for linearized Einstein equation.

$$S(\rho_A || \rho_A^{\text{vac.}}) = \Delta H_A - \Delta S_A = \int_\Sigma \xi^a T_{ab} \, \epsilon^b \tag{4.4.3}$$

which has a simple expression for spherically symmetric configurations, in which case one can evaluate it without detailed knowledge of the energy distribution in the bulk. This will be the way we extend the notion of bulk modular energy to excited states beyond perturbation theory.

The equation (4.4.3) is obtained by Stokes theorem. In general, variation on the holographic entanglement entropy can originate from two different sources. It can either come as the variation of the minimal surface or variation of the metric field on the surface. To stay in the domain of validity of the Stokes theorem the variation of the entanglement entropy with the minimal surface should vanish. This is the case for linear perturbations since entropy functional (area) is extremized on the same surface. Hence the perturbation theory should be truncated beyond this order. In the first order, entanglement entropy on the CFT gets contributions only from the expectation value of the boundary stress tensor. At the second order, all the one point functions start to contribute. Let us have a look at how the metric behaves near the boundary in the Fefferman-Graham expansion where the geometry near the boundary can be expanded in the following way,

$$g_{\mu\nu} = \frac{L^2}{z^2}(dz^2 + g_{\mu\nu}dx^\mu dx^\nu) \tag{4.4.4}$$

when the asymptotic boundary is flat $g_{\mu\nu} = \eta_{\mu\nu} + \delta g_{\mu\nu}$. In this expansion one can determine the behavior of the scalar fields near the boundary in terms of the one point functions on the CFT, $\phi \sim \gamma z^\Delta \langle \mathcal{O} \rangle + ...$ when fed back to Einstein equation with a scalar field, the boundary expansion of the metric is altered in the following way,

$$\delta g_{\mu\nu} \sim a z^d \langle T_{\mu\nu} \rangle + b z^{2\Delta} \langle \mathcal{O} \rangle \eta_{\mu\nu} + ... \tag{4.4.5}$$

where Δ is the dimension of the operator $\mathcal{O}$. Let us consider a case where both operators contribute at the same energy scale μ on the boundary theory, then contribution of each term to the entanglement entropy becomes $\langle \mathcal{O} \rangle \sim \mu^\Delta$ and $\langle T \rangle \sim \mu^d$. The dimensionless perturbation parameter becomes $\mu r \ll 1$ where r is the radius of the sphere CFT lives. Entanglement entropy takes a contribution $(\mu r)^d$ from stress tensor at the linear order and $(\mu r)^{2\Delta}$ from the one point functions at the quadratic order. Let us emphasize that the leading order contribution to the entanglement entropy from scalar operators comes at the quadratic level. In this case when the dimension of the operator satisfies $\frac{d}{2} - 1 < \Delta < \frac{d}{2}$ its contribution becomes the dominant one. In our case, we demand the stress energy contribution to be dominant and truncate the perturbation theory at the quadratic order of the stress energy contribution. In this case, the dimension of the scalar operator can take values between $\frac{d}{2} < \mathcal{O} < d$. Since the perturbation parameter is chosen to be the combination (μr) we have control on the entire bulk without restricting ourselves to near boundary regions. This is the weak field limit in the AdS. In the example of the conical defect this corresponds to small angle deficit limit $\delta\theta \ll 1$. In the weak field limit one can decode the matter stress distribution in the entire bulk via the relative entropy on the underlying theory. Mathematically this corresponds to inverse Radon transform. This is the weak field version of the near boundary tomography presented in [131].

4.4.2 Appearance of Radial Scale

In this part we will demonstrate an interesting observation on the contribution of conical defect to the relative entropy. As it is explained in the previous subsection the presence of a localized source increases the relative entropy only in the entanglement wedge that contains it. This is in agreement with the idea of entanglement wedge reconstruction, where one can reconstruct the bulk regions corresponding to the entanglement wedges of those regions. The entanglement wedge reconstruction idea has been studied mostly in cases that exclude backreaction on the geometry [132], although some speculations are made for the case involving backreaction, [133]. In general, entanglement wedge reconstruction considers the construction of the bulk fields around a classical background using the boundary CFT. Here, we provide further evidence that the conjecture should be valid even when backreaction is considered. The conical defect solution is again a suitable framework where we can explicitly find what the increase is, in the relative entropy due to presence of the defect. Interestingly the contribution of the conical defect can be expressed in terms of the ADM energy of the defect and radius of the region that includes defect or by the radius of the region that can not be probed by the boundary interval that excludes the defect. We have calculated in (4.2.9) the change of modular energy for the boundary interval A whose size $\alpha < \pi/2$ in the presence of a small defect. One can easily deduce the corresponding expression for the complementary boundary interval where $\alpha > \pi/2$.

$$\Delta H(\bar{\alpha}) = \Delta H(\pi - \alpha) = 2L\,(1 - (\pi - \alpha)\cot(\pi - \alpha))\,\Delta M_{\mathrm{ADM}} \tag{4.4.6}$$

The conical defect perturbation deforms the vacuum into a nearby pure state hence change of the entanglement entropies for the complementary regions should be equal. This allows us to calculate relative entropy for the boundary region $\bar{A}$ directly through the differences of the changes of modular energies of the complementary regions. For a perturbative pure state excitation that is excluded by one of the complementary regions, the first equality below always hold and the second equality is what we have obtained in the case of a conical defect.

$$S(\rho_{\bar{A}}||\rho_{\bar{A}}^{\mathrm{vac.}}) = \Delta\langle H(\bar{\alpha})\rangle - \Delta\langle H(\alpha)\rangle = 2\pi R M_{\mathrm{con}} \tag{4.4.7}$$

where M_{con} is the vacuum subtracted energy of the conical defect. R is the radius of the sphere that, observers having access to region A on the boundary, is blind to. In other words R is the scale in the bulk beyond which, one can not extract any information by having access to the boundary region A, (Figure:4.1). Since the perturbed state is homogeneous or translational invariant, all the regions with same size on the boundary have the same information content which can be denoted by a scale on the boundary. Interestingly there is a one to one map between the scale on the boundary and the radius of the bulk sphere. Boundary observers having access to subsystems of size 2α has no information regarding the sphere of radius $R = L\cot\alpha$, where L is the curvature radius of AdS. This expression is remarkable in the sense that it yields the information theoretic content of the bulk excitations in a non-local way in terms of the bulk quantities. A similar expression is used in, [73] to motivate the information theoretic effect of introducing matter onto spacetime. This effect appears as a reduction of entanglement entropy once it is postulated that surface area is the measure for entanglement entropy of the quantum state describing spacetime. In the next section we will give the derivation of the radial scale R in higher dimensional generalization of conical defects where excited state can be considered as a perturbation over the vacuum.

4.4.3 Higher Dimensional Generalizations for Perturbative Excitations

In this section we will extend the connection between relative entropy and the bulk modular energy to higher dimensions for spherically symmetric perturbative excitations. We will show how the bulk radial scale enter into the calculation which will later be used for the derivation of the Bekenstein bound in the bulk. $d+1$ AdS can be represented by the hyperboloid,

$$X_0^2 + X_{d+1}^2 - \sum_{i=1}^{d} X_i^2 = L^2 \tag{4.4.8}$$

embedded in $d+2$ dimensional flat space. Equation (4.4.8) can be solved by setting,

$$\begin{aligned} X_0 &= L\cosh\chi\cos\tau, \\ X_{d+1} &= L\cosh\chi\sin\tau, \\ X_i &= L\sinh\chi\,\Omega_i, \end{aligned}$$

where $\sum_{i=1}^{d-1}\Omega_i = 1$ and spans the trigonometric functions of θ, ϕ_i where i runs in $\{1...d-2\}$. The solution $(\chi \geq 0,\ 0 \leq \tau \leq 2\pi)$ covers the entire hyperboloid hence yields the global description of AdS_{d+1}, whose metric becomes,

$$ds^2 = L^2(-\cosh^2\chi d\tau^2 + d\chi^2 + \sinh^2\chi d\Omega_{d-1}^2) \tag{4.4.9}$$

We would like to compactify the solution such that boundary resides at a finite value of the radial direction. The casual structure of AdS_{d+1} can be studied by the following coordinate transformation,

$$\sinh\chi = \tan\rho, \qquad 0 \leq \rho \leq \pi/2, \tag{4.4.10}$$

the metric becomes,

$$ds^2 = \frac{L^2}{\cos\rho^2}\Big(- d\tau^2 + d\rho^2 + \sin^2\rho\,\underbrace{(d\theta^2 + \sin^{d-2}\theta\, d\Omega_{d-2}^2)}_{d\Omega_{d-1}^2}\Big) \tag{4.4.11}$$

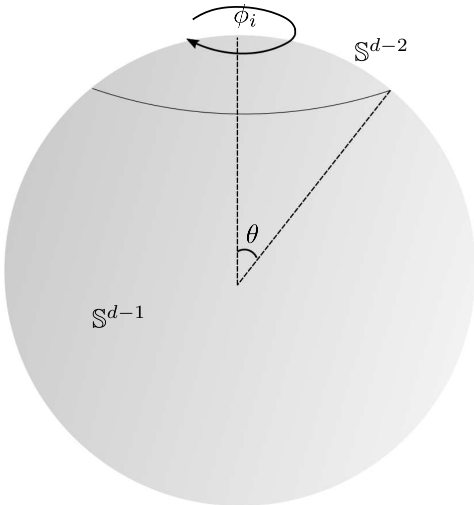

Figure 4.2: *Coordinate system used for constant time slice of AdS_{d+1}. Opening angle 2α of the boundary ball $\mathbb{S}^{d-2}$ is measured by the coordinate variable θ.*

In this coordinate system 4.2 boundary is located at $\rho = \pi/2$. Size of the $\mathbb{S}^{d-2}$ boundary ball, A is determined by the coordinate variable θ. The Killing field that generates the Rindler horizon in the global coordinates is given by,

$$\xi^a = \frac{\cos\tau \sin\rho \cos\theta - \cos\alpha}{\sin\alpha}\partial_\tau + \frac{\sin\tau \cos\rho \cos\theta}{\sin\alpha}\partial_\rho - \frac{\sin\tau \sin\theta}{\sin\rho \sin\alpha}\partial_\theta \tag{4.4.12}$$

ξ^a vanishes on the minimal surface that is homologous to the boundary ball A and it has unit surface gravity. α is the size of the boundary ball. Let us turn back to the problem of generalizing out observation on the conical defect to higher dimensions. Consider a spherically symmetric perturbative excitation in the AdS_{d+1}. Stress energy distribution characterizing this excitation depends only on the radial coordinate $T_{ab}^{\mathrm{bulk}} \equiv T_{ab}^{\mathrm{bulk}}(\rho)$ which corresponds to a uniform (spherically symmetric) energy density on the boundary. Let us pick the constant time slice to study the relative entropy around the vacuum whose geometric dual is given by the solution (4.4.11). Suppose there is no excitation in the casual wedge of the boundary region A *i.e.* all of the bulk excitation confined inside the complement $\bar{A}$. In this case the difference of the change of modular energies equal to the relative entropy of the complement and can be expressed as the modular integral of the bulk stress energy.

$$\begin{aligned} S(\rho_{\bar{A}}||\rho_{\bar{A}}^{\mathrm{vac.}}) = \Delta\langle H(\bar{\alpha}) - H(\alpha)\rangle &= 2\pi L \int_{\rho\le\rho_0} \frac{\sin\rho\cos\theta - \cos\alpha}{\sin\alpha} T^{\mathrm{bulk}}(\rho)\, d^dV \\ &= 2\pi L \cot\alpha\, \Delta M_{\mathrm{ADM}} \end{aligned} \tag{4.4.13}$$

Note that the expression is given in terms of curvature radius of the AdS_{d+1} as a necessity of dimensionless nature of relative entropy. The appearance of curvature scale plays an important role in the identification with the radial coordinates in the bulk. In the expression above ρ_0 represents the deepest point that can be probed via the boundary region A. Observers that have access to the smaller boundary can not decode the bulk beyond this point, hence it denotes a sphere of ignorance. Let us look at the physical interpretation of the factor $\cot\alpha$ from the bulk point of view. The equation of the minimal surface that are homologous to $(d-2)$ spheres on the boundary is given by,

$$\sin\rho\cos(\theta-\theta_0)=\cos\alpha \tag{4.4.14}$$

θ_0 denotes the center of the boundary ball in θ. The deepest point that the surface reach has the angular coordinate $\theta=\theta_0$. The radial coordinate of the point of ignorance becomes, $\sin\rho=\cos\alpha$. Let us represent the the radius of the the sphere of ignorance using the spherical coordinates. The radial coordinate sits in front of the angular directions in spherical coordinates as $R^2 d\Omega^2$. In the global AdS this corresponds to $\tan\rho$. Using the expression for the location of the tip in terms of α, we infer that radius of the sphere that observers having access to region A can not access becomes,

$$R_{scale}=L\cot\alpha \tag{4.4.15}$$

Therefore we have derived that in the perturbative regime the modular energy of spherically symmetric excitation can be seen as a non local contribution that depends on the size of the system. A similar result is used in [73] to motivate the idea that matter reduces the entanglement entropy of the spacetime in a way that is proportional to the radial scale of the hypothetical box that contains the excitation. Our result also indicate that it should be possible to construct modular modular hamiltonian for a spherical region in the bulk. To sum up we have obtained the following expression for the modular energy contribution of the bulk excitation to the entanglement wedge that includes the excitation,

$$\Delta\mathcal{E}^{\text{bulk}}\equiv 2\pi R_{scale}\Delta M_{\text{ADM}} \tag{4.4.16}$$

where bulk modular energy contribution to the entanglement wedge is defined as $\Delta\mathcal{E}^{\text{bulk}}\equiv\Delta\langle H_{\bar{A}}-H_A\rangle$. In the next section we will derive this relation completely through underlying theory and we will put some emphasis on the differences between pure and thermal state excitations.

4.5 Bekenstein bound and AdS/CFT

In the derivation of the expression (4.4.16), we used explicitly that the excited state is a pure state, in which the change of entanglement entropies of the complementary regions are equal. The purity of the excited state is used to relate the $\Delta\langle H_{\bar{A}} - H_A\rangle$ to the bulk modular integral of the stress energy tensor. To be explicit we employed the equality of $\Delta S_{\bar{A}}$ and $\Delta\langle H_A\rangle$ through their relation to ΔS_A. In the case of a thermal state, the change of entanglement entropies would not be equal anymore. Therefore, we could not derive the same expression when the excitation is a mixed state (4.3.10). In that case, bulk excitation would carry information that can not be deduced from the underlying state without access to auxiliary purification.

On the other hand, in the macroscopic description of the so-called first law, we have only specified change in the expectation value of the boundary stress tensor $\Delta\langle T_{\mu\nu}\rangle$ together with the knowledge of the purity of the perturbed state. The change in the expectation value of the boundary stress energy alone, does not specify the microscopic nature of the perturbation. Pure and mixed state perturbations are quite different although they may yield equal amount of change in the energy of the system. One important difference in their nature, as we have explained in section 4.3: a pure state perturbation with a non vanishing net energy increase can not take place at the linear level while that for a mixed state can. This is a very restrictive statement which implies that linearized Einstein equations with a classical source can not take place at the linear level from the point of microscopic theory, which as we have explained, took place at non linear level and source should be considered as the back-reaction of geometry on the stress energy tensor. This is expected, since bulk stress energy tensor vanishes for the perturbations at the linear level around the vacuum. However when the perturbation is a mixed state then there is no such constraint on the change of total energy of the system. To sum up when one only specifies the change of boundary stress tensor, one does not know the information theoretic content of the perturbation. A mixed state perturbation and a pure state one only differs in terms of their entanglement entropic content.

In the next section we will restrict ourselves to perturbative bulk excitations. In this regime one can use the covariant phase space formalism 2.2 and perform an analytical calculation in the bulk. In the perturbative regime the relation between boundary entanglement difference $|\delta S_{\bar{A}} - \delta S_A|$ and the entropy $\delta S^{\text{bulk}}_{\mathbb{S}_R}$ associated to the bulk region bounded by the area of the sphere of radius R_{scale} will be established. We will show that $\delta S^{\text{bulk}}_{\mathbb{S}_R} \leq |\delta S_{\bar{A}} - \delta S_A|$. The equality satisfies when all the excitation is confined to the sphere $\mathbb{S}_R$. In section 4.5.2, we extend our result to non perturbative regime. In this case a direct bulk computation is not possible however we observe the Bekenstein bound in the bulk purely from CFT.

4.5.1 Mixed state excitations at the linear level

Before observing the manifestation of the Bekenstein bound in the microscopic theory as the monotonicity of relative entropy on the CFT, let us study thermal states that are perturbations around the vacuum $\rho = \rho_0 + \delta\rho$ to see how different the observable $\delta S_{\bar{A}-A} \equiv \delta S_{\bar{A}} - \delta S_A$ behaves, which was vanishing for any pure state excitation.

In the case of a pure state, the change of entanglement for some region and its complement is equal, $\Delta S_A = \Delta S_{\bar{A}}$. Conical defect perturbation was an example of this case, where the defect was included only in one of the entanglement wedges, yet the information content of the complement is the same. We interpret this as by arguing that defect does not carry entropy in itself with respect to the underlying state containing it. If the perturbation is in the form of a mixed state, then one would expect totally different behavior. Note that we were also not allowed by the constraints, (4.3.7), to study the first law of entanglement for pure states which has $\delta E \neq 0$, which is not the case for thermal states.

Let us consider a mixed state perturbation at the linear order. It satisfies the first law of entanglement entropy, both for the boundary region A and its complement $\bar{A}$. Using the first law of entanglement, we can find the difference $\delta S_{\bar{A}} - \delta S_A$ which is considered as a measure of information associated to the entanglement wedge of $\bar{A}$, that can not be retrieved from A.

$$\delta S_{\bar{A}} - \delta S_A = \delta\langle H_{\bar{A}}\rangle - \langle\delta H_A\rangle \tag{4.5.1}$$

The change of entanglement entropy of the underlying state can be decomposed into two contributions. The area contribution and entanglement entropy of quantum fields in the bulk which emerges as quantum corrections to the underlying state. As we have explained in great detail in section 3, linearized perturbations on the underlying state can only change the total energy of the state if they are of thermal nature.

$$\delta S_{\bar{A}} = \frac{\delta A}{4G_{\mathrm{N}}} + \delta S^{\mathrm{bulk}}_{\Sigma_{\bar{A}}} \tag{4.5.2}$$

where Σ_A denotes the entanglement wedge of A. The contribution from bulk fields can be expressed as a local integral expression in the linear level [134].

$$\delta S^{\mathrm{bulk}}_{\Sigma_A} = \int_{\Sigma_A} \zeta^\mu \langle T_{\mu\nu}\rangle \, d\Sigma^\nu \tag{4.5.3}$$

when all the contribution is confined into the sphere of ignorance one can equate the bulk entanglement contribution of this region to the difference of change of modular energies. Because in this case even the perturbation is of mixed state nature, the change of areas would be equal due to extremal character of the surfaces. Hence,

$$\delta S^{\text{bulk}}_{\mathbb{S}_R} = 2\pi R \delta M_{\text{ADM}} \tag{4.5.4}$$

the bulk entanglement entropy resides in the sphere of ignorance is defined as

$$\delta S^{\text{bulk}}_{\mathbb{S}_R} = \int \Theta(R-r)\zeta^{\mu}\delta\langle T_{\mu\nu}\rangle \, d\Sigma^{\nu} \tag{4.5.5}$$

we obtain the entropic version of the (4.4.15). where $R = L\cot\alpha$ and α is the angular radius of the boundary region A. This is the maximum entropy that can be contained in the spherical region around the origin with radius, R for a system with energy M_{ADM}. It shows us that, the difference between entanglement entropies of complementary regions in a thermal perturbative excitation at the linear level is equal to saturation of Bekenstein bound for a system with energy M_{ADM} and size R. Indeed this is the region that the observer who has control on system A is blind to. Any deviation from thermal nature (mixture of thermal state and pure state as an ensemble), decreases the $\delta S_{\bar{A}-A}$ as it is zero in pure state.

Note that the expression diverges in the limit $\alpha \to 0$. In this case region $\bar{A}$ covers the whole boundary. How could we make sense of this expression, in this limit? There are two scales in the problem: that of α and β, the inverse temperature of the system. Although (4.5.4) does not contain β explicitly, it is absorbed into δE_{CFT} (4.5.8), which should be read in terms of the energy of lowest excited state weighted with the Boltzman constant and the degeneracy of the state $\delta E_{\text{CFT}} = \sum_i g_i E_i e^{-\beta E_i}$. This expression is sensitive to the order of limits and to make sense of it, one should consider the limit $\beta \to \infty$ before $\alpha \to \pi$ [129]. Although, order of limits can let us make sense of the expression in $\alpha \to 0$ limit, it is still an open question, at least to the author, how can we make sense of expression as an operator expression, since same expression can also be used to evaluate $\Delta\langle H_{\bar{A}}\rangle$ for non perturbative excited states.

4.5.2 Bekenstein Bound in the Bulk

Until now we have carried out a perturbative analysis using covariant phase space formulation. We computed the $\Delta\langle H_{\bar{A}-A}\rangle$ and $\delta S_{\bar{A}-A}$ as an integral of the bulk stress energy tensor using the fundamental theorem of the covariant phase space formalism. To emphasize again this is only valid when the geometric dual of the excited state can be seen as perturbation of the metric field around the AdS. In this section, we will show that one can go beyond perturbation theory solely using CFT modular Hamiltonian. Although the difference $\Delta\langle H_{\bar{A}} - H_A\rangle$ can be obtained only referring to the boundary quantities for non perturbative excitation, it has the similar bulk interpretations. Again the quantities involved in the expressions have different nature for pure and mixed states.

The modular Hamiltonian for a ball shaped region for the vacuum of the CFT is an operator expression (4.3.5). Hence it has no restriction on the state that the operator is evaluated. Until now we have used this expression for the states that are close to the vacuum, in that case there exists a dual bulk expression for the relative entropy when one of the entanglement wedges excludes any excitation in the bulk.

Let us consider an excited state $|\Psi\rangle$ which is orthogonal to the vacuum $\langle\Psi|0\rangle = 0$. One can study change in the modular energies by taking the expectation value of the stress energy around $|\Psi\rangle$. Considering the differences of the changes for complementary regions we obtain,

$$\Delta\langle H_{\bar{A}-A}\rangle = 2\pi \int_0^\pi r^{d-1}\, d\Omega_{d-2} \sin^{d-2}\theta \left(r\frac{\cos\theta - \cos\alpha}{\sin\alpha}\right) \Delta\langle \hat{T}_{00}\rangle(\vec{r}) \tag{4.5.6}$$

where $\Delta\langle\hat{T}_{00}\rangle \equiv \langle\hat{T}_{00}\rangle_\Psi - \langle\hat{T}_{00}\rangle_0$. For a pure state $\Delta\langle H_{\bar{A}-A}\rangle = S(\rho_{\bar{A}}|\rho^0_{\bar{A}}) - S(\rho_A|\rho^0_A)$. In this case one can not associate any entropy to the bulk excitation in the form of entanglement. Let us evaluate (4.5.6) for homogenous excitation where energy density of the state is given as $\epsilon = \frac{\Delta E}{r^{d-1}\Omega_{d-1}}$, in this case bulk dual of the state becomes spherically symmetric.

$$\begin{aligned}\Delta\langle H_{\bar{A}-A}\rangle &= 2\pi r \Delta E\, \frac{\mathrm{Vol}(S^{d-2})}{\mathrm{Vol}(S^{d-1})} \int_0^\pi \sin^{d-2}\theta \left(\frac{\cos\theta - \cos\alpha}{\sin\alpha}\right) \\ &= 2\pi r \cot\alpha\, \Delta E_{\mathrm{CFT}}.\end{aligned} \tag{4.5.7}$$

The expression is the higher dimensional non-pertubative generalization of the one that is obtained in (4.4.7) for conical defects. The expression is remarkable as it is valid in any dimension yet it is more interesting when one understand it in terms of the bulk quantities. Let us elaborate the identification with the bulk in more detail. We choose the cylindrical description $\mathbb{R}\times\mathbb{S}^{d-1}$ of AdS_{d+1}. The metric of the AdS_{d+1} is given in (4.4.11). Although one can fit the geometry on a finite piece of paper, the actual radius of the boundary sphere becomes infinite. However this is an overall conformal factor that can be removed such that volume of the boundary sphere becomes finite. Hence flow of time in the bulk and boundary descriptions are different as it is measured by the lapse function N $(=\sqrt{g_{tt}})$ in the ADM description. In the global coordinates $N \to R/L$ asymptotically.[4] The energy of the state in the CFT is proportional to the mass of the dual gravitational solution,

$$\Delta M_{\text{ADM}} = \frac{r}{L}\Delta E_{\text{CFT}} \tag{4.5.8}$$

where r is the curvature radius of $\mathbb{S}^{d-1}$ while L is the curvature of the AdS. Equality above also ensures a dimensionless identification in the information theoretic observables. If one identifies these two scales then energies of the theories are naturally identified. This identification is necessary to recognize the radial scale of the deepest point that can be probed by the state of A ($\text{Vol}(\bar{A}) \geq \text{Vol}(A)$) in CFT. Remember that the deepest point in the bulk that can be reached via the boundary region A was given in (4.4.15). Inserting these we see that the change in the full modular hamiltonian in the complementary regions becomes,

$$\Delta\langle H_{\bar{A}-A}\rangle = 2\pi R_{scale}\Delta M_{\text{ADM}} \tag{4.5.9}$$

This is entirely a bulk expression due to the natural identification between modular energies of the CFT and the gravitational dual. Once again, just like the perturbative case (4.4.13), we have observed that for spherically symmetric excitation, the difference in the entanglement energies of the complementary states of the underlying theory have an expression in terms of bulk quantities.

It has been emphasized along the chapter that when excitation is pure state, the change of entanglement for complementary regions are equal $\Delta S_{\bar{A}-A} = 0$. In this case one can not associate extra new correlation in the form of entanglement to the bulk excitation for scales less then the radial probing point of the boundary observes having access to A. In the language of bit threads [28] no additional thread (additional in the sense that comes by the excitation on top of the vacuum tread configuration) ends up in the bulk. Purity of the state constrains the amount of information that is missing beyond the scale R. On the other hand for a thermal state $\Delta S_{\bar{A}-A} \neq 0$. Let us focus to spherically symmetric thermal excitations again. For non-perturbative excitations we do not have the localized expression of the bulk stress tensor anywhere in the bulk which was possible for states that are dual to geometries that can be expressed as perturbations of the metric around the AdS. Positivity of relative entropy dictates that,

$$S(\rho_A||\rho_{0A}) \geq 0, \qquad \Longrightarrow \qquad \Delta\langle H_A\rangle - \Delta S_A \geq 0 \tag{4.5.10}$$

[4]The identification outlined above appears in a more rigorous way in the duality between $\text{AdS}_5 \times \mathbb{S}^5$ and $SU(N)$ Yang-Mills. The relation $L^3/G_{\text{N}} = 2N^2/\pi$ makes it possible to relate the mass to the dimension of the gauge group N. Then $rE_{\text{casimir}} = LM_{\text{casimir}} = 3(N^2-1)/16$.

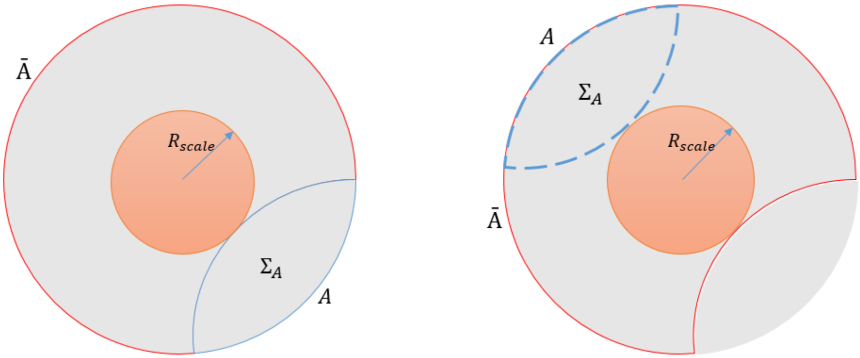

Figure 4.3: *The difference between change of entanglement entropies of the complementary regions in CFT$_d$ for a thermal state with respect to vacuum is bounded by the maximum entropy that can be contained in the region excluded by the smaller interval. Spherical symmetry allows one to translate the region A, which is pictured on the right. Monotonicity of relative entropy puts a bound on the difference of entanglement entropies, $\Delta S_{\bar{A}-A} \leq 2\pi R_{scale} \Delta M_{\rm ADM}$, which is the consequence of Bekenstein bound in the geometric description.*

Since, excitation is not perturbation around the vacuum, one can not equate the right hand side of the first equation to the modular integral of bulk stress energy. Possibly it is not even in the form of local expression. However assuming spherically symmetry for excited state one can understand the implications for the bulk physics in the non perturbative level. As a consequence of spherical symmetry one can translate the region A such that $\bar{A} \supseteq A$ without altering ΔS_A or $\Delta\langle \hat{H}_A\rangle$. Following rotations, Further impose the monotonicity of the relative entropy,

$$S(\rho_{\bar{A}}||\rho_{0\bar{A}}) \geq S(\rho_A||\rho_{0A}) \qquad \Longrightarrow \qquad \Delta S_{\bar{A}-A} \leq \Delta\langle H_{\bar{A}}\rangle - \Delta\langle H_A\rangle. \tag{4.5.11}$$

We have already calculated right hand side using the boundary expressions of the modular hamiltonian.

$$|\Delta S_{\bar{A}-A}| \leq 2\pi R_{\rm scale} \Delta M_{\rm ADM}\,. \tag{4.5.12}$$

The inequality is universal in the sense that it is independent of the details of the excitation, and how it is organized spatially in the gravitational theory apart from its spherical symmetry. Remember that we have encountered necessity of spherical symmetry in the derivation of the Bekenstein bound using positivity of relative entropy (2.5.3) also in the QFT [72]. One can deviate from spherical symmetry by considering $\mathcal{O}(1)$ deformations of the bound. Using the symmetry we argue that the entropy contained in the sphere of ignorance is bounded by the difference of vacuum subtracted entropies of the complementary regions on the boundary.

$$\Delta S^{\text{bulk}}_{\mathbb{S}_R} \leq \Delta S_{\bar{A}-A} \tag{4.5.13}$$

Remember that $\Delta S_{\bar{A}-A} \equiv \Delta S_{\bar{A}} - \Delta S_A$. The bound becomes and equality in the perturbative limit as shown in eq. (4.5.4). The symmetry of the system actually reduces the effective dimensions to one and allows us to represent entanglement entropy using the two scales of the system, namely energy and size of the box. In this one effective spatial dimensional information space, the difference between entanglement entopies of the complementary regions in the microscopic system, is bounded by the Bekenstein bound for a system with radius R and energy M in the bulk, which is indeed the region that is excluded by any observation on scale(A). If all the energy was contained in the radius $r_{\text{bulk}} < R_{\text{scale}}$ and organized to in a way to saturate the Bekenstein bound, then $|\Delta S_{\bar{A}-A}| = 2\pi R_{\text{scale}} \Delta M_{\text{ADM}}$. The inequality also holds perturbatively, at the linear order as we have shown. In the transition from a pure state to a thermal state, $|\delta S_{\bar{A}} - \delta S_A|$ is interpolates between 0 and $2\pi R_{\text{scale}} \Delta M_{\text{ADM}}$. Whenever some of the thermal energy is replaced by an equal amount of energy corresponding to a pure state, the difference between entanglement entropies decreases.

On the other hand, we should be careful using $\hat{H}_A$ on thermal states when the bulk dual is a black hole solution. In this case one observes a phase transitions [5] in the entanglement entropy along the continuous increment of the system size. These phase transitions are formulated as homology constraints for the minimal surfaces in the bulk [135]. Therefore the local expression of $\hat{H}_A$ is not valid for regions bigger than the critical size θ_{critical} beyond which phase transitions take place as formulated in homology constraints. The point where phase transition took place manifest itself as a sudden jump on the minimal surface. This is also the point where Araki-Lieb bound is saturated. In the next section we will study the relation between Araki-Lieb bound and the one we have derived via monotonicity of relative entropy.

4.5.3 Comparison with Araki-Lieb bound

The bound we have derived in (4.5.12) using the monotonicity and positivity of relative entropy for certain class of excitations has the same quantity with Araki-Lieb bound on the left hand side of the inequality. It is an interesting exercise to

[5] Phase transitions on the modular hamiltonian takes place also for the disjoint intervals depending on the distance of separation. Another example is the entanglement entropy in the conical defect geometry when it is seen as the entanglement entropy of the disjoint intervals in the parent theory. In both of these cases, the phase transition on entanglement entropy is due to a jump in the saddle point and mutual information between disjoint intervals is a probe of different phases.

study these two inequalities together and see whether the bound derived here is trivial when it is compared to Araki-Lieb bound.

The notion of entanglement entropy we had been referring along the chapter is von Neumann entropy, which quantifies the extent to which the state represented by ρ fails to be a pure state. The reason that von Neumann entropy serves as entanglement entropy is that when the state ρ is obtained from a pure state by tracing over part of the Hilbert space representing a subsystem, such as the one that can not be accessed by the observer, then von Neumann entropy measures the entanglement entropy between subsystem that is traced out and the rest. Suppose the Hilbert space of the full system $\mathcal{H}_{full}$ factorizes into Hilbert space of two subsystems, $\mathcal{H}_{full} = \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$. Then for each subsystem we have corresponding density matrices defined by tracing over the complementary subsystem $\rho_{A,\bar{A}} = \mathrm{Tr}_{\bar{A},A}(\rho_{\mathrm{full}})$. The entanglement entropies that are associated to each density matrix can be shown to satisfy following inequalities [136],

$$|S(\rho_A) - S(\rho_{\bar{A}})| \leq S(\rho_{\mathrm{full}}) \leq S(\rho_A) + S(\rho_{\bar{A}}). \tag{4.5.14}$$

The first part of the triangle inequality is usually referred as Araki-Lieb bound, while the second is known as subadditivity. The Araki-Lieb bound is derived from subadditivity. We have also derived an inequality that is similar to the first part of the triangle inequality. We have observed that the difference of entanglement entropies of the complementary subsystems follows the Bekenstein bound given in terms of the bulk quantities. Our bound becomes non trivial compared to Araki-Lieb bound when $2\pi MR \leq S(\rho_{\mathrm{full}})$. This happens when system sizes on the CFT approach each other. In this limit, Bekenstein bound takes over the Araki-Lieb. Let us compare these two bounds by considering d dimensional CFT at finite temperature and having geometric dual as AdS-Schwarzschild black hole.

The metric for $(d+1)$-dimensional static solution for asymptotically AdS spacetimes is given by,

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega_{d-1}^2, \qquad f(r) = 1 + \frac{r^2}{L^2} - \frac{\mu}{r^{d-2}} \tag{4.5.15}$$

where $\mu = \frac{16\pi G_{\mathrm{N}} M}{\Omega_{d-1}(d-1)}$. On this state Araki-Lieb bound can be expressed in terms of the black hole entropy.

$$|\Delta S_{A-\bar{A}}|(\alpha) \leq \frac{A(r_+)}{4G_{\mathrm{N}}} = \frac{r_+^{d-1}\Omega_{d-1}}{4G_{\mathrm{N}}} \tag{4.5.16}$$

where $\Delta S_{A-\bar{A}} = |\Delta S(\rho_A) - \Delta S(\rho_{\bar{A}})|$ and r_+ is the largest solution to the equation,

$$1 + \frac{r^2}{L^2} - \frac{\mu}{r^{d-2}} = 0. \tag{4.5.17}$$

We considered vacuum subtracted quantities on the right hand side. Since the vacuum entanglement entropies of the complementary states are equal, this subtraction does not change the difference. Since the solution is spherically symmetric we can use it to set Bekenstein bound. Let us calculate the mass of the solution in terms of the r_+. The calculation had been carried first by Hawking-Page [120] in $d = 3$ and later generalized to arbitrary dimensions by Witten [137]. The key point is the connection between action and the partition function $I = -\log Z$. Energy of the excitation can be calculated by change of the action with respect to inverse temperature $E = \partial_\beta I$. The action is calculated on shell, since on-shell configuration is the dominant contribution in path integral. On shell integral of the action for the regions outside of the horizon amounts to the volume of the spacetime.

$$I_{\text{on-shell}} = \frac{d}{8\pi G_{\text{N}}} \int \sqrt{g}\; d^{d+1}x \tag{4.5.18}$$

To calculate vacuum subtracted energy, one should calculate this integral for the AdS-Schwarzschild for region $r_+ \leq r \leq r_\infty$ and subtract the vacuum contribution by considering the same integral on AdS_{d+1} for $0 \leq r \leq r_\infty$.

$$I = \lim_{r_\infty \to \infty} \frac{d}{8\pi G_{\text{N}}} \left(\text{Vol}_{BH}(r_\infty) - \text{Vol}_{AdS}(r_\infty)\right) \tag{4.5.19}$$

Explicit calculation yields,

$$M = \frac{\partial I}{\partial \beta} = \frac{(d-1)\Omega_{d-1}}{16\pi G_{\text{N}}} \left(\frac{r_+^d}{L^2} + r_+^{d-2} \right) \tag{4.5.20}$$

which correctly reproduces the vacuum subtracted energy of the BTZ black hole $d = 2$ ($M_{\text{AdS}_3} = -1/8G_{\text{N}}$). Let us now use this expression to compare the two bounds and find the limit where Araki-Lieb takes over the modular energy bound.

$$\frac{d-1}{2} \left(\frac{r_+}{L} + \frac{L}{r_+} \right) \cot\alpha \leq 1 \tag{4.5.21}$$

Therefore when the condition above holds, Bekenstein bound sets a lower bound then the Araki-Lieb. For large black holes i.e. $r_+/L \gg 1$, the bound puts a more restrictive condition then Araki-Lieb when $\cot\alpha \leq \frac{2}{d-1}\frac{L}{r_+}$. In this case sphere of ignorance stays inside the black hole. Hence the entropy can not only be associated to the sphere $\mathbb{S}^{d-1}_{R_{\text{scale}}}$ as it is not confined in this region. On the other hand, when $r_+/L < 1$ thermal AdS solution dominates the canonical ensemble.

The metric of Euclidean thermal AdS solution is identical to empty Euclidean AdS apart from periodicity of time direction $t_E \sim t_E + \beta$. The difference with Euclidean Schwarzschild is that time circle in this solution does not cap off around the origin, in other words while space (r, t_E) is topologically a disc for BH solution, it is topologically equivalent to a cylinder in thermal AdS. Hence β is not fixed by any regularity condition and becomes a free parameter in this solution. The solution is represented as empty AdS therefore holographic entanglement entropy at the leading order is identical to vacuum entanglement entropy. One can calculate the entropy of the solution via the on-shell action calculated on $\mathbb{R}^{d+1} \times \mathbb{S}^1_\beta$, $S = (1 - \beta\partial_\beta)I = 0$. Thermal entropy on top of the vacuum contribution comes at the order $\mathcal{O}(G_{\rm N}^0)$. At the critical temperature where phase transition takes place it suddenly jumps to $\mathcal{O}(1/G_{\rm N})$. Below this phase transition, the thermal entropy can be fully confined inside the $\mathbb{S}^{d-1}_{R_{\rm scale}}$ as it has been studied for thermal states at the linear level. Boundary computation provides all orders to the perturbative excitation. Hence a valid interpretation of the Bekenstein bound on the entropy attributed to sphere of ignorance takes place below the Hawking-Page transition. This is in agreement with general understanding on the fact that Bekenstein bound is applicable to system having weak self gravitation.

To sum up when condition (4.5.21) is satisfied the Bekenstein bound proposed in this paper sets a lower bound then the Araki-Lieb bound. One can satisfy this condition in both sides of the critical temperature. When the temperature is above the critical temperature, black hole solutions are dominant in the phase space. In this case, the space of parameters that satisfy (4.5.21), have R_{scale} that falls into the black hole, $R_{scale} < L < r_+$. On the other regime, below the Hawking-Phase transition, the parameter space satisfy the condition when $r_+ \leq R_{scale}$. In this case one can confine all the excitation inside the bulk sphere $\mathbb{S}_R$. We think thermal AdS regime is more natural for Bekenstein bound interpretation in the bulk since one can push all thermal gas inside the sphere $\mathbb{S}_R$. Our understanding also agrees with the general idea that bound is valid for weakly self gravitating systems.

4.6 Conclusion and Discussion

In this chapter we have identified the *full modular Hamiltonian* from the bulk point of view. We have studied analogous quantity in the entanglement entropy and shown that it has distinct character for pure and mixed state excitation. In section 2 we show that purity of the state puts strict constraints on the allowed expectation value of the boundary stress energy tensor on the excited state. The bulk interpretation of the full modular Hamiltonian for certain class of excitation have a remarkably simple expression independent of the dimension of the space-time. The connection between the bulk expression of the full modular hamiltonian $(2\pi MR)$ and the change of area in a certain identification of the manifolds had been used to modify gravity at long distance scales [73]. In that proposal, this expression was the key component in the gravitational side where underlying state follows the area law. Here we have identified the expression as the full modular hamiltonian in the underlying theory. The main conclusions of this chapter can be listed in the following way.

- **Bekenstein bound in the bulk:** Using positivity together with the monotonicity of relative entropy, we have shown that the change of entanglement entropy for complementary states in spherically symmetric excitations are bounded by $2\pi M R_{\text{scale}}$. The expression is valid perturbatively as well as non perturbatively. In the perturbative regime, full modular hamiltonian can be expressed in the bulk as the integral of the local bulk stress energy tensor using covariant phase space approach. In this case it is clear that bound is saturated if all the excitation is hidden behind the sphere of ignorance defined with respect to the boundary region A. We have proposed that difference of the change of the entanglement entropies of complementary regions $(S_{\bar{A}} - S_A)$ in the boundary theory sets a bound for the entanglement entropy resides in the sphere of ignorance. This entanglement should be seen as the entanglement with respect to the purifying state, which would be zero for pure state excitation, which trivially satisfy the Bekenstein bound proposed here. In conclusion, the Bekenstein bound in the bulk manifest itself as the positivity together with the monotonicity of the relative entropy in the boundary CFT.

- **An example of UV-IR correspondence:** Bulk interpretation of the full modular Hamiltonian reflects the well known UV-IR correspondence [138]. This should be understood as follows; consider the change of full modular Hamiltonian, $\Delta\langle H^{\text{full}}(\alpha)\rangle$ in the boundary CFT as a function of the size of the boundary interval A, where $\alpha \in [0, \pi/2]$. In the boundary theory $\alpha \to 0$ limit identifies the short distance behaviour. On the other hand, as we have seen in the bulk the quantity amounts to $2\pi M R_{\text{scale}}$, where R_{scale} denotes the deepest point the bulk that can be probed from the boundary state A. The limit $\alpha \to 0$ corresponds to $R_{\text{scale}} \to \infty$ from the bulk point of view which yields one realization of the UV-IR correspondence in the AdS/CFT.

- **Black hole vs thermal gas limits:** Comparing with the Araki-Lieb bound we have seen that, Bekenstein bound sets a lower bound when the complementary regions are close to each other. In the case of large black holes, when the Bekenstein limit sets the lower bound with respect to Araki-Lieb, the sphere of ignorance to which we have associated the entropy $(S_{\bar{A}} - S_A)$ corresponds to the regions inside the black hole. In that case holographic bound is already satisfied due to the formation of the black hole. Hence holographic bound sets even a lower bound than the Bekenstein one in these cases. On the other hand, in the thermal gas limit, ($r_+/L < 1$), below the Hawking-Page transition, one can come up with a window, where Bekenstein bound becomes non-trivial with respect to Araki-Lieb and yet holographic bound is not saturated. This corresponds to limit where self gravitation of the excitation is weak, hence in agreement with the expectations that bound can be derived within the context of QFTs. On the other hand our derivation includes the backreaction on the geometry. We have also shown that in the weak field limit the full modular hamiltonian have well defined bulk expression which further justifies the proposals made in this chapter.

- **Boundary to bulk map and proof of the proposal:** We find it useful to emphasize that we have not provided the full proof of the derivation of the Bekenstein bound in the bulk via AdS/CFT. The relation between entropy associated to the bulk spheres $\Delta S^{\text{bulk}}_{\mathbb{S}_R}$ and boundary entropy difference $\leq \Delta S_{\bar{A}-A}$ is conjectured for a spherically symmetric state (4.5.13). We have proved this conjecture in the perturbative limit. Under such an assumption we show that Bekenstein bound in the bulk manifest itself as the information inequalities (positivity + monotonicity) in the underlying theory. It would be remarkable to find the exact map between entropy of the bulk regions and boundary regions to drop the assumption of spherical symmetry. That would also let us test the volume law conjecture in spacetime.

5

Pushing the first law of entanglement into the bulk in AdS_3/CFT_2

5.1 Introduction

The first observation on the connection between statistical physics of quantum nature and geometry of spacetime goes back to the Bekenstein-Hawking formula [14–16, 20, 139],

$$S_{\mathrm{BH}} = \frac{c^3 A_{\mathcal{H}}}{4 G_N \hbar}. \tag{5.1.1}$$

This observation opened a door to microscopic nature of gravity. Shortly after, it is understood that, black holes does not only possess entropy, they follow all the laws of thermodynamics [22]. If black holes are of thermodynamical nature, then natural questions would arise: what are the molecules of black holes, how could one count these, and where are they? String theory provided a partial answer to these question for some particular black holes [140–142]. However it is still not known for a generic black hole (such as Schwarzschild), where the information associated with black hole entropy resides. How does it encoded into the geometry? Is it located on the horizon or in the interior of the black hole and maybe it is spread non-locally and shared between the interior and exterior of the black hole [25, 27]. The problem of where the information resides and how it is returned back manifests itself by the firewall paradox [29].

While the microscopics of black hole entropy is still puzzling in many ways, together with the establishment of holographic principle [143, 144] via the duality of AdS/CFT [51], more doors opened regarding the information theoretic nature of gravity. AdS/CFT teaches that the relation between entropy and the area of spacetime is much more broad than event horizons [57]. Ryu-Takayanagi proposal states that the entanglement entropy of a region A in the conformal field theory corresponds to the area of the bulk surface S_A that is homologous to the boundary region A.

$$S(A) = \underset{S_A \sim A}{\text{ext}} \left(\frac{A_{S_A}}{4G_\text{N}} \right) \tag{5.1.2}$$

where $S_A \sim A$ indicates the homology relation between the surface S_A and A. This is the first indication that Bekenstein-Hawking type formulas apply to more general surfaces than black hole horizons. The formula is giving us strong hints on the emergent nature of gravity and its connection to information theory. Similar to the first law of black hole mechanics, holographic entanglement entropy follows a first law as a consequence of the positivity of relative entropy in the underlying theory. The first law of entanglement entropy in CFT reflects itself as the equality of the infinitesimal changes on the area of minimal surfaces and associated modular energy in the bulk. This is a strong manifestation of the information theoretic nature of the gravity.

The RT proposal is limited to the minimal (extremal) surfaces that extends to the boundary. Therefore it is still not entirely clear if there are information theoretic observables in the microscopic theory that corresponds to areas of arbitrary surfaces (non minimal) in the bulk and if there exists, how these quantities are embedded into the field theory. This is part of the localization problem in AdS/CFT.

The spacetime entanglement conjecture build on top of the Ryu-Takayanagi relation [60, 145] states that, in a theory of quantum gravity, any state describing a smooth spacetime geometry reflects the following property: for any sufficiently large region, there is a (finite) gravitational entropy which is characteristic of the entanglement between the degrees of freedom describing the given region and those describing its complement. Furthermore, the leading contribution to this entropy is given by the Bekenstein-Hawking formula (5.1.1) evaluated on the boundary of the region. That means, entanglement between degrees of freedom of the underlying theory reflects itself more naturally in the gravitational theory and only special cases of this entanglement, namely entanglement measured by minimal surfaces, is encoded spatially in the CFT description of the underlying theory. In this case one would also expect to observe a first law of entanglement on general surfaces.

How can we extend the first law of entanglement entropy that holds for regions on the boundary theory, to a general spacelike surface in the bulk? In three dimensional spherically symmetric asymptotically AdS geometries, the correspondence between areas of bulk surfaces and entanglement entropy of the subsystems in the CFT is made through differential entropy [68, 88] [1]. Using the notion of differen-

[1] Other examples where area of a bulk surface measures the amount entanglement are tensor

tial entropy, one can extent the first law of entanglement entropy for individual boundary regions, to a collection of intervals which in total yields the change in the differential entropy. In other words, the first law of Rindler horizons are extended to arbitrary bulk surfaces through the notion of differential entropy. The first law of entanglement in the bulk is also studied on a local neighborhood of the spacetime manifold [74]. We will not constrain ourselves to infinitesimal balls in the bulk.

We observe that, the first law of differential entropy is equivalent to the gravitostatics effects of the excitations on the surfaces according to Einstein equations in the bulk. The boundary first law also provides an entropic origin on the emergence of *entanglement shadows*. We have studied the first law around three class of solutions: AdS, conical defect solutions and planar BTZ.

Comparing area of an arbitrary region when spacetime is deformed always requires a scheme which involves fixing a parameter of the region defining the system. In [73] the geodesic distance in the weak field limit is fixed, while in [74] volume of the infinitesimal ball is kept fixed and in both cases an area deficit is observed due to the energy introduced into the system. We identify the regions of unperturbed vs perturbed manifolds, by keeping the number of degrees of freedom on the CFT is fixed, $A(S^1)/\epsilon$, where $A(S^1)$ is the size of the circle CFT lives and ϵ is the cutoff of the theory. In this case we always observe an area excess rather than deficit. However we should say that our result does not conflict with [73, 74]. Since we recognize the deficit terms in our expressions. It is rather a matter of choice what to keep fix, yet CFT provides a natural reference on what parameter should be fixed. Interestingly in the presence of conical defect we observe an amplification in the size of the area defect and when the state is thermal the term that causes area deficit even becomes an excess.

$$\delta S_{DE} = \delta \langle H_{DE} \rangle = -\Delta \frac{A_{\text{hole}}}{4G_{\text{N}}} + \dots \qquad \text{AdS}_3 \tag{5.1.3}$$

$$\delta S_{DE} = \delta \langle H_{DE} \rangle = -(n\Delta) \frac{A_{\text{hole}}}{4G_{\text{N}}} + \dots \qquad \text{AdS}_3/\mathbb{Z}_n \tag{5.1.4}$$

$$\delta S_{DE} = \delta \langle H_{DE} \rangle = \frac{\delta\beta}{\beta} \frac{A_{\text{hole}}}{4G_{\text{N}}} + \dots \qquad \text{Planar BTZ} \tag{5.1.5}$$

there is an additional term we omit to emphasize the structure of area deficit in different states. Δ is deficit introduced inside the hole or $2\pi\Delta$ measures the angle deficit. In the thermal case deficit is identified with variation of the inverse temperature. The result is independent of where the deficit is introduced inside the hole. In the presence of already existing defect the area deficit amplifies, and

network constructions [85, 86]. In holographic tensor network models, area of an arbitrary bulk surface is a measure for how much entanglement exists between both sides of the surface.

finally in a thermal system deficit becomes excess. The result above is presented for a spherical hole.

This paper organized as follows: in section 2.3 we review the concept of differential entropy and provide a novel derivation in the form of integral equations. Solution of the integral equation yields the measure in the kinematic space as a function of the entanglement entropy. We explain the method in general dimension and provide the solution in 3 dimensions. Spherical symmetry is sufficient for the derivation and hence it applies to excited states. In section 5.3 we construct the first law of differential entropy and show that differential energy measures the change of area in terms of the energy injected into the hole. We observed that for spherical holes, change of area is topological in the sense that it is independent of where the excitation is introduced inside the hole. In section 5.4, we derived a modular Hamiltonian for conical defect solutions at leading order in $1/N$. We provide consistency checks for the result and use the modular hamiltonian to study first law of differential entropy. Finally in 5.A we study thermal solutions in the same manner.

5.2 Differential entropy via integral geometry

In the previous section we have given the definition of differential entropy and show that it reproduces the area of the spherical surfaces. Here we will further prove that differential entropy always yields the area of the bulk surface whose tangent geodesics are the entanglement entropy of the intervals that goes into the definition of the discretized differential entropy. The connection between differential entropy and integral geometry is first observed in [146]. The derivation we will provide in the next section is novel and applies to any spherically symmetric excited state.

Integral geometry is basically the theory of measures on a geometrical space that is invariant under the symmetries of that space. The field initially emerged as an attempt to elucidate certain statements about geometric probability theory [147]. Recently integral geometry find its use in AdS/CFT as a formalism which connects geometry with information theory [146]. It has been observed that the notion of differential entropy actually corresponds to a classic result in integral geometry called Crofton formula [148] which relates the length of a curve to the number of times a random line intersect it.

Crofton formulas exist in higher dimensions both for flat and constant curvature spaces. In general, volume of a compact $m-$dimensional submanifold S embedded in $d-$dimensional constant curvature space (including flat geometry), can be expressed by an integral over the space of $r-$dimensional planes [149]

$$\mathrm{vol}(S) = \frac{O_r ... O_0 O_m}{O_n ... O_{n-r} O_{m+r-n}} \int dL_r \, \mathrm{vol}_{m+r-d}(S \cap L_r) \tag{5.2.1}$$

where dL_r is the volume form in the space of $r-$planes. L_r denotes the $r-$dimensional planes and O_i is the volume of $i-$dimensional unit sphere $\mathbb{S}^i$. Note that the integral is taken over the domain that intersect S in $(m + r - d)-$ dimensions, since other configurations has zero measure.

The connection with holography is a special case of (5.2.1). The formula is exploited for $r = 1$ and $m = 1$ in hyperbolic space $\mathbb{H}^2$. We will provide a method to obtain differential entropy which can be applied to derive differential entropy in higher dimensions, independent of the geometry of surface S. We will use the Crofton formula to express the areas of minimal surfaces. Hence left hand side of the equation (5.2.1) will be the entanglement entropy. For empty AdS the integral equation is given by,

$$S_{\mathrm{ent}}(\alpha_0) = \mathcal{N} \int K(\alpha, \alpha_0) \, f(\alpha) \, d\alpha \tag{5.2.2}$$

where α is the separation of the endpoints of the boundary anchored geodesic, α_0 is the angular radius of the boundary ball homologous to the minimal surface. $\mathcal{N}$

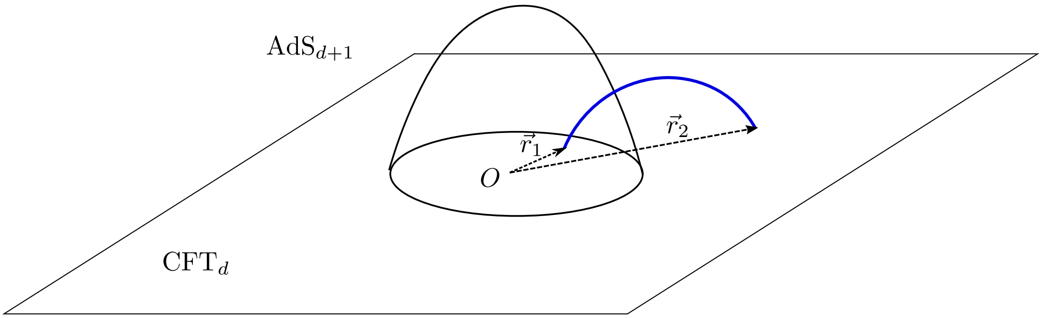

Figure 5.1: *Illustration of the parametrization for the geodesics that intersect a boundary ball of size R.*

is the normalization constant given in (5.2.1) and $f(\alpha)$ is the density of geodesics. The density $f(\alpha)$ will not depend on how geodesic is configured rotationally due to symmetry of the problem. As we will show, in higher dimensions Kernel is an integral of rotational degrees of freedom, Ω. Since overall space is $2(d-1)$ dimensional, Ω is $2d-3$ dimensional rotations. The integral over Ω is absorbed into the Kernel as measure jas an expression of the form $dL_1 = f(\alpha)d\alpha \wedge d\Omega$ due to spherical symmetry. The solution of $f(\alpha)$ in terms of $S_{\text{ent}}(\alpha_0)$ will be the generalization of differential entropy to higher dimensions. Since the derivation only depends on the spherical symmetry, it applies to any excitation of AdS with such symmetry. We will leave the solution of the integral in general dimensions to a future study and will provide the solution for $d=2$. But first let's derive the form of the Kernel in higher dimensions.

Let us apply the method suggested above and derive a kernel for the integral equation whose inverse yields an expressions of the density of geodesics in terms of the entanglement entropy of a totally geodesic surface.

Consider AdS$_{d+1}$ whose constant time slice is d dimensional hyperbolic space $\mathbb{H}^d$. Space of geodesics of $\mathbb{H}^d$ is $2(d-1)$ dimensional, since each two points on the boundary determines a geodesic. On the space of geodesics one can define a volume form which is denoted as dL_1, indicating that it is a measure for the space of 1-dimensional surfaces. We can only restrict ourselves to the coordinates of the boundary, and give the expression for the measure using spherical coordinates on the euclidean space. Let us express the measure dL_1 in terms of the endpoints of the geodesic $(\vec{r_1}, \vec{r_2})$ with respect to an arbitrary origin in euclidean space figure: 5.1. Density of geodesics does not depend on how the line is rotationally configured due to spherical symmetry. Therefore non-trivial dependence becomes only through the distance of endpoints.

$$dL_1 = f(|\vec{r_1} - \vec{r_2}|)\, d\vec{r_1}\, d\vec{r_2} \tag{5.2.3}$$

where $d\vec{r} = r^{d-2} dr d\Omega_{d-2}$ and $f(|\vec{r_1} - \vec{r_2}|)$ denotes the density of geodesics. Combining Crofton and Ryu-Takayanagi formula, we can express the entropy of a boundary ball in the following way.

$$S(R) = \int d\vec{r}_1 \int d\vec{r}_2\, f(|\vec{r_1} - \vec{r_2}|)\, (\Theta(R - |r_1|)\ \Theta(|r_2| - R)) \tag{5.2.4}$$

The equation indicates that geodesics intersecting the boundary ball has one endpoint inside the ball and the one outside. The center of the boundary ball is picked as the origin coordinate system. Going to the center of mass coordinates by defining,

$$\vec{r}_m = \vec{r}_1 + \vec{r}_2 \qquad\qquad \vec{r}_s = \vec{r}_1 - \vec{r}_2 \tag{5.2.5}$$

In these coordinates integral equation becomes,

$$S(R) = \int_0^\infty f(r_s) K(r_s, R)\, dr_s \tag{5.2.6}$$

and Kernel is given by

$$\begin{aligned} K(r_s, R) = \Omega_{d-2}\Omega_{d-3}\, r_s^{d-2} \int_0^\infty dr_m\, r_m^{d-2} \int_0^\pi d\theta\, (\sin\theta)^{d-3} \\ \Big(\Theta(R - (r_m^2 + r_s^2 + 2r_m r_s \cos\theta)^{1/2})\, \Theta((r_m^2 + r_s^2 - 2r_m r_s \cos\theta)^{1/2} - R)\Big) \end{aligned} \tag{5.2.7}$$

Solution of the integral equation (5.2.6) is given by the operator that inverts the kernel (5.2.7). Hence one finds an expression of the density of geodesics in terms of entanglement entropy of a minimal surface anchored to a boundary ball of size R. The power of the method sketched here is that it applies also to spherically symmetric excited states.

In the next section we will apply the method described here to AdS_{2+1} and derive well known expression for differential entropy.

5.2.1 Differential entropy for excited states

In this section we will give the derivation of differential entropy without referring to AdS_3. The derivation applies to all excited states with spherical symmetry. The correspondence between gravitational entropy and differential entropy will be naturally reflected via Crofton formula. The only assumption that goes into our derivation is existence of a well defined measure of the space of geodesics of the spacetime dual to an excited state.

Consider a $2d$ manifold $(\mathcal{M}, g)$ with boundary $\partial\mathcal{M}$. The manifold corresponds to constant time slice of spherically symmetric $3d$ spacetime with negative cosmological constant. For example in the case of AdS_3, $\mathcal{M}$ becomes $\mathbb{H}^2$. Let us extend every geodesic in $\mathcal{M}$ to infinity and denote each geodesic with two parameters depending on where they end up in $\partial\mathcal{M}$. One convenient choice is center θ and the angular separation between the endpoints on the boundary α. Let us denote the measure of the space of geodesics in $\mathcal{M}$ as dL_1. Because of the spherical symmetry the density of geodesics depends only on the angular width α.

$$dL_1 = f(\alpha)\, d\alpha \wedge d\theta \tag{5.2.8}$$

where $f(\alpha)$ is the density of geodesics. The Crofton formula can be used to compute the area of any surface. Let us use it to express the area of a minimal surface (geodesic)

$$S(\alpha_0) = \int_{\mathcal{K}} n(\theta_0, \alpha_0)\, f(\alpha)\, d\alpha \wedge d\theta. \tag{5.2.9}$$

$n(\theta_0, \alpha_0)$ determines the domain of integration where each point corresponds to the geodesics that intersects the minimal surface. For spherically symmetric spaces determining the domain is simple. The geodesics that intersects the surface has one end point in the region $[\theta_0 - \alpha_0, \theta_0 + \alpha_0]$ and other one is outside. Let's express the integral by dividing the domain into two regions $\alpha > \alpha_0$ and $\alpha \leq \alpha_0$,

$$S(\alpha_0) = \int_0^{\alpha_0} \int_{I_1(\theta)} f(\alpha)\, d\alpha \wedge d\theta + \int_{\alpha_0}^{\pi/2} \int_{I_2(\theta)} f(\alpha)\, d\alpha \wedge d\theta \tag{5.2.10}$$

these regions include all the geodesics that intersect the surface from right. Therefore total set includes a factor of 2.

$$I_1(\theta) = \{\theta \in [\theta_0 + \alpha_0 - \alpha, \theta_0 + \alpha_0 + \alpha]\}$$
$$I_2(\theta) = \{\theta \in [\theta_0 - \alpha_0 + \alpha, \theta_0 + \alpha_0 + \alpha]\}$$

One can take the θ integral as a consequence of the spherical symmetry and express the integration in the following way,

$$S(\alpha_0) = \int_0^{\pi/2} K(\alpha, \alpha_0) f(\alpha) d\alpha \tag{5.2.11}$$

where the kernel is $K(\alpha, \alpha_0) = \alpha\Theta(\alpha_0 - \alpha) + \alpha_0\Theta(\alpha - \alpha_0)$. $\Theta(.)$ is Heaviside step function. Once the inverse of kernel is known one can express the density

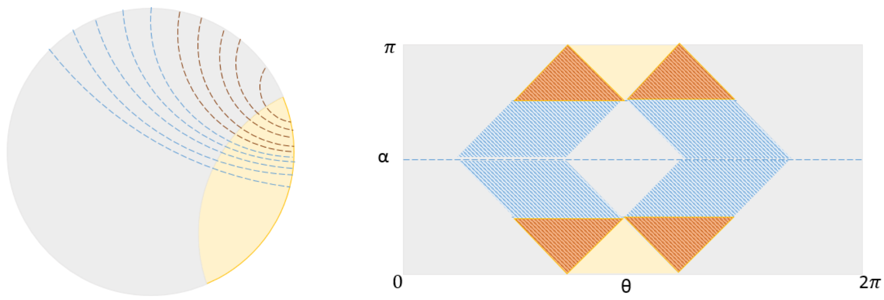

Figure 5.2: *Two classes of minimal curves according to their boundary width. On the right, the domain of these sets in the kinematic space are shown. The domain determines the kernel of the integration, whose evaluation yields the length of a curve via integral geometry. Inverse of the kernel gives the integral geometric proof of differential entropy.*

of geodesics in terms of the entanglement entropy. The operator that inverts the kernel is $-\frac{d^2}{d\alpha_0^2}K(\alpha, \alpha_0) = 4\delta(\alpha - \alpha_0)$. Hence,

$$dL_1 = -\frac{1}{4}\frac{d^2 S(\alpha)}{d\alpha}\, d\alpha \wedge d\theta \tag{5.2.12}$$

We extend the differential entropy without referring to AdS_3 so the correspondence between areas of surfaces and differential entropy applies to any spherically symmetric state that has a classical gravitational dual in the large c limit. Once we have expressed the volume form on the space of geodesics as a function of entanglement entropy, differential entropy is obtained by simply integrating out α.

$$S_{DE} = -\frac{1}{2}\int_0^{2\pi} d\theta \int_{\alpha(\theta)}^{\pi/2} \frac{d^2 S(\alpha)}{d\alpha}\, d\alpha \wedge d\theta = \frac{1}{2}\int_0^{2\pi} d\theta \left(\frac{dS(\alpha)}{d\alpha}\right)\bigg|_{\alpha(\theta)} \tag{5.2.13}$$

We should note that when the state is excited, minimal surfaces may not reach everywhere and there exists regions that can not be probed by entanglement entropy. These regions are known as entanglement shadows [92, 150]. The Crofton formula can also be used to derive the differential entropy for entanglement shadows but the measure can not be given in terms of entanglement entropy. We will not consider long geodesics having winding number.

5.3 The first law of differential entropy

In this section we will extend the first law of entanglement entropy for boundary intervals on a constant time slice to the notion of differential entropy. We will introduce the differential modular energy that measures the change of area in terms of the energy of the excitation introduced inside the hole.

5.3.1 Differential modular energy

The state of a subsystem A is described by reduced density matrix $\rho_A = \mathrm{tr}_{\bar{A}}\rho_{\mathrm{total}}$, where ρ_{total} is the density matrix describing the global state of the full system and $\bar{A}$ is the complement of the subsystem A. The entanglement of a subsystem with the rest can be quantified by the von Neumann entropy $S_A = -\mathrm{tr}\rho_A \log\rho_A$[2]. Reduced density matrix ρ_A is both hermitian and positive (semi)definite, hence it can be expressed as $\rho_A = \frac{e^{-H_A}}{\mathrm{tr}(e^{-H_A})}$, where the Hermitian operator H_A is known as the modular Hamiltonian. The denominator is included to ensure that the reduced density matrix has unit trace. Consider any infinitesimal variation to the state of the system. The first order variation of the entanglement entropy becomes $\delta S_A = -\mathrm{tr}(\delta\rho_A \log\rho_A) - \mathrm{tr}(\delta\rho_A)$ where the reduced density matrix has unit trace $\mathrm{tr}(\delta\rho_A) = 0$.

$$\delta S_A = \delta\langle H_A\rangle \tag{5.3.1}$$

Note that $\hat{H}_A$ is the modular Hamiltonian associated with the original unperturbed state. The first law satisfies for any state in any quantum mechanical theory.

We will apply the notion of first law of entanglement in the construction of differential entropy. In other words, the first law entanglement entropy is translated into the bulk as a first law relating the local excitation inside the hole to the change of area.

$$\delta S_{\mathrm{DE}} = \sum_{k=1}^{n} \left[\delta S(I_k) - \delta S(I_k \cap I_{k+1})\right] \tag{5.3.2}$$

Since the first law of entanglement entropy holds for individual intervals, one can elevate the first law type relation to the joint relation. Hence,

$$\delta S_{\mathrm{DE}} = \delta\langle \hat{H}_{\mathrm{DE}}\rangle \tag{5.3.3}$$

[2] von Neumann entropy does not always give the entanglement entropy between the partitions of a mixed state. However it can always be considered as the entanglement between a subsytem and its complement, where complement includes the purification.

When the perturbation around the ground state is considered, each RT surface becomes a Rindler horizon that is generated by particular linear combination of isometries of AdS. In this case, for each boundary interval the modular hamiltonian is an integral of a local expression. These integrals for each boundary ball together construct another local expression of energy momentum tensor integrated on the entire boundary which quantifies the change of area in terms of the energy of the perturbation. While the first law of differential entropy is a boundary notion, it manifest itself naturally in the bulk. The relation that emerges in the bulk is the one dictated by Einstein equations inside the hole. In the continuous limit the right hand side of the equation (5.3.3) becomes,

$$\delta\langle H_{DE}\rangle = \frac{1}{2}\int d\theta\, \left(d_\alpha \delta\langle H_A(\alpha)\rangle\right)_{\alpha=\alpha(\theta)} \tag{5.3.4}$$

Around the vacuum, change in differential entropy can be seen as the change of area when a conical defect is inserted into the hole. We will show that excitation has to be inserted inside the hole.

Let us give a general form of this expression for cases where there exist a local integral expression for the modular Hamiltonian. Consider the following expression for the modular Hamiltonian,

$$H_A(\theta_0,\alpha) = \int_{\theta_0-\alpha(\theta_0)}^{\theta_0+\alpha(\theta_0)} d\theta\, \zeta^t(\theta_0,\alpha,\theta)\, \hat{T}_{tt}(\theta) \tag{5.3.5}$$

where ζ conformal Killing vector when expression is considered in AdS. We evaluate the expression on a constant time slice. Killing field ξ that generates the Rindler horizons in AdS matches ζ near the boundary. If we plug this into the expression (5.3.4),

$$\hat{H}_{DE} = \frac{1}{2}\int_0^{2\pi} d\theta_0\, \frac{d}{d\alpha}\left(\int_{\theta_0-\alpha(\theta_0)}^{\theta_0+\alpha(\theta_0)} d\theta\, \zeta^t(\theta_0,\alpha,\theta)\, \hat{T}_{tt}(\theta)\right)_{\alpha=\alpha(\theta)} \tag{5.3.6}$$

The differentiation over α can be pulled inside the integral when the Kernels (ζ^t) vanishes on the boundary of the integration domain. $\zeta^t(\theta_0,\alpha,\theta)$ satisfies this condition in AdS as well as other known solutions. We propose the vanishing of the $\zeta^t(\theta_0,\alpha,\theta)$ at the boundary of each interval as a necessary condition to have a well defined local expression for the modular Hamiltonian

$$\zeta(\theta_0,\alpha,\theta)|_{\partial A} = 0. \tag{5.3.7}$$

Then one can define the differential modular energy as,

$$\hat{H}_{DE} = \frac{1}{2}\int_0^{2\pi} d\theta_0 \int_{\theta_0-\alpha(\theta_0)}^{\theta_0+\alpha(\theta_0)} d\theta\, \left(\frac{d}{d\alpha}\zeta^t(\theta_0,\alpha,\theta)\right)_{\alpha=\alpha(\theta)} \hat{T}_{tt}(\theta) \tag{5.3.8}$$

The differential entropy side gives the change of area of the hole in the bulk. We will see what differential energy provides. The expression (5.3.8) can be simplified further by taking one of the integrals. There are two cases where one can simplify this expression without explicit knowledge of $\alpha(\theta)$. The first one is when the expectation value of the stress energy is spherically symmetric in the perturbed state namely when $\delta\langle T_{tt}\rangle(\theta) = T$. In that case one can perform the inner integral. The other case is when the hole in the bulk is a spherically symmetric one $\alpha(\theta) = \alpha$. We will give an expression for spherical hole in the bulk, in this case one can replace the orders of integration using the symmetry of the operator. While it has periodicity 2π in AdS, it is $2\pi/n$ in the conical defect examples as we will see. Using the symmetry of the operator

$$\hat{H}_{DE} = C(\alpha)\int_0^{2\pi} d\theta\, \hat{T}_{tt}(\theta) \tag{5.3.9}$$

We will see what $C(\alpha)$ measures in the bulk. As you noticed the expression can be integrated and yields $\delta\langle\hat{H}_{DE}\rangle = C(\alpha)\delta E$ for all perturbations. Interestingly as soon as the excitation is inserted inside the hole the change does not depend on where exactly it is injected. This is the consequence of Euler invariants in $2d$.

5.3.2 Seeing inside the hole

Let us study what differential energy measures in the bulk. From the differential entropy side, it is clear that it measures the change of area. However we claim that the change is due to some local excitations that resides in the hole not outside of it. The reason becomes clear in the formalism a la Wald [64], where the first law can be seen as a consequence of the existence of a closed $d-2$-form, $\chi = \delta\mathbf{Q}[\xi] - \xi\cdot\mathbf{\Theta}$, on-shell.

$$\int_\Sigma d\chi_\xi = \int_\Sigma \left(\omega(g,\delta g,\mathcal{L}_\xi g) - 2\xi^a\delta E^g_{ab}\varepsilon^b\right) \tag{5.3.10}$$

where the equations of motion on the initial solution is assumed to satisfy. Σ denotes the spacelike hypersurface, E^g_{ab} stands for equation of motion derived by varying only the metric and ε^b for the volume form on Σ. The term on the left hand side yields the change of modular energy at the infinity and the change of area on the horizon. ω is the symplectic current and vanishes when ξ is generators of the isometry of unperturbed solution

$$\delta H_A = \int_{\partial\Sigma}(\delta\mathbf{Q}[\xi] - \xi\cdot\mathbf{\Theta}) - \int_\Sigma \epsilon_{dab}\,\xi^e\delta T^d_e. \tag{5.3.11}$$

$\partial\Sigma$ is the inner boundary which is the Rindler horizon in AdS. When the perturbation is sourced by the stress tensor located on Σ, there is an extra contribution which is the source term in the linearized equation of motion. Hence, when the origin of the perturbation is due to insertion of a localized source located between minimal surface and the boundary, one would expect a contribution from the stress energy of the perturbation. Therefore when the change of area is equal to change of modular energy, there can not be any localized sources between the boundary and minimal surface. Let us calculate the change explicitly.

We consider the hyperbolic slicing of AdS_3 (4.4.9). In these coordinates AdS_3 is represented as a product of the Poincare disk with an infinite time axis. This allows us to picture AdS spacetime as a cylinder over hyperbolic space. The two dimensional conformal field theory lives on the conformal boundary of this cylinder. In these coordinates, the Killing field ξ at $t = 0$ slice becomes.

$$\xi(\theta_0, \alpha) = \left(\frac{\cos(\theta - \theta_0)}{\sin\alpha} \frac{2r}{1 + r^2} - \frac{\cos\alpha}{\sin\alpha} \right) \partial_t \tag{5.3.12}$$

the normalization is determined by $\nabla_a \xi_b = \kappa \epsilon_{ab}$, where ϵ_{ab} is the bi-normal of the minimal surface. The surface gravity is normalized to 2π. The Killing field on the boundary at $t =$ constant becomes,

$$\zeta(\theta_0, \alpha) = \frac{(\cos(\theta - \theta_0) - \cos\alpha)}{\sin\alpha} \partial_t. \tag{5.3.13}$$

Using the first law of differential entropy $\delta S_{DE} = \delta \langle H_{DE} \rangle$ we can compute the change of area of the hole, namely differential entropy purely from the boundary data. In other words, we will use the local expression for the modular Hamiltonian on each boundary ball

$$\delta\langle H_{DE} \rangle = \frac{1}{2} \int_0^{2\pi} d\theta_0 \int_{\theta_0 - \alpha}^{\theta_0 + \alpha} d\theta \left(\frac{d}{d\alpha} \zeta(\theta, \theta_0, \alpha) \right) \Big|_{\alpha = \alpha(\theta)} \delta\langle T_{tt}(\theta) \rangle. \tag{5.3.14}$$

It is possible to simplify this expression for some special cases, including spherical hole around the origin for non-homogeneous perturbations of stress energy, a homogeneous perturbation for non-spherical hole as well as localized perturbations. We will simplify by calculating the change for a spherical region around the origin *i.e.* $\alpha(\theta_0) = \alpha$. By changing the order of integrals using the periodicity of the operator, we take the θ_0 integral

$$\delta\langle H_{DE} \rangle = \int_0^{2\pi} d\theta \left(\frac{-\cos\alpha}{\sin\alpha} + \frac{\alpha}{\sin^2\alpha} \right) \delta\langle T_{tt}(\theta) \rangle. \tag{5.3.15}$$

The overall factor can be taken out of the integral and we only have the integral of the change in the vacuum expectation value of the stress energy tensor, which is equal to the conformal weight for the state describing conical defect. The physical interpretation of the change in the differential entropy from the bulk point of view becomes clear when we study the conical defect solution.

5.3.3 Bulk interpretation of first law for differential entropy

It is a well known fact that in $3d$ there are no propagating degrees of freedom. This reflects itself in the representation of $3d$ gravity as a topological field theory [151] which is studied extensively as Chern-Simons gauge theory. If there are no propogating degrees of freedom then a natural question to ask is what linearized Einstein equations are in $3d$. Non existence of propogating degrees of freedom does not exclude the gravitostatic effects, which are described through Poisson equation. Therefore in $3d$, linearized gravity is purely gravitostatics. The conical defect solutions are also part of this phenomena. We will show that first law of differential entropy in the boundary theory realizes as the gravitostatics effects on the surface areas due to the appearance of conical defects in the bulk. This can be considered as the extension of the first law of Rindler horizons to general surfaces. The first law on general spacelike surfaces correspond to the first law of differential entropy on the boundary theory. Let us demonstrate on a simple case what differential energy measures in the bulk.

In order to understand the physical origin of the change from the bulk point of view, we consider the conical defect solution as the result of the perturbation.

$$ds^2 = -\left(\gamma^2 + \frac{R^2}{L^2}\right) dT^2 + \left(\gamma^2 + \frac{R^2}{L^2}\right)^{-1} dR^2 + R^2 d\theta^2 \tag{5.3.16}$$

where $0< \gamma <1$ and related to the deficit angle as $\delta\theta = 2\pi(1-\gamma)$. It turns out that the change of entanglement has two contributions. For perturbative excitations, the contributions are disentangles as the change of radius of the hole and the change due to deficit angle introduced by the source. Let us first give the expression for holographic entanglement entropy of a conical defect,

$$S(\alpha) = \frac{L}{2G_{\mathrm{N}}} \log\left(\frac{2L}{\gamma\epsilon}\sin(\gamma\alpha)\right) \tag{5.3.17}$$

which implies $\delta S = \frac{L}{4\pi G_{\mathrm{N}}}(1 - \alpha\cot\alpha)\delta\theta$ for a small deficit angle. Identification of the cut-offs is essential to be able to compare two states. Conventional wisdom would claim that the dependence on the cut-off in a logarithmic term is irrelevant, however when one compare the two states and look at the difference in entanglement entropies, identification of cut-offs is essential. It is equivalent to matching the number of degrees of freedom. In other words we are identifying the dimensionless quantity $\frac{A(S^1)}{\epsilon}$ where the numerator stands for the circumference of the boundary S^1.

Let's us first find the radius of the hole for fixed boundary intervals α. The distance of a tip of boundary-anchored geodesic from the center equals to the radius of a

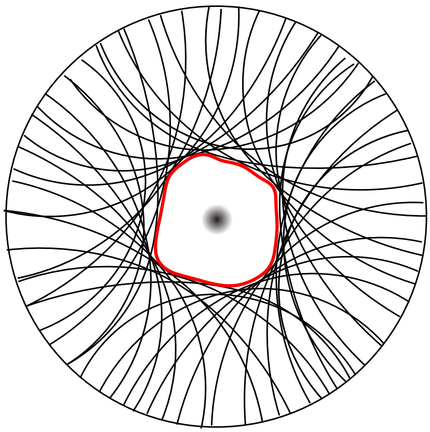

Figure 5.3: *First law of differential entropy corresponds to changes of areas due to excitation introduced into the hole. Although non of the observer can access inside the hole, their collective information can decode what happens inside the hole partially. Because gravity is a perfect conductor.*

hole. As we have noted, in hyperbolic slicing of AdS, (4.4.9) geodesics are circles intersecting the boundary. Hence in this coordinate system it is an easy task to find the location of the tip of the geodesic, $r_{\text{hole}} = \frac{1}{\cos\alpha} - \tan\alpha$. Using the relation between R and r

$$R_{\text{hole}} = L\frac{\cos\alpha}{\sin\alpha}, \qquad R'_{\text{hole}} = L\gamma\frac{\cos\alpha\gamma}{\sin\alpha\gamma} \tag{5.3.18}$$

R' stands for the radius of the hole in a conical defect solution. Hence for a small deficit, change in the differential entropy is given by,

$$\delta S_{DE} = 2\pi\frac{L}{4G}\left(\frac{\alpha}{\sin^2\alpha} - \frac{\cos\alpha}{\sin\alpha}\right)\delta + O(\delta^2). \tag{5.3.19}$$

There are two contributions to the change of area. The best way to observe the geometric origin of these contributions is to make a coordinate redefinition (5.3.16) $T' = T\gamma,\ R' = R/\gamma,\ \theta' = \theta\gamma$, which leads to global AdS, except for the different periodicity of θ, which has a deficit angle of $2\pi\delta$. Hence the origin of the second contribution becomes clear; it is the change of area due to the deficit $(2\pi\delta)R_{\text{hole}}$. The origin of the first term is purely bulk oriented in the sense that, in the presence of matter geodesics penetrate less into the bulk hence radius of the hole increases.

To sum up, the derivation of the first law for a region in spacetime always require a scheme of identification of the initial to the perturbed region. The change of areas or other quantities depend on how the two geometry is identified. There is no unique way for the identification and depends on the physics of a system.

For example the first law of black hole mechanics requires an identification on the horizon and infinity. In the derivation of the linearized Einstein equation [74] the volume of the infinitesimal ball and geodesic distance for radius of the ball is chosen as the parameter to identify two solution, while in [73] the geodesic distance is chosen as the parameter to be fixed. In these cases one observes a reduction in the area of a region. On the other hand, in our construction we have identified the number of degrees of freedom in the underlying theory by fixing the ratio of the boundary circle with the cutoff of the theory. Our identification is a very natural one and hard to refute. In this case we observe an area excess rather than a deficit.

5.3.4 The first law of differential entropy and emergence of gravitostatics

For every one parameter family of mixed states defined on a boundary according to the subsystem size $\{\rho(\alpha(\theta))\}$, there is an associated first law, which we defined as first law of differential entropy. In the bulk the relation corresponds to introducing stress energy inside the hole. There are infinitely many such relations foliate all the entire constant time slice in the bulk. Hence existence of the first law for any bulk curve is necessary condition for the linearized Einstein equations at every point in the bulk. The idea is similar to one presented in [62], where the linearized local field equations at each point in the bulk are equivalent to the first law of entanglement entropy for each boundary ball or first law for each Rindler wedge. Here we have replaced the boundary balls with one parameter family of boundary intervals and Rindler horizons with any closed surface in the bulk. Wald formalism is quite useful to lay out the relation between linearized equations and the first law.

$$\delta H_{DE}(\{\rho(\alpha(\theta_0)\}) - \delta S_{DE}(\{\rho(\alpha(\theta_0)\}) = \frac{1}{2}\int d\theta_0 \left(\frac{d}{d\alpha}\int_{C(\alpha,\theta_0)} d\chi_\xi\right)\Bigg|_{\alpha(\theta_0)} \tag{5.3.20}$$

where $C(\alpha, \theta_0)$ denotes the entanglement wedge for a single boundary interval of size α with center located at θ_0.

$$\delta H_{DE} - \delta S_{DE} = \frac{1}{2}\int d\theta_0 \left(\int_{\partial C(\alpha,\theta_0)} d_\alpha \chi_\xi + v_\alpha \cdot d\chi_\xi\right)\Bigg|_{\alpha(\theta_0)} \tag{5.3.21}$$

$\frac{d}{d\alpha}$ is denoted by d_α and v_α is the vector field generating the geodesic having boundary width, $\alpha + d\alpha$. We can express this relation in a more suggestive way as an operator acting on $d\chi_\xi$.

$$\delta H_{DE} - \delta S_{DE} = \frac{1}{2}\int d\theta_0 \left(\int_{C(\alpha,\theta_0)} (d_\alpha + \mathcal{L}_{v_\alpha})(d\chi_\xi) \right)\Bigg|_{\alpha(\theta_0)} \tag{5.3.22}$$

where we have used Stokes theorem for each term and the commutativity of $[d, d_\alpha] = 0$ for the first one and for the second term we have used the Cartan's magic formula together with $[d, \mathcal{L}_{v_\alpha}] = 0$. We know that the first law of differential entropy holds independent of the family of mixed states defined on a boundary, which ensures that left hand side of the equation vanishes. Therefore right hand side of the equation always vanishes independent of the choice of set of entanglement wedges. Using the fundemantal theorem, (5.3.10),

$$\delta H_{DE} - \delta S_{DE} = -\int d\theta_0 \left(\int_{C(\alpha,\theta_0)} (d_\alpha + \mathcal{L}_{v_\alpha})(\xi^a \delta E^g_{ab} \varepsilon^b) \right)\Bigg|_{\alpha(\theta_0)} \tag{5.3.23}$$

Since the left hand side vanishes independent of the surface $C(\alpha, \theta_0)$, we tempted to conclude that $\delta E^g_{ab} = 0$ at every point between the surface and infinity. Hence first law of entanglement entropy on bulk surfaces, implies the linearized Equations at every point between the boundary and the surface.

5.4 Modular Hamiltonian for conical defects

In section 5.2.1 we have shown that, notion of differential entropy not only applies to AdS_3 but to all spherically symmetric $3d$ asymptotically AdS geometries. Until now we have extended the first law of entanglement entropy from Rindler horizons to general closed surfaces in AdS_3. In the derivation we have introduced the differential modular energy that measures the change of area as a function of local excitation introduced inside the hole and observe an area excess.

Since differential entropy applies to spherically symmetric or translational invariant states, one can construct the first law for excited states as well. On the other hand, there are not many cases where there is a local expression for the modular Hamiltonian. One such case is the thermal state of the CFT on a flat background geometry where dual dual state is the planar BTZ. We study this case in the appendix (5.A) since it is very similar to the AdS_3 example. On the other hand, conical defect solutions also possess spherical symmetry hence differential entropy can be constructed in the same manner. Yet their modular Hamiltonian is not known explicitly. However it is possible to find an integral expression for the modular Hamiltonian in the leading order of N of a state dual to conical defect geometries in the bulk. We would like to study this case since it involves novel derivation of the approximate modular Hamiltonian.

5.4.1 Approximate modular hamiltonian of a state dual to conical defects

In this section, we will derive the modular Hamiltonian for a finite interval of a state dual to conical defects in 2d CFT living on a cylinder. The derivation is based on the description of conical defects as integer quotients of AdS.

We will find a local expression for the local modular Hamiltonian of the state dual to conical defect by uplifting it to covering theory and using the method of images that imposes $\mathbb{Z}_n$ invariance naturally. Before starting our derivation, we suggest reader to have a look at the review of conical defect solutions $\mathrm{AdS}_3/\mathbb{Z}_n$ and derivation of the holographic entanglement entropy in the parent theory, 2.6. This introduction is also the core of the derivation for approximate modular hamiltonian for the states dual conical defect geometries.

Consider a region A on the CFT_c. We would like to find an expression for the entanglement Hamiltonian of this region in the parent theory. We argue that modular Hamiltonian of a defect theory becomes a $\mathbb{Z}_n$ invariant vacuum modular Hamiltonian in a parent theory.

$$H_A = \sum_{i=1}^{n-1} g^i \tilde{H}_{\tilde{A}} \tag{5.4.1}$$

Method of images indeed let us ungauge the discreate symmetry and manifest it through the images of the uplifted region. Equivalent expression of the modular hamiltonian becomes the modular hamiltonian of the union of n evenly spaced region $\cup_i^n A_i$ of the state of the CFT dual to AdS_3. When intervals in the defect theory is sufficiently small ($2\alpha < \pi$) the modular Hamiltonian of the union can be expressed as the sum of the individual Hamiltonians for each interval in the leading order of $\mathcal{O}(1/c)$,

$$\tilde{H}_{\cup_{i=1}^n \tilde{A}_i} \approx \bigoplus_i^n \tilde{H}_{\tilde{A}_i}. \tag{5.4.2}$$

Each term on the right hand side can be expressed by the modular Hamiltonian of the vacuum of $\mathrm{CFT}_{\tilde{c}}$.

$$\tilde{H}_{\cup_{i=1}^n \tilde{A}_i} \approx 2\pi n \int_{\tilde{A}} \frac{\cos(\tilde{\theta}_0 - \tilde{\theta}) - \cos\tilde{\alpha}}{\sin\tilde{\alpha}} \hat{T}_{\text{parent}}\, d\tilde{\theta} \tag{5.4.3}$$

The relation between stress tensor of each theory is not trivial. Because, not all the generators in the Fourier decomposition descent to the defect theory. On a cylinder, stress energy tensor can be decomposed in a Fourier expansion through the generators of the Virasoro algebra,

$$\tilde{T}_{\text{cylinder}}(\tilde{\omega}) = -\sum_{k=-\infty}^{\infty} \tilde{L}_k e^{ik\omega} + \frac{\tilde{c}}{24}. \tag{5.4.4}$$

Consider the Fourier expansion of the stress tensor in the parent theory. When it is projected into the defect theory, not all the generators descent. Only the generators that are part of the subalgebra (2.6.2) will be in the defect theory. Hence the stress tensor of the defect theory will have Fourier decomposition in terms of these generators. Since $\tilde{L}_{nk} = nL_k$, the $\mathbb{Z}_n$ invariant part of the parent theory stress tensor will be,

$$\hat{\tilde{\tilde{T}}}_{\text{parent}/\mathbb{Z}_n} = n\hat{T}_{\text{defect}} \tag{5.4.5}$$

where subscript $\mathbb{Z}_n$ stands for the invariant sector of the operators. Rest of the generators are factored out. Now we can substitute this expression into (5.4.3) to obtain the correct expression on the defect theory where conical defect represent an excited state,

$$\hat{H}_A = 2\pi n \int_A \frac{\cos((\theta_0-\theta)/n) - \cos(\alpha/n)}{\sin(\alpha/n)}\, \hat{T}_{\text{defect}}\, d\theta + \mathcal{O}(1/c) \tag{5.4.6}$$

where the domain of the integral is given by $A \in [\theta_0-\alpha,\ \theta_0+\alpha]$ and $\theta \in [0, 2\pi)$. We also consider $\mathbb{Z}_n$ invariant part of the expression which is hidden in the definition of $\hat{T}_{\text{defect}}$ (5.4.5). Note that expression is correct for sufficiently small intervals in the defect theory which corresponds to $\alpha < \pi$. A similar expression can be obtained for cases where mutual information between the adjacent regions on the covering theory takes over the saddle point. In that case, one can send $\alpha \to \pi - \alpha$ and $\theta_0 \to \pi + \theta_0$ to obtain the modular Hamiltonian after the phase transition.

The bulk interpretation of the subleading term in this expression is the entanglement entropy of the quantum fields between the entanglement wedge Σ_A of the region A and its complement. The field theory in the bulk is not in its vacuum, as there is a classical source in the bulk. We left the details of the subleading term to a future study. Although derivation of the modular Hamiltonian strongly depends on the fact that n is integer, we are tempted to claim that n can be analytically continued to take real values.

Let us test this expression. First of all, a trivial check is, it yields modular Hamiltonian of the vacuum when $n \to 1$, since this is the limit where conical defect geometry becomes AdS_3.

The other limit, which provides a non trivial check is $n \to \infty$. This is the massless BTZ limit of conical defect geometry. In this limit the expression of the modular hamiltonian should match with the zero temperature limit of the modular hamiltonian for the thermal case up to the topological differences. The thermal state lives on an infinite line with periodic time direction in the euclidean space, however massless BTZ limit of conical defect is periodic on spatial slice. However, effective size of the boundary interval suppressed by n hence boundary becomes effectively an infinite line. Therefore we expect the zero temperature limit of the thermal modular Hamiltonian and $n \to \infty$ limit of conical defect to be similar up to topology. Firstly, in the zero temperature limit of the thermal state, we would expect to obtain the kernel of the modular Hamiltonian for the vacuum in Poincare patch. In $\beta \to \infty$ limit the expression, (2.4.8) becomes (2.4.3) as expected. We expect similar behaviour in $n \to \infty$ limit, which is a non trivial check for the conical defect modular hamiltonian.

$$\lim_{n\to\infty} H_A^{\text{con}} = 2\pi \int_A \left(\frac{\alpha^2 - (\theta - \theta_0)^2}{2\alpha} \right) T_{00}\, d\theta \tag{5.4.7}$$

The expression yields the modular hamiltonian of Poincare space, when α is identified with R and θ with x. Therefore the limit $n \to \infty$ matches the $T \to 0$ limit of the thermal case.

5.4.2 The first law of differential entropy for conical defect

We will apply the first law of differential entropy around the perturbations of a conical defect. Let us start by demonstrating the first law of entanglement entropy for a conical defect geometry using its modular Hamiltonian expression (5.4.6). The first law in this case is different than what we have demonstrated in the previous chapter, here we will look at the perturbations around an excited state with defect. The first law in the presence of a defect applies to one parameter family of solutions with changing angle deficit. Although we have derived the expression for integer n, for convenience we will analytically continue this parameter. Indeed such expression should exist as there is no quantization condition on the mass of conical defect.

Entanglement entropy for a boundary interval 2α in conical defect geometry in terms of the continuous parameter γ is given in (5.3.17). Let us express change in the entanglement through addition of extra defect,

$$\delta S^{\text{con}}(\alpha) = \frac{L\delta\theta}{4\pi G_{\text{N}}} \left(n - \alpha \cot(\alpha/n)\right) \tag{5.4.8}$$

where $\delta\theta = -2\pi\delta\gamma$ represent the amount of angle removed from the system due to increase in the angle deficit. Let us remind the relation $\gamma = \frac{1}{n}$. Before carrying out the modular energy side of the calculation lets present the dependence of ADM energy on the deficit angle The energy momentum tensor of the CFT is

$\delta\langle T_{tt}\rangle = \frac{-\delta(\gamma^2)}{16\pi G_{\rm N}}$. Hence change of ADM energy in terms of the new definition of the deficit angle becomes,

$$\delta M_{\rm ADM} = \frac{\delta\theta}{8\pi G_{\rm N} n}. \tag{5.4.9}$$

Let us carry the same calculation using the modular Hamiltonian expression derived in the previous section to provide even a further check on the expression (5.4.6),

$$\delta\langle H_A^{\rm con}\rangle = 2\pi n \int_{\theta_0-\alpha}^{\theta_0+\alpha} d\theta \frac{\cos\left((\theta_0-\theta)/n\right) - \cos(\alpha/n)}{\sin(\alpha/n)} \delta\langle T_{tt}(\theta)\rangle \tag{5.4.10}$$

$$= \frac{L\delta\theta}{4\pi G_{\rm N}}\left(n - \alpha\cot(\alpha/n)\right). \tag{5.4.11}$$

The radius of the circle where CFT lives is taken to be L. We perturbed the system by preserving the initial symmetries of the solution. In the first law of differential entropy we will not assume any spherical symmetry for the perturbation. Once the hole is chosen to be a sphere[3], the result is independent of the spatial configuration of the perturbation on CFT

$$\delta\langle H_{DE}^{\rm con}\rangle = n\pi \int d\theta_0 \int_{\theta_0-\alpha}^{\theta_0+\alpha} d\theta \left(\frac{1-\cos((\theta_0-\theta)/n)\cos(\alpha/n)}{\sin(\alpha/n)^2}\right)_{\alpha(\theta)=\alpha} \delta\langle T_{tt}\rangle \tag{5.4.12}$$

$$= 2\pi L\left(\frac{\alpha}{\sin(\alpha/n)^2} - n\cot(\alpha/n)\right)\int_S d\theta\, \delta\langle\hat{T}(\theta)\rangle. \tag{5.4.13}$$

We have used the implicit $\mathbb{Z}_n$ symmetry of the operator to swap the orders of integration. Comparing the result with the derivative of (5.4.8) with respect to α, we can easily notice the first law of differential entropy holds. In the AdS example we have realize that there are two contributions (5.3.19) to the entanglement entropy. The negative contribution was the one that corresponds to amount of area reduction due to point like mass introduced into the system. What happens if there is already a defect? The radius of the hole in the conical defect geometry is given by (5.3.18). Using this we can express the change in the differential modular energy in the following way,

$$\delta\langle H_{DE}^{\rm con}\rangle = \frac{\delta A_{\rm hole}}{4G_{\rm N}} = \frac{\delta\tilde{\theta}}{4G_{\rm N}}\left(\frac{\alpha}{(n\sin(\alpha/n))^2} - R_{\rm hole}\right) \tag{5.4.14}$$

where $\delta\tilde{\theta} = n\delta\theta$. Interestingly the reduction of area due to insertion of a defect amplified by n times in the presence of another existing conical defect of angle $1/n$.

[3]The hole is chosen to be a circle around the origin until now, yet one can always move it to any location through a conformal transformation.

5.5 Discussion

We have started this chapter by proposing higher dimensional generalization of differential entropy using integral geometry. Suggested generalization clarifies the necessity of spherical symmetry. By applying this method on $3d$ asymptotically AdS geometries we have shown that differential entropy can be extended to excited states having spherical symmetry. In the rest of the chapter we have used this definition to study first law of entanglement on general spacelike surfaces.

The first law of entanglement for the vacuum state of the CFT corresponds to first law of thermodynamics for a Rindler horizon in the bulk. The differential entropy uses minimal surfaces to construct a generic surface in the bulk. Using this construction we have seen that the first law of differential entropy corresponds to surface deformations of the 'hole' according to linearized Einstein equations. We have demonstrated that the linearized perturbations in the first law can be modeled as appearance of conical defects inside the hole. In the construction of the first law of differential entropy we have defined the concept of differential modular energy, which measures the change of area in terms of the energy density of the excitation introduced within the 'hole'. We have demonstrated the relation further for conical defect and planar BTZ geometries. To be able to demonstrate the first law for states dual to conical defect geometries we have derived the modular hamiltonian for a state dual to conical defect solutions in the bulk in section 5.4. This modular hamiltonian is approximate in $1/N$, since the mutual information in the parent theory is ignored.

The extension of the first law to general surfaces shows us that deformations of the surfaces and changes in the area of the surface due to local stress energy have information theoretic origin in the underlying theory. Although these phenomena are very simple and natural in the bulk perspective, the way that they are encoded in the boundary CFT might be complicated even in simple settings. Our demonstration puts an emphasis on this point.

We have observed again and again along this chapter and the previous one that studying a region in two different manifolds is a problematic task in many cases. It needs an identification scheme between two different manifolds even if they are infinitesimally close to each other along the deformation space. Our construction is free from such ambiguities since we have kept the number of degrees of freedom that each boundary observer can access to fixed. Therefore our results can be used to understand how to introduce energy associated to the matter onto holographic tensor codes [85, 86].

The information theoretic origin of area deformations has been studied in the literature recently [73, 74]. It has been observed that the excitations cause area deficit. Yet this observation depends on the comparison scheme of the manifolds as emphasized. In our construction we have observed an *area excess* rather than a deficit. The underlying reasons behind the area excess are: weak energy condition and the strong subadditivity of entanglement entropy. We do not think our result are in conflict with the existing literature on the area deficit. In our calculation one of the two contributions to the change of area comes as the area deficit term exactly same with the form observed in [73] and interpreted as the reduction of entanglement entropy due to matter. However we have also observed that the deficit term scales with the integer n in $\mathrm{AdS}_3/\mathbb{Z}_n$ geometries. In addition the sign of the term flips for a thermal state, *i.e.* the area deficit becomes area excess for perturbation around the planar BTZ (5.1.3).

As a final remark the author thinks that differential entropy is not the adequate way to study information theoretic notions in the bulk. The connection between integral geometry and differential entropy strongly indicates that the discrete version of the differential entropy is coincidental. The discrete expression (2.3.1) is simply due to the inverse of the kernel for integral equation (5.2.2) being a quadratic in $2d$. Moreover differential entropy strongly necessitates spherical symmetry. In the next chapter we will study the emergent aspects of the bulk in a more general and adequate way.

5.A First law of differential entropy for thermal states

Although the local expression for modular hamiltonian does not exist for any state, thermal states in 2d CFT admit such expression. In general, the states that are conformally equivalent to an annulus, admit local expression for modular Hamiltonians in 2d CFTs [127]. Without covering all such cases, we will focus on modular Hamiltonian for thermal state and use it to show explicitly that first law of differential entropy holds as it has to. Indeed one can see the first law of differential entropy as a necessary check on the local expressions of modular Hamiltonians. Let us first study briefly the derivation of modular Hamiltonian of thermal state.

We will use the expression (2.4.8) to demonstrate the first law for thermal state. The thermal states are dual to BTZ black holes above certain temperature. The transition from thermal gas to black hole is known as Hawking-Page transition [152]. While we will not restrict the energy distribution of the perturbation in the modular energy side of the calculation, in the differential entropy side we will consider homogeneous perturbations. Let us look at the contributions of the differential modular energy. Consider a perturbation on the state, then $\delta\langle H_{DE}\rangle$ is given by

$$\delta\langle H_{DE}\rangle = \left(-2\pi R\,\mathrm{csch}[2\pi R/\beta]^2 + \beta\coth[2\pi R/\beta]\right)\int dx\,\delta T_{00}(x) \tag{5.A.1}$$

Before giving an expression in terms of the radius of the hole. Let us look at the dual geometry. The metric for the non-rotating planar BTZ black hole can be written in Fefferman-Graham coordinate as

$$ds^2 = \frac{L^2}{z^2}\left[dz^2 + \left(1+\frac{\mu}{4}z^2\right)^2 dx^2 - \left(1-\frac{\mu}{4}z^2\right)^2 dt^2\right] \tag{5.A.2}$$

The coordinate z takes values in the interval $[0, z_e = 2/\sqrt{\mu}]$ covering the region outside the event horizon. In the CFT, it is a thermal state on Minkowski background. The energy density and temperature of the thermal state is given by,

$$\langle T_{tt}\rangle = \frac{\mu}{16\pi G_{\mathrm{N}} L}\ , \qquad\qquad T = \frac{\sqrt{\mu}}{2\pi L} \tag{5.A.3}$$

The von Neumann entropy of the state at temperature T and size R is given by,

$$S_A = \frac{c}{3}\log\left(\frac{\beta}{\pi\epsilon}\sinh(2\pi R/\beta)\right) \tag{5.A.4}$$

where ϵ is the short distance cut-off of the CFT. Using this expression we can immediately deduce the radius of the hole in the bulk,

$$R_{\mathrm{hole}} = L\sqrt{\mu}\coth\left(2\pi R/\beta\right) \tag{5.A.5}$$

Note that the expression is identical to the conical defect one (5.3.18), when one replaces $\sqrt{\mu}\to i\gamma$. Now we can express the the change in the area of the hole due to energy introduced within,

$$\frac{\delta A}{4G_{\mathrm{N}}} = \frac{1}{4G_{\mathrm{N}}}\left(R_{\mathrm{hole}} - R\,\mathrm{csch}[2\pi R/\beta]^2\right)\frac{\delta\beta}{\beta} \tag{5.A.6}$$

where we choose to keep the variation over β since it is a better measure for deficit angle then the total energy. One can immediately see the relation between conical defect geometries where $\beta/2\pi L$ takes the role of γ. It is interesting that there is a relative sign difference between two terms. While a conical defect reduces the area by introducing an area deficit, the analogous term in a thermal excitation introduces an area excess rather than a deficit.

6

The First Law of Spacetime Deformations

6.1 Introduction

Spacetime is a stage on which matter moves under the influence of forces. This is the common view on which most of our theories of the physical world are built. Yet there are well motivated reasons, coming both from theory as well as observations, to challenge this conventional point of view. From the observational side, the fact that 96% of our Universe appears to consist of mysterious forms of energy and matter, without a clear clue from theory about their nature or origin, should give sufficient motivation to reconsider these fundamental concepts. And from a theoretical perspective, besides the fact that dark energy and dark matter are still poorly understood, we start to see many indications that the standard concepts of matter and space time have a limited range of applicability. In particular, insights from black hole physics together with developments through AdS/CFT strongly suggest that our conventional notions of space time, matter, and forces are derived from an underlying microscopic description in the form of quantum information.

The first observation on the connection between information theoretic nature of the underlying microscopic description and geometry of spacetime comes through the Bekenstein-Hawking entropy [14–16, 20, 139]. Somehow gravity and horizons carry essential information about the truly fundamental constituents of Nature. It is a long standing problem: what are the fundamental constituents of Bekenstein-Hawking entropy for a generic black hole? Although there are answers for particular black holes through AdS/CFT [140–142], the generic black holes are not understood in terms of their microscopic constituents. Moreover, there are additional pieces of the puzzle on how these degrees of freedoms are represented on the spacetime or how they builds up the spacetime? These questions manifest themselves in the firewall paradox [29].

Fortunately black hole entropy is not the only observable where information theoretic nature of the underlying microscopic description and the geometry of spacetime comes together. It has been observed with the development of holographic principle in the context of AdS/CFT that, there is a deep connection between notion of area in spacetime and the entanglement entropy in the underlying microscopic state [57, 58].

$$S_{\text{ent.}} = \frac{A}{4G_{\text{N}}} \tag{6.1.1}$$

where area is the area of the minimal surface homologous to a boundary region. This is the first indication that Bekenstein-Hawking type formulas apply to more general surfaces than black hole horizons. This relation, known as Ryu-Takayanagi formula, gives us strong hints on the emergent nature of gravity and its connection to information theory.

The information theoretic nature of the microscopic constituents of black hole entropy also reflects itself in the emergent laws of black hole mechanics [22]. It seems that first law of thermodynamics or entanglement is a direct consequence of the emergent behaviour. The information theory behind the first law of thermodynamics is a universal behaviour that manifest itself almost in all the composite systems in physics. It is a universal emergent law independent of the details of the microscopic structure of the system. Even though different systems might have different definitions and expressions for the energy, the relation between information and energy at the linear level always manifest itself in the same manner. Since notion of entanglement entropy generalizes to generic spacelike surfaces through Ryu-Takayanagi (6.1.1) formula, it is natural to seek for an analogous first law for arbitrary spacelike surfacess. This is the main objective of this chapter.

Covariant phase space formulation [64, 65] of the geometric theories of gravity demonstrates the first law of black hole mechanics in a beautiful and deep way. It unifies all the geometric theories of gravity in the way that first law appears and gives a demonstration of the first law independent of the details of lagrangian of the theory. It also clarifies the connection between ADM energy that is defined at the infinity and the entropy of the black hole geometrically defined at the horizon. Wald formalism strongly suggest that it is the area law entanglement in spacetime that connects the perturbations of the conserved charges on homologous spacelike surfaces in locally vacuum spacetime. We have explored this suggestion by studying the formalism in a vector flow representation and by laying out its connections with the bit thread proposal [28].

Before introducing the content of this chapter let us note that this chapter is strongly connected with the next one 7. These two chapters will generalize what we have initiated in the previous chapter 5 using differential entropy. Namely we will be discovering emergent laws of spacetime mechanics on general spacelike surfaces. Our methodology can also be applied onto null surfaces, unfortunately within the duration of this thesis we couldn't find such opportunity. Let us give a brief summary and the main results of this chapter.

In the first section 6.2 we applied covariant phase space formalism 2.2 to general spacelike surfaces to seek for a generalization of the first law of black hole mechanics [22]. We show that on an arbitrary surface the formalism yields variations of the Brown-York quasi-local charge densities [65, 84]. In the presence of a Schwarzschild black hole the energetic cost of surface deformations given by Brown-York quasilocal charges corresponds to the change in the black hole entropy.

$$T\delta S_{BH} = \int_{S_n} N(\delta\boldsymbol{\varepsilon} + \delta\boldsymbol{\omega}) \tag{6.1.2}$$

where S_n is a spacelike codimension two surface, N is the redshift factor that provides coordinate independent way to identify surfaces along the deformation and weight factor of energy in terms of ADM energy. $\boldsymbol{\varepsilon}$ is the Brown-York surface energy density on an infinitesimal surface element $\boldsymbol{\varepsilon} = \epsilon dA$. $\delta\boldsymbol{\omega}$ is the work done on an infinitesimal surface given in terms of Brown-York surface stress tensor in the following way $\delta\boldsymbol{\omega} = \delta\sigma_{\mu\nu}\tau^{\mu\nu}dA$. Brown-York quasilocal energy density ϵ and surface stress $\tau^{\mu\nu}$ are defined as follows [84],

$$\epsilon = -\frac{1}{\sqrt{\sigma}}\frac{\delta S}{\delta N}, \qquad \tau^{\mu\nu} = \frac{2}{N\sqrt{\sigma}}\frac{\delta S}{\delta\sigma_{\mu\nu}} \tag{6.1.3}$$

where $\sigma_{\mu\nu}$ is the metric induced on the codimension 2 surface S_n and S is the action. Using the entropic correspondence of the surface deformations we have shown that this relation has an analog in elasticity theory known as the first law of elastic deformations. Using this analogy we interpret the eq. (6.1.2) as *the first law of spacetime deformations* in subsection 6.3. Although this interpretation is proposed based on the correspondence with thermal entropy we have further elaborated in the next chapter 7 that information theoretic quantity that corresponds to the left hand side of (6.1.2) is the volume entropy of the underlying phase space which is measured by the microcanonical action in the gravitational theory.

One of the main questions we have been asking in these two chapters is: what is the analog of black hole entropy on a general spacelike surface? We feel the urge of raising such a question due to recently developing paradigm: *spacetime is entanglement.* Drawing conclusions based on our findings on Bekenstein bound in the bulk 4 and evident distinction between locally vacuum spaces and thermalized regions (black holes) we seek for a different interpretation then thermal entropy (6.1.2) on a general surface. This leads us one of the main results of the next chapter. In a locally vacuum spacetime,

$$T\delta S_{BH} = \delta M_{\mathrm{ADM}} = -\omega\delta I_m \tag{6.1.4}$$

where ω is the inverse characteristic time scale of the underlying dynamics and I_m is the microcanonical action that corresponds to volume entropy of phase space in microcanonical ensemble. The adequate information theoretic interpretation on a general spacelike surface is this volume entropy. We will explain the reasons behind this interpretation in the next chapter.

The physics and mathematic behind why deformations on different surfaces are equal will be studied in 6.4, 6.5. The amount of deformations has an entropic correspondence on the horizon while an energetic (ADM) correspondence at infinity. We provide a flow line picture of the first law in subsections 6.4, 6.5 where incompressibility of the vector field naturally explains the correspondence between different surfaces. The vector field also illustrates the connectivity of the geometry in terms of the information flow. In subsection 6.5 we established the connection between the first law of deformation in spacetime with bit thread picture of holographic entanglement entropy. Although two framework is not entirely identical they have overlapping settings that we elaborate.

6.2 Application of Wald formalism to general spacelike surfaces

In this section we exploit the generic features of the formalism and apply it to codimension-two, spacelike surfaces in asymptotically flat, static solutions of general relativity. Our discussions can be completely adjusted to stationary solutions as well as asymptotically curved spaces because the only assumption that goes into the derivation of the first law in generic spacelike surfaces is the existence of a timelike Killing field. The application to stationary surfaces will provide an additional angular momentum term that we will elaborate later on. We have provided the details of the calculation in the appendix 6.A using frame field formalism. In the frame field formalism the calculation simplifies and the area term appears in a more observable fashion. To align it with the existing literature the expressions are presented in the metric formulation.

The key theorem in the covariant phase space formalism that leads to the first law of black hole mechanics is presented in (5.3.10). The theorem when executed on-shell connects the symplectic current with the exterior derivative of a $D-2$ form $\boldsymbol{\chi}$. The latter yields a change in the area and energy when integrated on the horizon and infinity respectively. On the other hand nothing prevents us from employing this theorem to more generic surfaces. Let us start by introducing the geometric setting of $(D-2)+1+1$ decomposition 6.1 where we will study $\boldsymbol{\chi}$.

Consider a static[1] solution of general relativity in D spacetime dimensions. The spacetime manifold is $\mathcal{M} = \Sigma \times I$, product of space manifold and a real line interval I. The manifold is endowed with a metric $g \coloneqq g_{\mu\nu} dx^\mu \otimes dx^\nu$. We introduce a scalar field $t(x^\mu)$ on the real line I such that $t =$ constant describes a family of non intersecting spacelike hypersurfaces Σ_t. The time function is arbitrary yet must be a single valued function of the coordinate system. The gradient of the time function defines a normal vector. All the normal vectors to the hypersurface Σ_t must be collinear to the gradient normal vector. Hence we define the future directed unit timelike normal to the hypersurface,

$$u_\mu = -N\partial_\mu t \qquad N \coloneqq (-\partial t \cdot \partial t)^{-1/2} \,. \tag{6.2.1}$$

Each spatial hypersurface Σ_t is also foliated by a family of $(D-2)$ dimensional spacelike, closed surface $(S_n)_{n\in\mathbb{R}}$ that is labeled by a real parameter n. In other words $\Sigma_t = \bigcup_{n(t)} S_{n(t)}$. For each point p in Σ_t there is a unique $S_{n(t)}$ that passes through p. We will suppress explicit t dependence of the surfaces $S_{n(t)}$ since we

[1] In general the setting explained here assumes one timelike Killing field. Generalization on the stationary solutions will be commented along the paper.

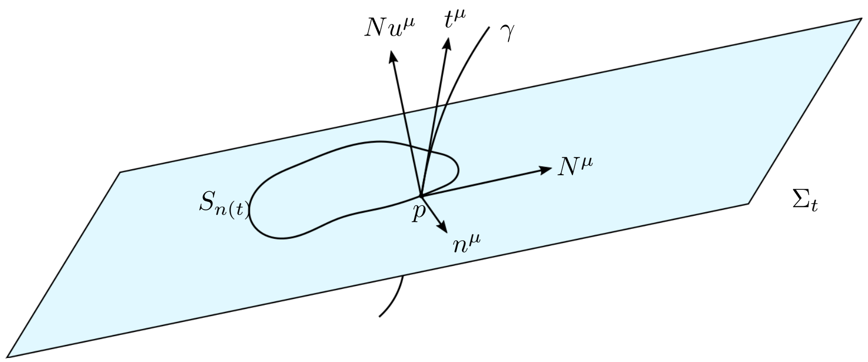

Figure 6.1: *Geometric setting for $(D-2)+1+1$ decomposition of gravitational theory is presented. The entire manifold is foliated by one parameter family of spacelike hypersurfaces Σ_t. Each spacelike hypersurface Σ_t is also foliated by one-parameter family of closed surfaces denoted by S_n. u^μ and n^μ is the normal vector of Σ_t and S_n respectively.*

restrict ourselves to one of the equivalent hypersurfaces. The $(D-2)$ dimensional surfaces S_n defines a natural normal co-vector through the equation of the surface $n_\mu \sim \partial_\mu S_{n(t)}(x)$. The tangent space of $\mathcal{M}$ at each point p decomposes into tangent vectors on S_n and a space of vectors orthogonal to S_n at p

$$\mathcal{T}_p(\mathcal{M}) = \mathcal{T}_p(S_n) \oplus \mathcal{T}_p(S_n)^\perp. \tag{6.2.2}$$

$\mathcal{T}_p(S_n)$ is the two dimensional space including the future pointing timelike unit vector field u^μ and the normal vector n^μ of the surface S_n. In the dual vector basis the metric function can be decomposed in the following form

$$g_{\mu\nu} = -u_\mu \otimes u_\nu + h_{\mu\nu} = -u_\mu \otimes u_\nu + n_\mu \otimes n_\nu + \sigma_{\mu\nu}. \tag{6.2.3}$$

$h_{\mu\nu}$ and $\sigma_{\mu\nu}$ are first fundamental forms induced on Σ_t and S_n respectively. It is this simplicity of the decomposition in the dual vector basis that we benefit from in the frame field formalism. The calculations via frame fields is presented in the appendix 6.A. Let us emphasize the difference between future directed unit normal u^μ and tangent vectors, $\partial_t = t^\mu \partial_\mu$ of the time coordinate. In general t^μ and u^μ are not collinear. They would be, only if the spatial coordinates (x^i) are chosen such that the lines $x^i =$ constant are orthogonal to the hypersurface Σ_t. Hence in general tangent vectors of the time coordinate would pick an additional component that is tangent to the hypersurface

$$t^\mu = Nu^\mu + N^\mu. \tag{6.2.4}$$

Since we will study the formalism on static solutions, in our case the relation simplifies to $t^\mu = Nu^\mu$. As we will comment later, this would only drop the angular momentum term from the first law. It will not effect the overall conclusions. Because of the static nature of the solution, the entire manifold can be foliated by a family of $D-1$ dimensional spatial hypersurfaces Σ_t that is orthogonal to the timelike Killing field. All such hypersurfaces will be equivalent.

Before presenting the result of the decomposition of $\boldsymbol{\chi}$ on S_n, let us comment on how we identify perturbed geometry with the unperturbed one. The geometries are identified such that infinities coincide. Moreover perturbation does not break the symmetries of the solution hence $\delta t^\mu = 0$. We do not assume any additional identification on S_n.

Let us present the expression for $\boldsymbol{\chi}$ evaluated on S_n calculated using the decomposition presented above

$$\delta \boldsymbol{Q}_t = \kappa \frac{\delta dA}{8\pi G_{\rm N}} + \delta\kappa \frac{dA}{8\pi G_{\rm N}} \tag{6.2.5}$$

$$t \cdot \boldsymbol{\Theta} = \frac{dA}{8\pi G_{\rm N}} \left(N\delta k + \frac{1}{2} N k^{\mu\nu} \delta\sigma_{\mu\nu} + \delta\kappa \right). \tag{6.2.6}$$

dA is the volume form on the surface S_n. κ is the surface gravity defined as $\kappa \equiv \frac{1}{2} b_{\mu\nu} \nabla^\mu t^\nu$ where $b_{\mu\nu} = 2u_{[\mu} n_{\nu]}$ is the bi-normal of the surface $S_{n(t)}$. The surface gravity κ can also be understood in terms of the four-acceleration of a test mass. It is the four-acceleration of a test mass in the normal direction to the $S_{n(t)}$ that is measured from infinity. In other words, if an observer at infinity would hold a unit test mass at a fixed orbit (*i.e.* by means of a long string) of the Killing vector field t^μ, then the amount of force that needs to be exerted is $F = \int_{S_{n(t)}} N n^\mu a_\mu dA = \int_{S_{n(t)}} \kappa dA$ where $a^\mu = u^\nu \nabla_\nu u^\mu$. N translates measurements of a local observer to ones measured from infinity. It quantifies how deep one is in the bulk with respect to infinity. It is well known that κ appears in the same way in the first law of black hole mechanics. Black hole temperature is defined as $T_{\rm BH} = \frac{\kappa}{2\pi}$. In our construction we have evaluated it on an arbitrary closed spacelike surface and observe the same structure. In this case it is the Unruh temperature due to gravitational acceleration.

Let us explain the terms in $t \cdot \boldsymbol{\Theta}$. $k_{\mu\nu}$ is the extrinsic curvature of the surface $S_{n(t)}$ embedded into Σ_t. It is the Lie drag of the first fundamental form $\sigma_{\mu\nu}$ along the unit normal n^μ, $k_{\mu\nu} \equiv \frac{1}{2} \mathcal{L}_n \sigma_{\mu\nu}$ [2]. We have presented the calculation of these quantities via frame field formalism extensively in the appendix. The trace of extrinsic curvature $k = \sigma^{\mu\nu} k_{\mu\nu}$ measures rate of expansion of the surface $S_{n(t)}$ along the direction n^μ. Hence for a minimal/extremal surface, the trace vanishes.

Before elucidating the connection between quasi-local surface quantities let us comment on the term $\delta\kappa \frac{dA}{8\pi G_{\rm N}}$. Usually $\xi \cdot \boldsymbol{\Theta}$ vanishes when the formalism is

[2] Lie derivative is defined with respect to $D-1$ dimensional manifold.

applied to bifurcation surfaces such as black hole horizons. Hence there are no contributions from $\xi \cdot \boldsymbol{\Theta}$ in the case of black holes (Note that, in general, the Killing field ξ does not have to be timelike Killing field, hence we have used the notation ξ, in these circumstances). In other words on the horizon generalized Newtonian potential $N \to 0$. However $\delta \boldsymbol{Q}_t$ also produces $\delta\kappa \frac{dA}{8\pi G_{\rm N}}$. Naively, because the black hole mass is changed, one would expect an accompanying change in the surface gravity. Then why it doesn't appear in the first law of black hole mechanics? Because in the derivation of the first law via covariant phase space approach [63] one also identifies the unit surface Killing field near the horizon, hence fixes the variations over the κ. In other words identification over the geometries in the presence of a black hole has one more restriction than assumed for a general surface S_n, which is the identifications of two solutions over the horizons. In equation (6.2.5), we observe that even if one does not identify the unit Killing field near the horizons, $\delta\kappa \frac{dA}{8\pi G_{\rm N}}$ terms cancel each other. This cancellation becomes much more apparent in the frame field formalism. $\delta \boldsymbol{Q}_t - t \cdot \boldsymbol{\Theta}$ can be cast into following form,

$$\boldsymbol{Q}_t - t \cdot \boldsymbol{\Theta} = -N\delta(kdA) - \frac{N}{2}\left(k^{\mu\nu} - \left(\frac{\kappa}{N} + k\right)\sigma^{\mu\nu}\right)\delta\sigma_{\mu\nu}dA \tag{6.2.7}$$

In this form of the expression, we immediately recognize Brown-York definition of quasilocal charges[3]. There is a relative sign between this expression and the one suggested in [83, 84] due to a relative sign difference in the definition of extrinsic curvature $k_{\mu\nu}$. Using the Brown-York definition of quasilocal quantities, we can rewrite the expression in the following form,

$$\boldsymbol{Q}_t - t \cdot \boldsymbol{\Theta} = N\left(\delta\boldsymbol{\varepsilon} + \boldsymbol{\tau}^{\mu\nu}\delta\sigma_{\mu\nu}\right) \tag{6.2.8}$$

$$= N(\delta\boldsymbol{\varepsilon} + \delta\boldsymbol{\omega}) \tag{6.2.9}$$

where $\boldsymbol{\varepsilon} = kdA$ is the energy on infinitesimal area on the surface $S_{n(t)}$. Similarly $\boldsymbol{\omega}$ is the amount of work that is done on the infinitesimal surface area due to deformations of the surface. Before we comment on the implications of this relation from the point of view of the first law on general surfaces let us have a look at the spherically symmetric case. In addition to reproducing the first law of black hole mechanics, we will also derive a more general relation that turns into first law when one moves the generic surface $S_{n(t)}$ to the horizon. The metric of such a solution is given by,

$$ds^2 = -N^2dt^2 + N^{-2}dr^2 + r^2 d\Omega^2_{D-2} \tag{6.2.10}$$

where

$$N^2 = e^{2\Phi} = 1 - \frac{16\pi G_{\rm N} M}{(D-2)\Omega_{D-2} r^{D-3}} \tag{6.2.11}$$

[3]Note that we do not have an angular momentum term and the extrinsic curvature $K_{\mu\nu}$ of Σ_t in $\mathcal{M}$ vanishes. This is consequence of static nature of the solution.

For this solution we will construct a novel first law type relation by employing again the fundamental theorem (5.3.10) of the covariant phase space formalism. The $(D-2)$-form $\boldsymbol{\chi}$ will be integrated on the surface at infinity and on an arbitrary spherical surface.

$$\int_{S(n)} \boldsymbol{\chi}_t = \int_{\infty} \boldsymbol{\chi}_t$$

Second term on the right yields the change in the total energy of the spacetime $\delta M_{\text{ADM}} = \int_\infty \boldsymbol{\chi}$ [64,65]. Let us evaluate the term on an arbitrary surface based on the quasilocal expressions we have derived. The extrinsic curvature in the normal coordinates is given by $k_{\mu\nu} = \frac{1}{2} N \partial_r \sigma_{\mu\nu}$. Hence,

$$k_{\mu\nu} = e^{\Phi} \frac{1}{r} \sigma_{\mu\nu}, \qquad k = (D-2) e^{\Phi} \frac{1}{r}. \tag{6.2.12}$$

Putting all pieces together and massaging the expression, one obtains the following the first law like relation,

$$\delta M_{\text{ADM}} = \frac{c^2}{8\pi G_{\text{N}}} \int_{S(n)} \partial_n e^{\Phi} \delta dA - \delta e^{\Phi} \partial_n dA \tag{6.2.13}$$

where the derivative in the surface normal direction is $\partial_n = n^\mu \partial_\mu$ and $\kappa = \partial_n e^{\Phi}$. On a surface of fixed radius $r =$ constant, one can integrate this relation,

$$\delta M_{\text{ADM}} = \frac{c^2}{8\pi G_{\text{N}}} \left(\partial_n e^{\Phi} \delta A - \delta e^{\Phi} \partial_n A \right) \tag{6.2.14}$$

We consider this equation a generalization of the first law of Schwarzschild black hole. Before commenting on that let us have a look at the two cases where surface S_n goes to horizon and infinity respectively. In the limit $S_n \to \mathcal{H}$, $e^{\Phi} \to 0$ and hence second term vanishes. Therefore we end up with $\delta M_{\text{ADM}} = c^2 \frac{\delta A}{8\pi G_{\text{N}}} \kappa$, which is the first law of black hole mechanics.

In the other limit where $S_n \to \infty$, the first term vanishes as $r \to \infty$ and one can express the second term $\delta e^{\Phi} \partial_n A = \frac{1}{16\pi G_{\text{N}}} \delta(e^{2\Phi}/r) A = -\delta M$, where expression (6.2.11) is used. Hence fully consistent with (6.2.14). Let us stop at this point for a moment and try to elaborate on the physics of the generalized first law of Schwarzschild black hole. Physically we can not simply say that this is a first law of thermodynamics on general surfaces since there is no thermalization process taking place. The solution could very well belong to a spherical body that is not a black hole and we are on an arbitrary surface. On this hypothetical surface we do not have any thermalization. One can define an Unruh temperature with respect to infinity. However this is not really a thermal entropy rather an ambiguity in the definition of the vacuum. Therefore the relation suggests another interpretation. It will be elaborated in 7.4 that on general surfaces measures the change in the phase space in the microcanonical ensemble multiplied by the frequency. We believe an entanglement interpretation is also suitable for the area term [60, 145].

$$S_{\text{ent}} = \frac{A}{4G_N}, \qquad\qquad \delta M = \kappa \delta S_{\text{ent}}|_{\phi} \tag{6.2.15}$$

This is a generalization (6.2.14) of the first law of entanglement to the bulk physics. At this point we do not have a strong microscopic interpretation for the second term but we should keep in mind that this term could be removed simply by identifying the surfaces (base and the perturbed) such that there is no variation on Φ. In other words one compares the surfaces having same potential N. That was what we have also executed in the first law of black hole mechanics. As we will encounter over and over along this paper when one compares the surface charges on a spacelike surfaces in two different manifold, there is no universal way to identify the surfaces. Now we will visit more general aspects of the decomposition of $\boldsymbol{\chi}$ on spacelike surfaces.

6.3 The first law of deformations in spacetime

In this section we will be studying the general expression (6.2.8) for the $(D-2)$-from $\boldsymbol{\chi}$ on spacelike surfaces for static solutions from the point of the thermodynamics of deformation in elastic bodies. Let us emphasize that our study can be completely generalized to stationary solutions and in that case one would observe the term $N^{\mu}\delta j_{\mu}$ where δj_{μ} is known as quasilocal angular momentum. For all practical purposes we will ignore this term and comment on that if the behaviour is different than expected in its presence.

Let us consider a solution of general relativity[4] that is thermalized through gravitational collapse and hence forms a black hole. The first law of black hole mechanics states that $\delta M_{\text{ADM}} = T\delta S_{BH}$ [22]. This fundamental relation can be obtained by Stokes theorem together with (5.3.10). On a hypersurface Σ_t integral at infinity yields δM_{ADM}. The infinity is the place where one can include all the energies associated to the spacetime such as binding energy. The integral at the horizon amounts to $T\delta S_{\text{BH}}$. Horizon is the place where all pixels[5] of the fabric of spacetime behaves identical. Energetic costs of deformation of spacetime from the point of infinity reduces only to a single term, change of area and associated cost is measured by surface gravity. In this setup we consider a third surface $S_{n(t)}$ between

[4] Although we refer to general relativity along the paper, covariant phase space formalism yields quasilocal charges for higher derivative theories as well [65].

[5] By pixel we refer to the quantization of the surface area in the units of Planck area.

horizon and infinity. One can relate the change of conserved charges on this surface to either the entropy of the black hole or to the total mass of the spacetime. In the former one has the following relation,

$$T\delta S_{BH} = \int_{S(n)} N \left(\delta \boldsymbol{\varepsilon} + \boldsymbol{\tau}^{\boldsymbol{\mu\nu}} \delta\sigma_{\mu\nu}\right). \tag{6.3.1}$$

The relation is very interesting as it is reminiscent of the fundamental thermodynamics relation of deformed bodies. Let us pause here for a moment and give a brief description of the thermodynamics of deformation in elasticity.

In the elastic materials the first law of thermodynamics of the deformation is given by $Tds = d\varepsilon + \delta\omega$. We have used lowercase notation to indicate that the relation is a local one and hence hold for unit volume. The work density term in the thermodynamical relation is expressed in the following way [153],

$$\delta W = \int F^i \delta u_i dV = \int \partial_j \tau^{ji} \delta u_i \, dV \tag{6.3.2}$$

where u_i is the displacement vector, that amounts to deformation of the body $u^i = x'^i - x^i$. Doing an integration by parts and sending the boundary term to infinity, one can obtain the expression in terms of the strain tensor which measures the change in the infinitesimal distance due to infinitesimal displacement

$$\delta W = \int \delta\omega dV = -\int \tau^{ji} \delta\varepsilon_{ij} \, dV. \tag{6.3.3}$$

where strain tensor is defined through the change in the infinitesimal length $dl'^2 = dl^2 + 2\varepsilon_{ij} dx^i dx^j$ hence $\varepsilon_{ij} = \partial_{(i} u_{j)}$. The overall negative sign in the work expression is due to the fact that this is the work done on the medium rather than the work it does. Hence local expression becomes,

$$Tds = d\varepsilon + \tau^{ji} \delta\varepsilon_{ij} \tag{6.3.4}$$

Let us go back and compare the fist law type relation we have derived for gravity with the first law of deformation in elasticity. The main motivation behind this comparison is the proposal [73] that our universe has an elastic component due to volume law entanglement. According to this proposal, the thermal nature of the spacetime is not localized on the horizons as suggested by general relativity but rather spread over the entire space. Hence universal information theoretic relation manifested in such a system should follow a volume law similar to the elasticity. Two phases of the underlying state, that are distinguished based on the entanglement behaviour, exists together and separated spatially. It is the competition between these two phases that causes additional forces that is interpreted as dark matter. Equations (6.3.1) and (6.3.4) can be useful to elaborate on the interaction of these two phases (area vs volume) away from spherical symmetry.

There are many differences between phenomena of elasticity and gravity. For example elasticity does not have same symmetry structure as theory of relativity or there is no spin two excitations in it. Yet there are also deep analogies when they are considered purely from an information theory perspective [73,75,77,81]. Here we want to point out a connection between these two phenomena based on their entropic nature. The equations (6.3.1) and (6.3.4), suggest a dictionary between two phenomena

$$\begin{array}{rcr} \text{Gravity} & & \text{Elasticity} \\ T\delta S_{\rm BH} & \longleftrightarrow & T\delta S \\ \delta\sigma_{\mu\nu} & \longleftrightarrow & \delta\varepsilon_{ij} \\ \dfrac{dA}{8\pi G_{\rm N}} & \longleftrightarrow & s_0 dV \end{array} \tag{6.3.5}$$

where s_0 is a parameter that has the inverse dimension of the volume such as entropy density.[6] Although we suggest a mapping between black hole entropy and entropy of the elastic material, this mapping will be elaborated in 7.4. While entropy is associated to black hole, corresponding qantity in a general surface is $\omega\,\delta I_m$, where ω is the characteristic frequency of the underlying d.o.f and δI_m is the change in the volume of the phase space in microcanonical ensemble. In the current setup $T\delta S_{BH}$ in (6.3.1) is equal to $-\omega\delta I_m$. Our main purpose is to interpret the gravitational side of the first law as a first law of deformation emerged from microscopic degrees of freedom.

Although it is harder to make an identification between unperturbed quantities due to coordinate invariance in gravity, it is meaningful to make the identification between strain and the dynamical fluctuations of the metric. Perturbation on the first fundamental form $\sigma_{\mu\nu}$ of the co-dimension two surface measures the change in the line element of the surface hence completely matching with the definition of strain tensor. The very interesting point about the identification is the dimensional reduction taking place in the gravity side. This dimensional reduction reflects itself (or as a consequence of) in the area law of black hole entropy. One interesting difference between the structure of the first laws is, while the first law in the elastic deformation can be cast into a local form through the relation of densities, the gravitational analog is a surface integral relation. This is a consequence of general coordinate invariance in general relativity. In the next section we will investigate whether a bi-local representation of the first law through a conserved vector flow is possible.

[6] In elasticity there are different types of parameters (elastic modulus), for our identification we only need a parameter that has the inverse dimension of volume. In [73] entropy density is identified with the $\approx \frac{1}{G_{\rm N}L}$ in the elastic phase of gravity.

Let us clarify further the motivation behind the map (6.3.5). The information theory behind the first law of thermodynamics reflects itself as a universal behaviour almost in all the systems in physics. It is a universal emergent law independent of the details of the microscopic structure of the system. Even though different systems might have different definitions and expressions for the energy, the relation between information and energy at the linear level always manifest itself in the same manner. We propose a map by exploiting the dimensionless quantities in the universal first law. It is the emergent nature of the first law that forces us to believe just like in the case of elasticity, the first law of deformations in gravity is also due to atoms of spacetime. These degrees of freedoms at the linear level behaves like threads.

Before finishing this section let us note that we could also establish a deformation first law between S_∞ and S_n. That would replace the left hand side of (6.3.1) with $\delta M_{\rm ADM}$ and would not affect the right hand side. In this case the cost of deformations on surface S_n equates to the change in the total energy of the system. Entropy and energy is different realization of the same underlying entity, information as advocated here [4]. In the gravitational system, change in energy and entropy of the system manifest itself on different holographic surfaces yet they are equal. However we do not identify the change of the surface charges on a generic surface with entropy or total mass of the spacetime. In other words on a generic cut on spacetime we do not associate any thermal entropy or total mass. Therefore we will be asking what is the unifying entity that amounts to the total change on every surface in spacetime. This unifying notion is the phase space in the microcanonical ensemble and the change we observe on every surface is the change on the phase space. We will close this section by stating first law of deformations in spacetime,

$$\omega\,\delta I_m = -\int N(\delta\boldsymbol{\varepsilon} + \delta\boldsymbol{\omega}) \tag{6.3.6}$$

6.4 A bilocal first law of gravity?

In the previous section we lay out the connection between elasticity and gravity from an information theoretic point of view. Intentionally we call it information theoretic since laws of thermodynamics can be derived from information theory [154], where energetic and entropic cost is the consequence of the manipulations of the information.

The main difference observed between gravity and elasticity is how the information is encoded spatially. The short range entanglement together with the increasing

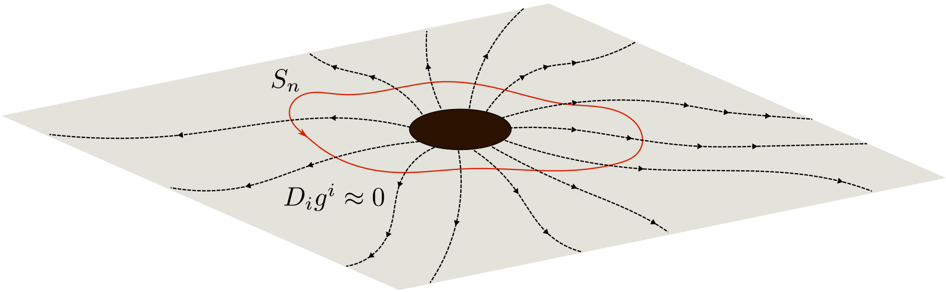

Figure 6.2: *Flow line representation of the first law. Due to incompressibility of the vector flow ($D_i \delta g^i \approx 0$), fluxes on homologous surfaces are equal. While on the horizon it amounts to the change in the black hole entropy, at the infinity it yields change of energy in the form of ADM mass. We show that on a generic surface S_n it measures change of volume entropy in the microcanonical ensemble, $\delta M_{ADM} = T\delta S = -\omega \delta I_m$*

.

number of degrees of freedom in the UV leads to area law in entanglement (6.2.15). We observe that the volume contributions in the elastic version of the first law is replaced by the area terms in gravity. In addition, we noticed that while the first law of elastodynamics or hydrodynamics exhibits a local form, it is a quasi-local expression in gravity. These quasi-local expressions can be localized by considering infinitesimal balls [74, 155] yet it still preserves its integral form. We do not prefer this sort of localization, rather we want to understand its global structure. Remember that the main identity that relates the surface integrals of $\boldsymbol{\chi}$ was the closedness of the form on shell when the symplectic current (2.2.2) vanishes. Hence there is no extra flux coming in between the two surfaces $S_{\mathcal{H}}$ and S_n.

In the presence of $(D-2)$-form $\boldsymbol{\chi}_t$, the space does not need to endowed with a metric structure to study the thermodynamics of deformations. In other words expressions of differential forms are less restrictive and has no reference to the metric structure of the manifold. This statement can be seen in the derivation of $\boldsymbol{\chi}_t$ via frame field formalism (6.B.1). On the other hand, in most cases the metric of the manifold is known and $\boldsymbol{\chi}_t$ is expressed through the perturbations of it. Given a metric, the $(D-k)$-form can be converted into a k-vector via Hodge star operator. Using the construction described in (6.2.3) for the static solutions, the $D-2$ form $\boldsymbol{\chi}_t$ can be converted into a vector field by considering the $(D-1)$ dimensional submanifold Σ_t of $\mathcal{M}$.

$$\delta g^i = h^{ij}(\star \boldsymbol{\chi}_t)_j \tag{6.4.1}$$

In a static system[7] t^μ is orthogonal to Σ_t and N^μ vanishes. Hence D dimensional manifold and it is dynamics can be represented by the metric of Σ_t induced by its

[7]Similar constructions can be applied to stationary solutions, where one chooses special

embedding in $\mathcal{M}$ and a real valued function N defined on the $(D-1)$-manifold $((\Sigma_t, h), N)$. The vector field δg^i when integrated along a surface on the unperturbed manifold $((\Sigma_t, h), N)$ amounts change of surface charges.

$$T\delta S = \int_{\mathcal{H}} \delta g^i dA_i \qquad\qquad \delta M_{\text{ADM}} = \int_{\mathcal{S}_\infty} \delta g^i dA_i \tag{6.4.2}$$

The fundamental theorem when considered on-shell reflects itself in the conservation of the vector field v^i with respect to covariant derivative defined on Σ.

$$D_i \delta g^i \approx 0 \tag{6.4.3}$$

where $\approx$ sign stands for on-shell condition including the linearized equation of the perturbation. D_i is the covariant derivative defined with respect to h_{ij} (6.2.3) and it is equal to projection of the full covariant derivative to the hypersurface $D_i v_j = h_i^k h_j^l \nabla_k v_l$. Divergenceless of the vector field (6.4.3) is equivalent to the linearized form of Gauss-Codazzi equations.

The vectorial representation of the first law is quite interesting since it provides a physical picture on why integrals on the two surfaces equal to each other. The fundamental theorem (5.3.10) implies the incompressibility of the vector flow and hence the vector field provides one to one map between change of surface quantities. It is appealing to give a bi-local interpretation to the fist law. However as you notice, δg^i has ambiguities in its definition. Namely any vector field that is divergenceless can be added to its definition without altering physics. This is the consequence of coordinate invariance of the theory.

$$\delta g^i \sim \delta g^i + d^i \qquad\qquad D_i d^i = 0 \tag{6.4.4}$$

The total derivative term vanishes when integrated on the closed surfaces hence has no affect on the total flux through closed surfaces. However it prevents us to give a canonical bilocal form to the first law. The same kind of ambiguity exist in the definition of $\boldsymbol{\chi}_\xi$ and extensively studied in [64]. Interestingly one particular form of ambiguity ($\boldsymbol{\chi}_\xi \to \boldsymbol{\chi}_\xi + \boldsymbol{Z}$ where $\boldsymbol{Z}$ is a closed form) is set to zero in AdS$_{d+1}$ to explicitly identify the changes on the CFT with the gravitational quantities. Because these ambiguities does not vanish in AdS$_{d+1}$ by itself since the minimal surfaces are not closed.

partial hypersurfaces such that Killing field t^μ is orthgonal to these surfaces. However such a coordinate system can not be used to foliate entire D dimensional manifold which is not our interest anyway.

6.5 Flow line representation of the first law and bit threads

Until now we have studied the Wald formalism on asymptotically flat solutions of general relativity, however almost everything stated so far is also true for asymptotically AdS solutions. In asymptotically AdS solutions, Wald formalism also reconciles the first law of holographic entanglement entropy and modular energy of the boundary CFT [62]. This entanglement first law can also be seen as the first law of topological black holes in asymptotically AdS solutions. On the other hand the use of formalism in the context of the first law of holographic entanglement entropy is very limited due to the necessity of a bifurcation Killing field. Only empty AdS_{d+1} posseses such Killing fields. These Killing fields ξ generates the Rindler horizons on which the $\boldsymbol{\chi}_\xi$ amounts to $\delta A/4G_\mathrm{N}$. Then the area of the Rindler horizon is identified by the entanglement entropy of the underlying theory via Ryu-Takayanagi formula [57, 58].

Ryu-Takayanagi formula is widely studied and applied in many context but its conceptual underpinnings are still under investigation. One possible explanation of why minimal/extremal surfaces yield the entanglement entropy of the region in the CFT homologous to the minimal surface in the bulk, comes via the bit thread picture of the information flow proposed in [28]. The physical picture provided through bit threads is promising as it provides a conceptual understanding on the Ryu-Takayanagi formula. The proposal provides an intuitive bulk based picture on why the von Neumann entropy on the boundary region A is provided by the minimal surface in the bulk. Bit threads naturally distinguish quantum correlations in the form of entanglement from classical ones. However the proposal is much less practical when it comes to calculate. We are going to make an observation on its connection to the vector flow representation of $\boldsymbol{\chi}_\xi$ that can expand the possible applications of the framework.

The bit thread construction starts with two main constraints on the flow line vector field. The vector field is defined to satisfy following two main property,

$$\nabla_a v^a = 0, \qquad\qquad |v| \leq C \tag{6.5.1}$$

where C is a positive constant. The first condition ensures that flow lines are decompressible and hence does not intersect each other anywhere. The second condition ensures that there is an upper bound in the number of flow lines one can put on a surface. The constant C takes the role of $4G_\mathrm{N}$ and provide a natural explanation on why entanglement entropy measured on the minimal surface is in the units of $4G_\mathrm{N}$. The von Neumann entropy on a boundary interval A is measured by the maximum possible flux configuration on A consistent with the two constraints above.

$$S(A) = \max_v \int_A v \tag{6.5.2}$$

The divergenceless condition implies that flux through A equals that through any homologous surface. Together with the maximization process this leads a saturation of the flow lines on the bottleneck of the submanifold having boundary A. In other words all the field lines that ends up in the region A needs to pass through the surface homologous to A and their density saturates the bound implied on the second constraint (6.5.1). In this surface all the flow lines are expected to be orthogonal and hence they yield the area of the minimal surface in the units of C. This is the bit thread picture behind the Ryu-Takayanagi formula. It should be noted that flow configuration is gauge dependent and there is a considerable freedom on the configuration. For example near the boundary one can insert flow lines that starts and ends on the boundary without intersecting the bottleneck. Such flow lines are allowed due to large volume near the boundary. The gauge dependence of the flow configuration is strongly interrelated with the diffeomorphisms in gravity. Note that the exact same gauge dependence is observed in the flow line representation of the $\boldsymbol{\chi}_\xi$ (6.4.4), which is the main obstacle in front of a canonical bi-local first law. Due to gauge dependence bit threads can not be physical observables by themselves but one can construct gauge independent quantities out of them such as flux of bit threads on a given surface.

Given a Riemannian manifold with a boundary one can come up with the flow satisfying the constraints above. The picture presented above is static in the sense that it does not provide dynamical evolution of the thread configuration nor it gives how the bit thread configuration modifies under the deformation of the manifold. Surely it should be connected to Einstein equations. One can deform the manifold through a deformation parameter and compare the flow lines in a fixed gauge, then change in the flow lines should be constrained under the Einstein equations.

Let us consider a manifold Σ endowed with a metric h_{ij} and associated flow lines v^i describing the entanglement of a boundary region A. Note that there is not just one flow configuration that describes the entire entanglement structure of the microscopic state, it is rather specific to a density matrix describing a subregion on the boundary. Remember that the vector field v^i measures the area homologous to the boundary region of interest. Therefore when one perturbs the metric linearly, location of the surface where the flow lines are maximized does not change due to extremality/minimality of the surface. One should only recalibrate the flow lines on the same surface such that they saturate the modified bound [156].[8] *i.e.* one

[8] As extensively studied in [156] the deformation of the area functional by a small parameter

can add and remove some flow lines (6.5.1).

Our goals is to see whether the vector field description of the Wald formalism can be connected with the flow lines of bit thread picture for a single boundary interval. The connection we propose is between the vector field δg^i and the linear change on the flow lines δv^i due to a perturbation of the metric. We will focus on what the constraint (6.4.3) implies for the bit tread vector field near the bottleneck. Studying the behaviour of the deformed flow in the neighborhood of the bottleneck will be suitable for our purposes since away from that one has so much freedom to cast the flow lines into a desired form. Let us call the minimal surface m homologous to the boundary region A. The deformed vector field and metric can be expanded around the undeformed ones,

$$h'_{ij} = h_{ij} + \lambda h^{(1)}_{ij} + ... \tag{6.5.3}$$

$$v^i = v^i_0 + \lambda\, v^i_1 + ... \tag{6.5.4}$$

where λ is the deformation parameter and v^i_0 is the flow on the initial manifold. We would like to see whether deformation on the vector flow v^i_1 satisfies the conservation equation (6.4.3). Both v^i and v^i_0 satisfies incompressibility condition with respect to manifolds they are defined. On the other hand the condition (6.4.3) is slightly different in the sense that it is the conservation of first order deformation of the vector field with respect to the initial manifold. Generally $D_i v^i_1$ does not vanish everywhere based on the divergenceless condition, however we will check whether v^i_1 satisfies the condition on a minimal surface. Once it satisfies the desired condition on the bottleneck whether it can be extended to entire manifold without violating the obstruction conditions [156] is a different problem. For now we will not be ambitious and stick to the condition on the minimal surface. Convenient coordinate system to study a hypersurface is Gaussian normal coordinates. In GNC, the metric of $D-1$ dimensional space can be expressed in the following way,

$$ds^2 = d\sigma^2 + \tilde{h}_{ab} dx^a dx^b \tag{6.5.5}$$

where σ is a parameter denoting hypersurfaces and $\tilde{h}_{ij}$ is the induced metric on the same hypersurface. On the initial manifold, v^i_0 is normal to the surface where it saturates the bound (6.5.1). Hence,

$$v^\lambda_0 = u, \qquad\qquad v^a_0 = 0 \tag{6.5.6}$$

namely, there is no tangential components of the vector field on the surface. Using divergenceless condition for v^i at the first order, we can show that

$$D_i v^i_1 + D^{(1)}_i v^i_0 = 0. \tag{6.5.7}$$

λ sets a new bound on the maximum number of threads.

$D_i^{(1)}$ denotes the linear term in covariant derivative in the expansion of λ. $D_i^{(1)} v_0^i = \delta\Gamma^j_{ji} v_0^i$. Since v_0^i vanishes in the tangential directions,

$$D_i v_1^i = -\delta\Gamma^j_{j\sigma} v_0^\sigma. \tag{6.5.8}$$

In GNC connections become expressions of extrinsic curvature of the surface $\Gamma^i_{j\lambda} = k^i_j$. Because of the minimality condition, $k = 0$ on the surface,

$$(D_i \delta v)_m = 0. \tag{6.5.9}$$

where subscript m indicates the condition satisfies on the surface. This is the same condition with $d\boldsymbol{\chi} \approx 0$. Yet whether the condition (6.5.9) can be extended everywhere on the manifold is different or whether it can be introduced as an additional constraint on the max flow min cut theorems for small deformations of the manifold according to Einstein equations. Let us now go back to the covariant phase space formalism and discuss in which setups it can be reconciled with the bit thread formulation.

When Wald formalism is applied to AdS_{d+1} using the Rindler horizon generating Killing fields ξ^μ then the equations (6.4.2) produces the change on the holographic entanglement entropy together and change in the modular energy.

$$\delta S_{\text{ent.}} = \int_m \delta g^i_\xi dA_i \qquad\qquad \delta H_{\text{mod.}} = \int_A \delta g^i_\xi dA_i \tag{6.5.10}$$

Moreover in this background one can perform the construction outlined in section (6.2). AdS_{d+1} is also static and can be foliated by the constant time slices. The Killing field becomes timelike on constant time hypersurface. In addition we describe the vector field δg^i in the same manner (6.4.1). In this background with bifurcation Killing field ξ, vector field δg^i matches the vector field δv^i describing the modification of the flow lines at the linear level due to a perturbation on the metric.

6.6 Conclusion and Discussion

In this chapter we study the covariant phase space formalism on codimension two spacelike surfaces and make observations regarding the emergent nature of gravitation at the linear order.

- **First law of spacetime entanglement:** We have generalized the first law of a Schwarzschild black hole to spherical holographic surfaces (6.2.15). In the limit that holographic surface approaches to black hole horizon one reproduces well known first law. We have advocated that areas of codimension two surfaces in the spacetime measure entanglement across regions separated by the surface. In other words the notion of area is result of dimensional reduction in entanglement in the underlying theory. Based on such postulate, we interpret eq. (6.2.15) as the first law of spacetime entanglement. On the other hand the relation does not hold for a generic surface and is limited to spherically symmetric systems. The interpretation for a general surface is done through the analogy to elasticity.

- **First law of spacetime deformations:** We have shown that $D-2$ form $\boldsymbol{\chi}_t$ yields changes on Brown-York quasi local charge densities on a general spacelike surface [65]. When the formalism applied to a general surface in the presence of a Schwarzschild black hole, one recognizes the connection between entropy of the black hole and quasi local charges defined on the surface. The relation (6.3.1) is very reminiscent to thermodynamics of elastic deformations. Based on their analogy through first law, we have provided a map between these two phenomena. Our map discovers the geometric correspondence between perturbations of metric and strain. It is interesting that we have discovered this analogy purely from information theoretic arguments. The main difference between the first laws is while elastic one has a local form gravitational analog possesses a qusilocal form due to general coordinate invariance. In addition the entropy and energy is measured on different surfaces in gravity, which brings us to the next conclusion.

- **Spacetime as a perfect conductor:** Wald formalism beautifully unifies different geometric theories of gravity and provides a canonical definition of the black hole entropy as Noether charge. In addition it provides a systematic derivation of the first law of black hole thermodynamics. The essential component of the derivation is on shell closedness of $\boldsymbol{\chi}_t$. This property allows one to equate deformations on homologous surfaces. We provide a vector flow representation of $\boldsymbol{\chi}$ via Hodge duality on $D-1$ dimensional spatial submanifold. In this representation, incompressibility of the flow lines connects the deformations on homologous surfaces. We observed that the vector field corresponds to gravitational field when $\boldsymbol{\chi}$ is considered with a timelike Killing field t and it coincides with the bit variations on the vector flow of the bit thread picture [28] when considered with horizon generating Killing field ξ in AdS. Therefore we concluded that the closedness of the form is strongly connected with the area law entanglement in the underlying theory.

6.A Conserved charges in frame field formalism

The frame field formalism, á la Palatini will provide us the simplicity and clarity to study covariant phase space approach of black hole thermodynamics in general spacelike surfaces for general relativity.

In the first-order orthonormal frame formalism, the Lagrangian of general relativity in D-dimensions is expressed in terms of the frame field 1-form e^a and connection 1-form ω^a_b which are both $SO(D-1,1)$ Lie algebra-valued. These are the independent dynamical variables of the theory. The spacetime metric in terms of the frame fields is given by, $g_{\mu\nu} = \eta_{ab} e^a_\mu e^b_\nu$, where η_{ab} is the Minkowski metric. The curvature 2-form is defined by $R^a_{\ b} = D\omega^a_{\ b} = d\omega^a_{\ b} + \omega^a_{\ c} \wedge \omega^c_{\ b}$.

The Lagrangian D-form for general relativity is a function of the frame and the connection 1-form via the curvature 2-form,

$$L(e,\omega) = \epsilon_{a\ldots bcd}\, e^a \wedge \ldots e^b \wedge R^{cd}$$

(we set $16\pi G = c = 1$)
$\omega(e)$ should be regarded as a function of the vielbein e. It is determined by the condition that the variation of the action with respect to ω vanishes. It is instructive to see the explicit variation of the Lagrangian before presenting equations motion and their geometric meanings.

$$\begin{aligned} \delta L &= \delta e^a \wedge \frac{\partial L}{\partial e^a} + D\delta\omega^{ab} \wedge \frac{\partial L}{\partial R^{ab}} \\ &= \delta e^a \wedge \frac{\partial L}{\partial e^a} + \delta\omega^{ab} \wedge D\frac{\partial L}{\partial R^{ab}} + d\left(\delta\omega^{ab} \wedge \frac{\partial L}{\partial R^{ab}}\right) \end{aligned} \tag{6.A.1}$$

From this equation one can read off equations of motion together with the symplectic potential. Equation of motion determining the ω is,

$$\epsilon_{abc\ldots df}\, e^c \wedge \ldots \wedge e^d \wedge De^f = 0$$

where D is Lorentz-covariant exterior derivative. The vanishing of this equation implies, $De^a = 0$, which is equivalent to the torsion free condition. The solution sets the connection, ω, in terms of frame fields, e. If one starts with the solution of ω at the level of Lagrangian then one has the second order frame formalism. To see the action of Lorentz-covariant exterior derivative, let's give the first structure equation explicitly.

$$De^a = de^a + \omega^a_{\ b} \wedge e^b = 0$$

The variation of the action with respect to the frame fields yield the other equation of motion.

$$\epsilon_{ab\ldots cde}\, e^b \wedge \ldots e^c \wedge R^{de} = 0$$

When the solution of the first structure equation is used for the equation of motion of the frame fields, it becomes equivalent to the vanishing of Ricci tensor so, one recovers the vacuum Einstein equations.

One, can further read off the symplectic potential from eq.(6.A.1).

$$\mathbf{\Theta}(\delta\omega) \equiv \delta\omega^{ab} \wedge \frac{\partial L}{\partial R^{ab}} \tag{6.A.2}$$

$$= \epsilon_{abd\ldots f}\, \delta\omega^{ab} \wedge e^{d} \ldots \wedge e^{f}$$

the final experssion is the symplectic potential for general relativity. Using the definition of Noether current (2.2.3), we write the explicit expression in the following way,

$$\boldsymbol{J}_\xi = d\left(\xi \cdot \omega^{ab} \wedge \frac{\partial L}{\partial R^{ab}}\right) - (\xi \cdot \omega^{ab}) \wedge D\frac{\partial L}{\partial R^{ab}} + \xi \cdot R^{ab} \wedge \frac{\partial L}{\partial R^{ab}} - \xi \cdot L$$

where we have used $\mathcal{L}_\xi \omega = \xi \cdot R + D(\xi \cdot \omega)$ in the expression for $\boldsymbol{J}$. The goal is to express the current as a closed form up to e.o.m. $D\frac{\partial L}{\partial R^{ab}}$ vanishes by the first structure equation. Explicitly, $\xi \cdot L$ is,

$$\xi \cdot L = \xi \cdot e^a \wedge \frac{\partial L}{\partial e^a} + \xi \cdot R^{ab} \wedge \frac{\partial L}{\partial R^{ab}}$$

this expression combined with the third term in $\boldsymbol{J}$ vanishes by the e e.o.m.

$$\boldsymbol{J}_\xi \approx d\left(\xi \cdot \omega^{ab} \wedge \frac{\partial L}{\partial R^{ab}}\right) \tag{6.A.3}$$

Now one can simply read the Noether charge,

$$\boldsymbol{Q}_\xi = \xi \cdot \omega^{ab} \wedge \frac{\partial L}{\partial R^{ab}} \tag{6.A.4}$$

Now we can give the explicit expression for, $D-2$-form $\boldsymbol{\chi}$,

$$\begin{aligned}\boldsymbol{\chi} = \delta \boldsymbol{Q}_\xi - \xi \cdot \mathbf{\Theta} &= \delta\left(\xi \cdot \omega^{ab} \wedge \frac{\partial L}{\partial R^{ab}}\right) - \xi \cdot \left(\delta\omega^{ab} \wedge \frac{\partial L}{\partial R^{ab}}\right) \\ &= \xi \cdot \omega^{ab} \wedge \delta\frac{\partial L}{\partial R^{ab}} + \delta\omega^{ab} \wedge \xi \cdot \frac{\partial L}{\partial R^{ab}}\end{aligned} \tag{6.A.5}$$

Integral of the $(D-2)$-form over the boundary of the manifold yields the first law of black holes, which is the special case of the variational identity obtained for generic spherical surfaces in a spherically symmetric static solution.

6.B Adiabatic variables in stationary systems

Consider a stationary solution of general relativity in D spacetime dimension. It possesses a timelike Killing vector field, ξ^μ. We pick a partial Cauchy surface on this solution such that ξ^μ is orthogonal on the boundaries of the surface. At first instance this may not seem plausible due to hypersurface orthogonality condition. However we consider only partial Cauchy surfaces and we are not interested in foliation of this surface along the orbits of ξ. For example, in the case of Kerr black hole, one can find such spacelike surface as soon as one stays outside of the horizon. This surfaces are also considered in [157] as zero angular momentum observer surfaces (ZAMO). In all the cases covered in this work, the codimension 2 surface stays outside of any event horizon.

Noether Charge on spacelike surfaces:

Let us start by studying the form of Noether charge on codimension-2 spacelike surfaces. Define the area element as the $D-2$-form

$$dA_{ab} \equiv \frac{\partial L}{\partial R^{ab}} = \epsilon_{abc\dots d} e^c \wedge \dots \wedge e^d$$

and the $(D-2)$-form χ is

$$\delta Q - \xi \cdot \Theta = i_\xi \omega^{ab} \delta dA_{ab} + \delta\omega^{ab} \wedge i_\xi dA_{ab} \tag{6.B.1}$$

We will study and extract the variables in δQ and $\xi \cdot \Theta$ on the surface of consideration. Let's start with the Noether charge,

$$Q = i_\xi \omega^{ab} dA_{ab}$$

On any $D-2$ dimensional surface residing on the partial Cauchy surface with a normal ξ^μ, Noether charge becomes surface acceleration times the area.

$$Q = -\kappa n^{ab} dA_{ab} = 2\kappa dA$$

where

$$n^{ab} = n^{\mu\nu} e^a_\mu e^b_\nu, \qquad \kappa = \frac{1}{2} n^{\mu\nu} \nabla_\mu \xi_\nu$$

are the bi-normal to S (converted to a Lorentz tensor) and the surface acceleration at S, respectively. We observe that the first part of the Noether charge indeed is

$$i_\xi (\omega^e)^a{}_b = e^\mu_b e^a_\nu \nabla_\mu \xi^\nu = -\frac{1}{2} n^a{}_b n^\mu{}_\nu \nabla_\mu \xi^\nu = -\kappa n^a{}_b$$

where e^μ_b is the inverse frame. Here we used the definition of n^{ab} and κ, and the fact that the bi-normal is normalized to $n^{\mu\nu} n_{\mu\nu} = -2$.

Before focusing our attention on the symplectic potential part of the variational relation. We would like to point out an interesting observation that is clear via frame field formalism.

In the Wald formalism, black hole entropy is given by the Noether charge, $\int \widetilde{\boldsymbol{Q}}_\xi$ and the variation on the entropy of the black hole in the first law is simply obtained by the variation of the Noether charge. The term including symplectic potential $\xi \cdot \boldsymbol{\Theta}(\delta\omega)$ vanishes since $\xi|_{\mathcal{H}} = 0$ as a property of Killing horizons. Then the first law reduces to the following variational relation,

$$\int_{\mathcal{H}} \delta \boldsymbol{Q} = \int_{\infty} \delta \boldsymbol{Q} - \xi \cdot \boldsymbol{\Theta} \tag{6.B.2}$$

while term on the right amounts to the variation on the ADM energy of spacetime, the term on left is not exactly equal to $T\delta S (= \kappa \delta A)$. One would also expect to see the variation on the surface gravity since the temperature of the perturbed system has also changed as it depends on the mass of the black hole. It has been justified in [REF: Black Hole entropy is Noether Charge] that the correct identification of the manifolds leads stationarity of the surface gravity. However, we will demonstrate that variation on the surface gravity actually is canceled when one considers the presypmlectic potential on the surface. The expression (6.A.5) includes terms having variation on the surface gravity, $\delta(\xi \cdot \omega^{ab}) \wedge \frac{\partial L}{\partial R^{ab}}$. Internal cancellation of the variation on the surface gravity between Noether charge and existence of presymplectic potential in $\boldsymbol{\chi}_\xi$ is the actual reason behind the stationarity of the surface gravity in the first law of black hole mechanics. This cancellation is even more important when one study variational approach of covariant phase space formalism on general spacelike surfaces since Killing field does not vanish on general surfaces. Now, let's focus our attention to the other term, $\xi \cdot \boldsymbol{\Theta}$.

Presypmlectic potential on spacelike surfaces:
The $D-2$ dimensional surface has two normal vectors, timelike one, ξ^μ and spacelike "radial" one, n^μ. Then on the surface, we have

$$e^t = 0 \quad \text{and} \quad e^n = 0$$

The statement is trivial in the sense that dual forms of normal vectors vanishes on the surface, in other words volume form on the surface does not contain the line elements along the normal directions. Then, the first structure equations take the form

$$\begin{aligned} de^{\theta_i} &= e^{\theta_j} \wedge \omega^{\theta_i}{}_{\theta_j} \\ 0 &= e^{\theta_j} \wedge \omega^{n}{}_{\theta_j} \\ 0 &= e^{\theta_j} \wedge \omega^{t}{}_{\theta_j} \end{aligned} \tag{6.B.3}$$

summation convension is used and θ_j denotes the elements on the surface. The geometric meaning of the first equation of (6.B.3), is the following. If we consider a surface as a $D-2$ dimensional Riemannian manifold, then the forms $\omega^{\theta_i}_{\theta_j}$ of this manifold are the same as those in the ambient manifold.
Consider now the second and third equations in (6.B.3). By Cartan's lemma, these equations imply the following linear relations.

$$\omega^{n,t}_{\ \ \theta_i} = k^{n,t}_{\ \ \theta_i\theta_j}\, e^{\theta_j} \tag{6.B.4}$$

by n, t we denote one equation for each normal direction. We will be particularly interested in $k^n{}_{\theta_i\theta_j}$. The second fundamental form of $D-2$ dimensional surface in ambient $D-1$ dimensional Riemannian manifold (because of the orthogonality of ξ^μ to the hypersurface, system can be considered in a $D-1$ dimensional manifold with Euclidean signature.) is given in the diagonalizing frame by

$$\Pi = k^n{}_{\theta_i\theta_i}\, e^{\theta_i} \otimes e^{\theta_i}$$

and $k^n{}_{\theta_i\theta_i}$ denotes the *principal curvatures*, eigenvalues of the second fundamental form. Remind yourself,

$$\begin{aligned}\xi \cdot \mathbf{\Theta} = -\delta\omega^{ab} \wedge i_\xi dA_{ab} &= 2N\epsilon_{tnbc\ldots d}\,\delta\omega^{nb} \wedge e^c \wedge \ldots \wedge e^d \\ &= 2N \sum_{\theta_i} \epsilon_{tn\theta_1\ldots\theta_i..\theta_{D-2}}\, e^{\theta_1} \ldots \wedge \delta(k_{\theta_i} e^{\theta_i}) \ldots \wedge e^{\theta_{D-2}}\end{aligned} \tag{6.B.5}$$

where we have denoted principal curvatures with a single index, k_{θ_i}. $\xi \cdot \mathbf{\Theta}$ measures the change in each principal curvature together with the line element of the *curvature lines*. The terms $k_{\theta_i} e^{\theta_i}$ should be considered as canonically conjugate variables to e^{θ_i}. For spherically symmetric solution we obtain the desired expression.

$$\xi \cdot \mathbf{\Theta} = \delta(e^{2\phi})\partial_r dA$$

To sum up in general dimensions we have the following expression

$$\boldsymbol{\chi} = \sum_{\theta_i} e^{\theta_1} \ldots \wedge \left(2\kappa\, \delta e^{\theta_i} - 2N\delta(k_{\theta_i} e^{\theta_i})\right) \ldots \wedge e^{\theta_{D-2}} \tag{6.B.6}$$

It is clear that we have $D-2$ pairs of independent variables. These are the variables of an adiabatic change. The first group of variables are the variations of the frame fields. All of them together yields the change of area, which we consider as a measure of entanglement entropy of the spacetime. It is a measure of connectivity in the macroscobic sense, which is translated as entanglement entropy in the microscopic realm. The other set of variables, $(k_{\theta_i} e^{\theta_i})$ are the reminiscent of what we see as the change of Newtonian potential, $\delta e^{2\phi}$, in the spherically symmetric case. Simply because $(k_{\theta_i} e^{\theta_i}) = -e^{2\phi}$ for each principal direction when the system is spherically symmetric.

6.B.1 The fundamental variational identity in frame field formalism

Let's start by giving the off-shell expression for the Noether current.

$$\boldsymbol{J}_\xi = d\boldsymbol{Q}_\xi + \xi^a \boldsymbol{C}_a \tag{6.B.7}$$

where $\boldsymbol{C}_a$ are the constraint equations on a fixed-time slice. That is,

$$\boldsymbol{C}_a = \sum_\phi \left(\sum_{i=1}^{r} (E^\phi)^{b_1 \ldots b_s}_{c_1 \ldots a \ldots c_r} \phi^{c_1 \ldots c_i \ldots c_r}_{b_1 \ldots b_s} \boldsymbol{\varepsilon}_{ci} - \sum_{i=1}^{s} (E^\phi)^{b_1 \ldots b_i \ldots b_s}_{c_1 \ldots c_r} \phi^{c_1 \ldots c_r}_{b_1 \ldots a \ldots b_s} \boldsymbol{\varepsilon}_{b_i} \right) \tag{6.B.8}$$

where ϕ is a type (r, s) tensor and the $D-1$-from $\boldsymbol{\varepsilon}_{b_i}$ is the natural volume form on the co-dimension one surfaces. The expression for the off-shell current in the frame field formulation is given by,

$$\boldsymbol{J}_\xi = d\boldsymbol{Q}_\xi - (\xi \cdot e^a) \wedge \frac{\partial L}{\partial e^a} - (\xi \cdot \omega^{ab}) \wedge D \frac{\partial L}{\partial R^{ab}} \tag{6.B.9}$$

The demonstration of the fundamental theorem in the Cartan formulation is as follows,

$$d(\delta \boldsymbol{Q}_\xi - \xi \cdot \boldsymbol{\Theta}) = \delta \boldsymbol{J}_\xi + \delta \left((\xi \cdot e^a) \wedge \frac{\partial L}{\partial e^a} \right) + \delta \left((\xi \cdot \omega^{ab}) \wedge D \frac{\partial L}{\partial R^{ab}} \right) - d(\xi \cdot \boldsymbol{\Theta}) \tag{6.B.10}$$

$$= \delta \boldsymbol{\Theta}(\mathcal{L}_\xi \phi) - \delta(\xi \cdot L) + \delta \left((\xi \cdot e^a) \wedge \frac{\partial L}{\partial e^a} \right) + \delta \left((\xi \cdot \omega^{ab}) \wedge D \frac{\partial L}{\partial R^{ab}} \right) - d(\xi \cdot \boldsymbol{\Theta}) \tag{6.B.11}$$

$$= \boldsymbol{\omega}(\phi, \delta\phi, \mathcal{L}_\xi \phi) + \left(\delta e^a \wedge \xi \cdot \frac{\partial L}{\partial e^a} \right) + \left(\xi \cdot e^a \wedge \delta \frac{\partial L}{\partial e^a} \right) + \left(\delta \omega^{ab} \wedge \xi \cdot D \frac{\partial L}{\partial R^{ab}} \right) + \left(\xi \cdot \omega^{ab} \wedge \delta\, D \frac{\partial L}{\partial R^{ab}} \right) \tag{6.B.12}$$

In the second line definition of the Noether current is used, (2.2.3) and in the third, we plugged in the expression for $\delta(\xi \cdot L)$, (6.A.1).

To sum up, we have demonstrated explicitly that exterior derivative of $(D-2)$-form $\boldsymbol{\chi}$ yields presymplectic potential together with e.o.m. and linearized ones.

$$d\boldsymbol{\chi} = \boldsymbol{\omega}(\phi, \delta\phi, \mathcal{L}_\xi \phi) + \delta\phi^a\, \xi \cdot E^\phi_a + \xi \cdot \phi^a\, \delta E^\phi_a \tag{6.B.13}$$

7

Adiabatic Principle and Origin of Inertia

You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, nobody knows what entropy really is, so in a debate you will always have the advantage.
— John von Neumann *to Claude Shannon.*

7.1 Introduction

In the previous chapter, we have generalized the first law of black hole mechanics to general spacelike surfaces using covariant phase space formalism. We have interpreted this generalization as the first law of deformations in spacetime based on the analogies with elasticity. To be able to draw analogies with elasticity we kept the thermal entropy in the picture. On the other hand we have explicitly commented on the fact that, on a codimension two spacelike surface the concept of thermal entropy doesn't adequately reflect the physics. Although one can advocate the existence of the thermal bath in a non inertial frame due to Unruh effect, we view it as an ambiguity in our theoretical description rather than an actual thermalization process. Moreover one does not observe a degeneration in the separation of time scales while approaching a generic spacelike surface. We will further discuss which other properties force us to distinguish the horizon from other spacelike surfaces. We will propose the entropic measure of the volume of the phase space in the microcanonical ensemble as the corresponding notion which is measured by the microcanonical action in the gravitational theory.

The previous chapter is about the emergent aspects of linearized gravity. It explores the information theoretic realizations of perturbative gravitational excitations on a given background geometry. In this chapter we will be asking a different question: how does a probe explores so called 'background geometry', what is the principle behind geodesic motion? These questions are very different in nature then ones answered in the previous chapter. The difference between setup studied in the previous chapter and this one finds its roots in the question whether Einstein field equations imply geodesic motion, or geodesic motion is an additional postulate that is independent of field equations. Einstein himself was also aware of this distinction and pointed it out that geodesic motion is introduced into the theory as an additional postulate [158, 159].

In this chapter we propose an underlying principle for geodesic equation. It is proposed that underlying our universe there exists a microscopic fast dynamical system, from which our usual macroscopic concepts such as matter and forces have to be derived. Specifically, we will argue that all inertial forces, including gravity, are a consequence of the fact that the phase space volume of the underlying system is influenced by macroscopic variables such as the positions of material objects. When a fast dynamical system is driven by one or more slow variables, the fast system reacts back on the slow variables, and creates a reaction force. This reaction force can be studied in a systematic expansion using the small parameter that controls the separation of time scales between the slow and the fast variables. In leading order, the reaction force on the slow system follows from the adiabatic principle that the phase space volume of the fast system is preserved. In quantum mechanics this principle is a stated in the Born-Oppenheimer approximation. The existence of the inertial and gravitational force thus follows from general physical principles, and as we will see, requires very little information about the underlying dynamical system except its phase space volume, its density of states and the typical time scale of its dynamics, all as a function of the slow variables. In subsection 7.2 we have demonstrated adiabatic principle and emergent reaction forces for an ergodic system consists of N d.o.f.

The first order correction to the adiabatic reaction force is due to the Berry phase [160], which is described by an abelian connection on the space of slow variables. In specific situations the Berry connection can be naturally generalized to a non-abelian gauge field. The next order reaction force is represented in terms of a metric on the space of slow variables, while in higher order other tensor fields are expected to appear [161]. Given these facts, it is natural to propose that the other forces of nature (as well as the the full relativistic form of gravity) can all be understood as reaction forces caused by the same underlying microscopic system. In this thesis we don't discuss these corrections and their possible relation with gauge forces any further, but instead we will focus our attention on the leading adiabatic reaction force and its connection with gravity and inertia 7.3.

The concept of an adiabatic reaction force is closely related to that of an entropic force [11]. The main difference between these two types of emergent forces is the degree in which the time scales of the fast dynamical system and the slow macroscopic variables are separated. Our conventional space time description works well as long as the dynamics of the underlying system is much faster than that of the observed phenomena. When the separation of time scales becomes small the fast dynamical will start to behave as a thermal heath bath for the slow variables. We argue that this is precisely what happens near event horizons. It appears that the separation of the time scales between the microscopic and macroscopic variables is controlled by the red shift factor, and thus starts to breaks when one approaches a surface of infinite redshift (=horizons). As a result in the near horizon region gravity slowly degenerates from an adiabatic reaction force in the Born-Oppenheimer regime into a genuine entropic force in the usual thermodynamic sense. At the horizon the separation of time scales completely breaks down, and gravity, matter and all other notions that involve space and time cease to exist as well defined concepts since their definitions relied in an essential way on the separation of scales. This distinction between entropic and adiabatic regimes let us interpret the information theoretic correspondence of deformations on a spacelike surface with the notion of volume entropy[1] which measures the volume of the phase space in microcanonical ensemble. This volume entropy is measured by the microcanonical action in the gravitation theory 7.4. Hence we present the main result of this chapter,

$$\delta_{\boldsymbol{X}} M_{\text{ADM}} = -\omega \delta_{\boldsymbol{X}} I_m - F^g_\mu \delta \boldsymbol{X}^\mu \tag{7.1.1}$$

where $\boldsymbol{X}$ is the slow variable corresponding to the location of the macroscopic objects on spacetime. I_m is the microcanonical action and corresponds to the volume entropy of Gibbs and ω is level spacing of the energy levels in the underlying fast system.

In sections 7.5, we have studied the Newtonian limit of the covariant version of the adiabatic principle (7.1.1). The surface expression of the gravitational force is reproduced. In the Newtonian regime we show that the entropy (I_m) becomes entanglement entropy and the level-spacing ω is given by the Unruh temperature of the non inertial frame. Hence we conclude that our proposal reproduces the previous hypothesis on the entropic origin of inertia [11]. In the Newtonian regime, the principle behind the inertia can be summarized as the invariance of entanglement structure of the underlying state with respect to the local deformations of the macroscopic slow variable.

To sum up, we will study the conceptual implications of what we developed so far

[1]Volume entropy introduced here should not be confused with the idea introduced in [73]. In our context it is an entropic measure on the volume of the phase space and it follows area law in spacetime.

to the emergent gravity paradigm. We will elaborate some of the ideas presented in [11] and clarify some of the misunderstandings that originated thereafter. We will put forward the adiabatic principle as the underlying microscopic mechanism through which inertia emerges. Our discussion will be presented in a covariant fashion when it comes to general relativity. We will enquire the following conceptual problems.

- What is the microscopic principle behind inertia?
- What is the information theoretic notion that amounts to the integral of $\boldsymbol{\chi}_t$ (5.3.10) on general surfaces? Is it the change of entropy or total energy?
- What is adiabatic principle, adiabatic invariance, and adiabatic reaction force and what distinguishes it from an entropic one?

We will answer these questions in general relativity. For us general relativity is the true theory in its domain of validity and everything proposed here should be demonstrated in it. We think it is most beneficial to start by explaining concept of adiabatic reaction force.

7.2 Adiabatic Principle

Adiabatic principle is one of the fundamental principles of nature that reflect itself in many different branches of physics. In essence it is about the fact that when two systems which are separated by a gap in the timescale hierarchy then these two systems can not efficiently communicate. In other words, a system adjust itself to an affect that is slow compared to the timescale of the system. This readjustment leaves some relevant attributes of the system invariant approximately. The adiabatic principle that will be considered here follows from its quantum mechanical consideration. We state it as follows:

A physical system stays in its instantaneous eigenstate when it is deformed slowly with respect to the characteristic time scale of the system.

As a result of the coupling of two system having distinct characteristic time scales, adiabatic reaction forces arise. Let us refer these two systems as *slow* and *fast* systems according to the characteristic time scales of their dynamics. In the adiabatic averaging approximation, the energy of the fast system calculated for frozen slow variables, acts as a potential in which the slow system moves. Coupling of two systems manifests itself in the fast system as the parametric dependence of its Hamiltonian on the position of the slow system. Due to the separation of time scales, the fast system remains in same energy eigenstate with respect to the deformations of the slow variable. In the limit of infinitely slow variations, this fact can be used to determine the reaction force entirely from the principle that the phase space below an energy eigenstate remains constant. The quantities that remain approximately constant are called *adiabatic invariants*. Entropic force is also a form of adiabatic reaction force where the entropy of the system serves as the adiabatic invariant. In general the volume of the phase space of the fast system becomes the adiabatic invariant.

We will discuss the adiabatic principle and the reaction forces deriving from it on systems having ergodic behaviour. The main purpose of this section is to illustrate the connection of the principle of adiabatic invariance to emergence of inertial forces through a first law type relation.

Adiabatic reaction force in an ergodic system

In this section we will discuss the generalization of adiabatic reaction forces to many degrees of freedom N. The special case is where degrees of freedom are decoupled and system is integrable. In this case one has N adiabatic invariants and total phase space is the sum of phase spaces of individual d.o.f. This kind of a system is trivial extension of the adiabatic principle for a single harmonic oscillator.

We will consider the generalization of the adiabatic principle to a fast ergodic system that consists of many coupled degrees of freedom. Slow system will be coupled to microcanonical ensemble of the fast one.

Let us consider that slow system is coupled to the fast system through it is position on the space denoted by $\boldsymbol{X}^i$ where i runs in $\{1, 2...(D-1)\}$. We denote the $2N$ dimensional phase space of the fast system by the parameters (q^a, p^a) where a runs in $\{1, 2...N\}$. Consider the Hamiltonian of the fast system parametrically depends on the slow variable $\boldsymbol{X}^i$.

$$H_{\text{fast}} \equiv H(p^a, q^a; \boldsymbol{X}^i) \tag{7.2.1}$$

Since adiabatic affection of the system takes place very slowly one can study its affects on the distribution function of the phase space perturbatively. The energy eigenvalues below a given adiabatically varying eigenstate becomes constant, at least to a very good approximation. This statement is a consequence of the fact that under very slow variations the energy eigenvalues do not cross. The statement becomes exact in the limit of infinitely slow variations. According to the Bohr-Sommerfeld quantization rule, the number of states below an energy eigenstate is given by the volume of the phase space contained inside the corresponding energy surface measured in units of $\hbar^N$. Given the fact that this number stays constant, we arrive at the following adiabatic invariant for ergodic systems [161]

$$\Omega(E, \boldsymbol{X}) = \frac{1}{(2\pi\hbar)^N} \int ... \int dp^N dq^N \ \Theta(E(\boldsymbol{X}) - H(p, q; \boldsymbol{X})) \tag{7.2.2}$$

where Θ is the Heaviside step function. This is the generalization of the action variable (1.1.5) to ergodic systems. It counts the number of energy eigenvalues below a certain value. Although we have emphasized that we are in the microcanonical ensemble this quantity does not count the number of possible states at a certain energy rather all the levels below a certain energy. The microcanonical form of the ensemble will become clear in the definition of the expectation values of the observables and density of states. Using the expression of adiabatic invariant one can extend the definition of action variable to ergodic case. This is exact analog of thermodynamical entropy S_T yet we prefer to call it volume entropy S which makes it clear that it is the entropic measure of the volume of the phase space under a certain energy level. This measure of entropy is also known as the Gibbs entropy in the microcanonical ensemble [8].

$$S = \log \Omega(E, \boldsymbol{X}) \tag{7.2.3}$$

This entropy will be used as an adiabatic invariant. We will show that gravitational analog of the entropy is the microcanonical action. One can derive the first law of adiabatic principle via variation of the slow variable,

$$dE = TdS + F_i d\boldsymbol{X}^i \tag{7.2.4}$$

where one can express the force in the usual way as the gradient of the energy.

$$F_i = -\partial_{\boldsymbol{X}} E(S, \boldsymbol{X})_S \tag{7.2.5}$$

where subscript S denotes the stationarity of the entropy. This is the adiabatic reaction force in a system of N degree of freedom at the leading order. What is denoted as temperature should be considered a measure of the level spacing of the energy levels, and describes the response of the energy under small changes in the phase space volume. However, it is important to note that there is no actual change of phase space volume in an adiabatic process. One can obtain further relations from (7.2.4) that relates the entropy and the force. This is typically the notion of entropic force,

$$F_i = \partial_{\boldsymbol{X}} S(E, \boldsymbol{X})|_E \qquad \frac{1}{T_V} = \partial_E S(E, \boldsymbol{X})|_x. \tag{7.2.6}$$

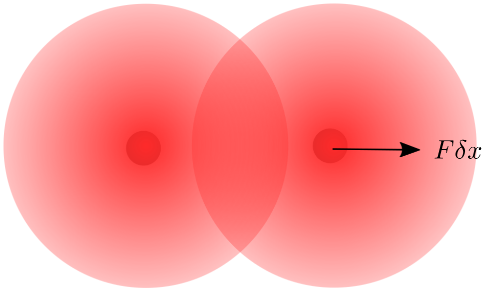

Figure 7.1: *Influence of the nuclei on the phase space of the electrons causes the molecular force between each molecule known as Bohr-Oppenheimer force. Virtual replacement of a molecule encounters a reaction force that emerges due to invariance of the phase space volume in an adiabatic deformation.*

Contrary to the common understanding it does not lead to an increase in the entropy rather energy of the system adjust itself such that entropy is invariant.

The expression (7.2.5) is also known as Born-Oppenheimer force 7.1. The Born-Oppenheimer approximation can be turned into a systematic approach by parametrizing the separation of time scales between the slow and the fast system using a small parameter. In leading order the quantum state of the total system can be factorized into a product of the energy eigenstate of the fast system with the state of the slow system. In this regime the dynamics of the slow system can be described by an effective Hamiltonian that includes a potential term equal to the energy eigenvalue of the fast system. The resulting adiabatic reaction force is thus equal to the gradient of this energy eigenvalue with respect to the slow variable and is known as the Born-Oppenheimer force.

An important aspect of the Born-Oppenheimer approximation is that as long as the time scales are widely separated the slow system can be treated fully quantum mechanically and maintains all its quantum properties, even though it is coupled to a system with many degrees of freedom.

Let us close this section on the discussion of why the ensemble of the fast system is considered to be microcanonical even though phase space of the system is calculated by counting the states below a certain energy rather than near that energy. Any quantity derived from S necessarily involves differentiation with respect to E, $\boldsymbol{X}$ or possibly some other variable and therefore is given by an expression defined only on the energy surface [162]. Moreover density of states by which one can calculate the ensemble average is defined in the following way,

$$\rho_0(p,q) = \frac{\delta\big(E(X) - H(p,q;\boldsymbol{X})\big)}{\partial_E \Omega(E,\boldsymbol{X})}. \tag{7.2.7}$$

This is the density of states frozen with respect to the slow variable. Also it represent the zeroth order density of states in the pertubative analysis of adiabatic modification of the system. This is the measure where Born-Oppenheimer force is defined as an ensemble average.

$$\vec{F} = -\langle \partial_{\boldsymbol{X}} H(p, q; \boldsymbol{X}) \rangle \tag{7.2.8}$$

where

$$\langle .... \rangle = \int dp^N dq^N \; .... \; \rho_0 \tag{7.2.9}$$

This is the standard definition of the average in the micro-canonical ensemble as used in textbooks on statistical mechanics. Note that for an ergodic system the time average of any physical quantity is given by the average over the energy surface.

To sum up, it is explain that, adiabatic principle yields similar expressions to the first law of thermodynamics, yet their origin is different. It is also emphasized that the emergent force is due to existing of an adiabatic invariant whose invariance causes a reaction force to the deformations by the slow dynamics.

7.3 Inertia as an Adiabatic Reaction Force

In the previous section we have discussed the deformation of a fast system by a slow variable and emerging reaction force exerted by the fast system on the slow one. In other words the fast system builds a potential for slow system. In essence this aligns with the current understanding of emergent gravity, where the moduli space of the underlying gauge theory corresponds to emergent space. In other words, the integrated out matrix degrees of freedom builds a potential which becomes the spacetime for the probe. The setups where one can carry out exact calculations starting from the gauge theories are mostly supersymmetric. There are also attempts to go beyond supersymmetry [163]. We will also assume that the underlying fast system is a $U(N)$ gauge theory, yet we will not refer to the details of it. It is also not necessary from the point of adiabatic reaction forces, since it can be derived independent of the details of the fast dynamics.

In this section, using the tools introduced in the first chapter we will demonstrate the adiabatic first law in gravity by means of generic surfaces on spacetime. We will elaborate on the adiabatic invariant in gravitational systems and show its connection to microcanonical action in gravity. We will show that the integral of $\boldsymbol{\chi_t}$ on spacelike surfaces measures the change in the phase space of the microcanonical ensemble (7.2.2). As a consequence of the area law in gravity, change of phase space includes terms proportional to area.

The general philosophy we adhere to is that any geometric solution $\mathcal{M}$ in a well defined semi-classical gravitational theory corresponds to a state $|\mathcal{M}\rangle$ in thermodynamic limit of ordinary quantum mechanical systems. The geometric solution $\mathcal{M}$ is the emergent potential derived by integrating out the fast degrees of freedom of the corresponding state $|\mathcal{M}\rangle$. On this quantum mechanical state we introduce a probe that takes the role of a slow parameter. Although adiabatic approximation preserves the coherence of a quantum mechanical probe, for all practical purposes, the classical probe will suffice.

The underlying state $|\mathcal{M}\rangle$ is stationary in the decoupling limit of the fast-slow systems. Hence the geometric description of it corresponds to a stationary solution of a gravitational theory possessing a timelike Killing field t^{μ}[2]. Although $|\mathcal{M}\rangle$ has an exact timelike Killing symmetry without the slow variable,(or in the leading order of the adiabatic approximation) the symmetry becomes an approximate notion due to the coupling of the underlying Hamiltonian to the slow parameter (7.2.1). The corresponding statement in the underlying quantum mechanical system is,

$$\rho = \rho_0 + \sum_{n=1}^{\infty} \lambda^n \rho_{(n)}, \qquad\qquad [\hat{H}_{\text{fast}}, \rho_0] = 0 \tag{7.3.1}$$

where ρ is the density matrix of the fast system in the adiabatic approximation. λ is the parameter of the approximation and should be on the order of the ratio of the characteristic scales. The invariant phase space is measured with respect to ρ_0.

The adiabatic first law will be demonstrated in the emergent geometric setup but the principle behind is a microscopic one. The starting point is the fundamental theorem of covariant phase space formulation (5.3.10). We consider the snapshot of the $|\mathcal{M}\rangle$ at a point in time $t =$ constant that corresponds to a constant time hypersurface Σ_t in the geometric description. Energy of the state is measured from the point of infinity and defined by the ADM mass of the solution. On this solution we consider a holographic screen, a codimension two surface in $\mathcal{M}$, which captures the microscopic information of the state partitioned according to holographic surface $\Sigma = \Sigma_{in} \cup \Sigma_{out}$ and associated subalgebras. We prefer to define holographic screen through the subalgebra associated to Σ_{in} since factorization of Hilbert space in quantum gravity is problematic [164, 165].

According to the holographic paradigm, the subalgebra associated to Σ_{in} can be encoded on the surface S_n that separates Σ_{in} and Σ_{out}. The entanglement

[2] Although we have carried out explicit computations particularly for general relativity in the first part, our arguments apply to higher derivative theories as well.

entropy is also defined through the subalgebra without explicitly referring to the factorization of the state. The dimensional reduction (holographic principle) is the consequence of the area law entanglement in the underlying state $|\mathcal{M}\rangle$. Let us consider a third surface surrounding the probe Σ_p. The fundamental theorem of covariant phase space formulation implies,

$$\int_{\Sigma_{out}/\Sigma_p} \boldsymbol{\omega}(\phi, \delta\phi, \delta_t\phi) = \oint_{S_n} \boldsymbol{\chi}_t + \oint_{S_p} \boldsymbol{\chi}_t - \oint_{S_\infty} \boldsymbol{\chi}_t \tag{7.3.2}$$

where Σ_{out}/Σ_p indicates the region Σ_{out} excluding the region of the probe Σ_p. The location of the probe corresponds to the slow variable explained in (7.2). Perturbation on the underlying state is caused by the change in the slow variable $\delta\phi \equiv \delta_{\boldsymbol{X}}\phi$. The presymplectic current vanishes since $\delta_t\phi = 0$. The integral around infinity amounts to a change in the energy of system δM_{ADM} in the form of ADM mass.

$$\delta M_{\text{ADM}} = \oint_{S_\infty} \boldsymbol{\chi}_t \tag{7.3.3}$$

We first replace the surface integral around the probe by a volume integral using the Stokes theorem. Then assuming that the initial geometry satisfies field equations, we are left with the linearized field equations in (6.B.13):

$$d\boldsymbol{\chi}_t = -2t^\mu(\delta_{\boldsymbol{X}} E_{\mu\nu})d\Sigma^\nu \tag{7.3.4}$$

where $d\Sigma^\nu$ is the volume element of Σ_p. We will replace the linearized equations for the gravitational field by the change in the location of the probe. The perturbation is due to a virtual displacement of the location of the probe

$$d\boldsymbol{\chi}_t = -t^\mu(\delta_{\boldsymbol{X}} T_{\mu\nu})d\Sigma^\nu. \tag{7.3.5}$$

Without loss of generality the form of stress energy tensor is assumed to be $T^{\mu\nu} = \rho u^\mu u^\nu + t^{\mu\nu}$. There are two equivalent ways to describe the situation. In the first case, we can assume a Dirac delta distribution for the probe $\rho(x) \equiv \rho\delta^{D-1}(x^i - \boldsymbol{X}^i)$. In this snapshot of the $D-1$ dimensional hypersurface, one can vary the location of the probe $\boldsymbol{X}'^i = \boldsymbol{X}^i + \delta\boldsymbol{X}^i$ then show that integral around the probe amounts to the force term.

We will demonstrate the force term in a different way which reflects the fictitious character of the deformation more naturally. Suppose one just starts with the background geometry. In this background we first introduce the probe in the location $\boldsymbol{X}^i$ which causes a change in the total energy of the system δM_{ADM}. We consider another perturbation on the initial unperturbed geometry by introducing the source to the location $\boldsymbol{X}^i + \delta\boldsymbol{X}^i$ which yields a change in the total energy of

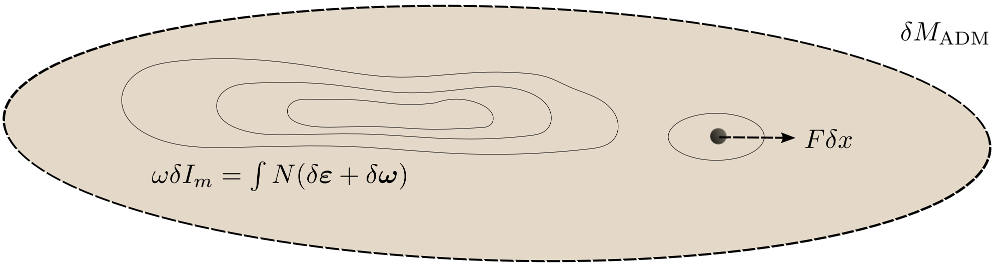

Figure 7.2: *Adiabatic principle in gravity. The equation on the left is interpreted as the first law of deformation in spacetime. The invariance of the phase space with respect to deformations that does not become part of the system, causes reaction forces on the probe. The adiabatic principle in gravitation is formulated as* $\delta M_{\text{ADM}} = \omega\delta I_m - F\delta x$.

the system $\delta M'_{\text{ADM}}$. The difference of energies of two perturbation amounts to virtual change in the location of the probe

$$\oint_{S_p} \boldsymbol{\chi}_t = -\left(\int_{\Sigma_p} \rho\,(\nabla_\mu N) u_\nu d\Sigma^\nu\right) \delta \boldsymbol{X}^\mu = -F^g_\mu \delta \boldsymbol{X}^\mu. \tag{7.3.6}$$

Note that $d\Sigma^\nu = u^\mu dV$ and $u \cdot u = -1$. The additional minus sign appears in the definition of reaction force. Hence we end up with following relation,

$$\delta_{\boldsymbol{X}} M_{\text{ADM}} = -\omega\delta_{\boldsymbol{X}} I_m - F^g_\mu \delta \boldsymbol{X}^\mu \tag{7.3.7}$$

where we defined $\omega\delta_{\boldsymbol{X}} I_m \equiv \oint_{S_n} \boldsymbol{\chi}_t$ and I_m stands for the microcanonical action attributed to the region Σ_{in}. The connection to the microcanonical action will be elaborated in the next section. The (7.3.7) is interpreted as follows: invariance of the microcanonical action against deformations of the probe that do not become part of the phase space associated to Σ_{in} causes a change in the energy of the underlying state such that a reaction force on the probe emerges. The entropic origin is also hidden in this statement *i.e.* it is the invariance of the entropy of the volume (7.2.3) of the phase space that causes reaction force.

Let us continue the discussion by the following thought experiment on black holes to observe the invariance of the phase space against perturbations that are not merged into the system or cross the surface S_p.

Physical process version of the first law and invariance of the phase space

First law of black hole mechanics [22] when considered for a Schwarzschild black hole states that the change in the entropy of the black hole is proportional to change in the energy of the system.

$$T\delta S_{BH} = \delta M_{\text{ADM}} \tag{7.3.8}$$

This is a simple yet profound relation that yields the thermodynamic nature of the black holes. The natural way to understand this relation is considering two black hole solutions such that their mass is infinitesimally close to each other. These two solutions have different horizon areas again infinitesimally close to each other. The variation on the mass and on the area of the horizon are related to each other according to the first law.

On the other hand, one can demonstrate the first law by lowering mass and finally dropping it into the black hole [166]. The picture provided in such a thought experiment is giving more information about what happens during the process. Let us briefly go over this thought experiment. For our purposes we will simplify the experiment by considering Schwarzschild black holes. In the same way the box dropped into the black hole will have no electric charge.

One starts the experiment by introducing a probe having mass m which will be later dropped into the black hole having mass M_{ADM} and entropy S_{BH}. When a probe is introduced into the system, the total energy of the system changes by an amount of the energy of the probe. The increase in the energy of the system depends at which point of the geometry the mass is introduced, simply because ADM mass is measured at infinity. The increase in the ADM mass in terms of the mass m is given by Nm. At the point of time where the mass is introduced to the system, the total energy of the system changes by an amount $\delta M(= Nm)$. It would be absurd to think that the black hole entropy changes at the moment the mass m is introduced into the system. It would violate causality. Hence in the beginning there is an increase in the energy of the system, yet the black hole entropy stays the same.

Again by applying Stokes theorem on the initial hypersurface Σ, it becomes obvious that $\int_{S_{\mathcal{H}}} \boldsymbol{\chi} = 0$ and $\delta M_{\text{ADM}} = \int_{\Sigma} d\boldsymbol{\chi}$. Using the closedness of $d\boldsymbol{\chi}$ one can replace the integral over the spacelike hypersurface Σ with the null surface on the black hole horizon, $\int_{\Sigma} d\boldsymbol{\chi} = \int_{\mathcal{H}} d\boldsymbol{\chi}$. Assuming that all the matter eventually will falls into the black hole, one can show that $\int_{\mathcal{H}} d\boldsymbol{\chi} \sim \kappa\delta A$ using the Raychaudhuri equation.

Rather than the process itself, we are interested in the point where the probe mass is introduced into the system. It is easy to see that black hole entropy does not change on this hypersurface. More importantly integral $\boldsymbol{\chi}$ vanishes on any closed surface between the horizon $S_{\mathcal{H}}$ and infinity which doesn't include the probe. The moment probe passes through a holographic surface, one expects an increase on volume of the phase space equal to an amount of Nm, this is an example of non adiabatic transition. As we will show in the next section, the change in the entropy is measured by the change in the microcanonical action of gravity.

7.4 Microcanonical Action as the Adiabatic Invariant

Information about the observables in a quantum field theory can be extracted from the Euclidean generating functional

$$Z(\hbar, J) = \int \mathcal{D}\phi \, e^{-1/\hbar \, (S[\phi] + \int J \cdot \phi)}. \tag{7.4.1}$$

The form of the generating functional manifestly covariant. The generating functional can be interpreted as a partition function with the following correspondence,

$$\hbar \leftrightarrow T, \qquad\qquad \text{Energy} \leftrightarrow S[\phi]. \tag{7.4.2}$$

The canonical ensemble of the statistical physics studies systems under fixed temperatures. It is not most general or fundamental formulation of the statistical mechanics. Rather it can be derived from the microcanonical formulation. Fundamental postulates of statistical mechanics, ergodicity and conservation of energy manifest itself naturally in the microcanonical formulation. In renormalizable quantum field theories canonical and microcanonical formulations are equivalent to each other [167]. However distinction becomes more important for the systems that does not have any temperature. Another distinction between the microcanonical description of ordinary quantum field theories and gravity is that fixing the energy in a quantum field theory corresponds to restricting the integral of the Hamiltonian on the entire manifold while in gravity this can be done simply by restricting the boundary data. Hence microcanonical ensemble in gravity corresponds to fixing the energy flux on the boundary in the form of quasilocal energy density and angular momentum [168]. The change of boundary data amounts to a Legendre transformation between the energy density and the inverse temperature, which are thermodynamically conjugate variables.

In this section we will comment on the the physical interpretation of the surface charge $\int \boldsymbol{\chi}_t$ from a microscopic point of view. We interpret that its invariance leads to the adiabatic reaction forces. In section 7.2 adiabatic invariant amounts to the volume of the phase space in microcanonical ensemble. We will now show that the adiabatic invariant in gravity is also connected to the volume of the phase space in microcanonical ensemble. Let us start by reminding ourself the expression of $\boldsymbol{\chi_t}$ in terms of the change in The Noether charge $\boldsymbol{Q}_t$ and symplectic potential

$\boldsymbol{\Theta}$.

$$\int_S \boldsymbol{\chi}_t = \int_S \delta \boldsymbol{Q}_t - t \cdot \boldsymbol{\Theta} \tag{7.4.3}$$

The symplectic potential appears as a boundary term in the variation of the Lagrangian (2.2.1). One can add additional boundary terms that corresponds to changing the ensemble where the observables are calculated in. The boundary term that transform the ensemble to microcanonical is the Noether charge

$$I_m = \int_{\mathcal{M}} \boldsymbol{L} - \int_{\partial \mathcal{M}} dt \wedge \boldsymbol{Q}_t. \tag{7.4.4}$$

This is the form of microcanonical action derived by Brown and York [169] and elaborated by Wald within the covariant phase space formalism [65]. The variation of this action yields the form $\boldsymbol{\chi}_t$ as a boundary term

$$\delta I_m = \int_{\mathcal{M}} \delta \boldsymbol{E}^{ab} \delta g_{ab} - \int_{\partial \mathcal{M}} dt \wedge (\, \delta \boldsymbol{Q}_t - t \cdot \boldsymbol{\Theta}) \tag{7.4.5}$$

$$= \int_{\mathcal{M}} \delta \boldsymbol{E}^{ab} \delta g_{ab} - \int_{\partial \mathcal{M}} dt \wedge \boldsymbol{\chi}_t \tag{7.4.6}$$

Fixing this data on the boundary corresponds to fixing the total energy of the system. The relative minus sign is due to the direction of the boundary integral. While we have considered the integral of $\boldsymbol{\chi}_t$ on the surfaces with outward directed surface normal n_μ, the surface normal of the boundary is defined on the opposite direction (7.3.2).

When the equations of motion holds for a region of spacetime, the change in the action that is associated to the same region, becomes equal to the change in the boundary data in the microcanonical ensemble. For a stationary solution the boundary integral along the time axis can be carried out. We assume that there is a characteristic time scale for the underlying fast system where change in the action due to the deformations of the slow parameter is localized within (7.3.7) this scale [3]. Hence we can safely assume that the contribution to the integral will come within this range of time in the case of an adiabatic affection of the system. Therefore the adiabatic invariant in the first law (7.3.7) is identified with the microcanonical action in gravity

$$\omega \, \delta I_m = - \int_{S(n)} \boldsymbol{\chi}_t \tag{7.4.7}$$

[3] For thermal systems, time coordinate becomes periodic in inverse tempreture, β to prevent conical singularities near the origin. In this case volume of the phase space becomes finite dimensional and proportional to inverse tempreture.

Frequency or inverse level spacing is defined as the inverse characteristic time scale of the system $\omega = \frac{1}{T_{\text{char.}}}$. This is also the time scale for the system to adjust itself to a non adiabatic transition. In terms of the microscopics, we propose that microcanonical action counts the volume entropy of Gibbs. Therefore even there is no degeneracy for a particular macrostate, the entropy does not vanish as it counts the number of states below a certain energy eigenstate. We propose that,

$$\delta S \equiv -\delta I_m \qquad\qquad T_V = \omega \tag{7.4.8}$$

With this identification the adiabatic first law (7.3.7) becomes,

$$\delta_{\boldsymbol{X}} E = T_V(\delta_{\boldsymbol{X}} S) - F^g_\mu \delta \boldsymbol{X}^\mu \tag{7.4.9}$$

7.5 Newtonian limit

It is an interesting exercise to study what $\boldsymbol{\chi}_t$ yields on generic surfaces in the Newtonian limit of general relativity. Newtonian limit is interesting in the sense that one can see the backreaction of the probe on the geometry. Therefore when the energy of the system is conserved, namely the probe is moved by the system rather then an external virtual affection of the system, we should be able to recognize force expression on general surfaces. In other words when $\delta E = 0$, the change in the phase space is equal to the derived force $\omega \delta I_m = F \delta x$. This was also how the entropic force as the origin of inertia is first presented [11].

We use the following solution as the Newtonian limit of general relativity in D dimensions

$$ds^2 = -\left(1 + \frac{2\Phi}{c^2}\right) dt^2 + \left(1 - \frac{2\Phi/c^2}{D-3}\right) d\vec{x}^2. \tag{7.5.1}$$

Note that the solution does not need to be spherically symmetric. We will exponentiate the solution to simplify the calculation. This is only for computational convenience, by no means we demand exponentiated metric to be a solution of general relativity

$$ds^2 = -e^{2\Phi/c^2} dt^2 + e^{-2\Phi/(D-3)c^2} d\vec{x}^2. \tag{7.5.2}$$

The above ansatz reduces to the Newtonian limit of general relativity in the leading order and that is the only requirement. Spatial section of the metric can be expressed in the following form

$$h_{ij} = e^{2\sigma} \eta_{ij} \tag{7.5.3}$$

where $\sigma = -\frac{\Phi/c^2}{D-3}$ and η_{ij} is flat metric. We will make use of the fact that h_{ij} is related to the flat space through a Weyl transformation. We will plug this solution into the expressions (6.2.5),

$$\delta E = \frac{c^4}{8\pi G} \oint \kappa\, \delta dA - N(\delta k + k^{ij}\delta\sigma_{ij})dA - F\delta x \tag{7.5.4}$$

The surface charges of the metric (7.5.1) is related to the one defined in the flat space in the following way,

$$\tilde{\sigma}_{ij} = e^{2\sigma}\sigma_{ij} \tag{7.5.5}$$

$$\tilde{k}_{ij} = e^{\sigma}\left(k_{ij} + \sigma_{ij}\partial_\eta\sigma\right) \tag{7.5.6}$$

$$\tilde{k} = e^{-\sigma}\left(k + (D-2)\partial_\eta\sigma\right). \tag{7.5.7}$$

σ_{ij} and k_{ij} are the first and second fundamental forms of $(D-2)$ surface embedded into flat space which serves as a reference manifold. $\partial_\eta = \eta^i\partial_i$ is the normal derivative of codimension two surface in flat space. ∂_η will be fixed under deformations. The problem we had encountered in the fully relativistic theory namely how to identify deformed manifold with the unperturbed one does not exist in the Newtonian limit. One can use the flat space as the reference manifold and only vary Newtonian manifold. Therefore variation is only on the Newtonian potential Φ and not on σ_{ij}, k_{ij} or any quantity defined with respect to the flat geometry η_{ij}. The philosophy we follow here is similar to the derivation of the adiabatic first law of inertia (7.3.7). We perturbed the initial manifold in two different ways and compare these two different configuration. Newtonian solution is already a perturbation therefore changing the potential in the base solution is equivalent, in the leading order, to compare it with a different configuration of the potential. The final expression will be given in terms of the geometry of the initial Newtonian metric, not in terms of the quantities defined on the flat reference manifold. We will drop all the factors of $1/c^2$ and $8\pi G_{\mathrm{N}}$ for now and put them back at the end. Based on this setup we will give the expressions for $\delta Q - t\cdot\Theta$ using the general expression,

$$\delta Q = e^{-\sigma}(\partial_\eta N)\underbrace{(D-2)\delta\sigma dA}_{\delta dA} = (\partial_\eta\Phi)(D-2)\delta\sigma dA_\eta \tag{7.5.8}$$

$$t\cdot\Theta = Ne^{-\sigma}(D-2)\partial_\eta(\delta\sigma)\, dA = \partial_\eta(\delta\sigma)\, dA_\eta \tag{7.5.9}$$

where $N = e^{\Phi/c^2}$ and all the subscripts η indicates that the quantity is calculated with respect to the reference flat manifold. These expressions should be given in terms of the quantities of unperturbed space which is not flat and related to the flat space through a Weyl transformation (7.5.5). We will give the expressions in terms of $\partial_n = e^{-\sigma}\partial_\eta$ which is the normal derivative in the undeformed solution

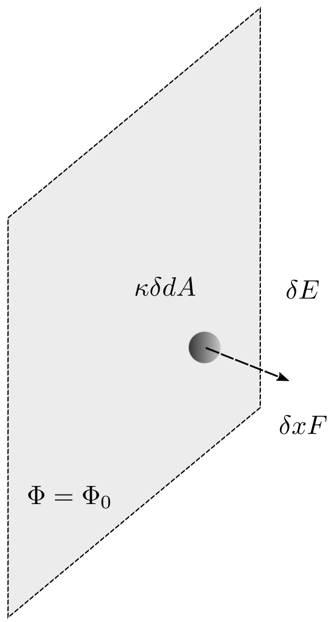

Figure 7.3: *In the Newtonian limit, entropic measure of the microcanonical phase space is identified with entanglement entropy of the underlying state. Invariance of the entanglement structure of the underlying theory causes the emergent reaction force,* $\delta E = \int \kappa \delta dS_{\text{ent}} - F_i \delta x^i$.

(7.5.2) and the area element of the surface embedded in it, which is denoted by dA.

$$\delta Q = e^{\Phi}(\partial_n \Phi) \underbrace{\left(-\frac{D-2}{D-3}\right) \delta\Phi dA}_{\delta dA} \tag{7.5.10}$$

$$t \cdot \Theta = e^{\Phi} \left(-\frac{D-2}{D-3}\right) \partial_n \delta\Phi \, dA \tag{7.5.11}$$

Change in the energy of the system is given by

$$\delta E = \frac{c^4}{16\pi G} \oint \delta Q - t \cdot \Theta - F\delta x = \frac{c^2}{8\pi G} \oint e^{\Phi}(\partial_n \Phi)\delta dA + e^{\Phi} \left(\frac{D-2}{D-3}\right) \partial_n \delta\Phi \, dA. \tag{7.5.12}$$

If we expand e^{Φ} in powers of $1/c^2$, we see order by order contribution to the change of energy. Note that change of area implicitly at the order $1/c^2$ (7.5.10), so that term does not appear in the leading order,

$$\delta E = \frac{c^2}{8\pi G} \oint \left(\frac{D-2}{D-3}\right) \partial_n \delta\Phi \, dA + \frac{1}{8\pi G} \oint (\partial_n \Phi) \delta dA + \left(\frac{D-2}{D-3}\right) \Phi \, \partial_n \delta\Phi \, dA - F\delta x. \tag{7.5.13}$$

The first term on the right hand side is Gauss's law for gravity (or integral version of the Poisson equation), that measures the change of energy due to addition of gravitational charges inside the surface. The second one corresponds to work done by the force on a test mass, m outside of the region. One can use electrostatic analogy to observe that this term indeed gives the gravitational force in the most general case presented in three dimensions,

$$F_i \delta x^i = \frac{-1}{4\pi G_{\mathrm{N}}} \oint \left(\delta_x \Phi \partial_n \Phi - \Phi \partial_n \delta_x \Phi\right) \, dA_n \tag{7.5.14}$$

In other words, one can redistribute the entire gravitational charge contained inside the region over the surface without changing the forces on the particles. Considering variations on the potential caused by changing the location of the probe (the change of the potential that test mass feels due to virtual change in its location) which can be obtained from the Green's function for the Laplacian, we obtain,

$$F_i \delta x^i = \frac{1}{8\pi G_{\mathrm{N}}} \oint (\partial_n \Phi) \delta dA + \left(\frac{D-2}{D-3}\right) \Phi \, \partial_n \delta\Phi \, dA. \tag{7.5.15}$$

This is the generalized force expression expressed as a surface integral. It is interesting that in the Newtonian limit we end up with an expression having change of area due to fictitious relocation of a probe. The connection between fixed spacetime and dynamical one is made through the weak field limit of general relativity (7.5.1) which open us a window to understand the underlying physics.

Although the surface expression $\omega \delta I_m$ (which is shown to be proportional to the change in the volume of phase space) looks like changing this is an illusion due to conservation of energy. It would change if we let system evolve and let the gravitational waves propagate through the surface and let the probe pass the surface. But remember we were moving the probe fictitiously on a frozen moment of time. In other words the whole setup is at a constant time slice and on this time slice we are probing the hamiltonian virtually. Indeed the surface expression (7.5.15) is invariant under fictitious deformations. It is the change of energy that corresponds to work done by the force. In other words we use the relation between action and slow variable by keeping the energy fixed (1.1.8).

7.5.1 Microscopics of the Newtonian regime

It is explained that when the system is deformed fictitiously at a constant time [4], the phase space of the underlying dynamics stays invariant. This invariance reflected itself as a reaction force that one can calculate simply by $F = -\partial_x E(x)|_{I_m}$. In general when system continues to evolve, the probe will pass the surface and non adiabatic transitions will occur. Therefore general relation that also governs the non adiabatic transitions, manifests itself in the form of a first law,

$$\delta E = -\omega\delta I_m - F\delta x \tag{7.5.16}$$

The entropic version of the force $F = T_V \partial_x S(x)|_E$ can be interpreted through the conservation of energy in a closed system in contrast to the manipulation of the system from outside by fictitious affections.

Let us look at a special case where the $D-1$ dimensional manifold is foliated by equipotential surfaces. This is a natural foliation since Newtonian potential naturally determines the energy scale with respect to the infinity where energy of the entire system is well defined. When the surface is chosen to be an equipotential one $\Phi = \Phi_0$, the Φ in the second term in (7.5.15) can be pulled out of the integral and since $\delta\Phi$ is sourced by only the particles outside the screen, the remaining integral just gives zero. Let us study the result from a microscopic perspective, where the change amounts to a change in the phase space having energy levels ω

$$\omega\delta I_m = \frac{-1}{8\pi G_\mathrm{N}} \oint (\partial_n \Phi)\, \delta dA \tag{7.5.17}$$

In the Newtonian limit one can localize this identity by dividing surface into area elements. This sort of localization is not unique in general relativity due to general coordinate invariance, as explained in section 6.4. The changes in phase space of the underlying theory in the Newtonian limit becomes solely determined by the area of the holographic screen. We propose the following identification,

$$\omega = \frac{\partial_n \Phi}{2\pi} \qquad\qquad \delta dN = \delta \frac{dA}{4G_\mathrm{N}}. \tag{7.5.18}$$

To align our language with [11], we replaced $-I_m = S = N$. Frequency of the harmonic oscillators on the holographic surface is given by the Unruh temperature associated to the screen, $\omega = T_\mathrm{U}$. The change in the phase space can also be expressed in terms of the potential of the probe (7.5.10),

$$\delta dN = -\frac{D-2}{D-3}\delta\Phi dN \tag{7.5.19}$$

[4] One can also consider a deformation which is not taking place at a constant time slice but happens in a much slower than the time scale of the fast dynamics.

Negative sign indicates that when probe passes the holographic screen, there is an increase in the volume of the phase space. Newton potential keeps track of the changes of information per unit bit. The equation (7.5.19) together with (7.5.18) indicates that the change in the phase space that is equal to the area is proportional to change in the Newtonian potential of the surface.

7.6 Conclusion and Discussion

In this chapter we revisited the proposal that inertial forces have entropic origin [11]. We reconsidered the proposal from a relativistic point of view using covariant phase space formalism. The idea of entropic force is elaborated and generalized as adiabatic reaction force which follows from the principle of adiabatic invariance. We also proposed that on a general spacelike surface, volume of the phase space is the correct notion corresponding to adiabatic invariant. We argue that in a geometric theory of gravity, change in the entropy of the volume of the phase is measured by the microcanonical action which agrees with the general definition of the volume entropy as the measure a of volume of a phase space in microcanonical ensemble. We provided reasons to distinguish it from thermal entropy. Let us summarize our conclusions on this chapter:

- **Adiabatic first law:** In this chapter we clarify the distinction between the first law that corresponds to fluctuations of gravitational field and the first law type relation that emerges due to deformations caused by a probe. The former is studied in the previous chapter 6. We start by introducing adiabatic first law in an ergodic system 7.2. The principle relates energy, force and volume of the phase space. The realization of the adiabatic first law in gravity is demonstrated in section 7.3. We clarified some of the misunderstandings that originated thereafter the entropic gravity proposal. In particularly, the concept of reaction force makes it clear that it is not the change of entropy that causes inertial forces, rather it is the invariance of entropy that requires energy to adjust itself such that a reaction force emerges as gradient of the energy. Because of the distinction between locally vacuum spaces and thermalized regions we proposed the notion of phase space volume entropy. It counts the number of energy eigenvalues below a certain state. We have proposed that change in the entropy is measured by the microcanonical action in a geometric theory of gravity. Our interpretations apply also to higher derivative theories of gravity.

- **Microcanonical action is entanglement entropy in the Newtonian limit:** In the Newtonian limit, on equipotential surfaces, the microcanonical action that measures volume of the phase space δI_m becomes change of area of the surface. In this limit, one can localize the quasilocal expressions by referencing to the flat manifold. In this local form of the expression we deduce that level-spacing between energy levels corresponds to Unruh temperature. Moreover, we have derived Gauss's law and well known surface expression of the gravitational force as the Newtonian limit of Brown-York quasilocal charges.

- **Adiabatic invariance and LOCC:** Invariance of phase space with respect to local deformations of the slow variable is consequence of the fact that local operations and classical communications can't increase/decrease entanglement. Hence entropic origin of inertia should not be considered as change of entropy, it is rather caused by change of energy under the invariance of entropy.

- **Emergence gravitational waves and gauge fields:** Finally we would like to emphasize that adiabatic reaction force is the zeroth order affect of the underlying fast system. We interpret this zeroth order affect as inertia or geodesic motion. One can also study subleading contributions to reaction force in a systematic expansion over the ratio of characteristic time scales of the slow and fast systems [161]. These subleading terms come in the form of a symmetric and anti-symmetric tensors. Therefore reaction of the fast system provides space to generate gravitational fields and gauge fields. We leave studying these subleading terms in a future study. It would be a strong check of the hypothesis provided here.

Bibliography

[1] I. Ilgin and I.-S. Yang, "Causal Patch Complementarity: The Inside Story for Old Black Holes," *Phys. Rev.* **D89** no. 4, (2014) 044007, arXiv:1311.1219 [hep-th].

[2] I. Ilgin, "Bekenstein bound in the bulk and AdS/CFT," arXiv:1809.05770 [hep-th].

[3] I. Ilgin, E. Verlinde, and M. Visser, "Emergent Laws of Spacetime Mechanics,".

[4] I. Ilgin and I.-S. Yang, "Energy carries Information," *Int. J. Mod. Phys.* **A29** no. 20, (2014) 1450115, arXiv:1402.0878 [hep-th].

[5] G. Deleuze and F. Guattari, *What is Philosophy?* Columbia University Press, 1991.

[6] J. A. Wheeler, "Information, physics, quantum: The search for links," in *Proceedings III International Symposium on Foundations of Quantum Mechanics*, pp. 354–358. 1989.

[7] R. Clausius, *The Mechanical Theory of Heat: with its Applications to the Steam Engine and to Physical Properties of Bodies*. 1867.

[8] D. Frenkel and P. B. Warren, "Gibbs, boltzmann, and negative temperatures," *American Journal of Physics* **83** no. 2, (2015) 163–170, arXiv:1403.4299.

[9] P. Ehrenfest, "Xlviii. adiabatic invariants and the theory of quanta," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **33** no. 198, (1917) 500–513.

[10] L. de Broglie, "The Thermodynamics of the isolated particle ,".

[11] E. P. Verlinde, "On the Origin of Gravity and the Laws of Newton," *JHEP* **04** (2011) 029, arXiv:1001.0785 [hep-th].

[12] M. B. Plenio and S. Virmani, "An Introduction to entanglement measures," *Quant. Inf. Comput.* **7** (2007) 1–51, arXiv:quant-ph/0504163 [quant-ph].

[13] W. H. Zurek, "Decoherence, einselection, and the quantum origins of the classical," *Rev. Mod. Phys.* **75** (May, 2003) 715–775.

[14] J. D. Bekenstein, "Black holes and the second law," *Lett. Nuovo Cim.* **4** (1972) 737–740.

[15] J. D. Bekenstein, "Black holes and entropy," *Phys. Rev.* **D7** (1973) 2333–2346.

[16] J. D. Bekenstein, "Generalized second law of thermodynamics in black hole physics," *Phys. Rev.* **D9** (1974) 3292–3300.

[17] R. Penrose, "Gravitational collapse: The role of general relativity," *Riv. Nuovo Cim.* **1** (1969) 252–276. [Gen. Rel. Grav.34,1141(2002)].

[18] D. Christodoulou, "Reversible and irreversible transformations in black-hole physics,".

[19] D. Christodoulou and R. Ruffini, "Reversible transformations of a charged black hole," *Phys. Rev. D* **4** (Dec, 1971) 3552–3555.

[20] S. W. Hawking, "Particle Creation by Black Holes," *Commun. Math. Phys.* **43** (1975) 199–220.

[21] S. W. Hawking, "Black holes and thermodynamics," *Phys. Rev. D* **13** (Jan, 1976) 191–197.

[22] J. M. Bardeen, B. Carter, and S. W. Hawking, "The four laws of black hole mechanics," *Comm. Math. Phys.* **31** no. 2, (1973) 161–170.

[23] A. Strominger and C. Vafa, "Microscopic origin of the Bekenstein-Hawking entropy," *Phys. Lett.* **B379** (1996) 99–104, arXiv:hep-th/9601029 [hep-th].

[24] J. L. Cardy, "Operator content of two-dimensional conformally invariant theories," *Nuclear Physics B* **270** (1986) 186 – 204.

[25] S. B. Giddings, "Nonviolent nonlocality," *Phys. Rev.* **D88** (2013) 064023, arXiv:1211.7070 [hep-th].

[26] S. B. Giddings, "Hawking radiation, the StefanBoltzmann law, and unitarization," *Phys. Lett.* **B754** (2016) 39–42, arXiv:1511.08221 [hep-th].

[27] E. Verlinde and H. Verlinde, "Black Hole Information as Topological Qubits," arXiv:1306.0516 [hep-th].

[28] M. Freedman and M. Headrick, "Bit threads and holographic entanglement," *Commun. Math. Phys.* **352** no. 1, (2017) 407–438, arXiv:1604.00354 [hep-th].

[29] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, "Black Holes: Complementarity or Firewalls?," *JHEP* **02** (2013) 062, arXiv:1207.3123 [hep-th].

[30] S. W. Hawking, "Breakdown of predictability in gravitational collapse," *Phys. Rev. D* **14** (Nov, 1976) 2460–2473.

[31] T. Jacobson, "Black hole evaporation and ultrashort distances," *Phys. Rev.* **D44** (1991) 1731–1739.

[32] G. 't Hooft, "On the Quantum Structure of a Black Hole," *Nucl. Phys.* **B256** (1985) 727–745.

[33] G. 't Hooft, "The Scattering matrix approach for the quantum black hole: An Overview," *Int. J. Mod. Phys.* **A11** (1996) 4623–4688, arXiv:gr-qc/9607022 [gr-qc].

[34] L. Susskind and L. Thorlacius, "Gedanken experiments involving black holes," *Phys. Rev.* **D49** (1994) 966–974, arXiv:hep-th/9308100 [hep-th].

[35] L. Susskind, L. Thorlacius, and J. Uglum, "The Stretched horizon and black hole complementarity," *Phys. Rev.* **D48** (1993) 3743–3761, arXiv:hep-th/9306069 [hep-th].

[36] D. A. Lowe, J. Polchinski, L. Susskind, L. Thorlacius, and J. Uglum, "Black hole complementarity versus locality," *Phys. Rev.* **D52** (1995) 6997–7010, arXiv:hep-th/9506138 [hep-th].

[37] G. 't Hooft, "Quantum gravity without space-time singularities or horizons," *Subnucl. Ser.* **47** (2011) 251–265, arXiv:0909.3426 [gr-qc].

[38] D. N. Page, "Information in black hole radiation," *Phys. Rev. Lett.* **71** (1993) 3743–3746, arXiv:hep-th/9306083 [hep-th].

[39] D. N. Page, "Time Dependence of Hawking Radiation Entropy," *JCAP* **1309** (2013) 028, arXiv:1301.4995 [hep-th].

[40] S. B. Giddings and Y. Shi, "Effective field theory models for nonviolent information transfer from black holes," *Phys. Rev.* **D89** no. 12, (2014) 124032, arXiv:1310.5700 [hep-th].

[41] S. B. Giddings, "Possible observational windows for quantum effects from black holes," *Phys. Rev.* **D90** no. 12, (2014) 124033, arXiv:1406.7001 [hep-th].

[42] E. Verlinde and H. Verlinde, "Passing through the Firewall," arXiv:1306.0515 [hep-th].

[43] K. Papadodimas and S. Raju, "Black Hole Interior in the Holographic Correspondence and the Information Paradox," *Phys. Rev. Lett.* **112** no. 5, (2014) 051301, arXiv:1310.6334 [hep-th].

[44] K. Papadodimas and S. Raju, "An Infalling Observer in AdS/CFT," *JHEP* **10** (2013) 212, arXiv:1211.6767 [hep-th].

[45] S. Banerjee, J.-W. Bryan, K. Papadodimas, and S. Raju, "A toy model of black hole complementarity," *JHEP* **05** (2016) 004, arXiv:1603.02812 [hep-th].

[46] S. Raju, "A Toy Model of the Information Paradox in Empty Space," arXiv:1809.10154 [hep-th].

[47] J. Maldacena and L. Susskind, "Cool horizons for entangled black holes," *Fortsch. Phys.* **61** (2013) 781–811, arXiv:1306.0533 [hep-th].

[48] G. T. Horowitz and J. M. Maldacena, "The Black hole final state," *JHEP* **02** (2004) 008, arXiv:hep-th/0310281 [hep-th].

[49] D. Harlow and P. Hayden, "Quantum Computation vs. Firewalls," *JHEP* **06** (2013) 085, arXiv:1301.4504 [hep-th].

[50] T. Jacobson, "Thermodynamics of space-time: The Einstein equation of state," *Phys. Rev. Lett.* **75** (1995) 1260–1263, arXiv:gr-qc/9504004 [gr-qc].

[51] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity," *Int. J. Theor. Phys.* **38** (1999) 1113–1133, arXiv:hep-th/9711200 [hep-th]. [Adv. Theor. Math. Phys.2,231(1998)].

[52] E. Witten, "Anti-de Sitter space and holography," *Adv. Theor. Math. Phys.* **2** (1998) 253–291, arXiv:hep-th/9802150 [hep-th].

[53] O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri, and Y. Oz, "Large N field theories, string theory and gravity," *Phys. Rept.* **323** (2000) 183–386, arXiv:hep-th/9905111 [hep-th].

[54] S. El-Showk and K. Papadodimas, "Emergent Spacetime and Holographic CFTs," *JHEP* **10** (2012) 106, arXiv:1101.4163 [hep-th].

[55] D. Dieks, J. van Dongen, and S. de Haro, "Emergence in Holographic Scenarios for Gravity," *Stud. Hist. Phil. Sci.* **B52** (2015) 203–216, arXiv:1501.04278 [hep-th].

[56] P. W. Anderson, "More is different," *Science* **177** no. 4047, (1972) 393–396, http://science.sciencemag.org/content/177/4047/393.full.pdf.

[57] S. Ryu and T. Takayanagi, "Holographic derivation of entanglement entropy from AdS/CFT," *Phys. Rev. Lett.* **96** (2006) 181602, arXiv:hep-th/0603001 [hep-th].

[58] S. Ryu and T. Takayanagi, "Aspects of Holographic Entanglement Entropy," *JHEP* **08** (2006) 045, arXiv:hep-th/0605073 [hep-th].

[59] V. E. Hubeny, M. Rangamani, and T. Takayanagi, "A Covariant holographic entanglement entropy proposal," *JHEP* **07** (2007) 062, arXiv:0705.0016 [hep-th].

[60] M. Van Raamsdonk, "Building up spacetime with quantum entanglement," *Gen. Rel. Grav.* **42** (2010) 2323–2329, arXiv:1005.3035 [hep-th]. [Int. J. Mod. Phys.D19,2429(2010)].

[61] H. Casini, M. Huerta, and R. C. Myers, "Towards a derivation of holographic entanglement entropy," *JHEP* **05** (2011) 036, arXiv:1102.0440 [hep-th].

[62] T. Faulkner, M. Guica, T. Hartman, R. C. Myers, and M. Van Raamsdonk, "Gravitation from Entanglement in Holographic CFTs," *JHEP* **03** (2014) 051, arXiv:1312.7856 [hep-th].

[63] R. M. Wald, "Black hole entropy is the Noether charge," *Phys. Rev.* **D48** no. 8, (1993) R3427–R3431, arXiv:gr-qc/9307038 [gr-qc].

[64] 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, arXiv:gr-qc/9403028 [gr-qc].

[65] 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, arXiv:gr-qc/9503052 [gr-qc].

[66] T. Faulkner, A. Lewkowycz, and J. Maldacena, "Quantum corrections to holographic entanglement entropy," *JHEP* **11** (2013) 074, arXiv:1307.2892 [hep-th].

[67] S. Popescu, A. J. Short, and A. Winter, "Entanglement and the foundations of statistical mechanics," *Nature Physics* **2** (2006) 754–758.

[68] V. Balasubramanian, B. Czech, B. D. Chowdhury, and J. de Boer, "The entropy of a hole in spacetime," *JHEP* **10** (2013) 220, arXiv:1305.0856 [hep-th].

[69] M. Headrick, R. C. Myers, and J. Wien, "Holographic Holes and Differential Entropy," *JHEP* **10** (2014) 149, arXiv:1408.4770 [hep-th].

[70] J. D. Bekenstein, "A Universal Upper Bound on the Entropy to Energy Ratio for Bounded Systems," *Phys. Rev.* **D23** (1981) 287.

[71] J. D. Bekenstein, "Black holes and information theory," *Contemp. Phys.* **45** (2003) 31–43, arXiv:quant-ph/0311049 [quant-ph].

[72] H. Casini, "Relative entropy and the Bekenstein bound," *Class. Quant. Grav.* **25** (2008) 205021, arXiv:0804.2182 [hep-th].

[73] E. P. Verlinde, "Emergent Gravity and the Dark Universe," *SciPost Phys.* **2** no. 3, (2017) 016, arXiv:1611.02269 [hep-th].

[74] T. Jacobson, "Entanglement Equilibrium and the Einstein Equation," *Phys. Rev. Lett.* **116** no. 20, (2016) 201101, arXiv:1505.04753 [gr-qc].

[75] P. Millette, "Elastodynamics of the spacetime continuum,".

[76] H. Kleinert, "Gravity as a theory of defects in a crystal with only second gradient elasticity," *Annalen der Physik* **499** no. 2, 117–119.

[77] M. O. Tahim, R. R. Landim, and C. A. S. Almeida, "Spacetime as a deformable solid," *Mod. Phys. Lett.* **A24** (2009) 1209–1217, arXiv:0705.4120 [gr-qc].

[78] R. A. Battye and J. A. Pearson, "Massive gravity, the elasticity of space-time, and perturbations in the dark sector," *Phys. Rev. D* **88** (Oct, 2013) 084004.

[79] A. Tartaglia, "Four-dimensional elasticity and general relativity," *Grav. Cosmol.* **1** (1995) 335–338, arXiv:gr-qc/9509043 [gr-qc]. [Submitted to: Grav. Cosmol.(1995)].

[80] J. A. Pearson, "Material models of dark energy," *Annalen Phys.* **526** (2014) 318–339, arXiv:1403.1213 [astro-ph.CO].

[81] T. Padmanabhan, "Gravity as elasticity of spacetime: A Paradigm to understand horizon thermodynamics and cosmological constant," *Int. J. Mod. Phys.* **D13** (2004) 2293–2298, arXiv:gr-qc/0408051 [gr-qc].

[82] M. Bucher and D. N. Spergel, "Is the dark matter a solid?," *Phys. Rev.* **D60** (1999) 043505, arXiv:astro-ph/9812022 [astro-ph].

[83] J. D. Brown, S. R. Lau, and J. W. York, Jr., "Action and energy of the gravitational field," arXiv:gr-qc/0010024 [gr-qc].

[84] J. D. Brown and J. W. York, Jr., “Quasilocal energy and conserved charges derived from the gravitational action,” *Phys. Rev.* **D47** (1993) 1407–1419, arXiv:gr-qc/9209012 [gr-qc].

[85] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter, and Z. Yang, “Holographic duality from random tensor networks,” *JHEP* **11** (2016) 009, arXiv:1601.01694 [hep-th].

[86] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, “Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence,” *JHEP* **06** (2015) 149, arXiv:1503.06237 [hep-th].

[87] B. Swingle, “Entanglement Renormalization and Holography,” *Phys. Rev.* **D86** (2012) 065007, arXiv:0905.1317 [cond-mat.str-el].

[88] V. Balasubramanian, B. D. Chowdhury, B. Czech, J. de Boer, and M. P. Heller, “Bulk curves from boundary data in holography,” *Phys. Rev.* **D89** no. 8, (2014) 086004, arXiv:hep-th/1310.4204 [hep-th].

[89] R. C. Myers, J. Rao, and S. Sugishita, “Holographic Holes in Higher Dimensions,” *JHEP* **06** (2014) 044, arXiv:1403.3416 [hep-th].

[90] B. Czech, P. Hayden, N. Lashkari, and B. Swingle, “The Information Theoretic Interpretation of the Length of a Curve,” *JHEP* **06** (2015) 157, arXiv:1410.1540 [hep-th].

[91] 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** no. 2, (1986) 207–226.

[92] V. Balasubramanian, B. D. Chowdhury, B. Czech, and J. de Boer, “Entwinement and the emergence of spacetime,” *JHEP* **01** (2015) 048, arXiv:1406.5859 [hep-th].

[93] E. Keski-Vakkuri, “Bulk and boundary dynamics in BTZ black holes,” *Phys. Rev.* **D59** (1999) 104001, arXiv:hep-th/9808037 [hep-th].

[94] V. Balasubramanian, P. Kraus, and M. Shigemori, “Massless black holes and black rings as effective geometries of the D1-D5 system,” *Class. Quant. Grav.* **22** (2005) 4803–4838, arXiv:hep-th/0508110 [hep-th].

[95] J. de Boer, M. M. Sheikh-Jabbari, and J. Simon, “Near Horizon Limits of Massless BTZ and Their CFT Duals,” *Class. Quant. Grav.* **28** (2011) 175012, arXiv:1011.1897 [hep-th].

[96] C. Agn and T. Faulkner, “Quantum Corrections to Holographic Mutual Information,” *JHEP* **08** (2016) 118, arXiv:1511.07462 [hep-th].

[97] S. W. Hawking, "Breakdown of Predictability in Gravitational Collapse," *Phys. Rev.* **D14** (1976) 2460–2473.

[98] S. D. Mathur, "The Information paradox: A Pedagogical introduction," *Class. Quant. Grav.* **26** (2009) 224001, arXiv:0909.1038 [hep-th].

[99] S. L. Braunstein, S. Pirandola, and K. Zyczkowski, "Better Late than Never: Information Retrieval from Black Holes," *Phys. Rev. Lett.* **110** no. 10, (2013) 101301, arXiv:0907.1190 [quant-ph].

[100] S. G. Avery, B. D. Chowdhury, and A. Puhm, "Unitarity and fuzzball complementarity: 'Alice fuzzes but may not even know it!'," *JHEP* **09** (2013) 012, arXiv:1210.6996 [hep-th].

[101] B. D. Chowdhury, "Cool horizons lead to information loss," *JHEP* **10** (2013) 034, arXiv:1307.5915 [hep-th].

[102] R. Bousso, "Violations of the Equivalence Principle by a Nonlocally Reconstructed Vacuum at the Black Hole Horizon," *Phys. Rev. Lett.* **112** no. 4, (2014) 041102, arXiv:1308.3697 [hep-th].

[103] S. Lloyd and J. Preskill, "Unitarity of black hole evaporation in final-state projection models," *JHEP* **08** (2014) 126, arXiv:1308.4209 [hep-th].

[104] L. Susskind, "The Census taker's hat," arXiv:0710.1129 [hep-th].

[105] J. Garriga and A. Vilenkin, "Holographic Multiverse," *JCAP* **0901** (2009) 021, arXiv:0809.4257 [hep-th].

[106] R. Bousso, "Complementarity in the Multiverse," *Phys. Rev.* **D79** (2009) 123524, arXiv:0901.4806 [hep-th].

[107] R. Bousso and I.-S. Yang, "Global-Local Duality in Eternal Inflation," *Phys. Rev.* **D80** (2009) 124024, arXiv:0904.2386 [hep-th].

[108] K. S. Thorne, R. H. Price, and D. A. MacDonald, "Black holes: The membrane paradigm,".

[109] G. 't Hooft, "The black hole interpretation of string theory," *Nucl. Phys.* **B335** (1990) 138–154.

[110] L. Susskind, "String theory and the principles of black hole complementarity," *Phys. Rev. Lett.* **71** (1993) 2367–2368, arXiv:hep-th/9307168 [hep-th].

[111] E. P. Verlinde, "Black hole evaporation and complementarity," in *Mathematical physics. Proceedings, 11th International Congress, Paris,*

France, July 18-22, 1994, pp. 668–673. 1994. arXiv:hep-th/9503120 [hep-th]. [,668(1994)].

[112] O. Lunin and S. D. Mathur, "Statistical interpretation of Bekenstein entropy for systems with a stretched horizon," *Phys. Rev. Lett.* **88** (2002) 211303, arXiv:hep-th/0202072 [hep-th].

[113] S. D. Mathur, "The Fuzzball proposal for black holes: An Elementary review," *Fortsch. Phys.* **53** (2005) 793–827, arXiv:hep-th/0502050 [hep-th].

[114] R. Bousso, "Complementarity Is Not Enough," *Phys. Rev.* **D87** no. 12, (2013) 124023, arXiv:1207.5192 [hep-th].

[115] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford, and J. Sully, "An Apologia for Firewalls," *JHEP* **09** (2013) 018, arXiv:1304.6483 [hep-th].

[116] D. Marolf and J. Polchinski, "Gauge/Gravity Duality and the Black Hole Interior," *Phys. Rev. Lett.* **111** (2013) 171301, arXiv:1307.4706 [hep-th].

[117] R. Bousso, "Firewalls from double purity," *Phys. Rev.* **D88** no. 8, (2013) 084035, arXiv:1308.2665 [hep-th].

[118] D. A. Lowe and L. Thorlacius, "Comments on the black hole information problem," *Phys. Rev.* **D73** (2006) 104027, arXiv:hep-th/0601059 [hep-th].

[119] R. Schutzhold and W. G. Unruh, "On Quantum Correlations across the Black Hole Horizon," *Phys. Rev.* **D81** (2010) 124033, arXiv:1002.1844 [gr-qc].

[120] S. W. Hawking and D. N. Page, "Thermodynamics of Black Holes in anti-De Sitter Space," *Commun. Math. Phys.* **87** (1983) 577.

[121] L. Hui and I.-S. Yang, "Complementarity plus backreaction is enough," *Phys. Rev.* **D89** no. 8, (2014) 084011, arXiv:1308.6268 [hep-th].

[122] J. J. Bisognano and E. H. Wichmann, "On the Duality Condition for Quantum Fields," *J. Math. Phys.* **17** (1976) 303–321.

[123] A. Lewkowycz and J. Maldacena, "Generalized gravitational entropy," *JHEP* **08** (2013) 090, arXiv:1304.4926 [hep-th].

[124] J. de Boer, M. P. Heller, R. C. Myers, and Y. Neiman, "Holographic de Sitter Geometry from Entanglement in Conformal Field Theory," *Phys. Rev. Lett.* **116** no. 6, (2016) 061602, arXiv:1509.00113 [hep-th].

[125] D. Blanco, H. Casini, M. Leston, and F. Rosso, "Modular energy inequalities from relative entropy," *JHEP* **01** (2018) 154, arXiv:1711.04816 [hep-th].

[126] C. P. Herzog, "Universal Thermal Corrections to Entanglement Entropy for Conformal Field Theories on Spheres," *JHEP* **10** (2014) 28, arXiv:1407.1358 [hep-th].

[127] J. Cardy and E. Tonni, "Entanglement hamiltonians in two-dimensional conformal field theory," *J. Stat. Mech.* **1612** no. 12, (2016) 123103, arXiv:1608.01283 [cond-mat.stat-mech].

[128] S. Deser, R. Jackiw, and G. 't Hooft, "Three-Dimensional Einstein Gravity: Dynamics of Flat Space," *Annals Phys.* **152** (1984) 220.

[129] J. Cardy and C. P. Herzog, "Universal Thermal Corrections to Single Interval Entanglement Entropy for Two Dimensional Conformal Field Theories," *Phys. Rev. Lett.* **112** no. 17, (2014) 171603, arXiv:1403.0578 [hep-th].

[130] B. Swingle and M. Van Raamsdonk, "Universality of Gravity from Entanglement," arXiv:1405.2933 [hep-th].

[131] J. Lin, M. Marcolli, H. Ooguri, and B. Stoica, "Locality of Gravitational Systems from Entanglement of Conformal Field Theories," *Phys. Rev. Lett.* **114** (2015) 221601, arXiv:1412.1879 [hep-th].

[132] X. Dong, D. Harlow, and A. C. Wall, "Reconstruction of Bulk Operators within the Entanglement Wedge in Gauge-Gravity Duality," *Phys. Rev. Lett.* **117** no. 2, (2016) 021601, arXiv:1601.05416 [hep-th].

[133] T. Faulkner and A. Lewkowycz, "Bulk locality from modular flow," *JHEP* **07** (2017) 151, arXiv:1704.05464 [hep-th].

[134] S. Leichenauer, "Thermal Corrections to Entanglement Entropy from Holography," *JHEP* **09** (2015) 014, arXiv:1502.07348 [hep-th].

[135] V. E. Hubeny, H. Maxfield, M. Rangamani, and E. Tonni, "Holographic entanglement plateaux," *JHEP* **08** (2013) 092, arXiv:1306.4004 [hep-th].

[136] H. Araki and E. H. Lieb, "Entropy Inequalities," *Commun.Math. Phys.* **18** (1970) 160–170.

[137] E. Witten, "Anti-de Sitter space, thermal phase transition, and confinement in gauge theories," *Adv. Theor. Math. Phys.* **2** (1998) 505–532, arXiv:hep-th/9803131 [hep-th].

[138] L. Susskind and E. Witten, "The Holographic bound in anti-de Sitter space," arXiv:hep-th/9805114 [hep-th].

[139] S. W. Hawking, "Black hole explosions," *Nature* **248** (1974) 30–31.

[140] A. Strominger, "Black hole entropy from near horizon microstates," *JHEP* **02** (1998) 009, arXiv:hep-th/9712251 [hep-th].

[141] J. M. Maldacena and L. Susskind, "D-branes and fat black holes," *Nucl. Phys.* **B475** (1996) 679–690, arXiv:hep-th/9604042 [hep-th].

[142] J. M. Maldacena, A. Strominger, and E. Witten, "Black hole entropy in M theory," *JHEP* **12** (1997) 002, arXiv:hep-th/9711053 [hep-th].

[143] G. 't Hooft, "Dimensional reduction in quantum gravity," *Conf. Proc.* **C930308** (1993) 284–296, arXiv:gr-qc/9310026 [gr-qc].

[144] L. Susskind, "The World as a hologram," *J. Math. Phys.* **36** (1995) 6377–6396, arXiv:hep-th/9409089 [hep-th].

[145] E. Bianchi and R. C. Myers, "On the Architecture of Spacetime Geometry," *Class. Quant. Grav.* **31** (2014) 214002, arXiv:1212.5183 [hep-th].

[146] B. Czech, L. Lamprou, S. McCandlish, and J. Sully, "Integral Geometry and Holography," *JHEP* **10** (2015) 175, arXiv:1505.05515 [hep-th].

[147] L. A. Santaló, *Integral geometry and geometric probability.* Cambridge university press, 2004.

[148] M. W. Crofton, "On the theory of local probability, applied to straight lines drawn at random in a plane; the methods used being also extended to the proof of certain new theorems in the integral calculus," *Philosophical Transactions of the Royal Society of London* **158** (1868) 181–199.

[149] E. Gallego and G. Solanes, "Integral geometry and geometric inequalities in hyperbolic space," *Differential Geometry and its Applications* **22** no. 3, (2005) 315 – 325.

[150] B. Freivogel, R. A. Jefferson, L. Kabir, B. Mosk, and I.-S. Yang, "Casting Shadows on Holographic Reconstruction," *Phys. Rev.* **D91** no. 8, (2015) 086013, arXiv:1412.5175 [hep-th].

[151] E. Witten, "Three-Dimensional Gravity Revisited," arXiv:0706.3359 [hep-th].

[152] S. W. Hawking and D. N. Page, "Thermodynamics of black holes in anti-de sitter space," *Comm. Math. Phys.* **87** no. 4, (1982) 577–588.

[153] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, vol. 7. Elsevier, 3 ed., 1986.

[154] J. M. Parrondo, J. M. Horowitz, and T. Sagawa, “Thermodynamics of information,”.

[155] M. Parikh, S. Sarkar, and A. Svesko, “A Local First Law of Gravity,” arXiv:1801.07306 [gr-qc].

[156] J. Harper, M. Headrick, and A. Rolph, “Bit Threads in Higher Curvature Gravity,” arXiv:1807.04294 [hep-th].

[157] A. V. Frolov and V. P. Frolov, “Rigidly rotating zero-angular-momentum observer surfaces in the Kerr spacetime,” *Phys. Rev.* **D90** no. 12, (2014) 124010, arXiv:1408.6316 [gr-qc].

[158] A. Einstein and N. Rosen, “The particle problem in the general theory of relativity,” *Phys. Rev.* **48** (Jul, 1935) 73–77.

[159] M. Tamir, “Proving the principle: Taking geodesic dynamics too seriously in einstein’s theory,” *Studies in History and Philosophy of Modern Physics* **43** (May, 2012) 137–154.

[160] M. V. Berry, “Quantal phase factors accompanying adiabatic changes,” *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **392** no. 1802, (1984) 45–57.

[161] M. V. Berry and J. M. Robbins, “Chaotic classical and half-classical adiabatic reactions: geometric magnetism and deterministic friction,” . **442** no. 1916, (1993) 659–672, http://rspa.royalsocietypublishing.org/content/442/1916/659.full.pdf.

[162] E. P. Verlinde, “The Hidden Phase Space of Our Universe,” *unpublished notes* .

[163] F. Ferrari, “Gauge Theories, D-Branes and Holography,” *Nucl. Phys.* **B880** (2014) 247–289, arXiv:1310.6788 [hep-th].

[164] S. B. Giddings, “Hilbert space structure in quantum gravity: an algebraic perspective,” *JHEP* **12** (2015) 099, arXiv:1503.08207 [hep-th].

[165] D. Harlow and D. Jafferis, “The Factorization Problem in Jackiw-Teitelboim Gravity,” arXiv:1804.01081 [hep-th].

[166] S. Gao and R. M. Wald, “The ’Physical process’ version of the first law and the generalized second law for charged and rotating black holes,” *Phys. Rev.* **D64** (2001) 084020, arXiv:gr-qc/0106071 [gr-qc].

[167] A. Strominger, “Microcanonical quantum field theory,” *Annals of Physics* **146** no. 2, (1983) 419 – 457.

[168] J. D. Brown and J. W. York, Jr., "The Microcanonical functional integral. 1. The Gravitational field," *Phys. Rev.* **D47** (1993) 1420–1431, arXiv:gr-qc/9209014 [gr-qc].

[169] J. D. Brown and J. W. York, Jr., "Jacobi's Action and the Recovery of Time in General Relativity," *Phys. Rev.* **D40** (1989) 3312–3318.

Contributions to Publications

Below I will describe my personal contributions to the publications on which this thesis is based. Note that in theoretical high energy physics, the order of the list of authors is alphabetical, so it does not reflect the contribution of the individual authors.

[1] I. Ilgin and I-S. Yang
Causal Patch Complementarity: The Inside Story for Old Black Holes
Phys. Rev. **D89 no. 4**, 044007 (2014), arXiv:1311.1219 [hep-th].

I was responsible for all the calculations in this paper under the guidance of I-Sheng Yang who was the main author behind the idea. I have also contributed to conceptual content of the paper

[4] I. Ilgin and I-S. Yang
Energy carries Information
Int. J. Mod. Phys. **A29 no. 20**, 1450115 (2014), arXiv:1402.0878 [hep-th].
I have carried out the analysis that provides the content of section two and section three in this paper. I-Sheng Yang focused on the application to black hole physics. The original idea on the relation between energy and information was a joint effort.

[2] I. Ilgin
Bekenstein bound in the bulk and AdS/CFT
on submission, arXiv:1809.05770 [hep-th].
I am responsible both for the calculations and content behind this work.

[3] I. Ilgin, E. Verlinde and M. Visser
Emergent Laws of Spacetime Mechanics
to appear.

The content of this paper is yet to be determined, hence I will provide my contributions on the chapters 6, 7 that will be used in this paper.
The idea that entropic force can be elaborated as an adiabatic reaction force belongs to Erik Verlinde. In addition he proposed that covariant phase space

formalism can be applied to general spacelike surfaces. Manus Visser applied the formalism to spherically symmetric surfaces on static spherically symmetric backgrounds and obtain the result (6.2.14)*. All the rest of the calculations and interpretations used in this thesis are result of my work.*

Samenvatting

In deze scriptie, hebben we de hypothese onderzocht dat zwaartekracht een *emergent fenomeen* is. We beginnen eerst met het onderzoeken van wat een emergent verschijnsel betekend vanuit een algemeen perspectief, en vervolgens verklaren we de betekenis van het concept, middels het onderzoek in deze scriptie.

De moderne fysica, met de komst van de kwantumfysica, is gewend geraakt om de wereld in hirarchien van complexiteiten te zien. Elke klasse binnen de hirarchie is het resultaat van de relaties en interacties van de bouwstenen van de hirarchien binnen die klasse en eromheen. Emergentie is het proces waarin een radicaal anders en samenhangend systeem ontstaat, als gevolg van de relaties van zijn componenten. De fysica van de emergerende laag in een hirarchie drukt zich uit in een nieuwe taal die geen verwijzingen hoeft te hebben naar de bouwstenen. Hoewel het duidelijk is dat elke structuur binnen de hirarchie voortkomt uit de interacties van zijn bestanddelen binnen en met de omgeving, is het betwistbaar of men de eigenschappen van de emergerende hirarchie kan construeren en af kan leiden uitgaande van onderliggende structuren. Wij geloven dat een dergelijke reductionistische constructie in principe mogelijk is wanneer men rekening houdt met de omstandigheden.

Het fenomeen emergentie is overal aanwezig. Water kwam als medium voort uit de interacties van moleculen die bestaan uit twee waterstofatomen en een zuurstofatoom. Verschillende takken van de wetenschap kunnen gekaderd worden door de relatie van emergentie; chemie komt voort uit de natuurkunde, terwijl biologie voortkomt uit chemie, enzovoort. Misschien is een van de mooiste voorbeelden van emergente verschijnselen, die van thermodynamica, die vooral interessant is voor deze scriptie. Toen de wetten van de thermodynamica geformuleerd werden, wisten fysici niet dat die wetten het resultaat waren van het collectieve en het statistische gedrag van de onderliggende microscopische structuren. De ontwikkeling van statische fysica legde de details van de emergentie van thermodynamica bloot.

Hoewel opkomst een enorme rol speelt in de fysica, heeft het geen fundamenteel belang voor de beschrijving van elementaire deeltjesfysica. Natuurkundigen geloven dat er een aantal krachten en deeltjes zijn die niet kunnen worden herleid tot de onderliggende bestanddelen. Zwaartekrachtveld wordt lang beschouwd als een dergelijke fundamentele, elementaire beschrijving. Alle pogingen om de zwaartekracht te kwantificeren via veldtheoretische benaderingen en tot zinloos. Er was geen consistente renormalisatieprocedure om oneindigheden uit de theorie te verwijderen.

Het verband tussen zwarte gaten en thermodynamica, en de komst van de snaartheorie, met name ijktheorie / zwaartekracht (AdS / CFT) dualiteit, leidde tot nieuwe perspectieven over de oorsprong van zwaartekracht. Het paradigma verschoof naar een richting waarin de klassieke ruimte-tijdbeschrijving over zwaartekracht niet meer als een elementaire beschrijving wordt beschouwd. Het was eerder grof en gemiddeld, een emergente beschrijving van enkele onderliggende microscopische structuren. Met andere woorden, wat wordt waargenomen als het weefsel van ruimte-tijd, vergelijkbaar met water, bestaat uit atomen, of eigenlijk, kwantumbits.

Laten we de nadruk leggen op dat, hoewel het emergent verschijnsel van de zwaartekracht een van de dominante paradigma's is met betrekking tot de oorsprong van, hoe en waaruit zwaartekracht ontstaat, varieert op basis van onderzoeksrichtingen. Een dominant perspectief met een solide basis is het overwegen van emergentie door dualiteit. Klassieke zwaartekracht is de emergente beschrijving die voldoet aan de onderliggende conforme veldtheorie wanneer kwantumeffecten worden onderdrukt door grote aantallen volgens AdS / CFT. We zullen dit perspectief hier niet verder onderzoeken, omdat dit in detail is onderzocht in deze scriptie. Een andere perspectief, dat we in deze scriptie in detail hebben onderzocht, is het identificeren van de grove- emergente- waarnemingen binnen de klassieke beschrijving van zwaartekracht, namelijk ruimte-tijd, zonder expliciet te verwijzen naar een duale theorie. Hoewel deze poging enigszins verschilt van een op dualiteit gebaseerde emergentie, maakt het de nodige conceptuele sprongen, gebaseerd op de lessen van AdS / CFT. Daarom hebben we geconcludeerd dat dit laatste perspectief op de emergente zwaartekracht geen sterke basis heeft, zoals vanuit het perspectief van de gauge. Het is echter zeer belangrijk, omdat het ons mogelijk maakt interessante fenomenen (zoals donkere materie en donkere energie) in ons universum te begrijpen, dat geen anti-Sitter-ruimte is en daarom niet goed in het de beschrijving van het holografische paradigma past.

In deze scriptie hebben we ons vooral gericht op het toepassen van de verbindingen tussen de kwantuminformatie en geometrie die worden waargenomen in de fysica van zwarte gaten, en de holografie op meer algemene ruimteachtige oppervlakken. We zullen uitweiden over wat de conclusies en resultaten van deze scriptie zijn.

Het eerste hoofdstuk start het onderzoek naar het verband tussen kwantuminformatie en zwaartekracht via de informatieparadox in het zwarte gat. Of de informatie van objecten die in een zwart gat vallen al dan niet kan worden teruggewonnen tijdens de verdamping van het zwarte gat, is een al lang bestaande vraag. In de jaren 90 werd gepostuleerd dat de Hilbert-ruimte geassocieerd met het interieur van een zwart gat complementair is aan de Hilbert-ruimte die de Hawking-straling ondersteunt. Hoewel een wiskundig solide formulering van de dualiteit niet was geconstrueerd, werd voorgesteld dat het complementariteitsbeginsel de informatieparadox oplost. In 2012 werd vastgesteld dat het complementariteitsbeginsel een intern conflict in zijn axioma's zou kunnen hebben, namelijk tussen effectieve veldtheorie en het gelijkwaardigheidsbeginsel. In deze scriptie beschouwen we een meer verifieerbare en daarom strikte versie van het principe dat bekend staat als causal patch-complementariteit. In deze versie kan schending van de principes van lokaliteit en de toepasbaarheid van effectieve veldtheorie alleen worden geverifieerd door waarnemers, die elkaar overlappende causale patches hebben. Met andere woorden, waarnemers kunnen een dergelijke paradox alleen verifiren als ze de mogelijkheid hebben om te communiceren. Een super-waarnemer die toegang heeft buiten zijn lichtkegel, mag de inconsistenties van het principe in deze benadering niet kaderen. In de causale patch-complementariteit hebben we aangetoond dat twee waarnemers, een infalling en een externe, de experimenten niet kunnen uitvoeren binnen de geldigheid van effectieve veldtheorie. Daarom heeft een operationele, verifieerbare versie van de complementariteit geen last van interne consistenties, omdat deze niet kan worden uitgevoerd binnen de effectieve veldtheorien.

Hoewel we een analyse voorleggen van een op de waarnemers gebaseerde versie van de complementariteit, onthult onze analyse niet hoe informatie over het zwarte gat wordt gedragen via het verdampingsproces. Op dit punt denken we dat het de absolute plaats is wat ten onrechte wordt aangenomen. Als we hier even mogen speculeren, zouden we kunnen stellen dat de informatieoverdracht plaatsvindt op een niet-lokale manier, tot het niveau van infrarood vrijheidsgraden, of tot lange golflengten in vergelijking met de lokale kromming van ruimte-tijd, en daarom kan het niet worden onderzocht door lokale waarnemers. Er zijn veel ideen over hoe absolute lokaliteit kan worden gewijzigd. We behandelen deze voorstellen in detail in deze scriptie en zullen er daarom hier niet nogmaals naar verwijzen. De richting van deze scriptie wordt sterk benvloed door de mogelijkheid dat effectieve veldtheorie kan worden aangepast vanwege niet-lokaal gecodeerde informatie, die verborgen is voor de lokale fysica. Een soortgelijk effect kan worden waargenomen in de elasticiteit van verstrengelde polymeren, waar dergelijke systemen twee verschillende dynamieken regelen die worden ontkoppeld vanwege de scheiding van tijdschalen, maar toch kan worden waargenomen hoe lange afstand stringy soorten vrijheidsgraden de lokale fysica beinvloeden als gevolg van het resultaat van langdurig elastisch gedrag.

Met deze motieven zoeken we informatietheoretische waarnemingen, afgeleid van de onderliggende theorie. Een van de doelen die in deze scriptie worden nagestreefd, is het uitwerken van het microscopisch mechanisme, waardoor macroscopische objecten traag worden. Het is een uiterst moeilijke taak om deze waarnemingen eenvoudig te identificeren met behulp van de relativiteitstheorie, omdat het geen microscopische beschrijving biedt. Daarom volgen we een strategie door het toepassen van lessen uit de holografie, waarbij gravitatietheorie een overeenkomstige dubbele microscopische beschrijving heeft. Het holografische paradigma stelt dat verstrengelingentropien van de dubbele microscopische beschrijving zijn gecodeerd als segmenten met minimale oppervlakken in de gravitatiebeschrijving. Hoewel dualiteit beperkt is tot de minimale oppervlakken, is het waarschijnlijk dat informatietheoretische correspondenties van meer algemene ruimteachtige oppervlakken op een meer gecompliceerde manier worden gecodeerd in de microscopische theorie. Dit idee weerspiegelt zich in de notie van differentile entropie, evenals de verbinding met integrale geometrie.

We hebben verder onderzocht hoe materie met ruimtetijd samenwerkt vanuit een informatietheoretisch perspectief. We hebben gezien dat je de Bekenstein-gebonden in de zwaartekracht kan waarnemen, als de gebonden in de entropie tussen complementaire regio's in de onderliggende theorie toeneemt. We hebben ook kwantitatieve verschillen aangegeven tussen pure toestanden en gemengde toestanden en hebben hun impact op de gravitatietheorie beoordeeld. Wij geloven dat onze bevindingen een belangrijke rol spelen bij het begrijpen van de entropie van de zwaartekracht in de Sitter-ruimte, zoals onlangs bepleit door Verlinde. Helaas zijn de meeste van onze conclusies beperkt tot sferisch symmetrische systemen. In zekere zin is dit te verwachten, omdat de Bekenstein-binding ook stiekem deze symmetrie in zijn formulering aanneemt. Met andere woorden, de factor 2π in de binding is het resultaat van sferische symmetrie. Dit is duidelijk geverifieerd in de QFT-derivatie van de binding met behulp van de positiviteit van relatieve entropie. De nieuwigheid van onze benadering is het verband tussen de bulkversie van de gebonden en de dubbele conforme veldtheorie.

In opkomende fenomenen definieert het collectieve gedrag van onderliggende bestanddelen een begrip van entropie. In de thermodynamica kwantificeert entropie bijvoorbeeld onwetendheid met betrekking tot de toestand van het microscopische systeem. Op een lineair niveau hangen energie en entropie nauw samen; deze relatie wordt de eerste wet genoemd. Het verband tussen zwaartekracht en emergent gedrag werd voor het eerst waargenomen door de ontdekking van de entropie van zwart gaten, kort nadat de eerste wet werd geformuleerd van de mechanica van zwart gaten, als de relatie tussen massa en het oppervlak van het zwarte gat. In deze scriptie streven we naar een generalisatie van de eerste wet op algemene ruimteachtige oppervlakken. De reden achter een dergelijk onderzoek is dat binnen de holografische context wordt begrepen dat de eerste wet van verstrengelingentropie in de onderliggende microscopische theorie zichzelf weerspiegelt als de eerste wet op minimale oppervlakken. Een natuurlijke vraag die volgde, was of zo'n relatie kan worden uitgebreid tot elk oppervlak.

Een terugkerend probleem dat we tijdens ons onderzoek tegenkwamen, was de moeilijkheid om ruimtetijden te vergelijken van verschillende toestanden die verbonden zijn door een energetische vervorming. Deze obstructie is een natuurlijk gevolg van de diffeomorfisme invariante formulering van relativiteit. Er is geen natuurlijk referentiepunt om een oppervlak te identificeren en de verandering ervan onder vervormingen te bestuderen. Deze kwestie is minder problematisch in het kader van holografie, omdat men de grens van ruimtetijd als een natuurlijk referentiepunt kan gebruiken. Recente ontwikkelingen binnen het holografische paradigma maken het mogelijk om niet alleen grensverankerde minimale oppervlakken te identificeren, maar ook algemene ruimteachtige oppervlakken. Deze opties bieden ons de mogelijkheid om de eerste wet van verstrengelings entropie in een grens te duwen als een eerste wet van differentile entropie in de bulk. Deze eerste wet manifesteert zich als een reactie van ruimte op energie. Bovendien hebben we geconstateerd dat, in tegenstelling tot de algemene mening over energetische excitaties die gebiedstekorten veroorzaken, hebben we aangetoond dat er, wanneer naar de grens wordt verwezen, oppervlakte-overschotten zijn. Het geeft aan dat een gebiedstekort een enigszins willekeurig begrip is, omdat het afhangt van het identificatieschema van een vervormde oplossing in relatie tot het oorspronkelijke.

Nadat we hebben aangetoond dat de eerste wet van verstrengeling kan worden verplaatst van minimale oppervlakken naar generieke ruimteachtige oppervlakken via integrale geometrie in 3D, veranderen we onze koers naar meer algemene oplossingen in alle dimensies, in plaats van anti de Sitter-space.

De wiskunde achter de eerste wet van zwarte gaten mechanica kan het best worden begrepen binnen de covariante fase ruimteformalisme. Het formalisme verenigt verschillende theorien van zwaartekracht vanuit het perspectief van de emergente wetten van zwarte gaten mechanica. Verder maakt het het mogelijk om eerste wettype relaties te bestuderen op andere oppervlakken dan de horizons van zwarte gaten. Dit is een vrijheid die we in deze scriptie uitgebreid hebben benut. We hebben de volgende twee vragen geformuleerd: ten eerste, wat zouden de consequenties zijn van het toepassen van deze formulering op algemene oppervlakken vanuit het oogpunt van emergente wetten? En ten tweede, wat zou dit ons leren over het begrip inertia? In het laatste deel van deze scriptie hebben we deze vragen beantwoord. We hebben de entropie van zwaartekracht opnieuw geformuleerd en verder uitgewerkt. In deze herformulering wordt de emergentie van traagheid gemodelleerd als de reactiekracht op een adiabatische vervorming van het systeem.

Met andere woorden, het systeem behoudt zijn toestand door op langzame vervormingen te reageren. Het idee dat zwaartekracht een emergent fenomeen is, stelt natuurlijk een onderscheid voor tussen tijdschalen van het onderliggende microscopisch systeem en de macroscopische sonde die ermee in wisselwerking staat. Een microscopisch systeem bestaat uit Planckiaanse vrijheidsgraden. Daarom is de dynamiek veel sneller in vergelijking met een macroscopische sonde. Het is deze scheiding van tijdschalen die het mogelijk maakt dat een snel systeem reageert op veranderingen, veroorzaakt door een langzaam systeem, en de energie van het systeem aanpast, zodat de zogenaamde adiabatische invarianten van het systeem behouden blijven. Wanneer de locatie van de macro-sonde als een parameter in de toestand van het systeem wordt beschouwd, hebben we aangetoond dat de traagheidswet van Newton gelijk is aan de adiabatische eerste wet. De adiabatische eerste wet werpt een relatie tussen verandering in de energie, microcanonieke entropie en kracht, op een manier vergelijkbaar met de eerste wet van de thermodynamica.

Bovendien hebben we gepostuleerd dat men een maat voor het volume van een faseruimte van ruimte-tijdregio's die zijn omgeven door gesloten oppervlakken kan associren. Interessant is dat deze maatregel overeenkomt met de micro-canonieke werking van de zwaartekracht. Het is de invariantie van het volume van de faseruimte die een reactiekracht oplevert, vanwege de aanpassing van de energie van het systeem. We hebben aangetoond dat covariant faseruimte-formalisme een canonieke methode aanbiedt om veranderingen in het volume van een faseruimte van regio's in ruimtetijd te meten. Deze notie generaliseert de entropie van het zwarte gat.

Een andere interessante bevinding van ons onderzoek is dat de eerste wet van zwarte gaten mechanica een bijzondere, speciale vorm is van een meer algemene eerste wet, die we de eerste wet van ruimte-tijd vervormingen noemden. Het is bekend dat zwaartekracht en elasticiteit intutieve overeenkomsten hebben, met name dat de algemene relativiteitstheorie zwaartekrachtmodelleerd als een elastische vervorming van ruimte-tijd. Onze bevindingen weerspiegelen deze eigenschappen vanuit een meer informatie-theoretisch perspectief. We hebben waargenomen dat zwaartekracht dimensionaal geprojecteerde elasticiteit is op twee co-dimensionale oppervlakken. Wij geloven dat de reden achter een dergelijke dimensionale reductie het principe van lokaliteit en de lokale Lorentz-symmetrie van de theorie is. Bovendien dwingt diffeomorfisme-invariantie dat verweven is in de algemene relativiteitstheorie deze elastische aard een integrale vorm te hebben in plaats van een lokale.

Toekomstige richtingen:
Door dit onderzoek, terwijl we streven naar een generalisatie van de eerste wet van

zwarte gaten mechanica op algemene oppervlakken, belanden we in een analogie tussen zwaartekracht en elasticiteit, puur gebaseerd op informatie-theoretische argumenten. Het belangrijkste kwantitatieve informatietheoretische verschil tussen beide is dat elasticiteit entropie schaalt met het volume van het systeem, terwijl het schaalt met het gebied in zwaartekracht. Ons universum heeft een thermisch karakter, maar het is niet duidelijk of deze thermische natuur gecentraliseerd is aan de kosmologische horizon of verspreid is in de bulk. Als het het laatste is, kan men beweren dat het vanwege zijn volume, zoals entropie, elastisch gedrag regelt. Wij geloven dat ons onderzoek methodes en kaders kan aanbieden om de relatie tussen elastische en zwaartekrachtsfasen op een covariante manier te bestuderen.

Een andere interessante richting voor onderzoek zou de tweede wet van ruimtetijd-mechanica kunnen zijn. Met andere woorden, wat de generalisatie van de tweede wet van zwarte gaten mechanica naar algemene oppervlakken zou zijn. Als we hier een moment over mogen speculeren, wijzen we op ons vermoeden met betrekking tot het verband tussen de tweede wet en het feit dat de klassieke oplossing van de actie een extremum is in de ruimte van mogelijke veldconfiguraties. De hint ligt in de identificatie van microcanonieke actie met de volume-entropie van Gibbs. We willen deze mogelijkheid graag onderzoeken in toekomstig onderzoek.

Summary

In this thesis, we have investigated the hypothesis that gravity is an *emergent phenomena*. We start by examining what emergent phenomena means from a general point of view, and then explain the meaning of the concept, as it is studied in this thesis. Modern physics, with the advent of quantum physics, has become habituated to seeing the world in hierarchies of complexities. Each class within the hierarchy is the result of the relations and interactions of the building blocks of the hierarchies within the class and around it.

The emergence of a phenomenon is the process in which a radically different and coherent system arises, as a result of the relations of its components. The physics of the emergent layer in a hierarchy is understood in a new language that does not need to refer to its building blocks. Although it is clear that each structure within the hierarchy arises from the interactions of its constituents within and with the environment, it is arguable whether one can construct and deduce the properties of the emergent hierarchy starting from underlying structures. We believe such a reductionist construction, in principle, is possible when one takes the environment into account.

The phenomena of emergence is everywhere in existence. Water as a medium emerges from the interactions of molecules that consist of two hydrogen atoms and an oxygen atom. Different branches of science can also be viewed through the relation of emergence; chemistry emerges from the world of physics, while biology emerges from chemistry and so on. Perhaps one of the most beautiful examples of emergent physics is that of thermodynamics, which is of particular interest for our thesis. When the laws of thermodynamics had been formulated, physicists were unaware that those laws were a result of the collective and statistical behaviour of the underlying microscopic structures. It was the development of statistical physics that uncovered the details of the emergence of thermodynamics.

Although emergence plays a tremendous role in physics, it has not been of fundamental importance for the description of elementary particle physics. Physicists believe that there are some forces and particles which cannot be reduced to its underlying constituents. Gravitational field has long been considered to be such

a fundamental, elementary description. However, all the attempts to quantize gravity through field theoretic approaches end up futile. There was no consistent renormalization procedure to remove infinities from the theory.

The connection between black holes and thermodynamics and the advent of string theory, particularly gauge theory/gravity (AdS/CFT) duality led to new perspectives emerging regarding the origin of gravity. The paradigm shifted towards a direction, in which the classical space-time description of gravity is not considered as an elementary description. Rather it was coarse-grained and averaged-out, an emergent description of some underlying microscopic structures. In other words, what is observed as the fabric of space-time, similar to water, is made out of atoms or, in fact, quantum bits.

Let us emphasize that, although emergence of gravity is one of the dominant paradigms regarding the origin of gravity, how and from what gravity emerges varies, based on research directions. One dominant perspective with strong foundations is considering emergence through duality. Classical gravity is the emergent description of underlying (dual) conformal field theory when quantum effects are suppressed by a big number according to AdS/CFT. We will not explore this perspective further here, as it has been examined in detail in the thesis. The other perspective, which we have investigated in detail in the thesis is identifying the coarse-grained - emergent - observables within the classical description of gravity, namely space-time, without explicitly referring to a dual theory. Although this attempt slightly differs from a duality-based emergence, it makes the necessary conceptual jumps, based on the lessons from AdS/CFT. Therefore we concluded that the latter perspective on the emergent gravity, does not have as strong a foundation as it does from the perspective of the gauge/gravity description. However, it is highly important, as it might allow us to understand interesting phenomena (dark matter, dark energy) of our universe, which is not an anti-de Sitter space, and hence, does not fit nicely into holographic paradigm as of today.

In this thesis, we have mainly focused on how to apply the connections between the quantum information and geometry observed in black hole physics, and the holography to more general space-like surfaces. Let us expand further what the conclusions and results of the thesis are.

The first chapter initializes the investigation on the connection between quantum information and gravity through the black hole information paradox. Whether or not the information of objects that fall into a black hole can be recovered during the evaporation of the black hole, has been a long-standing question. In the 90s, it was postulated that the Hilbert space associated to the interior of a black hole is complementary to the Hilbert space that supports the Hawking radiation. Although mathematically solid formulation of the duality was not constructed, it was proposed that the complementarity principle resolves the information paradox. In 2012, it was found that the principle of complementarity might have an internal conflict in its axioms, namely between effective field theory and the equivalence principle. In this thesis we consider a more verifiable and, hence, strict version of the principle known as causal patch complementarity. In this version, violation of the principles of locality and the applicability of effective field theory can only be verified by observers, who have intersecting causal patches. In other words, observers can verify such a paradox, only if they have the possibility to communicate. A super-observer that has access outside of its light cone, is not allowed to frame the inconsistencies of the principle in this approach. In the causal patch complementarity, we have shown that two observers, an infalling observer and an external one, cannot conduct the experiments within the validity of effective field theory. Hence an operational, verifiable version of the complementarity does not suffer from internal consistencies, as it cannot be performed within the effective field theories.

Although we put forward an analysis on the observer-based version of the complementarity, our analysis does not reveal how information of the black hole is carried out through evaporation process. At this point, we believe it is the absolute locality that is falsely presumed. If we are allowed to speculate here for a moment, we would state that the information transfer takes place in a non-local way, to infrared degrees of freedom, or to long wavelengths compared to the local curvature of space-time, and hence cannot be probed by local observers. There are many ideas on how absolute locality can be modified. We have covered these proposals in detail in the thesis and hence will not be referring to them once again here. The direction of the thesis is influenced strongly by the possibility that effective field theory can be modified due to non-locally encoded information, which is hidden from local physics. A similar effect can be observed in the elasticity of entangled polymers, where such systems govern two different dynamics that are decoupled due to separation of timescales, yet one can observe how long-range stringy types of degrees of freedom alter the local physics as a result of long time-scale elastic behaviour.

Having these motives, we seek information theoretic observables, derived from underlying theory. One of the goals pursued in the thesis is elaborating the microscopic mechanism, by which macroscopic objects gain inertia. It is an extremely hard task to identify these observables simply via theory of relativity, as it does not provide any microscopic description. Hence, we follow a strategy on applying lessons learned from holography, where gravitational theory has a corresponding dual microscopic description. Holographic paradigm states that entanglement entropies of the dual microscopic description have been encoded as areas of minimal surfaces in the gravitational description. Although duality is limited to the minimal surfaces, it is likely that information theoretic correspondences of more general space-like surfaces are encoded in a more complicated manner in the microscopic theory. This idea reflects itself in the notion of differential entropy, as well as its connection to integral geometry.

We further investigated how matter interacts with space-time from an information theoretic perspective. We have observed that the Bekenstein bound in gravity becomes apparent, as the bound on the entropy between complementary regions in the underlying theory increases. We have also indicated quantitative differences between pure states and mixed state excitations and assessed their impact on the gravitational theory. We believe our findings play an important role for understanding the entropic modifications of gravity in de Sitter space, as advocated recently by Verlinde. Unfortunately, most of our conclusions are limited to spherically symmetric systems. In a way, this is to be expected, as the Bekenstein bound also secretly assumes this symmetry in its formulation. In other words, the factor of 2π in the bound is the result of spherical symmetry. This has been clearly verified in the QFT derivation of the bound using the positivity of relative entropy. The novelty of our approach is the connection of the bulk version of the bound and the dual conformal field theory.

In emergent phenomena, the collective behaviour of underlying constituents define a notion of entropy. For example, in thermodynamics, entropy quantifies ignorance regarding the state of the microscopic system. At a linear level, energy and entropy are tightly related; this relation is called the first law. The connection between gravity and emergent behaviour was first observed through the discovery of black hole entropy, soon after the first law of black hole mechanics was formulated as being the relation between mass and the surface area of the black hole. In this thesis, we seek a generalization of the first law on general space-like surfaces. The reason behind such an investigation is that within the holographic context, it is understood that the first law of entanglement entropy in the underlying microscopic theory reflects itself as the first law on minimal surfaces. A natural question that followed was whether such a relation can be extended to any surface.

One recurring problem we encountered during our research was difficulty in comparing space-times of different states that are connected through an energetic deformation. This obstruction is a natural result of the diffeomorphism invariant formulation of relativity. There is no natural reference point to identify a surface and study its change under deformations. This issue is less problematic in the framework of holography, as one can use the boundary of spacetime as a natural reference point. Recent developments within the holographic paradigm make it possible to identify not only boundary anchored minimal surfaces, but also general space-like surfaces. These possibilities provides us with the opportunity to push the first law of entanglement entropy in the boundary as a first law of differential entropy into the bulk. This first law manifests itself as the reaction of area to the energy. Moreover, we observed that contrary to the general opinion on energetic excitations causing area deficits, we have shown that when referenced to the boundary, there are area excesses. It indicates that an area deficit is a somewhat arbitrary notion, as it depends on the identification scheme of a deformed solution with the initial one.

After demonstrating that the first law of entanglement can be moved from minimal surfaces to generic space-like surfaces via integral geometry in 3D, we change our direction to more general solutions in any dimensions, rather than anti de Sitter space.

The mathematics behind the first law of black hole mechanics is best understood within the covariant phase space formalism. The formalism unifies different theories of gravity from the perspective of emergent laws of black hole mechanics. Further, it allows one to study first law-type relations on surfaces other than black hole horizons. This is a freedom we have exploited heavily in this thesis. We have formulated the following two questions: firstly, what would be the consequences of applying this formulation on general surfaces from the point of view of emergent laws? And secondly, what would this teach us regarding the notion of inertia? In the last part of the thesis, we have answered these questions. We reformulated and elaborated on the entropic gravity. In this reformulation, the emergence of inertia is modelled as the reaction force to an adiabatic deformation of the system. In other words, the system preserves its state by reacting to slow deformations. The idea of gravity being an emergent phenomena, naturally proposes a distinction between timescales of underlying a microscopic system and the macroscopic probe interacting with it. A microscopic system consists of Planckian degrees of freedom. Hence, its dynamics is much faster, when compared to a macroscopic probe. It is this separation of timescales that allows a fast system to react to changes caused by a slow system, and adjust its energy, such that the so-called adiabatic invariants of the system are preserved. When the location of the macro-probe is considered to be a parameter in the state of the system, we have shown that Newtons law of inertia is equivalent to the adiabatic first law. The adiabatic first law casts a relation between change in the energy, micro-canonical entropy and force, in a way similar to the first law of thermodynamics.

In addition, we have postulated that one can associate a measure for the volume of a phase space of space-time regions enclosed by closed surfaces. Interestingly, this measure corresponds to the micro-canonical action of gravity. It is the invariance of the volume of the phase space that yields a reaction force, due to the adjustment on the energy of the system. We have shown that covariant phase space formalism provides a canonical method to measure changes in the volume of a phase space of regions in space-time. This notion generalizes the black hole entropy.

Another interesting finding of our research is that the first law of black hole mechanics is a particular, special form of a more general first law, which we named the first law of space-time deformations. It is known that gravity and elasticity have intuitive similarities, specifically that general relativity models gravity as an elastic deformation of space-time. Our findings reflect these properties from a more information theoretic perspective. We observed that, gravity is dimensionally projected elasticity on two co-dimensional surfaces. We believe the reason behind such dimensional reduction is the principle of locality and the local Lorentz symmetry of the theory. Moreover, diffeomorphism invariance embedded in general relativity enforces this elastic nature to be in the integral form, rather than a local one.

Future directions:
Through this research, while seeking a generalization of the first law of black hole mechanics onto general surfaces, we end up in an analogy between gravity and elasticity, purely based on information theoretic arguments. The main quantitative information theoretic difference between them is that entropy scales with the volume of the system in entropy, while it scales with area in gravity. Our universe has thermal character, yet it is not clear whether this thermal nature is localized on the cosmological horizon or distributed in the bulk. If it is the latter, one can argue that due to its volume, like entropy, it governs elastic behaviour. We believe our research may provide methods and frameworks to study the relation between elastic and gravitational phases in a covariant fashion.

Another interesting direction for investigation would be the second law of space-time mechanics. In other words, what the generalization of the second law of black hole mechanics to general surfaces would be. If we allow ourselves to speculate here for a moment, we would point out our suspicion regarding the connection between the second law and the fact that the classical solution of the action is an extremum in the space of possible field configurations. The hint resides in the identification of micro-canonical action with the Gibbs volume entropy. We would like to investigate this possibility in future research.

Acknowledgement

It would not have been possible to see the end of this journey without the support of many others.

Firstly, I would like to thank my supervisor, Erik Verlinde, for giving me this opportunity. I am grateful to him for sharing his knowledge, enthusiasm and joy while doing research and for giving me freedom to follow my own interests and opening his research program to me. I am indebted for his efforts on encouraging me to be more independent and strong person.

I would also like to thank my first collaborator I-Sheng Yang, who engaged me into a conceptual yet clear and solid research project. I am truly thankful for his patience and honesty towards me and to our research.

During my time in physics, perhaps Nava is the person that I am most strongly coupled to. He was a great colleague and friend. I am grateful for our discussions, for his proof readings on most of this thesis and my papers, and mostly being a strong support on difficult times. I have learned much from him not only about physics but also about life.

It was always a motivation boost to talk about physics with Fotis. It is a great pleasure to still have him as a close friend. I am thankful for his relaxed attitude. Two other close friends I am truly indebted are Laurens and Francesca. Both of them provided their genuine support in most difficult times. It was always a relief and joy to talk with them. I am thankful for their presence.

I would also like to thank Manus Visser for his collaboration.

As a not very organized person, I would like to give a special thanks to the support staff at IOP, especially Anne-Marieke, Joost, Natalie, Fatima, and Yocklang, who provide their help with care and kindness regarding the issues that I had encountered during daily life in the institute.

I would like to thank Reisa for helping me in the dutch translation of the summary of this thesis.

I would like to thank, Ata Aydin for his advises and always encouraging me on this path.

I would like to thank my mother and brother, whose support I had always felt along this journey.

Finally I would like to thank my partner in life, Ceren. I can not express how grateful I am for her presence in my life.