Momentum-space entanglement and the gravity of entanglement in AdS/CFT

by

Michael M^{c}Dermott

B.Sc., Simon Fraser University, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Physics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

September 2014

© Michael M^{c}Dermott 2014

Abstract

In the first part of this thesis we explore the entanglement structure of relativistic field theories in momentum space. We discuss a Wilsonian path integral formulation and a perturbative approach. Using perturbation theory we obtain results for specific quantum field theories. These are understood through scaling and decoupling properties of field theories. Convergence of the perturbation theory taking loop diagrams into account is also discussed. We then discuss the entanglement structure in systems where Lorentz invariance is broken by a Fermi surface. The Fermi surface helps the convergence of perturbation theory and entanglement of modes near the Fermi surface is shown to be amplified, even in the presence of a large momentum cutoff.

In the second part of this thesis we explore the connection between entanglement and gravity in the context of the AdS/CFT correspondence. We show that there are certain thermodynamic-like relations common to all conformal field theories, which when mapped via the AdS/CFT correspondence to the bulk are tantamount to Einstein's equations, to lowest order in the metric.

Preface

This thesis contains some of the work published by the author of this dissertation which has previously appeared in the peer reviewed journals Physical Review D and the Journal of High Energy Physics.

The subject of Chapter 2 was published in Physical Review D [74] under the title "Momentum-space entanglement and renormalization in quantum field theory". This work, which has been edited, was a collaboration between the author of this dissertation, the author's supervisor Mark Van Raamsdonk and Vijay Balasubramanian, a professor at the University of Pennsylvania.

The subject of Chapter 3 was published in the Journal of High Energy Physics [24] under the title "Momentum-space entanglement for interacting fermions at finite density." This work, which has been edited, was a collaboration between the author of this dissertation, the author's supervisor Mark Van Raamsdonk and Leo Hsu, an undergraduate student at the time.

The subject of Chapter 4 was published in the Journal of High Energy Physics [26] under the title "Gravitational Dynamics From Entanglement "Thermodynamics"." This work, which has been edited, was a collaboration between the author of this dissertation, the author's supervisor Mark Van Raamsdonk and Nima Lashkari, a post-doctoral researcher.

Table of Contents

Abstract ii

Preface iii

Table of Contents iv

List of Tables viii

List of Figures ix

Glossary xi

Acknowledgements xii

Dedication xiii

1 Introduction 1

1.1 Entanglement 2

1.1.1 Quantum entanglement 2

1.1.2 Measures of entanglement 3

1.1.3 Entanglement in field theory 12

1.2 AdS/CFT 17

1.2.1 Holography 17

1.2.2 Why AdS? 20

1.2.3 Matching degrees of freedom 21

1.2.4 Partition functions 22

1.2.5 Scalar field on AdS 24

1.2.6 Geometric entropy in AdS/CFT 27

1.3 Motivation 31
1.4 Outline of thesis 35

2 Momentum-space entanglement and renormalization in quantum field theory 38
2.1 Introduction 38
2.2 The low-energy density matrix 40
2.3 Measures of entanglement 44
2.3.1 Entanglement observables in perturbation theory . . 46
2.3.2 Entanglement observables in quantum field theory . . 49
2.4 Scalar field theory: entanglement between scales 52
2.4.1 The ϕ^3 theory in 1+1 dimensions 54
2.4.2 The ϕ^3 theory in higher dimensions 57
2.4.3 ϕ^4 theory 59
2.4.4 General remarks 60
2.5 The extent of entanglement between scales 63
2.5.1 An aggregate measure of the range of entanglement . 63
2.5.2 Single mode entanglement 66
2.5.3 Mutual information between individual modes 68
2.5.4 Convergence and validity of leading order expressions. 69
2.6 Entanglement in the Thirring model 71
2.6.1 The correspondence 72
2.6.2 The ground state 72
2.6.3 Thirring ground state 74
2.6.4 Checks 75
2.6.5 Scaling 78
2.6.6 Direct calculation 81
2.7 Comments 86

3 Momentum-space entanglement for interacting fermions at finite density 91
3.1 Introduction 91
3.2 Momentum-space entanglement in perturbation theory . . . 94

3.3 Entanglement entropy for lattice fermions with nearest neighbor interactions . . . 97
3.4 Entanglement entropy for continuum non-relativistic fermions 101
3.4.1 Mutual information between modes . . . 103
3.5 Entanglement entropy for relativistic fermions . . . 105
3.6 Discussion . . . 113

4 Gravitational Dynamics From Entanglement "Thermodynamics" . . . 114
4.1 Introduction . . . 114
4.2 Entropy-energy relation . . . 116
4.3 Gravitational implications of $dS = dE$ in holographic theories 118
4.3.1 Gravitational calculation of dS . . . 118
4.3.2 Gravitational calculation of dE . . . 120
4.4 Derivation of linearized Einstein's equations from $dE = dS$. 121
4.4.1 Proof that $\delta S = \delta E$ for solutions of Einstein's equations . . . 122
4.4.2 Proof that $\delta S = \delta E$ implies the linearized Einstein's equations . . . 124
4.5 Discussion . . . 128

5 Outlook . . . 131

Bibliography . . . 133

Appendices

A . . . 142
A.1 Momentum-space entanglement and correlators . . . 142
A.2 Entanglement entropy in a fermionic system . . . 143

B . . . 145
B.1 Two qubit system example . . . 145

C 150
C.1 Alternative derivation of linearized Einstein's equations from $\delta E = \delta S$ 150
C.1.1 Expansion of $\delta E = \delta S$ in powers of R 150
C.1.2 Checking that solutions of Einstein's equations satisfy $\delta S = \delta E$ 151

List of Tables

2.1 Spatial dimensions where momentum space mutual information and entanglement entropy converge. The results apply for any bounded regions A and B in momentum space. 71

List of Figures

1.1 Mutual information 5
1.2 Entanglement entropy in a two spin system. 9
1.3 Density matrix path integral boundary conditions 14
1.4 Reduced density matrix path integral boundary conditions . . 14
1.5 Diagram depicting Bekenstein-Hawking entropy as maximal . 18
1.6 Diagram illustrating some requirements for a holographic theory 19
1.7 Diagram illustrating the holographic dual of geometric entropy 27
1.8 Diagram illustrating area law divergence 30
1.9 Diagram used in strong subadditivity proof 32
1.10 Ryu-Takayanagi surface for global AdS 33
1.11 UV/IR connection in AdS/CFT 34
1.12 Figure of possible momentum space Ryu-Takayanagi relation 35
1.13 Regions of spacetime separating due to loss of entanglement . 36

2.1 Leading contributions to $S(\mu)$ for ϕ^3 theory in 1+1 dimensions. Full result for $S(\mu)$ is proportional to $\lambda^2(\log(1/\lambda^2)+1)$ times bottom function plus λ^2 times top function. 57
2.2 (**A**) Integration regions for ϕ^3 theory in 2+1 dimensions. (**B**) The function $F(x)$ appearing in the entanglement entropy for ϕ^4 theory in $1+1$ dimensions. 58
2.3 Ratio of first and second terms in (2.53) vs $\mu = (\mu_2 - \mu_1)/m$ for (**A**) $\mu_1 = 1$ and (**B**) $\mu_1 = 4$. This is a measure of the range of entanglement in ϕ^4 theory in $1+1$ dimensions. We have taken the mass to be $m = 1$. 66

2.4 Single-mode entanglement entropy vs magnitude of mode momentum for ϕ^3 field theory in 1+1 (bottom), 2+1 (middle), and 3+1 (top) dimensions. The entropies are normalized by their values at $p = 0$. 68

3.1 Leading perturbative contribution to single mode entanglement entropy $S(k)$ as a function of mode momentum for lattice fermions at half filling. This diverges logarithmically at the Fermi points $k = \pm\pi/2$. 100
3.2 Leading perturbative contribution to single mode entanglement entropy $S(p_F k)$ as a function of mode momentum (as a fraction of the Fermi momentum) for weakly interacting continuum non-relativistic fermions. This diverges logarithmically at the Fermi points. 102
3.3 Regions of (k, l) space with different behaviors for $\mathcal{I}(k, l)$. Results for the unshaded regions may be obtained from the results for the shaded regions using the indicated symmetries of $\mathcal{I}$. 104
3.4 Mutual information $\mathcal{I}(p_F k, p_F l)$ between individual modes for non-relativistic fermions with Fermi momentum p_F. 106
3.5 Mutual information $\mathcal{I}(p_F k, p_F l)$ vs l for non-relativistic fermions with $k = 0$ (top left), $k = 0.75$ (top right), $k = 0.95$ (bottom left), and $k = 1.25$ (bottom right). The overall scale for $\mathcal{I}$ is arbitrary 107
3.6 Mutual information $\mathcal{I}(p, q)$ vs q for relativistic fermions with $\mu = 10m$ and $p = 0$ (top left), $p = 7.5$ (top right), $p = 9.5$ (bottom left), and $p = 12.5$ (bottom right). Momenta are given in units of m, so the Fermi points are at ± 10. The overall scale for $\mathcal{I}$ is arbitrary. 111
3.7 Mutual information $\mathcal{I}(p, q)$ as a function of p and q for relativistic fermions with $\mu = 0$. Momenta are given in units of m. 112

Glossary

QFT	Quantum field theory
CFT	Conformal field theory
AdS	Anti-de-Sitter
AdS/CFT	The correspondence between gravitational theories on asymptotically AdS spacetime and CFT on the boundary
DOF	Degrees of freedom
IR	Infrared
UV	Ultraviolet
LHS	Left hand side
RHS	Right hand side
RT	Ryu-Takayanagi
BH	Bekenstein-Hawking
EPR	Einstein, Podolski and Rosen

Acknowledgements

I wish to sincerely and gratefully thank my supervisors Mark Van Raamsdonk and Gordon Semenoff.

It is to Gordon's energy and brilliance that I owe my original impetus to explore string theory, his joy in physics is delightfully contagious.

Without Mark's intellectual force this thesis would not have been possible. His patient guidance and lucid explanations have been more than indispensable, they have been appreciated.

I would also like to express gratitude to Louis Deslauriers, Moshe Rozali and Ariel Zhitnitsky.

Dedication

To my Family.

Chapter 1

Introduction

The concept of entanglement and the oddities it entails were famously considered in a paper by Einstein, Podolsky and Rosen (EPR) [16]. Making use of an entangled state they argued that a quantum mechanical description of reality could not be considered complete. It was thus their view that even though it accurately captured some aspects of physical reality, there were others that it was *in principle* incapable of describing.

It wasn't until 29 years later that Bell [4] studied what a completion of quantum mechanics would imply. He termed a theory with completeness in the sense demanded by EPR a local hidden variable model. Such a theory is local so that spacelike separated events and observations are independent, it has definite physical properties (some of which are "hidden" from quantum mechanics) *before* measurement and furthermore these are unaffected by the measurement. Using these assumptions he was able to show that a local hidden variable model would give rise to experimental results inconsistent with quantum mechanics. Interestingly, the difference in predictions was entirely due to the possibility of entanglement in the quantum description.

Many experiments, e.g [2], have since been performed, excluding the possibility of local hidden variable models. It would thus appear that entanglement is a physical aspect of reality. This means that systems which are far apart really can display "spooky action at a distance" [1] whereby the measurement of systems separated by incredibly large distances can "have an effect" on each other. Furthermore, complete knowledge of the state of a physical system does not in fact imply knowledge of the state of its constituent subsystems.

Bizarre as its properties may be, entanglement has become a very im-

[1] This sentiment was expressed by Einstein in a correspondence with Max Born.

portant topic in physics in the last 20 years and has found many useful applications. A few key areas where entanglement plays a fundamental role are quantum cryptography [5], quantum teleportation [6], and in a very general sense quantum information theory. It has been used as a tool for understanding disordered systems with long ranged interactions [15], superconductivity [42], and properties of frustrated system ground states [13]. It also serves as an order parameter in understanding quantum phase transitions, see e.g. [43] and references therein.

The main motivation for the study of entanglement in this thesis, however, comes in the context of AdS/CFT. This will be discussed at the end of the chapter after a brief review of some salient features of entanglement and AdS/CFT. Throughout the thesis we use units where $c = \hbar = k = 1$.

1.1 Entanglement

1.1.1 Quantum entanglement

Quantum mechanics as a theory is incredibly successful, however, it applies to scales far outside the realm of human experience. As a result, it has many counter-intuitive features such as quantization, superposition, tunnelling and summing over paths. Not least of these features is quantum entanglement, which has no classical counterpart. To gain an understanding of this consider a physical system which may be divided into multiple (k) subsystems.

As a concrete classical example this could be k isolated boxes of identical gas, each with number of particles n_k. Then we can specify the classical state of the system by giving the position and momenta of all the particles in each box. So denoting the list of coordinates in box j as $\{\vec{q}_i, \vec{p}_i\}_j$, the total state of the system would be $[\{\vec{q}_{i_1}, \vec{p}_{i_1}\}_1, \{\vec{q}_{i_2}, \vec{p}_{i_2}\}_2, ..., \{\vec{q}_{i_k}, \vec{p}_{i_k}\}_k]$.

The quantum mechanical description of k subsystems will, however, be quite different. Suppose that the j-th subsystem has a complete basis of states $|s^{(j)}{}_i\rangle$. Then because of the superposition principle an arbitrary state

$|\phi\rangle$ of the system may be written as:

$$|\phi\rangle = \sum_{i_1,\ldots,i_k} c_{i_1,\ldots,i_k} |{s^{(1)}}_{i_1}\rangle \otimes |{s^{(2)}}_{i_2}\rangle \otimes \ldots \otimes |{s^{(k)}}_{i_k}\rangle. \tag{1.1}$$

The first difference to note between the classical and quantum case is that, for the classical case the total state is just a product of the states of the subsystem. The quantum system will generally not be writable as a product of states of the subsystem $|\phi\rangle = |\phi_1\rangle \otimes |\phi_2\rangle \otimes \ldots \otimes |\phi_k\rangle$. We are unsure about the state of the subsystems, the state is entangled.

The second thing to note, which is closely related to the first, is that in the classical system knowledge of the state of the larger physical system implies knowledge of the state of the subsystems. In an entangled quantum state, however, we run into the strange situation where complete knowledge of the large physical system (i.e. specifying its state as in (1.1)) *does not* imply knowledge of the states of the subsystems.

1.1.2 Measures of entanglement

Shannon entropy

The Shannon entropy is a classical entropy for a system with random variables X. Thus event X_i occurs with probability p_i, and the Shannon entropy is defined as:

$$H(X) = -\sum_i p_i \log p_i. \tag{1.2}$$

It measures the amount of information gained about X, on average, when an event occurs. Equivalently, and this view of entropy will generally be more useful to us, the Shannon entropy quantifies the uncertainty in X we have before an event occurs. [2] The statement that the Shannon entropy measures the average amount of information gained when we learn the value of a random variable X can be heuristically "derived" from the following

[2] A nice illustration to have in mind is a dice thrown under a box. The entropy measures our lack of knowledge about what the dice will read when we lift the box.

arguments. First one demands continuity: we want a measure of gained information $G(p)$ to depend continuously on the probabilities of the random variable. Second one wants additivity of information: the amount of information gained from knowledge of two independent events simultaneously (with probability $p_1 p_2$) should be the same as the sum of the information gained from individual events separately:

$$G(p_1 p_2) = G(p_1) + G(p_2). \tag{1.3}$$

From this it follows immediately that $G(p) = G(1) + G(p)$ so that $G(1) = 0$, which is saying that the amount of information gained from an event certain to happen is 0. Next, taking $p_2 = c/p_1$ (where $0 \leq c \leq p_1$ is introduced to keep $p_2 \leq 1$) we have $G(p) = -G(1/p)$. Making use of continuity we also have

$$\begin{aligned} G'(p)dp =& G(p + dp) - G(p) && (1.4)\\ =& G(1 + \frac{dp}{p}) && (1.5)\\ =& G(1) + \frac{G'(1)}{p}dp && (1.6)\\ =& \frac{G'(1)}{p}dp && (1.7) \end{aligned}$$

Integrating both sides with $G(1) = 0$ gives $G(p) = G'(1)\log(p)$. Note that it makes sense to gain more information from an unlikely event. To get the average gain we simply weight each gain by the probability of it occuring:

$$\langle G \rangle = G'(1) \sum_i p_i \log(p_i). \tag{1.8}$$

Thus we see that the Shannon entropy is proportional to the average gain in information when we learn a random variable.

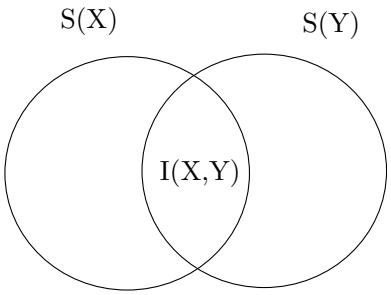


Figure 1.1: Diagram illustrating the mutual information.

Classical mutual information

Since mutual information deals with two sets of random variables X and Y we define the joint probability $p(x, y)$ (probability of obtaining both x and y). Then the joint Shannon entropy (information gained on average by obtaining a joint event) is simply:

$$H(X, Y) = -\sum p(x, y) \log p(x, y), \tag{1.9}$$

the sum being over all possible events.

The mutual information between two random variables X and Y can then be defined as:

$$I(X, Y) = H(X) + H(Y) - H(X, Y). \tag{1.10}$$

Any information which is gained in the independent regions is canceled out (counted twice in $H(X)$ together with $H(Y)$ and twice in $H(X, Y)$) and any information gained in the mutual region is counted once in total (twice in $H(X)$ together with $H(Y)$ and once in $H(X, Y)$). Thus the mutual information is a measure of how much information is in common between X and Y, see Figure 1.1.

To illustrate mutual information consider an example involving two 3-sided dice.

1. The two dice are thrown independently. Then the probability distri-

butions are independent $p(x, y) = p(x)p(y)$ which means $H(X, Y) = H(X) + H(Y)$. Thus in this case the mutual information $I_1(X, Y)$ vanishes.

2. The two dice are attached in such a way that they always come up with the same value. In other words they are the "same". The joint options are $(1, 1), (2, 2)$ and $(3, 3)$. In terms of probabilities, this is the same as the individual options $(1), (2)$ and (3), and so $H(X, Y) = H(X) = H(Y)$. We then have $I_2(X, Y) = H(X) = H(Y) = H(X, Y) = \log(3)$. Thus the mutual information in the case where the two random variables are perfectly correlated (and so are basically the same) is just the total information.

3. The two dice are arranged in such a way that whenever a 1 is rolled on one die, the other also rolls a 1. There are 5 options for the joint distribution: $(1, 1), (2, 2), (2, 3), (3, 2), (3, 3)$. Then $I_3(X, Y) = H(X) + H(Y) - H(X, Y) = \log(3) + \log(3) - \log(5) = \log(9/5)$. Thus in the intermediary case where the two dice aren't independent but aren't completely correlated either, the mutual information is between that of the two limiting cases: $I_1 < I_3 < I_2$.

Von Neumann entropy

In classical mechanics systems which are in a statistical mixture of classical states x are described by a probability distribution $p(x)$. Expectation values of functions of these classical variables are simply obtained as:

$$\langle F(x) \rangle = \sum_x p(x)F(x). \tag{1.11}$$

The situation is different in quantum mechanics. Say a physical system is in a mixture of quantum states $|\phi_i\rangle$ which form a complete basis, each with probability p_i. Then if O_i is the expectation value of some physical observable O in the state i, then its expectation value in the statistically mixed state is:

$$\langle O\rangle = \sum_i p_i O_i \tag{1.12}$$
$$= \sum_i p_i \langle\phi_i|O|\phi_i\rangle \tag{1.13}$$
$$= \sum_i p_i \langle\phi_i|O| \sum_n |n\rangle\langle n|\phi_i\rangle \tag{1.14}$$
$$= \sum_n \langle n| \sum_i p_i |\phi_i\rangle\langle\phi_i|O|n\rangle \tag{1.15}$$
$$= \mathrm{tr}\,(\rho O) \tag{1.16}$$

where we have defined $\rho = \sum p_i|\phi_i\rangle\langle\phi_i|$. Thus a quantum mixed state is described by a density matrix ρ, which encodes the classical probability p_i of a system being in a quantum mechanical state $|\phi_i\rangle$.

We can associate an entropy to the mixed state, the Von Neumann entropy:

$$S(\rho) = -\,\mathrm{tr}\,\rho\log\rho = -\sum p_i \log p_i. \tag{1.17}$$

Notice that in the eigenbasis, this is the Shannon entropy for the eigenvalues. Thus in the eigenbasis of the density matrix, the Von Neumann entropy quantifies how uncertain we are about which state the system is in.

Entanglement entropy

Consider a physical system which can be decomposed into two subsystems A and its complement $\bar{A}$, so that its Hilbert space is a product $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$. In such a case we can form a reduced density matrix ρ_A^R for degrees of freedom in the subsystem A by tracing the full density matrix $\rho_{A\bar{A}}$ over degrees of freedom in the subsystem $\bar{A}$:

$$\rho_A^R = \mathrm{tr}_{\bar{A}}\,\rho_{A\bar{A}}. \tag{1.18}$$

Note that if $\rho_{A\bar{A}}$ satisfies all the properties of a density matrix, then so will ρ_A^R. Physically, ρ_A^R is the density matrix for an observer who only has access

to degrees of freedom in A. [3]

We are now ready to define the entanglement entropy. Assuming the full system starts in a pure state $|\psi\rangle$, so that $\rho_{A\bar{A}} = |\psi\rangle\langle\psi|$, then *the entanglement entropy is the Von Neumann entropy for the reduced density matrix* $\rho_A = \mathrm{tr}_{\bar{A}} |\psi\rangle\langle\psi|$.

As an example consider a system of two sites with spin 1/2. Prepare the system in the one parameter family of states:

$$|\psi\rangle = \sqrt{u}\,|\uparrow\uparrow\rangle + \sqrt{1-u}\,|\downarrow\downarrow\rangle\,, \tag{1.19}$$

with $u \in [0,1]$. Notice that at $u = 0$ and $u = 1$ the system is in a product state and so is not entangled. At $u = 1/2$ each subsystem is equally likely to be spin up or down and the system is maximally entangled.

Let us calculate the entanglement entropy. The reduced density matrix (for either spin) is:

$$\rho_{s_{1/2}} = \begin{pmatrix} u & 0 \\ 0 & 1-u \end{pmatrix} \tag{1.20}$$

Thus the entanglement entropy is

$$S(\rho_{s_{1/2}}) = -u\log u - (1-u)\log(1-u) \tag{1.21}$$

This is plotted in Figure 1.2. The entanglement entropy vanishes for $u = 0$ and $u = 1$ and attains its maximum at $u = 1/2$, which is what we expected from considering the entanglement in the original state.

Entanglement between mixed states

In order to discuss mutual information let us first discuss the concept of entanglement between mixed states. Two subsystems A and B in mixed

[3] An example of such a scenario is an accelerated Minkowski observer. Due to the acceleration there is an entire half of the space from which the observer may never receive light signals, and as such the state is properly described by a reduced density matrix. This has an observable consequence in the Unruh effect, where the observer detects radiation at a temperature proportional to his/her acceleration [40] .

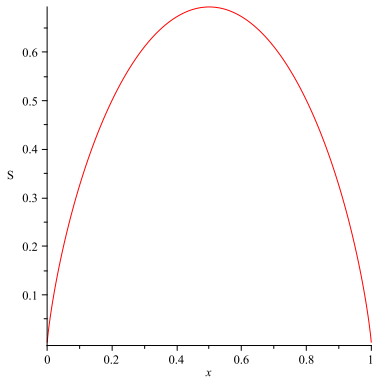


Figure 1.2: Entanglement entropy in a two spin system.

states ρ_A and ρ_B are said to be entangled with each other if we *cannot* write the joint state ρ_{AB} as [23]:

$$\rho_{AB} = \sum_i p_i \rho_A^i \otimes \rho_B^i, \tag{1.22}$$

where p_i are probabilities, and ρ_A^i and ρ_B^i are some density matrices for degrees of freedom in A and B respectively.

This can be understood as follows [44]. Consider a system with Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. An uncorrelated state is one where each subsystem is in a state independent of the other, i.e.

$$\rho = \rho_A \otimes \rho_B \tag{1.23}$$

where ρ is the full state. An uncorrelated state has the property that expectation values of joint observables $\mathcal{O} = \mathcal{O}_A \otimes \mathcal{O}_B$ factorize:

$$\begin{aligned} \langle \mathcal{O} \rangle &= \mathrm{tr}(\rho \mathcal{O}) && (1.24)\\ &= \mathrm{tr}(\rho_A \mathcal{O}_A \otimes \rho_B \mathcal{O}_B) && (1.25)\\ &= \langle \mathcal{O}_A \rangle_A \ \langle \mathcal{O}_B \rangle_B && (1.26) \end{aligned}$$

We can create a situation where the states of the two subsystems are correlated. Simply have a random generator output numbers $i = 1, ..., n$ with probability p_i. When i is generated, prepare the system in a state $\rho_A^i \otimes \rho_B^i$. In this situation the expectation for $\mathcal{O}$ should be:

$$\begin{aligned} \langle \mathcal{O} \rangle &= \sum_i p_i \, \mathrm{tr}(\rho_A^i \mathcal{O}_A) \, \mathrm{tr}(\rho_B^i \mathcal{O}_B) && (1.27)\\ &= \mathrm{tr}(\sum_i p_i \rho_A^i \otimes \rho_B^i \ \mathcal{O}_A \otimes \mathcal{O}_B) && (1.28)\\ &= \mathrm{tr}(\rho_c \mathcal{O}). && (1.29) \end{aligned}$$

Thus we conclude the system is in a correlated state:

$$\rho_c = \sum_i p_i \rho_A^i \otimes \rho_B^i. \tag{1.30}$$

The important point to note is that the device used is simply a classical number generator (it could have been rolling a weighted n-side die) and so these correlations are classical. Now we see why density matrices which cannot be written as ρ_c are entangled. Density matrices which can be written as ρ_c are termed separable.

Quantum mutual information

We are now ready to define mutual information. Take a system admitting a Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$. Then the mutual information between degrees of freedom in A and B is defined analogously to the classical case, in terms of the entanglement entropy as:

$$I(A, B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}). \tag{1.31}$$

Now, for any regions R_1 and R_2 for which entanglement entropy S can be defined, it satisfies certain inequalities, one of the more useful being strong subadditivity which states that:

$$S_{R_1} + S_{R_2} \geq S_{R_1 \cup R_2} + S_{R_1 \cap R_2} \tag{1.32}$$

from which directly follows subadditivity of the entropy

$$S_{R_1} + S_{R_2} \geq S_{R_1 \cup R_2}. \tag{1.33}$$

This last inequality implies that $I \geq 0$. It is also true that the mutual information saturates the bound $I = 0$ iff $\rho_{AB} = \rho_A \otimes \rho_B$ [4], i.e. the

[4] This follows as a simple application of Klein's inequality, which states that for two density matrices ρ and d we have

$$\operatorname{tr} \rho \log \rho - \operatorname{tr} \rho \log d \geq 0, \tag{1.34}$$

with the bound saturated iff $\rho = d$. To prove subadditivity set $\rho = \rho_{AB}$ and $d = \rho_A \otimes \rho_B$.

mutual information vanishes only when there is no classical correlation or entanglement in the state ρ_{AB}. Thus the mutual information measures all the correlations between A and B, both classical and quantum mechanical. This has been shown rigorously in [19].

1.1.3 Entanglement in field theory

Normalized replica method

Consider an unnormalized density matrix ρ, the normalized density matrix being $\hat{\rho} = \rho/\operatorname{tr}\rho$. Then

$$-(\frac{d}{dn} - 1)\ln \operatorname{tr}\rho^n\Big|_{n=1} = -\frac{1}{\operatorname{tr}\rho^n}\operatorname{tr}(\ln\rho\ \rho^n)\Big|_{n=1} + \ln\operatorname{tr}\rho \tag{1.39}$$

$$= -\operatorname{tr}(\hat{\rho}\ln\hat{\rho}) \tag{1.40}$$

We thus have

$$S(\hat{\rho}) = -(\frac{d}{dn} - 1)\ln\operatorname{tr}\rho^n\Big|_{n=1}. \tag{1.41}$$

This is a very useful relation since it allows us to determine the entropy of $\hat{\rho}$ from $\operatorname{tr}\rho^n$ (since we are differentiating with respect to n, we will need this as a function of n). This is known as the replica method for computing the entropy of a density matrix. Notice the other nice feature which is that we need not worry about the normalization of the density matrix, since the quantity in (1.41) always computes the entropy for the corresponding normalized state.

Then the Klein inequality gives

$$S(\rho_{AB}) \leq -\operatorname{tr}\rho_{AB}\log\rho_A\otimes\rho_B \tag{1.35}$$

$$= -\operatorname{tr}\rho_{AB}\log\rho_A - \operatorname{tr}\rho_{AB}\log\rho_B \tag{1.36}$$

$$= -\operatorname{tr}\rho_A\log\rho_A - \operatorname{tr}\rho_B\log\rho_B \tag{1.37}$$

$$= S(\rho_A) + S(\rho_B). \tag{1.38}$$

Thus $S(\rho_A) + S(\rho_B) \geq S(\rho_{AB})$. The equality condition is $\rho = d$ which yields $\rho_{AB} = \rho_A\otimes\rho_B$.

Let us now see how the ground state entanglement entropy between spatial regions can be calculated using this method. Locality guarantees that the Hilbert space factorizes for spatial regions. Consider for simplicity a scalar field theory in one spatial dimension with some action $\mathcal{S}(\varphi)$. The system starts in the ground state:

$$\rho = |\Omega\rangle \langle\Omega| . \tag{1.42}$$

In position basis this is

$$\rho(\varphi_a, \varphi_b) = \langle\varphi_a| \Omega\rangle\langle\Omega |\varphi_b\rangle . \tag{1.43}$$

We can represent the transition amplitude as a path integral (after Wick rotation to Euclidean space)

$$\langle\varphi_a| \Omega\rangle = \int_{t=-\infty}^{\varphi(t=0,x)=\varphi_a(x)} D\varphi \; e^{-\mathcal{S}(\varphi)} \tag{1.44}$$

The total density matrix is thus a path integral over fields φ taking values in $\mathbb{R}^2$, and with different boundary conditions at $t = 0$ depending on wether we approach from below ($\varphi(t = 0^-, x) = \varphi_a(x)$) or above ($\varphi(t = 0^+, x) = \varphi_b(x)$). This is depicted in Figure 1.3.

Now divide space into two regions A and $\bar{A}$ with x_A consisting of all the points $x \in A$ and $x_{\bar{A}}$ consisting of all the points $x \in \bar{A}$. We obtain the reduced density matrix for the region x_A by tracing out the fields in $\bar{A}$. This amounts to setting the fields equal on $(t = 0, x_{\bar{A}})$ and integrating over them in the path integral, see Figure 1.4. This path integral can be written in terms of functional delta functions as: [5]

$$\rho(\varphi_a(x_A), \varphi_b(x_A)) = \int D\varphi\delta\left(\varphi(0^-, x_A) - \varphi_a(x_A)\right) \delta\left(\varphi(0^-, x_A) - \varphi_a(x_A)\right) e^{-\mathcal{S}}. \tag{1.45}$$

[5] We need not keep track of the normalization by the vacuum partition function in this case, since our expression for entropy is independent of the normalization.

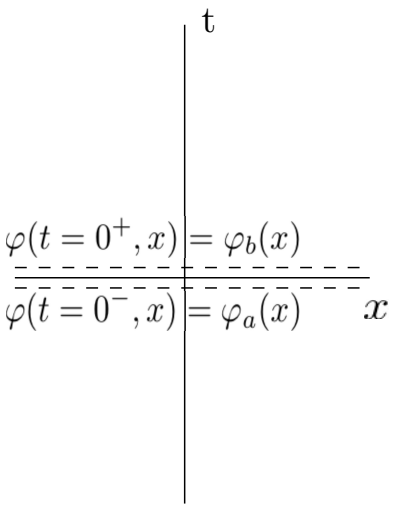


Figure 1.3: Boundary conditions for the path integral representation of the ground state density matrix.

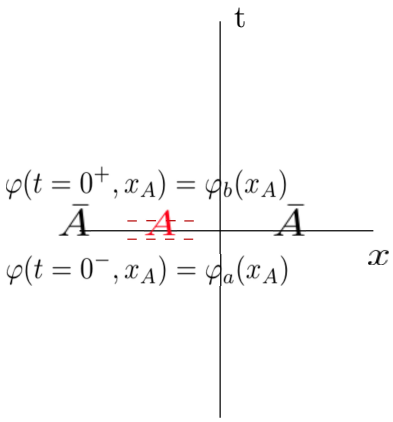


Figure 1.4: Boundary conditions for the path integral representation of the reduced density matrix.

Now in order to compute $\mathrm{tr}\,\rho_A^n$ we need:

$$\mathrm{tr}\,\rho_A^n = \int \prod_i^n D\varphi_i^A \ \rho(\varphi_1^A, \varphi_2^A)\rho(\varphi_2^A, \varphi_3^A)...\rho(\varphi_n^A, \varphi_1^A). \tag{1.46}$$

Each $\rho(\varphi_i^A, \varphi_{i+1}^A)$ is a path integral on $\mathbb{R}^2$, and the total path integral (1.46) relates them by requiring that the upper ($t = 0^+$) boundary condition for the j integral be the same as the lower ($t = 0^-$) on the $j + 1$ integral.

We can thus instead think of the full path integral (1.46) as being over a surface Σ_n, made up of n sheets, connected by these boundary conditions:

$$\mathrm{tr}\,\rho_A^n = \int_\Sigma D\varphi e^{-\mathcal{S}(\varphi)}. \tag{1.47}$$

One important point to note is that if one were to stand at the region separating A from $\bar{A}$ then one would need to travel in a circle n times to come back to the original spot. There is thus a deficit angle of $2\pi - 2\pi n$ at such locations.

Area law

Entanglement entropy in local quantum field theory, between a connected region A and its complement $\bar{A}$ is a divergent quantity, owing to the fact that in any region, no matter how small, a field theory has infinitely many degrees of freedom. If we cut off the theory in the UV by placing it on a lattice with size a then one finds [8, 33] that the entanglement entropy diverges proportionally to the area ∂A bounding the region A:

$$S(\rho_A) = k \cdot \frac{\partial A}{a^d} + O\left(\frac{1}{a^{d-1}}\right) \tag{1.48}$$

where d is the spatial dimension of the theory [6]. This result can be understood heuristically as arising because the entanglement between two spatial

[6] Note that the mutual information is a nice quantity in this regard since it will in general not suffer from this divergence. Recall that $I(A, B) = S(A) + S(B) - S(A \cup B)$, thus as long as the regions A and B don't share any area in common and divergence proportional to the area cancels out.

regions is concentrated near the boundary region dividing them.

There is, however, a violation of this area law in 1+1 dimensional conformal field theory. There one finds a logarithmic dependence on the size of the region A. If the region size is L then in a theory with central charge c one has [9, 22]:

$$S = \frac{c}{3} \log L/a. \tag{1.49}$$

A nice scaling argument [38] which captures these features of the entanglement entropy exists. The idea is to add up all the contributions to the entanglement entropy from each distance scale x. Thus we start with a microscropic Hamiltonian at some cutoff scale H_a, then we imagine changing the scale of interest to x, under which the Hamiltonian renormalizes to H_x. The theory can be considered local for scales larger than the renormalizion scale (i.e. for a lattice cutoff the theory is local at the lattice scale). Because of locality the theory will be entangled accross the boundary, but because we are considering scales larger than x this sets a minimum scale for degrees of freedom. Thus the boundary may only contribute as $\partial A^{d-1}/x^{d-1}$ to the entanglement. Couple this with the fact that in renormalization group theory the canonical measure that always appears is $d\mu/\mu$ or dx/x one gets that each scale contributes to the entropy as:

$$dS = \frac{\partial A^{d-1}}{x^{d-1}} \frac{dx}{x}. \tag{1.50}$$

The total contribution from all scales is then:

$$S = \int_{x_{UV}}^{x_{IR}} \frac{\partial A^{d-1}}{x^d} dx. \tag{1.51}$$

Taking $x_{UV} = a$ and x_{IR} large this reproduces the area divergence for $d \geq 2$. It also reproduces the conformal result for $d = 1$ since in that case the only available length scale (the system is infinite in extent) for x_{IR} is L.

1.2 AdS/CFT

Much of what is discussed in this chapter is based on excellent reviews such as [1, 14, 29–31].

1.2.1 Holography

Although the specifics of the so-called AdS/CFT duality can be quite involved, and the literature on the subject is already quite substantial and ever growing, the basic idea is very simple. It states that a quantum theory of gravity (on asymptotically Anti-de-Sitter spacetime) is exactly the same as a regular quantum field theory (with conformal symmetry) on its boundary. It relates a $d+1$ dimensional gravitational theory to at d dimensional field theory. As such it is an example of a more general idea called the holographic principle [35, 39]: that our physical universe is a "hologram" which may in fact be described by a theory in one less dimension.

The canonical example of holography, AdS/CFT, is provided by string theory. The logic of holography, however, does not rely on string theory. [7]

The first motivation for holography comes from considering the entropy of a black hole. Black holes were known to obey analogous laws to the four laws of thermodynamics, but it wasn't until Hawking [20] proved that black holes radiate a thermal spectrum that they were shown to be more than just an analogy; blackholes really were thermodynamic objects. Along with this came the famous result for the black hole entropy, the Bekenstein-Hawking entropy:

$$S_{BH} = \frac{A}{4G_N}, \tag{1.52}$$

where A is the area of the black hole. Because this is a true entropy it suggests that the degrees of freedom for a black holes scale with its *area* instead of its volume.

Furthermore the Bekenstein-Hawking entropy is actually the maximum

[7] It should be noted that in practice all known examples of holographic dualities known end up having interesting connections to string theory.

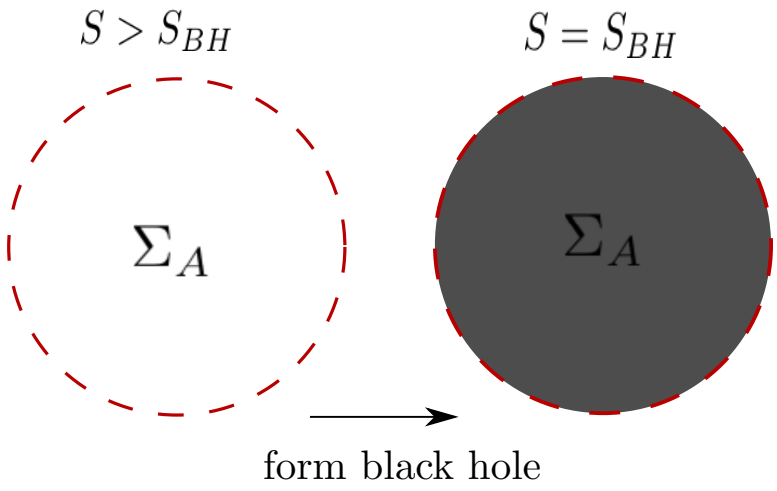


Figure 1.5: Diagram depicting Bekenstein-Hawking entropy as maximal.

entropy for a region Σ_A delimited by a sphere of area A. To see this, assume to the contrary that there is something else inside the spherical region Σ_A bounded by area A and which has entropy $S > S_{BH}$ (Figure 1.5). The starting configuration then cannot be a larger black hole since that wouldn't fit inside Σ_A, it can only be a smaller black hole or something with less energy. Thus since the starting configuration must have less energy than the black hole of size Σ_A, we can always just throw in energy until we create the black hole which fits. This process, however, can only increase the entropy, but since we end up with $S = S_{BH}$ this is a contradiction.

Then $S_{BH} = A/4G_N$ gives the maximum entropy for a region, and so, somewhat astoundingly, we conclude that the number of degrees of freedom we need to describe a gravitating region Σ_A scales with its *area* rather than its volume.

A second argument for holography can be given as follows [28]. Holography relates fields in a bulk region to fields on the boundary of this region. This on its own is perhaps not surprising since in order to describe what is happening at point A Alice can send out light signals, which will eventually

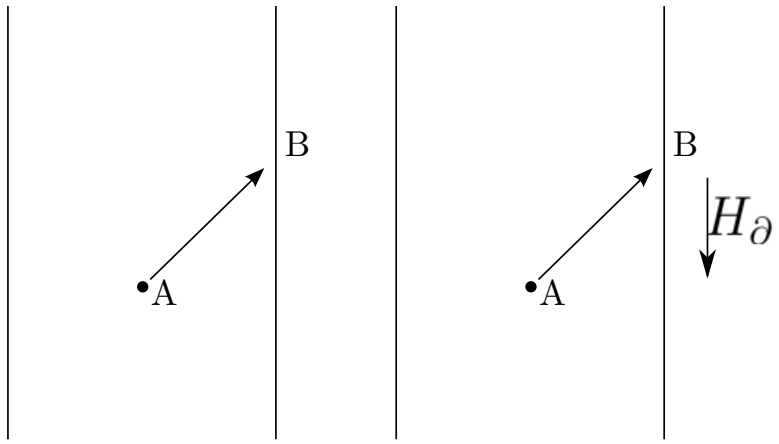


Figure 1.6: Diagram depicting some requirements for a holographic theory.

reach the boundary at point B [8] where Bob can be sitting to collect the signals. Bob can then account for the interactions with the signal on the way to him to reconstruct Alice's message. In this way we can write boundary quantities in terms of bulk ones (this is depicted in the left side of Figure 1.6).

Holography, however, demands more than this. It demands that quantities in the bulk can be written in terms of quantities on the boundary at the *same* time. This is because the boundary theory has its own Hamiltonian H_∂ which can be used to write fields at later times in terms of earlier ones (this is depicted in the right side of Figure 1.6). Typical field theories will certainly not have this stronger property [9].

Theories of quantum gravity, however, may have this property. This is because the gravitational Hamiltonian H_{QG} is known to reduce to a boundary term on shell. If this remains true in the full quantum gravity then the stronger step required by holography is easily accomplished, we can simply

[8] This is of course except in the case where Alice is inside an eternal black hole. Since we are assuming a unitary theory with no loss of information, even if Alice started in a black hole, as long as it evaporates the information would escape.

[9] Classical theories of gravity won't either since we can specify initial conditions on a time slice, so that quantities on the boundary are independent of those in the bulk.

take the observable at time t_B where Bob was and map it back to an earlier time t:

$$\mathcal{O}(t_B) = e^{iH_{QG}(t_B-t)}\mathcal{O}(t)e^{-iH_{QG}(t_B-t)}. \tag{1.53}$$

In this way one can relate boundary observables at any time to those at time t.

1.2.2 Why AdS?

In this section we wish to further motivate the correspondence by asking why it is anti-de-Sitter space in particular which is related to a conformal field theory on the boundary. Given that the $d+1$ dimensional CFT is in Minkowski space, it has the Poincaré invariant metric:

$$ds^2_{CFT} = -dt^2 + d\vec{x}^2. \tag{1.54}$$

The most general higher dimensional metric (z being the extra dimension) which respects the Poincaré symmetry is

$$ds^2_{AdS} = f(z)\left(-dt^2 + d\vec{x}^2\right) + g(z)dz^2, \tag{1.55}$$

which after a coordinate redefinition can be written:

$$ds^2_{AdS} = c(z)\left(-dt^2 + d\vec{x}^2 + dz^2\right). \tag{1.56}$$

Demanding that this be invariant under conformal transformations $(t, \vec{x}, z) \to \lambda(t, \vec{x}, z)$ (since z has dimensions of length it should transform along with boundary coordinates) requires $c(z) \propto 1/z^2$. Inserting a length scale L to make $c(z)$ dimensionless gives:

$$ds^2_{AdS} = \frac{L^2}{z^2}\left(-dt^2 + d\vec{x}^2 + dz^2\right), \tag{1.57}$$

which is the metric for pure AdS spacetime in $d+2$ dimensions. [10]

[10] This metric actually only covers the part of AdS spacetime visible to certain accelerated coordinates, which is termed the Poincaré patch. This is similar to the Rindler

Another check that it is AdS which should correspond to the CFT is that the symmetries of the theories match. In this case the isometries (transformations which leaves the metric invariant) of pure AdS_{d+2} form a $SO(2, d+1)$ symmetry, which is indeed the same as the conformal group in one temporal and d spatial dimensions. [11]

Also note that from (1.57) we can immediately obtain the so-called UV/IR correspondence in AdS/CFT which states that high energy excitations in CFT are related to length scales far away from the interior of AdS (near the boundary). From (1.57) we have

$$\frac{1}{E_{AdS}} \sim \frac{1}{z}\frac{1}{E_{CFT}}. \tag{1.58}$$

E_{CFT} is scaling as $1/z$ so that approaching the boundary $z \to 0$ corresponds to higher energies in the boundary theory (and vice versa).

1.2.3 Matching degrees of freedom

In order for the AdS/CFT duality to be correct there must be a matching of degrees of freedom between the two theories. Field theories boast infinitely many degrees of freedom, however, and so one must impose an IR cutoff by limiting the spatial extent of the system to R (i.e. we put the field theory in a box of size R) and a UV cutoff by imposing a lattice cutoff a.

On the $d+1$ dimensional field theory side the number of degrees of freedom N_{CFT} is simply the number of lattice sites N_l times the degrees of freedom per site c. [12] Thus

$$N_{CFT} = c \cdot N_l = c \cdot \frac{R^d}{a^d}. \tag{1.59}$$

We can calculate the degrees of freedom in AdS by using the Bekenstein-

coordinates in flat spacetime.

[11] The conformal group in $3+1$ dimensions consists of 15 symmetries. The Poincaré group: 4 translations, 6 boosts and rotations, then the leftover conformal part has 1 dilation and 4 special conformal transformations.

[12] The nomenclature comes from the fact that in a conformal field theory the central charge c measures the number of degrees of freedom per site.

Hawking entropy:

$$N_{AdS} \approx S_{BH} = \frac{A_{Bound}}{4G_N}, \tag{1.60}$$

where A_{Bound} is the area of the boundary of AdS. Using the metric (1.57), with the UV cutoff at $z = a$ (since z is related to energy scales in the field theory) we have:

$$A_{Bound} = \int_{z=a} d^d x \sqrt{g} = \int d^d x \frac{L^d}{a^d} = R^d \frac{L^d}{a^d}, \tag{1.61}$$

where the boundary coordinates were cutoff by the size of the box, R. Thus up to overall constants, the matching $N_{CFT} = N_{AdS}$ requires

$$c = \frac{L^d}{G_N} = \frac{L^d}{\ell_P^d}, \tag{1.62}$$

where ℓ_P is the planck length.

A particularly important limit for us will be when the classical gravity theory is a good approximation. This is the regime where general relativity is valid, i.e. when the length scale of the spacetime is much larger than the Planck length. This means that $\frac{L^d}{\ell_P^d} \gg 1$. Thus in the CFT, according to (1.62) this means that $c \gg 1$. So classical gravity will be a good approximation when the number of degrees of freedom per lattice site in the CFT is very large. [13]

1.2.4 Partition functions

One can package all the information in a quantum theory into its partition function, and so it shouldn't be surprising that a statement of the AdS/CFT duality is [45]:

$$Z_{AdS} = Z_{CFT}. \tag{1.63}$$

[13] Note that for an $SU(N)$ gauge theory $N_{CFT} = c \sim N^2$, and thus classical gravity is a good description at large N.

More precisely, to every field φ^i [14] in the gravity side there corresponds an operator $\mathcal{O}^i$ in the CFT which couples to the field φ^i at the boundary, as it would to a source. In other words the presence of a field φ^i on the gravity side with boundary conditions $\varphi^i(x, z = 0) = \varphi_0^i(x)$ gives rise to a source term $\int \varphi_0^i \mathcal{O}_i$ for the CFT. Thus

$$Z_{AdS}|_{\varphi^i \to \varphi_0^i} = \langle e^{\int \varphi_0^i \mathcal{O}_i} \rangle_{CFT} \equiv Z_{CFT}[\varphi_0^i]. \tag{1.64}$$

We will be particularly interested in the limit where classical gravity applies so that the gravity partition function is dominated by a saddle point, and thus well approximated by its classical solution:

$$Z_{AdS}|_{\varphi^i \to \varphi_0^i} = \sum_{\text{all fields}} \Big|_{\varphi^i \to \varphi_0^i} e^{S_G} \approx e^{S_G^{classical}} \Big|_{\varphi^i \to \varphi_0^i}. \tag{1.65}$$

So with (1.64) we now have a way to compute quantum correlation functions in terms of a classical gravity action as

$$\langle \mathcal{O}_k(x_1) ... \mathcal{O}_j(x_n) \rangle = \frac{\delta}{\delta \varphi_0^k(x_1)} ... \frac{\delta}{\delta \varphi_0^j(x_n)} \log Z_{CFT}[\varphi_0^i] \tag{1.66}$$

$$= \frac{\delta}{\delta \varphi_0^k(x_1)} ... \frac{\delta}{\delta \varphi_0^j(x_n)} \; S_G^{classical} \Big|_{\varphi^i \to \varphi_0^i}. \tag{1.67}$$

One should be careful when performing these manipulations since generically the gravitational action and the fields at the boundary are divergent. To cure these divergences one typically adds a counter term action and extracts the finite part of the boundary behaviour of the fields. An example of this is shown for a scalar field in the next section.

[14] We represent any index or set of indices under which the field φ transforms (such as spinor or vector) as i, with inner product denoted by contraction between raised and lowered indices.

1.2.5 Scalar field on AdS

In this section we wish to calculate the expectation value of an operator $\mathcal{O}$ in the CFT vacuum, dual to a scalar field φ in pure AdS_{d+2} gravity (we will use the Poincaré coordinates (1.57)).

The action for a free massive field in Euclidean AdS_{d+2} is:

$$S = -\frac{1}{2}\int dz d^{d+1}x\sqrt{g}\left(\partial_\mu\varphi\partial^\mu\varphi + m^2\varphi^2\right) \tag{1.68}$$

Since we are assuming φ to be a classical solution we may integrate by parts and use the equations of motion to get a boundary term:

$$S^{classical} = \frac{1}{2}\int d^{d+1}x\sqrt{g}\left(\varphi g^{zz}\partial_z\varphi\right)_{z=\epsilon}, \tag{1.69}$$

where we have cutoff the boundary at $z=\epsilon$. In order to evaluate this one can easily find the behaviour of the field φ near the boundary by considering the equations of motion obtained from 1.68 at small z. After a Fourier transform

$$\varphi(z,x) = \int \frac{d^{d+1}k}{(2\pi)^{d+1}} e^{ikx}\varphi_k(z) \tag{1.70}$$

one finds that near the boundary $z=0$ the field φ behaves as:

$$\varphi_k(z) \sim C_1(k)z^{d+1-\Delta} + C_2(k)z^{\Delta}, \tag{1.71}$$

where $\Delta = (d+1)/2 + \sqrt{(d+1)^2/4 + m^2L^2}$.

Substituting into (1.69) gives at $z=\epsilon$

$$S^{classical} = \frac{L^d}{2}\int \frac{d^{d+1}k}{(2\pi)^{d+1}}\left[\epsilon^{-(2\Delta-(d+1))}(d+1-\Delta)C_1(-k)C_1(k) + (d+1)C_1(-k)C_2(k)\right], \tag{1.72}$$

where the $C_2(-k)C_2(k)$ term is of order $z^{2\Delta-(d+1)}$ and so vanishes as $z\to 0$ for scalar fields which satisfy the Breitenlohner-Freedman (BF) bound. [15]

[15] Since $\Delta = (d+1)/2 + \sqrt{(d+1)^2/4 + m^2L^2}$, we have

$$2\Delta - (d+1) = 2\sqrt{(d+1)^2/4 + m^2L^2}. \tag{1.73}$$

Now we notice that the expression for $S^{classical}$ diverges as $\epsilon \to 0$. To remedy this we should add a counterterm action S_{ct}. Like any good action it should be local in the field φ. Note that to cancel the divergence the action must be a boundary action, and it must be quadratic in the field in order to be able to cancel the term of the form $C_1 C_1$. A natural guess for the counter-term action is then

$$S_{ct} \sim \int d^{d+1}x \sqrt{\gamma_\epsilon}\, \varphi^2, \tag{1.75}$$

γ_ϵ being the induced metric for the $z = \epsilon$ boundary. This action indeed diverges as $\epsilon^{-(2\Delta-(d+1))} C_1(-k)C_1(k)$, and so we just need to find the right prefactor to cancel the divergence in $S^{classical}$. This is

$$S_{ct} = -\frac{d+1-\Delta}{2L} \int d^{d+1}x \sqrt{\gamma_\epsilon}\, \varphi^2. \tag{1.76}$$

In this way we finally arrive at the finite piece of the action $S^{renorm} = S^{classical} + S_{ct}$, evaluated on the classical solution:

$$S^{renorm} = \frac{L^d(2\Delta-(d+1))}{2} \int \frac{d^{d+1}k}{(2\pi)^{d+1}} C_1(-k)C_2(k). \tag{1.77}$$

So now to get the expection value, from (1.67) we need:

$$\langle \mathcal{O} \rangle = \frac{\delta}{\delta \varphi_0} S^{classical}. \tag{1.78}$$

We already removed the divergences in $S^{classical}$ by using S^{renorm}, but there is still the issue of the field φ diverging at the boundary.

From (1.71) we see that φ diverges as $z^{d+1-\Delta}$ near the boundary, and so we can remove this divergence (and the divergence in the action) by instead

This quantity is thus positive (or 0) whenever

$$m^2 \geq -\left(\frac{d+1}{2L}\right)^2, \tag{1.74}$$

which is known as the BF (Breitenlohner-Freedman) bound on the mass.

using the expression:

$$\langle \mathcal{O}(x) \rangle = \frac{\delta}{\delta \left(z^{-(d+1-\Delta)} \varphi_0(x) \right)} S^{renorm}. \tag{1.79}$$

Then using (1.71), along with a Fourier transform we have

$$\langle \mathcal{O}(k) \rangle = \frac{\delta}{\delta C_1(-k)} S^{renorm}. \tag{1.80}$$

Combining this with (1.77) we can finally find the expectation value of the operator in the CFT to be [16]

$$\langle \mathcal{O}(k) \rangle = L^d (2\Delta - (d+1)) C_2(k). \tag{1.84}$$

Thus, for a scalar field in pure AdS, the expectation value of the dual operator in the CFT is related to the subleading behaviour of the field at the boundary.

If one similarly calculates the 2-point function, one finds

$$\langle \mathcal{O}(x_1) \mathcal{O}(x_2) \rangle \sim \frac{1}{|x_1 - x_2|^{2\Delta}}. \tag{1.85}$$

This is precisely the result for correlation function between operators of dimension Δ in a CFT.

[16] Actually when one solves the equations of motion for the full solution $\varphi_k(z) = C_1(k) f_1(k,z) + C_2(k) f_2(k,z)$, one finds that demanding that it stays finite as $z \to \infty$ fixes the ratio $C_1(k)/C_2(k) = R(k)$. Thus when taking the functional derivative (1.80) one has

$$\begin{aligned} \frac{\delta}{\delta C_1(-k)} \int dq C_1(-q) C_2(q) =& C_2(k) + \int dq C_1(q) \frac{\delta}{\delta C_1(-k)} C_2(-q) && (1.81) \\ =& C_2(k) + C_1(k) R(k) && (1.82) \\ =& 2C_2(k) && (1.83) \end{aligned}$$

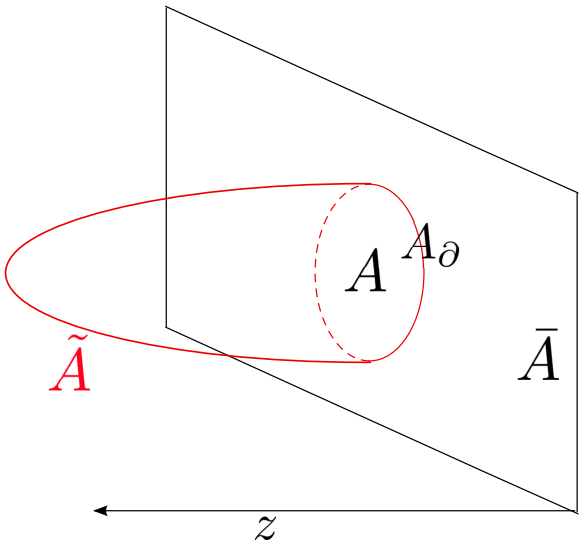


Figure 1.7: Diagram illustrating the holographic dual of geometric entropy.

1.2.6 Geometric entropy in AdS/CFT

Ryu-Takayanagi formula

Suppose we divide up a $d+1$ dimensional CFT into two spatial regions A and $\bar{A}$, with boundary between the two A_∂. We can then calculate the entanglement entropy S_A between degrees of freedom in A and $\bar{A}$. In the context of AdS/CFT a natural question to ask is "how do we compute S_A in the dual gravity theory?". Since we are working in the classical gravity limit we may expect something geometric, which also satisfies a least action (or saddle point) condition. The correct dual quantity is given by the area $\tilde{A}$ of the minimal surface whose boundary is given by A_∂ [17], see Figure 1.7. The precise result, known as the Ruy-Takayanagi formula [32] is:

$$S_A = \frac{\tilde{A}}{4G_N}. \tag{1.86}$$

[17] We will only be concerned with static metrics and entanglement entropies, thus we can think everything taking place at a given time slice.

The numerical factors are the same as in the Bekenstein-Hawking entropy, as is required for consistency in the thermal state.

Sketch of proof

The formula (1.86) was proved in [27]. An outline of the proof can be given as follows. We know from the discussion in section 1.1.3 that we should compute the CFT partition function on a space Σ_n composed of n sheets of $\mathbb{R}^{1,d}$ glued together in the region A, and with deficit angles $2\pi(1-n)$ at the dividing surface A_∂. To compute this partition function the holographic duality states we need to compute the partition function of the corresponding gravity solution. Assuming that the dual gravity solution turns out to be one with a surface σ_A ending on A_∂ which preserves the deficit angle (such that σ_A also carries a deficit angle of $2\pi(1-n)$). This can be shown to induce a curvature singularity [17] on σ_A, $R \sim 4\pi(1-n)\delta(\sigma_A)$. Since the gravitational action integrates over the curvature R, this induces a term proportional to the area of the surface σ_A, which will be the only term which contributes to the entanglement entropy. [18] At this point we see that since the area term appears in the action it should be minimal at the saddle point, so that $Area(\sigma_A) = \tilde{A}$, the minimal area extending into the bulk with boundary A_∂. So we finally find

[18] The gravitational action with cosmological constant is

$$S^{(n)}_{grav} = -\frac{1}{16\pi G_N}\int d^{d+2}x\sqrt{g}\,(R+\Lambda)\,, \tag{1.87}$$

so R having a delta function singularity on some surface produces a term proportional to the area of that surface. Also

$$\mathrm{tr}\,\hat{\rho}^n = \frac{Z_n}{Z_1^n}, \tag{1.88}$$

where Z_n is the partition function on the n-sheeted space Σ_n. Thus any term in the total partition function Z_n which is just n copies of single copy partition function cancels out. This is why only the piece which comes from the deficit angle is important.

$$S_A = -\frac{d}{dn} \ln \operatorname{tr} \hat{\rho}^n \Big|_{n=1} \tag{1.89}$$

$$= -\frac{d}{dn} \ln \frac{Z_n}{Z_1^n} \Big|_{n=1} \tag{1.90}$$

$$= -\frac{d}{dn} \ln \frac{Z_{AdS^{(n)}}}{Z_{AdS}^n} \Big|_{n=1} \tag{1.91}$$

$$= -\frac{d}{dn} \frac{4\pi(1-n) Area(\sigma_A)}{16\pi G_N} \Big|_{n=1} \tag{1.92}$$

$$= \frac{\tilde{A}}{4G_N}, \tag{1.93}$$

which is the Ryu-Takayanagi formula.

Consistency checks

Area law The first consistency check comes from our knowledge that in a CFT_{d+1} the geometric entropy obeys an area law divergence (see section 1.1.3):

$$S_A \sim k \cdot \frac{A_\partial}{a^{d-1}} \tag{1.94}$$

for some small distance cutoff a.

On the gravity side a small distance cutoff a corresponds to a cutoff on the spacetime at $z = a$, see Figure 1.8. Set up $d-1$ coordinates σ_i on the boundary, such that the surface A_∂ is parameterized as $x(\sigma)$, with induced metric on the surface d then

$$\int_{A_\partial} d^{d-1}\sigma \sqrt{d} = A_\partial, \tag{1.95}$$

where we are using A_∂ to describe both the surface and its area. In the bulk (with pure AdS metric (1.57)), however, on a slice of constant z the area is modified by the induced metric h as well:

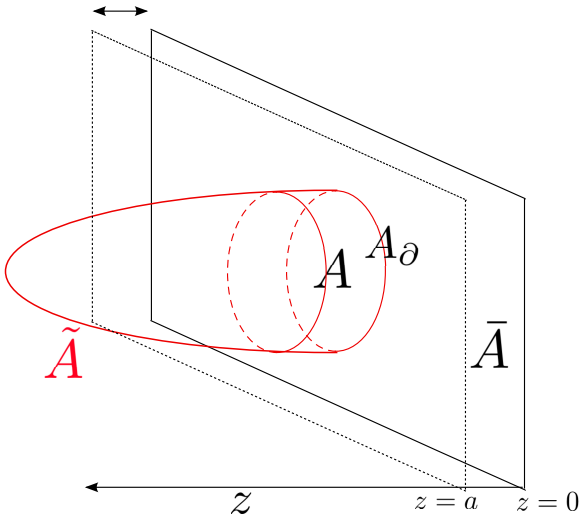


Figure 1.8: Diagram illustrating area law divergence.

$$A_{bulk}(z) = \int_{A_\partial} d^{d-1}\sigma\sqrt{d}\sqrt{h}, \tag{1.96}$$

$$= \frac{L^d}{z^d}\int_{A_\partial} d^{d-1}\sigma\sqrt{d}, \tag{1.97}$$

$$= \frac{L^d}{z^d}A_\partial. \tag{1.98}$$

Then near the boundary the contribution the the area $\tilde{A}$ can be obtained by integrating over the physical area $A_{bulk}(z)$ close to the cutoff. The divergent piece of this comes from the cutoff surface itself so that

$$\tilde{A} \sim \int^a dz A_{bulk}(z) \sim L^d \frac{A_\partial}{a^{d-1}}, \tag{1.99}$$

and so we have

$$S_A = \frac{\tilde{A}}{4G_N} \sim \frac{L^d}{G_N}\frac{A_\partial}{a^{d-1}}. \tag{1.100}$$

Thus the Ryu-Takayanagi formula reproduces the known area law (1.94) in

CFT.

Entropic inequalities It is also possible to check that the Ryu-Takayanagi formula reproduces the strong subadditivity relation [21]:

$$S_A + S_B \geq S_{A\cup B} + S_{A\cap B}. \tag{1.101}$$

Denote the area of a holographic surface with boundary A as Σ_A, if it is minimal we'll call it σ_A. Then from Figure 1.9 it is clear that

$$\begin{aligned} \sigma_A + \sigma_B =& \Sigma_{A\cup B} + \Sigma_{A\cap B}, & (1.102)\\ \geq& \sigma_{A\cup B} + \sigma_{A\cap B}. & (1.103) \end{aligned}$$

In going from the first to second line we have gone from surfaces which are not necessarily minimal to ones which are. Since the Ryu-Takayanagi formula in this context simply states that

$$S_A = \frac{\sigma_A}{4G_N}, \tag{1.104}$$

the area inequalities (1.103) imply strong subadditivity of the entanglement entropy (1.101).

1.3 Motivation

As we have seen, AdS/CFT (see Section 1.2) relates a gravitational theory on asymptotically Anti-de-Sitter (AdS) space to a conformal field theory (CFT) on its boundary. Since our main interest lies in understanding the interplay between geometry and entanglement, we will assume that we are in a regime in the gravitational theory where the general relativistic space-time picture is valid.[19] In this regime there is a very nice geometric way of quantifying the entanglement between two complementary spatial regions - A and $\bar{A}$ - in the CFT. This is captured by the Ryu-Takayanagi (RT) formula

[19] This is analogous to a regime in Quantum Electrodynamics where Maxwell's equations apply.

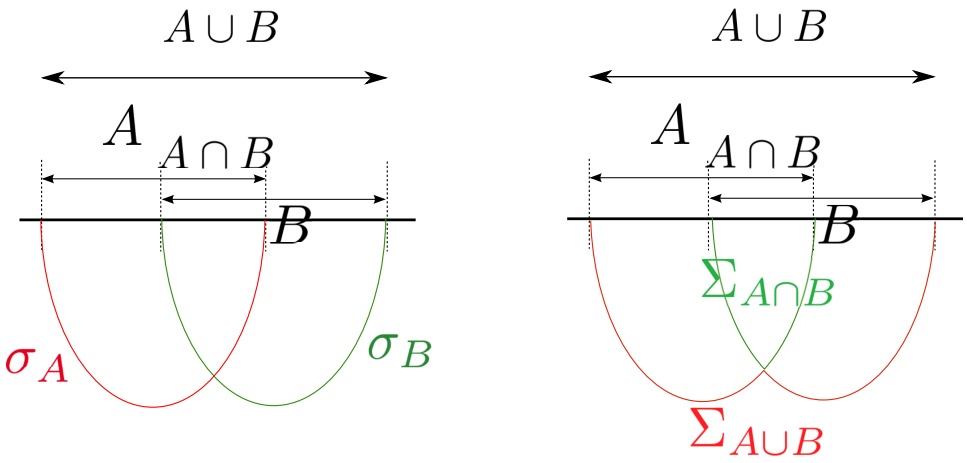


Figure 1.9: Diagram used in strong subadditivity proof.

(see Section 1.2.6). A measure of how entangled degrees of freedom in A and $\bar{A}$ are in a given state is computed by the entanglement entropy S_A (see Section 1.1). The RT formula states that S_A is the area of a minimal surface in the gravity theory (having as its boundary the region A_∂ separating A and $\bar{A}$), see Figure 1.10 :

$$S_A = \frac{\tilde{A}}{4G_N}. \tag{1.105}$$

In this way a purely quantum mechanical phenomenon in one theory is related to an entirely classical quantity in another. Roughly speaking, in Figure 1.10, we can think of $\tilde{A}$ as separating that region $\mathcal{M}_A$ in the bulk corresponding to A on the boundary, from the region $\mathcal{M}_{\bar{A}}$ in the bulk corresponding to $\bar{A}$ on the boundary. [20]

Now in AdS/CFT there is also a relation between the length scales on the boundary and "radial" distance in the gravity, the so-called UV/IR relation (see Section 1.2.2). It states that the physics of small length (large) scales on the boundary CFT is understood by studying regions near (far away from) the boundary in the gravity theory. This can actually be seen in the RT

[20] This correspondence was made more precise in [12].

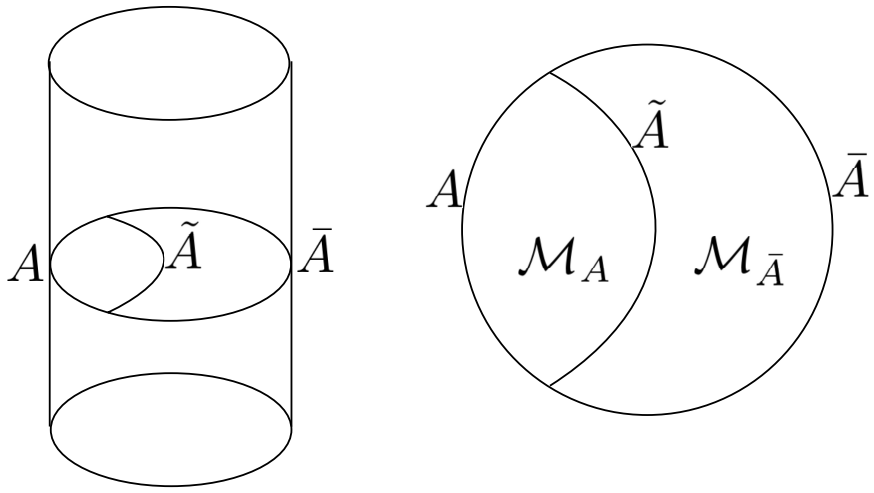


Figure 1.10: Diagram displaying the Ryu-Takayanagi surface $\tilde{A}$. The cylinder on the left is global AdS spacetime (with time running vertically). On the right is a constant time slice where we consider the Ryu-Takayanagi surface for S_A, separating spacetime into regions $\mathcal{M}_A$ and $\mathcal{M}_{\bar{A}}$.

formula (see Figure 1.11). If A is a small region, then the minimal surface $\tilde{A}$ stays relatively near the boundary. As we increase the size of A, $\tilde{A}$ dips deeper and deeper into the bulk and so probes longer and longer distance scales in the gravity side.

A natural question then is: *what region in the bulk (if any) corresponds to degrees of freedom on the boundary CFT below or above a certain length scale?* Divide up the boundary into degrees of freedom below and above a length scale ℓ as $\ell^<$ (UV) and $\ell^>$ (IR) and denote the corresponding bulk regions as $\mathcal{M}_{\ell^<}$ and $\mathcal{M}_{\ell^>}$. Now we can think of $\mathcal{M}_A$ and $\mathcal{M}_{\bar{A}}$ as corresponding to A and $\bar{A}$ respectively, and $\mathcal{M}_A$ and $\mathcal{M}_{\bar{A}}$ are obtained by considering the entanglement between regions A and $\bar{A}$. Thus we argue that the regions $\mathcal{M}_{\ell^<}$ and $\mathcal{M}_{\ell^>}$ should be related to the entanglement in the CFT between UV and IR degrees of freedom.

Fortuitously field theories provide a natural basis of states to study one measure of the UV/IR entanglement; the Fock basis. This basis of states can be written as a product of occupation number n_{p_i} of modes with momentum

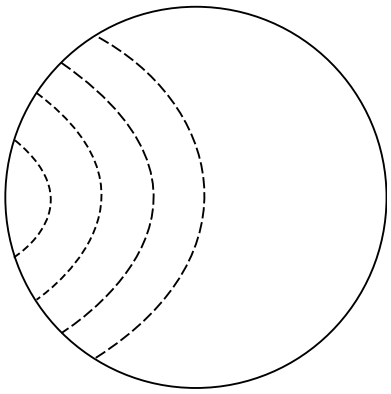


Figure 1.11: Diagram displaying a manifestation of the UV/IR connection with RT surfaces.

p_i below a certain scale μ (IR) $|p_i| < \mu$, and occupation n_{P_i} of modes with momentum P_i above that scale (UV) $|P_i| > \mu$. This gives us a way to calculate the entanglement entropy S_μ between UV and IR degrees of freedom. Studying the question in enough detail we may hope to find a momentum space RT-like geometric picture in the gravity theory, depicted in Figure 1.12.

The RT result also leads us to consider the "connectedness" of spacetime regions $\mathcal{M}_A$ and $\mathcal{M}_{\bar{A}}$ in terms of the entanglement between A and $\bar{A}$. This was explored in [41]. Consider subregions $a \subset A$ and $\bar{a} \subset \bar{A}$. For a general state correlation functions between operators at spacelike points x_a and $x_{\bar{a}}$ are exponentially decaying; $\langle \mathcal{O}_a \mathcal{O}_{\bar{a}} \rangle \sim e^{-s(x_a, x_{\bar{a}})/\xi}$ where $s(x_a, x_{\bar{a}})$ is the proper length of the minimal geodesic between x_a and $x_{\bar{a}}$, and ξ is some length scale. A measure of the entanglement between degrees of freedom in a and $\bar{a}$ is given by the mutual information $I(a, \bar{a})$ (it also includes classical correlations, see Section 1.1.2). This measure obeys a certain inequality [55]:

$$I(a, \bar{a}) \geq \frac{1}{2} \left(\frac{\langle \mathcal{O}_a \mathcal{O}_{\bar{a}} \rangle}{|\mathcal{O}_a||\mathcal{O}_{\bar{a}}|} \right)^2 \tag{1.106}$$

for any operators $\mathcal{O}$ referring to regions a and $\bar{a}$ with vanishing expectation value. Thus if the entanglement as measured by $I(a, \bar{a})$ vanishes, so must

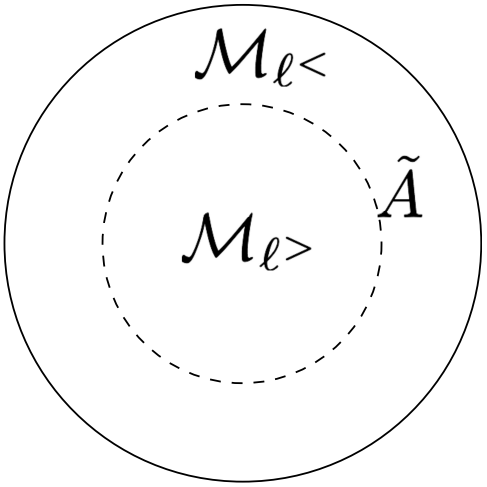


Figure 1.12: Figure of possible momentum space Ryu-Takayanagi relation.

the correlations between operators $\mathcal{O}$ on these regions, and this implies that the proper length between these regions diverges.

We thus arrive at the following picture (see Figure 1.13) of what happens as we remove the entanglement between A and $\bar{A}$: the bounding surface $\tilde{A}$ separating $\mathcal{M}_A$ and $\mathcal{M}_{\bar{A}}$ shrinks in accordance with RT, and the boundaries A and $\bar{A}$ get further and further away from each other as measure by the proper distance or geodesics connecting them. So we get a situation where $\mathcal{M}_A$ and $\mathcal{M}_{\bar{A}}$ are pinching off from one another. From this it is apparent that entanglement in the CFT is serving to hold the dual space-time together. A natural question to ask is: *precisely how do changes in the entanglement (perhaps subject to certain conditions) affect the dual geometry* ?

1.4 Outline of thesis

The structure of this thesis is as follows. In Chapter 2 we introduce and explore entanglement in momentum space. We begin by giving a path integral description of the reduced density matrix for low momentum degrees of freedom. We then develop a perturbative approach for computing entanglement

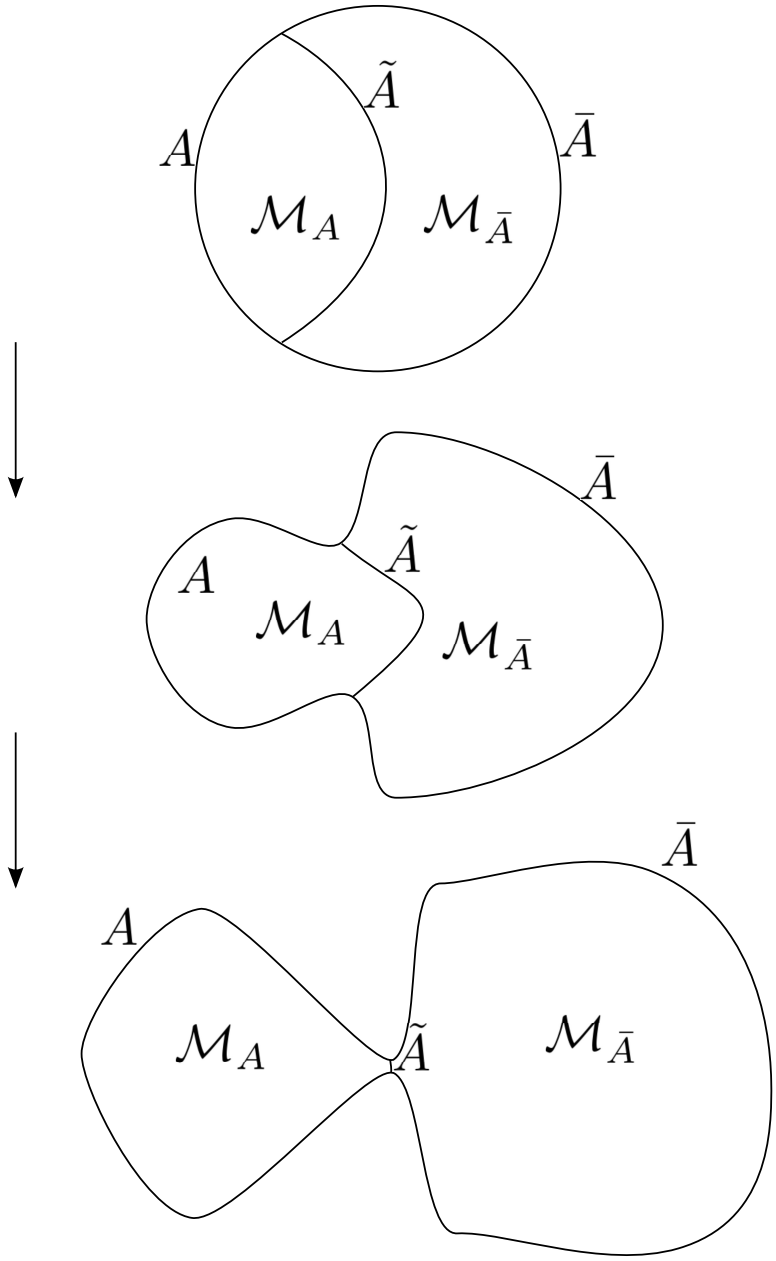


Figure 1.13: Diagram illustrating the separation of spacetime due to loss of entanglement in the CFT.

entropy and mutual information and apply these to example field theories. Entanglement properties are understood in terms of scaling arguments and related to properties of decoupling. Conditions under which our measures are well defined in perturbation theory are also discussed.

We change gears from Lorentz invariant theories to exploring systems with Fermi surfaces in Chapter 3. We study the entanglement between single modes in a lattice theory and its continuum version. Surprisingly the presence of a Fermi surface allows to compute the entanglement entropy in dimensions where the Lorentz invariant version would have diverged. Interesting behaviour near the Fermi surface is shown not to depend on high energy details.

In Chapter 4 we discuss how gravity "emerges" from certain thermodynamic-like relations on the boundary in AdS/CFT. We show how one arrives at the relation $dS = dE$ in the CFT and explore the implications of each side of this equation on the AdS gravity side. The equation of the two quantities on the AdS side can be shown (under suitable conditions) to imply Einstein's equations.

Chapter 2

Momentum-space entanglement and renormalization in quantum field theory

2.1 Introduction

A quintessential property that distinguishes quantum mechanics from classical mechanics is the entanglement of otherwise distinct degrees of freedom. When certain degrees of freedom are entangled with the rest of a quantum system, it is not possible to describe them by a pure state. Rather, the most complete description of a subsystem A at a particular time is via the reduced density matrix obtained by tracing over the degrees of freedom in the complement, $\rho_A = \mathrm{tr}_{\bar{A}}(|\Psi\rangle\langle\Psi|)$, where $|\Psi\rangle$ is the state of the entire system. The entropy constructed from the reduced density matrix, $S(\rho_A) = -\mathrm{tr}(\rho_A \log(\rho_A))$, quantifies the amount of entanglement between A and its complement. The entanglement entropies corresponding to reduced density matrices for diverse subsets of degrees of freedom provide a rich characterization of the quantum state for systems with many degrees of freedom.[21]

In physical systems, we typically only have access to a subset of the degrees of freedom, namely the low-energy or long-wavelength modes that are

[21] For a basic review of density matrices, entanglement entropy, and related concepts, see for example [68].

directly accessible to experiments. In an interacting theory, it will generally be true that these degrees of freedom are entangled with the inaccessible high-energy degrees of freedom. Thus, the long-wavelength modes will be described by a density matrix. A more familiar description of low-energy degrees of freedom is due to Wilson [47] – one carries out the complete path integral over the inaccessible degrees of freedom, arriving at an effective action capturing the dynamics of the remaining system. Here we will index the degrees of freedom by their spatial momenta and consider integrating out modes of high spatial frequency. We show that the resulting Wilsonian prescription is compatible with the description in terms of a density matrix: given a Wilsonian effective action (defined more precisely below) we can canonically associate the corresponding density matrix via equation (2.5) below.

For continuous physical systems described by interacting quantum field theories, understanding the variation with scale of the Wilsonian effective action $S_W(\mu)$ provides key insights into the nature of the quantum field theories, revealing a striking insensitivity of the low-energy physics to the details of the ultraviolet description. Correspondingly, it is natural to consider the variation with scale of the density matrix $\rho(\mu)$ for the degrees of freedom with momentum $|\vec{p}| < \mu$ and the associated entanglement entropy $S(\mu)$. To make our considerations concrete, we derive a formula for this low-energy entanglement entropy in perturbative quantum field theory, and apply it to scalar field theories with ϕ^n potentials in various dimensions. The scale-dependence of the entropy $S(\mu)$ in such theories can be understood in terms of the variation of the coupling and number of degrees of freedom with scale.

To study entanglement between scales in greater detail, we can consider the entanglement entropy associated with any subset of the allowed momenta, or the mutual information between any two subsets of the allowed momenta, for example between individual modes with momenta $\vec{p}$ and $\vec{q}$. These measures characterize the extent of entanglement between specific scales in field theory, and we compute the rate at which this entanglement declines as the scales separate. This fall-off may give an alternative charac-

terization of the property of decoupling in quantum field theories. In theories that do not enjoy the property of decoupling, e.g. noncommutative gauge theories [48] and theories of gravity, the entanglement between degrees of freedom at different scales may play an especially important role.

Entanglement entropy in quantum field theory has been considered previously, but almost all previous work has focused on entanglement between degrees of freedom associated with spatial regions (e.g. [69, 70]). The notion of a density matrix for low-momentum modes or entanglement between different momentum modes has appeared earlier in the context of cosmology and condensed matter physics (e.g. [51, 53, 71]), but there is little overlap with the present work.

2.2 The low-energy density matrix

A quantum system with many degrees of freedom has a Hilbert space of the form $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots$. Given a subset of these degrees of freedom (A, with complement $\bar{A}$), we can write $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$ where $\mathcal{H}_A$ is the tensor product of Hilbert spaces for the degrees of freedom in A. If ρ is the density matrix for the full system (which may be in a pure state), a reduced density matrix for A is defined by tracing over $\bar{A}$: $\rho^A = \mathrm{tr}_{\bar{A}}(\rho)$ (or, given components in a specific basis, $\rho^A_{mn} = \sum_N \rho_{mN,nN}$). Expectation values of operators that act only on A can be computed as $\mathrm{tr}\left(\rho\left(\mathcal{O}_A \otimes \mathbb{1}\right)\right) = \mathrm{tr}_A(\rho^A \mathcal{O}_A)$. If A is entangled with its complement, ρ^A will have a finite entropy: $S = -\mathrm{tr}_A(\rho^A \log \rho^A) > 0$.

In this construction, the Hilbert space can be be decomposed into a tensor product in any convenient way. For example, the Hilbert space for two identical oscillators could be decomposed either as a product of the Hilbert spaces for the individual oscillators, or as a product of the Hilbert spaces of even and odd normal modes. A reduced density matrix could be computed in either case – good choices of decomposition are dictated by the structure of the interactions and restrictions on which degrees of freedom are accessible to measurements.

In quantum field theories, locality makes it natural to associate independent degrees of freedom with disjoint spatial domains. Hence, given a

spatial region A (and complement $\bar{A}$), one can decompose the Hilbert space as $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$ and trace over $\bar{A}$ to derive the reduced density matrix of A. But since the Hamiltonians of free field theories are diagonalized by modes of fixed momentum, it is in many ways more natural to use the Fock space decomposition, $\mathcal{H} = \otimes_{\vec{p}} \mathcal{H}_{\vec{p}}$, where $\mathcal{H}_{\vec{p}}$ is the Hilbert space of modes of momentum $\vec{p}$.[22] While this decomposition is motivated by considering the case of free field theory, it applies equally well once we turn on interactions, and is indeed the standard setting for computations in perturbative quantum field theory.[23]

In free field theory, the vacuum state is a tensor product of the Fock space vacuum states for each independent field mode – there is no entanglement between the field modes at different momenta. But in an interacting theory, the full vacuum state will be a superposition of Fock basis states – hence the modes of different momenta will generally be entangled. In this case, the reduced density matrix corresponding to a subset of the degrees of freedom (A) will necessarily have a finite entropy, indicating that A is effectively in a mixed state if the rest of system is traced over. Now, one is most often interested in the physics of the "infrared" degrees of freedom that are accessible to experiment, i.e., the degrees of freedom with momenta below some scale μ. The present discussion shows that tracing over the ultraviolet, i.e. degrees of freedom with momenta above μ, should lead to an infrared effective description in terms of a *low-momentum density matrix* corresponding to a mixed state with finite entropy.

[22] Here, it is clearest to define the field theory as a limit of a theory on finite volume so that the tensor product is over a discrete set of allowed momenta. For a general field theory, the factors would be labeled by field type and spin/polarization in addition to momentum.

[23] There is a formal sense in which turning on interactions takes one out of the original Hilbert space. However, by placing a cutoff at some energy scale much higher than any scale of interest, the Hilbert space structure of the interacting theory will be the same as that of the free theory, and a density matrix for low-momentum modes can be precisely defined. Furthermore, as we will see later, various observables related to the spectrum of the density matrix have a well-defined limit as we take the cutoff to infinity.

Relation between low-energy density matrix and low-energy effective action

The standard way of studying the low-energy theory is through the Wilsonian effective action. How is this quantity related to the low-energy density matrix? To begin, consider a bare action S_Λ defined with a cutoff Λ such that $|p| \leq \Lambda$. Associated to this, we have a Hamiltonian H_Λ, which will have some ground state $|\Psi^0_\Lambda\rangle$ and corresponding density matrix $\rho^0_\Lambda = |\Psi^0_\Lambda\rangle\langle\Psi^0_\Lambda|$. This density matrix can be written as a Euclidean path integral by taking the $T \to 0$ limit of the finite temperature density matrix

$$\langle\hat{\phi}_y|\rho^T_\Lambda|\phi_y\rangle = \frac{1}{Z}\langle\hat{\phi}_y|e^{-\beta H_\Lambda}|\phi_y\rangle = \frac{1}{Z}\int_{\phi(\tau=-\beta/2)=\phi_y}^{\phi(\tau=\beta/2)=\hat{\phi}_y} \mathcal{D}\phi(\tau)\, e^{-S^E_\Lambda}\,, \qquad (2.1)$$

where $\{\phi_y\}$ is a basis of field configurations indexed by y, $\beta = 1/T$ is the inverse temperature, S^E_Λ is the Euclidean action, and Z is the partition sum that normalizes the path integral.

Given a subset of degrees of freedom A (with complement $\bar{A}$) and the tensor product structure of the Hilbert space, we can split the parameter y which indexes the basis states as $y = (a, \bar{a})$, and a pick a basis with $\phi_y = \phi_a \times \phi_{\bar{a}}$. To define a reduced density matrix for A by tracing over $\bar{A}$ we write

$$\langle\hat{\phi}_a|\rho^T_A|\phi_a\rangle = \int \mathcal{D}\phi'_{\bar{a}}\, \langle\hat{\phi}_a\phi'_{\bar{a}}|\rho^T_\Lambda|\phi_a\phi'_{\bar{a}}\rangle = \frac{1}{Z}\int_{\phi_A(-\beta/2)=\phi_a}^{\phi_A(\beta/2)=\hat{\phi}_a} \mathcal{D}\phi_A(\tau)\,\mathcal{D}\phi_{\bar{A}}(\tau) e^{-S_E}\,. \qquad (2.2)$$

In the last expression, the fields $\phi_{\bar{A}}$ are periodic in the range $[-\beta/2, \beta/2]$, which is implied after substituting (2.1) into the trace in the middle expression.

Now define a conventional thermal effective action for the the subsystem A:

$$e^{-S^T_W(\phi_A)} = \int_{-\beta/2\leq\tau\leq\beta/2} \mathcal{D}\phi_{\bar{A}}(\tau)\, e^{-S_E(\phi_A,\phi_{\bar{A}})}\,. \qquad (2.3)$$

In terms of this, the density matrix for A is

$$\langle\hat{\phi}_a|\rho_A^T|\phi_a\rangle = \frac{1}{Z}\int_{\phi_A(-\beta/2)=\phi_a}^{\phi_A(\beta/2)=\hat{\phi}_a}\mathcal{D}\phi_A(\tau)\,e^{-S_W^T(\phi_A)} = \frac{1}{Z}\int_{\phi_A(\tau=0^-)=\phi_a}^{\phi_A(\tau=0^+)=\hat{\phi}_a}\mathcal{D}\phi_A(\tau)\,e^{-S_W^T(\phi_A)}\ . \tag{2.4}$$

In the last expression we translated time to put the discontinuity in the integral at $\tau = 0^{\pm}$ and the fields are taken to be periodic at $\tau = \pm\beta/2$. The reduced density matrix for A in the ground state $|\Psi_\Lambda^0\rangle$ for the entire system is extracted by taking $\beta \to \infty$.

We now specialize to the case where A is the subset of degrees of freedom with spatial momenta $|p| < \mu$ for any scale μ which is lower than the ultraviolet cutoff Λ. The reduced density matrix $\rho_{|p|<\mu}$ obtained by tracing over the degrees of freedom with momenta in the range $\mu < |p| \leq \Lambda$ is thus given by:

$$\langle\hat{\phi}_{|p|<\mu}|\rho_{|p|<\mu}|\tilde{\phi}_{|p|<\mu}\rangle \;=\; \frac{1}{Z}\int_{\phi_{|p|<\mu}(\tau=0^-)=\tilde{\phi}_{|p|<\mu}}^{\phi_{|p|<\mu}(\tau=0^+)=\hat{\phi}_{|p|<\mu}}\mathcal{D}\phi_{|p|<\mu}\,e^{-S_W(\phi_{|p|<\mu})} \tag{2.5}$$

where now (having taken $\beta \to \infty$,) S_W is a Wilsonian effective action obtained by integrating out the degrees of freedom with large spatial momenta:[24]

$$e^{-S_W(\phi_{|p|<\mu})} = \int\mathcal{D}\phi_{|p|>\mu}(\tau)\,e^{-S_E(\phi_{|p|<\mu},\phi_{|p|>\mu})}\ . \tag{2.6}$$

Equation (2.5) is our final result for the low-energy density matrix. In

[24] In relativistic quantum field theories, it is perhaps more common to define a Wilsonian effective action by performing the Euclidean path integral over all field variables $\phi(p^\nu)$ with $|p^\nu p_\nu| > \mu$. This is convenient, since it leads to an effective action that is manifestly Lorentz invariant. However, the remaining variables in the path integral correspond to a restricted set of frequencies for each field mode $\phi(\vec{p})$. Hence, this Wilsonian action does not represent the full effective action for a particular subset of degrees of freedom, but rather the action for a restricted set of configurations of a subset of degrees of freedom. In our case, while the path integral is still Euclidean, we are integrating out all modes with spatial momenta $|\vec{p}| > \mu$ and leaving all modes with $|\vec{p}| \leq \mu$, regardless of frequency. The result is an effective action that describes all possible configurations of a subset of degrees of freedom for the theory, the field modes with $|\vec{p}| \leq \mu$. This type of Wilsonian action is more commonly discussed in field theories without Lorentz invariance, such as those describing condensed matter systems (see, for example, [54]).

particular, if $\mathcal{O}$ is an observable built out of the low-momentum modes at $\tau = 0$, it follows from (2.5) that

$$\mathrm{tr}(\hat{\mathcal{O}}\rho) = \frac{1}{Z}\int [d\phi_{|p|<\mu}]\mathcal{O}e^{-S_W(\phi_{|p|<\mu})}\ , \tag{2.7}$$

so the standard calculation using the effective action is equivalent to a calculation using the density matrix. Of course, the full Wilsonian effective action contains more information than the density matrix associated with the vacuum state of the field theory. The former is a functional of time-dependent field configurations, while the latter depends only on a pair of *time-independent* field configurations.

The description of low-energy degrees of freedom via a density matrix may seem unfamiliar and one may ask why we cannot simply associate a pure vacuum state to the low-energy degrees of freedom based on the effective action. The reason is that S_W will typically contain terms with higher time derivatives, so there is no way to associate to S_W a Hamiltonian H_μ expressed exclusively in terms of the low-momentum variables and their conjugate momenta. Thus, there is no canonical way to associate a pure state of the low-momentum part of the Hilbert space to the full ground state of the theory. As we have seen, the object that can be canonically associated to a Wilsonian effective action for these low-momentum degrees of freedom is a density matrix.

2.3 Measures of entanglement

What observables quantify the amount of entanglement between the degrees of freedom in different ranges of momenta? In this section we begin by discussing such quantities in generality and conclude by constructing perturbative expressions for such observables in weakly coupled field theories.

First, for any density matrix ρ, the von Neumann entropy

$$S(\rho) = -\,\mathrm{tr}(\rho\log(\rho)) \tag{2.8}$$

measures the classical uncertainty associated with the mixed state described by ρ. When ρ represents a microcanonical or canonical ensemble, the von Neumann entropy gives the thermodynamic entropy. When ρ is the reduced density matrix describing a subsystem A of a quantum system that is in a pure state, S quantifies the entanglement between A and its complement ($\bar{A}$). In this case the *entanglement entropy* of A is equal to that of $\bar{A}$, a fact that follows from a stronger result that the spectrum of eigenvalues of ρ_A matches the spectrum of eigenvalues of $\rho_{\bar{A}}$.

When the Hilbert space for the theory can be decomposed into a tensor product with three or more factors, the quantum entanglement and classical correlations between pairs of these subsystems are jointly quantified by the *mutual information*. For instance, if the Hilbert space is of the form $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \otimes \cdots$, the *mutual information* between A and B is defined as

$$I(A, B) = S(A) + S(B) - S(A \cup B) \, . \tag{2.9}$$

where $S(X)$ is the von Neumann entropy of the reduced density matrix of subsystem X. Mutual information is always greater than or equal to zero, with equality if and only if the density matrix for the AB subsystem is a tensor product of the reduced density matrices for subsystems A and B. In other words, mutual information is zero if and only if there is neither any entanglement nor any classical correlation between the two subsystems.[25] Mutual information provides an upper bound on all correlators between the two regions: for any bounded operators $\mathcal{O}_A$ and $\mathcal{O}_B$, acting only on the subsystems A and B, we have [55]

$$I(A, B) \geq \frac{(\langle \mathcal{O}_A \mathcal{O}_B \rangle - \langle \mathcal{O}_A \rangle \langle \mathcal{O}_B \rangle)^2}{2|\mathcal{O}_A|^2 |\mathcal{O}_B|^2} \, . \tag{2.10}$$

If the Hilbert space consists of three factors $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ and

[25] The systems A and B are said to be entangled if and only if the density matrix for the AB subsystem cannot be written as $\sum_i p_i \rho_A^i \otimes \rho_B^i$. *Separable* density matrices of this form represent states which have no quantum entanglement, but may have classical correlations. The mutual information for such a state can be nonzero.

the complete system is in a pure state it follows from the definitions that

$$I(A \cup B, C) = I(A, C) + I(B, C) \, . \tag{2.11}$$

But if A, B, and C together comprise only a part of the system, another interesting observable is the *tripartite information* which quantifies the extent to which the mutual information between $A \cup B$ and C is determined by the pairwise mutual informations $I(A, B)$ and $I(B, C)$:

$$I(A, B, C) = I(A \cup B, C) - I(A, C) - I(B, C) \, . \tag{2.12}$$

In general, this quantity can be positive, negative, or zero. For a pure state of the full system, $I(A, B, C)$ is symmetric between the four subsystems A,B,C, and $\overline{A \cup B \cup C}$.

2.3.1 Entanglement observables in perturbation theory

For weakly coupled quantum field theories, we can use perturbation theory methods to calculate the entanglement observables described in the previous section. To begin, it is useful to derive a set of perturbative results that apply more broadly.

Consider a general quantum system whose Hilbert space may be decomposed into a tensor product $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, and start with a Hamiltonian of the form

$$H = H_A \otimes \mathbb{1} + \mathbb{1} \otimes H_B \, . \tag{2.13}$$

Denote the energy eigenstates of H_A by $|n\rangle$ and the energy eigenstates of H_B by $|N\rangle$, with energies E_n and $\tilde{E}_N$ respectively. Before adding interactions, the ground state is

$$|0, 0\rangle \equiv |0\rangle \otimes |0\rangle \, . \tag{2.14}$$

Now, perturb the Hamiltonian by an interaction λH_{AB}, where λ is a small parameter. The perturbed ground state may be written (before normaliza-

tion) as

$$|\Omega\rangle = |0,0\rangle + \sum_{n\neq 0} A_n|n,0\rangle + \sum_{N\neq 0} B_N|0,N\rangle + \sum_{n,N\neq 0} C_{n,N}|n,N\rangle\,, \qquad (2.15)$$

where A,B, and C are coefficients starting at order λ that may be computed in perturbation theory. To normalize, we should multiply by $1/\mathcal{N}^{\frac{1}{2}}$, where $\mathcal{N} = 1 + \sum |A_n|^2 + \sum |B_N|^2 + \sum |C_{n,N}|^2$. Now, the density matrix corresponding to the subsystem A is

$$\rho_A = \frac{1}{1+|A|^2+|B|^2+|C|^2}\begin{pmatrix} 1+|B|^2 & A^\dagger + BC^\dagger \\ A + CB^\dagger & AA^\dagger + CC^\dagger \end{pmatrix}\,, \qquad (2.16)$$

where the elements of this matrix correspond to $|0\rangle\langle 0|$,$|0\rangle\langle n|$,$|m\rangle\langle 0|$,$|m\rangle\langle n|$ terms respectively. By a symmetry transformation $\rho \to M\rho M^{-1}$, we can simplify the form to

$$\hat{\rho}_A = \begin{pmatrix} 1-|C|^2 & 0 \\ 0 & CC^\dagger \end{pmatrix} + \mathcal{O}(\lambda^3) \qquad (2.17)$$

where we are using the fact that A, B, and C start at order λ. (See below for why this is possible.)

Up to corrections of order λ^3, the eigenvalues of this matrix are $\lambda^2 a_i$ and $1-\lambda^2\sum a_i$, where $\{a_i\}$ (normalized to be of order λ^0) are the eigenvalues of the matrix $CC^\dagger/\lambda^2$. Thus, the entanglement entropy is

$$\begin{aligned} S_A = -\operatorname{tr}(\rho_A \log(\rho_A)) &= -(1-\lambda^2\textstyle\sum a_i)\log(1-\lambda^2\sum a_i) - \sum \lambda^2 a_i \log(\lambda^2 a_i) \\ &= -\lambda^2\log(\lambda^2)\textstyle\sum a_i + \lambda^2 \sum a_i(1-\log a_i) + \mathcal{O}(\lambda^3) \end{aligned} \qquad (2.18)$$

Now, an explicit expression for the C matrix using standard perturbation theory is

$$C_{nN} = \lambda\frac{\langle n,N|H_{AB}|0,0\rangle}{E_0+\tilde{E}_0-E_n-\tilde{E}_N} + \mathcal{O}(\lambda^2)\,. \qquad (2.19)$$

Using this, the leading term in the entanglement entropy for small λ is

explicitly

$$S_A = -\lambda^2 \log(\lambda^2) \sum_{n\neq 0, N\neq 0} \frac{|\langle n, N|H_{AB}|0,0\rangle|^2}{(E_0 + \tilde{E}_0 - E_n - \tilde{E}_N)^2} + \mathcal{O}(\lambda^2)\,. \tag{2.20}$$

Interestingly, the entanglement entropy is not analytic in λ at $\lambda = 0$. Also, the leading order perturbative result (up to order λ^2 terms which are not written explicitly) depends only on matrix elements of the interaction Hamiltonian between the vacuum and states with both subsystems excited. This follows since (2.15) can be written as

$$|\Omega\rangle = (|0\rangle + \sum_{n\neq 0} A_n|n\rangle) \otimes (|0\rangle + \sum_{N\neq 0} B_N|0,N\rangle) + \sum_{n,N\neq 0} (C_{n,N} - A_n B_N)|n\rangle \otimes |N\rangle\ . \tag{2.21}$$

In this expression, the entanglement would be zero without the second term, and in this term, C_{nN} starts at order λ while $A_n B_N$ starts at order λ^2. The A and B coefficients do appear in the order λ^3 contributions to the entanglement entropy.

Mutual information

By a similar calculation, starting from a pure state in a theory with $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, the leading contribution to $I(A,B)$ in perturbation theory is

$$I(A,B) = -\lambda^2 \log(\lambda^2) \left\{ 2 \sum_{N_A\neq 0, N_B\neq 0, N_C = 0} + \sum_{N_A\neq 0, N_B\neq 0, N_C\neq 0} \right\} \frac{|\langle N_A, N_B, N_C|H_{int}|0,0\rangle|^2}{(E_0 - E_{N_i})^2} \tag{2.22}$$

Similarly, when $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \otimes \mathcal{H}_D$, at leading order in perturbation theory, the tripartite information $I(A,B,C)$ is

$$I(A,B,C) = +\lambda^2 \log(\lambda^2) \sum_{N_i\neq 0} \frac{|\langle N_A, N_B, N_C, N_D|H_{int}|0,0\rangle|^2}{(E_0 - E_{N_i})^2} + \mathcal{O}(\lambda^2) \tag{2.23}$$

While $I(A, B, C)$ can in general be positive, negative or zero, we see that the leading perturbative result for $I(A, B, C)$ is always less than or equal to zero. This implies that to leading order in perturbation theory,

$$I(A \cup B, C) \leq I(A, C) + I(B, C) \ . \tag{2.24}$$

This result is not true for general systems.[26] Note also that if the matrix element of the interaction Hamiltonian between the free vacuum and states with all four subsystems excited is zero[27] then we will have

$$I(A \cup B, C) = I(A, C) + I(B, C) \ . \tag{2.25}$$

to leading order in perturbation theory. In this case, the leading order contribution to mutual information between any two subsystems is completely determined from the mutual information between any pair of minimal subsystems.[28]

2.3.2 Entanglement observables in quantum field theory

For all of the observables discussed above, the essential quantities we have to compute are the reduced density matrices of the various subsystems. Given these quantities, we can compute the associated von Neumann entropies and mutual informations. In local quantum field theory, recent discussions of entanglement have focused on the density matrices associated with bounded spatial regions. These are well-defined because (by locality) there are independent degrees of freedom in disjoint spatial domains, so the Hilbert space factorizes as required. The associated spatial entanglement entropy is typically divergent, even in free field theory, because in the continuum limit any spatial region contains an infinite number of degrees of freedom at arbitrarily short wavelengths. These divergences require regularization (e.g.

[26] In particular, if A and B are completely uncorrelated, $\rho_{AB} = \rho_A \otimes \rho_B$, the opposite inequality, $I(A \cup B, C) \geq I(A, C) + I(B, C)$ follows from strong subadditivity of entanglement entropy.

[27] In field theory, this will be true for theories with only cubic interaction terms.

[28] In field theory, such minimal subsystems will be the Hilbert spaces associated with modes at a single momentum.

by including an ultraviolet cutoff) and some care is needed to extract finite regularization-independent data [70].

Now, as discussed above, it is often more natural in quantum field theory to organize degrees of freedom by momentum (or wavelength). Corresponding to any bounded subset of momenta in a field theory there are a finite number of degrees of freedom per unit spatial volume.[29] As a result, the entanglement entropy associated with such a subset should be *finite* for a finite volume system, increasing with the volume considered. For a translation-invariant state, we expect that the momentum space entanglement entropy will be an extensive quantity with a finite *density* S/V. We will verify this below.

What observables can we compute? We can define the entanglement entropy $S(P)$ associated to any subset P of the allowed momenta[30], the mutual information $I(P,Q)$ between any two subsets of momenta, or the tripartite entanglement $I(P,Q,R)$ for any three subsets of momenta. We will focus on

- $S(\mu)$, the entanglement entropy between all degrees of freedom with momenta above and below the scale μ.
- $S([\mu_1,\mu_2])$, the entanglement entropy for a shell of momenta $\mu_1 \leq |p| \leq \mu_2$.
- $S(\vec{p})$, the entanglement entropy for a single mode with momentum p.
- $I(\{|p|<\mu_1\},\{|p|>\mu_2\})$, the mutual information between degrees of freedom with momenta below a scale μ_1 and degrees of freedom above the scale μ_2.
- $I(\vec{p},\vec{q})$, the mutual information between modes with momenta $\vec{p}$ and $\vec{q}$.

[29] For a field theory at finite volume there will be a finite number of degrees of freedom in a bounded range of momenta. In the infinite volume limit, the set of allowed momenta becomes continuous. While there are now an infinite number of degrees of freedom with momenta in a finite region of momentum space, the number per unit spatial volume remains finite.

[30] More generally, P could represent a subset of the allowed single particle states.

These quantities probe the strength and extent of entanglement in momentum space.

In free field theory, the Hamiltonian does not couple degrees of freedom with different momenta and thus all these measures of entanglement in momentum space should vanish. Adding a weak interaction term that couples degrees of freedom with different momenta modifies the ground state and should introduce a small amount of entanglement between the various field modes. We can characterize this entanglement in perturbative quantum field theory by adapting the general results derived above. For the calculation of entanglement entropy, the two subsystems correspond to two complementary subsets A and $\bar{A}$ of the allowed momenta. The eigenstates $|n, N\rangle$ of the unperturbed Hamiltonian are elements of the Fock space basis $|\{n_i\}, \{N_I\}\rangle$, where n_i and N_I are occupation numbers for particle states in the two subsets. The interaction Hamiltonian takes the form

$$H_I = \int d^d x \, \mathcal{H}_I(x)$$

for some local Hamiltonian density $\mathcal{H}_I(x)$ that is polynomial in the fields and their derivatives. The matrix elements

$$\langle \{n_i\}, \{N_I\} | H_I | 0, 0 \rangle \tag{2.26}$$

may be computed by expanding the interaction Hamiltonian density in terms of creation and annihilation operators. The sum in (2.20) is now over all states with at least one particle having momentum in the subset A and at least one particle having momentum in the subset $\bar{A}$. The nonzero matrix elements (2.26) in the sum involve states with at most k momenta, where k is the number of fields in the interaction, and the momenta must add to zero since translation-invariance of the interaction Hamiltonian leads to a momentum-conserving delta function.

The leading contribution (2.20) to the entanglement entropy can be rewritten in terms of a projected two-point correlator of the interaction Hamiltonian (see Appendix A.1). Below we will work directly with the

expression (2.20).

2.4 Scalar field theory: entanglement between scales

To develop some concrete examples of the general formalism, we will calculate momentum space entanglement entropy in $d+1$ dimensional scalar theories with action:

$$S = \int d^{d+1}x (\frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{n!}\phi^n) \ . \tag{2.27}$$

For ease of formulation, we will first take the theory to be defined in a box of size L with periodic boundary conditions, and assume a UV cutoff at a scale Λ.

We will compute the entanglement entropy $S(\mu)$ of degrees of freedom with momenta $|p| < \mu$ with the high-momentum modes. Denoting by p_i and P_i the allowed momenta below and above μ, the Fock space basis elements are written as $|\{n_{p_i}\}\rangle \otimes |\{n_{P_i}\}\rangle$. To use the general formula (2.20) for the leading contribution to the entanglement entropy, we need matrix elements of the interaction Hamiltonian between the free vacuum and the states with both low and high momenta excited. Recall that the fields can be expanded in terms of creation and annihilation operators as

$$\phi(x) = \frac{1}{L^{\frac{d}{2}}} \sum_p \frac{1}{\sqrt{2\omega_p}} (a_p e^{-ip\cdot x} + a_p^\dagger e^{ip\cdot x}) \tag{2.28}$$

where $\omega_p = \sqrt{p^2 + m^2}$. The nonzero matrix elements for the interaction Hamiltonian between the Fock space vacuum and states with n particles

excited[31] are

$$\langle p_1 \cdots p_n | H_I | 0 \rangle = \frac{1}{2^{\frac{n}{2}} L^{d(\frac{n}{2}-1)}} \frac{\delta_{p_1+\cdots+p_n}}{\sqrt{\omega_1 \cdots \omega_n}} \tag{2.29}$$

From (2.20), we then have

$$S(\mu) = -\lambda^2 \log(\lambda^2) \sum_{\{p_i\}_\mu} \frac{\delta_{p_1+\cdots+p_n}}{2^n L^{d(n-2)} \omega_1 \cdots \omega_n (\omega_1 + \cdots + \omega_n)^2} + \mathcal{O}(\lambda^2) \tag{2.30}$$

where the sum is over distinct sets of spatial momenta such that at least one momentum is below the scale μ and at least one momentum is above the scale μ. More generally, the entanglement entropy for the field modes with momenta in some set A is given by the same formula with the sum over distinct sets of momenta such that at least one momentum is in a set A and at least one momentum is in $\bar{A}$.

It is straightforward to take the limit of infinite volume. By the usual replacements

$$\sum_{\vec{p}} \to \left(\frac{L}{2\pi}\right)^d \int d^d p \qquad\qquad \delta_{\vec{p}} \to \left(\frac{2\pi}{L}\right)^d \delta^d(p)$$

we find that the entanglement entropy density $S(\mu)/V$ has a well-defined limit:

$$S(\mu)/L^d = -\lambda^2 \log(\lambda^2) \frac{1}{(2\pi)^{d(n-1)} 2^n} \int_{\{p_i\}_\mu} \prod d^d p_i \frac{\delta(p_1 + \cdots + p_n)}{\omega_1 \cdots \omega_n (\omega_1 + \cdots + \omega_n)^2} + \mathcal{O}(\lambda^2) \ . \tag{2.31}$$

Here, the integral is again over distinct sets of momenta such that at least one momentum is below the scale μ and at least one momentum is above the scale μ.

[31] We are only interested in matrix elements between the vacuum and states with at least one low-momentum particle and at least one high-momentum particle. For ϕ^3 and ϕ^4 field theory, the only such non-zero matrix elements have 3 and 4 particles excited respectively. For ϕ^n theory with $n > 4$, matrix elements with $n - 2k \geq 3$ particle states can also contribute, but for these theories we must also include ϕ^{n-2k} counterterms in the action. For now, we focus on the case of ϕ^3 and ϕ^4 theory.

In practice, it is often simplest to calculate the derivative $dS/d\mu$, since the μ-dependence comes only in the domain of integration, and this domain for $S(\mu + d\mu)$ is almost the same as for $S(\mu)$. In the difference

$$S(\mu + d\mu) - S(\mu)$$

the only contributions that don't cancel between the two terms are a positive contribution in which one momentum is in the range $[\mu, \mu + d\mu]$ and all the other momenta have magnitude larger than μ, and a negative contribution in which one momentum is in the range $[\mu, \mu+d\mu]$ and all the other momenta have magnitude smaller than μ.

2.4.1 The ϕ^3 theory in 1+1 dimensions

The simplest example is the ϕ^3 theory in 1+1 dimensions.[32] From (2.31),

$$\begin{aligned} S(\mu)/V &= -\lambda^2 \log(\lambda^2) \frac{1}{32\pi^2} \int_{\{p_i\}_\mu} \prod dp_i \frac{\delta(p_1 + p_2 + p_3)}{\omega_1\omega_2\omega_3(\omega_1 + \omega_2 + \omega_3)^2} + \mathcal{O}(\lambda^2) \\ &\equiv -\lambda^2 \log(\lambda^2) \frac{1}{32\pi^2} I(\mu) \ . \end{aligned}$$

Letting

$$J(p_1, p_2, p_3) = \frac{1}{\omega_1\omega_2\omega_3(\omega_1 + \omega_2 + \omega_3)^2} \tag{2.32}$$

we find that

$$\frac{1}{2}\frac{dI}{d\mu} = \int_\mu^\infty dp\, J(\mu, p, -p - \mu) - \int_{-\mu/2}^0 dp\, J(\mu, p, -p - \mu) \ . \tag{2.33}$$

[32] We work with a massive ϕ^3 theory, so the theory is perturbatively stable. We can assume higher order interaction terms ϕ^{2n} which stabilize the theory non-perturbatively but do not affect our leading-order perturbative calculations.

Evaluating the right hand side analytically for large and small μ, we find that[33]

$$I(\mu) \to \begin{cases} \frac{\mu}{m^4}(\pi - \frac{8\sqrt{3}}{27}\pi - \frac{4}{3}) & \mu \ll m \\ \frac{1}{12\mu^3}\left\{\frac{23}{12} + \ln\left(\frac{\mu^2}{m^2}\right)\right\} & \mu \gg m \end{cases} \tag{2.34}$$

As discussed further below, the linear behaviour for small μ is related to the linear growth in the number of degrees of freedom below scale μ, while the falloff at large μ is related to fact that a ϕ^3 interaction is relevant in 1+1 dimensions so that the physics at large scales approaches that of the free theory, for which there is no entanglement between modes at different momenta.

Order λ^2 terms

Above we computed the $O(\lambda^2 \log(\lambda^2))$ term in the entanglement entropy that dominates at infinitesimal λ. At small, but finite λ, the $O(\lambda^2)$ term in (2.18) could compete with this. To calculate this term we must determine the eigenvalues (and not just the trace) of the matrix $CC^\dagger/\lambda^2$, where

$$C_{\{p_i\},\{P_i\}} = -\frac{\lambda}{L^{\frac{1}{2}}2^{\frac{3}{2}}} \frac{\delta_{\sum p_i + \sum P_i}}{(\sum \omega_{p_i} + \sum \omega_{P_i})\sqrt{\prod \omega_{p_i} \prod \omega_{P_i}}} \tag{2.35}$$

and the sets $\{p_i\}$ and $\{P_i\}$ must have either one and two elements or two and one elements. Thus the matrix $M = CC^\dagger/\lambda^2$ has nonzero elements of the form $M_{p,q}$ and $M_{\{p_1,p_2\},\{q_1,q_2\}}$. We have

$$M_{p,q} = \delta_{p,q} \frac{1}{8L} \sum_{P,Q} \frac{\delta_{p+P+Q}}{\omega_p \omega_Q \omega_P (\omega_p + \omega_Q + \omega_P)^2} \ . \tag{2.36}$$

Thus, for each p with $|p| < \mu$ we have one eigenvalue

$$a_p = \frac{1}{8L} \sum_{|P|>\mu, |Q|>\mu} \frac{\delta_{p+P+Q}}{\omega_p \omega_P \omega_Q (\omega_p + \omega_P + \omega_Q)^2} \ . \tag{2.37}$$

[33] To find $I(\mu)$ we evaluate the expression for $dI/d\mu$ and then integrate with respect to μ. The constant of integration is fixed by requiring that the entanglement entropy vanish as $\mu \to 0$.

The remaining block of the matrix M has entries

$$M_{\{p_1,p_2\},\{q_1,q_2\}} = \frac{\delta^*_{p_1+p_2,q_1+q_2}}{8L} \frac{1}{\sqrt{\omega_{p_1}\omega_{p_2}\omega_{q_1}\omega_{q_2}}\omega_{p_1+p_2}(\omega_{p_1}+\omega_{p_2}+\omega_{p_1+p_2})(\omega_{q_1}+\omega_{q_2}+\omega_{p_1+p_2})}$$

where δ^* indicates that we must have $|p_1+p_2|>\mu$ for a nonzero result. To find the remaining eigenvalues, we put M in block diagonal form, with one block for each P with $|P|>\mu$, where $p_1+p_2=q_1+q_2=P$. For the block labeled by P, we can label the matrix entries by p_1 and q_1, with

$$M_{p_1,q_1} = \frac{1}{8L}\frac{1}{\sqrt{\omega_{p_1}\omega_{P-p_1}\omega_{q_1}\omega_{P-q_1}}\omega_P(\omega_{p_1}+\omega_{P-p_1}+\omega_P)(\omega_{q_1}+\omega_{P-q_1}+\omega_P)} = \frac{V(p_1)V(q_1)}{8L}$$

where

$$V(p) = \frac{1}{\sqrt{\omega_p\omega_{P-p}\omega_P}(\omega_p+\omega_{P-p}+\omega_P)} \ . \tag{2.38}$$

A matrix of this form has only one nonzero eigenvalue, equal to

$$a_P = \frac{1}{8L}\sum_{|p|<\mu,|q|<\mu} \frac{\delta_{p+q+P}}{\omega_p\omega_q\omega_P(\omega_p+\omega_q+\omega_P)^2} \ . \tag{2.39}$$

We have one such eigenvalue for each P with $|P|>\mu$.

Having found all the eigenvalues of $CC^\dagger/\lambda^2$, we can use (2.18) to write an expression for $S(\mu)$ including the order λ^2 term. Recall that $S(\mu) = \lambda^2(1-\log(\lambda^2))\sum a_i - \lambda^2\sum a_i\log(a_i)$. Taking $L\to\infty$,

$$\sum a_i/L = \int \frac{dp_1}{2\pi} I(p_1) \tag{2.40}$$

and

$$\sum a_i\log(a_i)/L = \int \frac{dp_1}{2\pi} I(p_1)\log(I(p_1)) \tag{2.41}$$

where

$$I(p_1) = \int_* \frac{dp_2dp_3}{16\pi}\frac{\delta(p_1+p_2+p_3)}{\omega_1\omega_2\omega_3(\omega_1+\omega_2+\omega_3)^2} \ . \tag{2.42}$$

Here, the asterisk indicates that $p_2<p_3$, and that p_2 and p_3 must have magnitude less than μ if p_1 has magnitude greater than μ, while p_2 and p_3

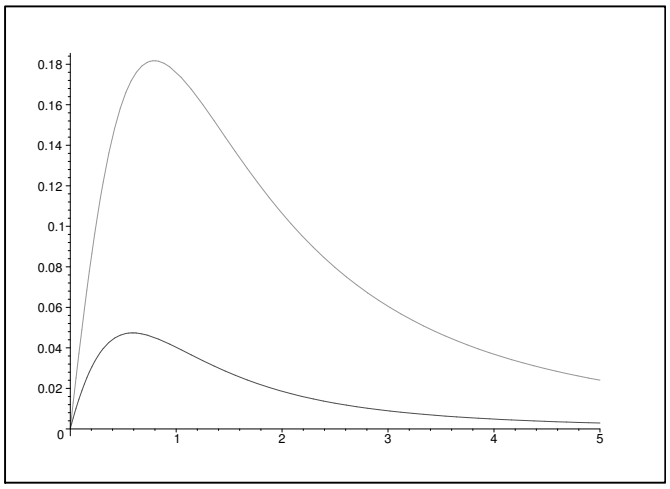


Figure 2.1: Leading contributions to $S(\mu)$ for ϕ^3 theory in 1+1 dimensions. Full result for $S(\mu)$ is proportional to $\lambda^2(\log(1/\lambda^2)+1)$ times bottom function plus λ^2 times top function.

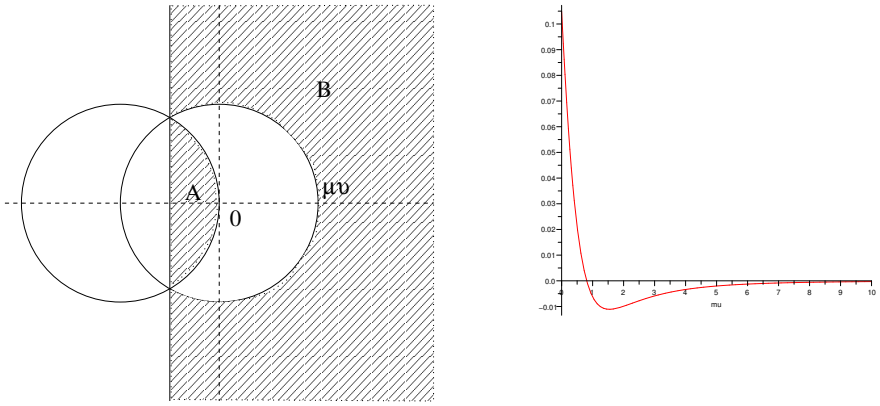
must have magnitude greater than μ if p_1 has magnitude less than μ.

We have plotted the two leading contributions (2.41) and (2.40) in figure 2.1. We see that the two terms have a qualitatively similar behavior. In detail, the term (2.41) falls off slightly more slowly for large μ, behaving as $1/\mu^3 \log^2(\mu^2/m^2)$ compared with $1/\mu^3 \log(\mu^2/m^2)$ for (2.40). Thus, for fixed λ and sufficiently large μ (of order m/λ), the $\mathcal{O}(\lambda^2)$ term will be larger than the $\mathcal{O}(\lambda^2 \log(\lambda^2))$ term, although the qualitative behavior is similar. In this work, our focus is on the physics in the limit of small λ, so in the remainder of the discussion we will concentrate on $\mathcal{O}(\lambda^2 \log(\lambda^2))$ terms which dominate as long as we stay below the parametrically large energies of order $1/\lambda$ relative to the mass.

2.4.2 The ϕ^3 theory in higher dimensions

In general dimensions, the entanglement entropy for the modes below scale μ in the ϕ^3 field theory is given by

$$S(\mu)/L^d = -\lambda^2 \log(\lambda^2) \frac{1}{8(2\pi)^{2d}} \int_{\{p_i\}_\mu} d^d p_1 d^d p_2 d^d p_3 \frac{\delta(p_1+p_2+p_3)}{\omega_1\omega_2\omega_3(\omega_1+\omega_2+\omega_3)^2} + \mathcal{O}(\lambda^2)$$



(**A**) (**B**)

Figure 2.2: (**A**) Integration regions for ϕ^3 theory in 2+1 dimensions. (**B**) The function $F(x)$ appearing in the entanglement entropy for ϕ^4 theory in $1+1$ dimensions.

$$\equiv \quad -\lambda^2 \log(\lambda^2) \frac{1}{8(2\pi)^{2d}} I_d(\mu) \ . \tag{2.43}$$

It is more convenient to compute

$$\frac{1}{\omega_{d-1}\mu^{d-1}} \frac{dI_d}{d\mu} = (\int_B - \int_A) d^2p J((\mu,0), \vec{p}, -(\mu,0) - \vec{p}) \tag{2.44}$$

where J is defined in (2.32), and A and B are the regions shown in Fig. 2.2A (symmetric between the vertical axis shown and the directions not depicted in the case $d > 2$). Here $\omega_d = 2\pi^{(d+1)/2}/\Gamma((d+1)/2)$ is the volume of the unit d-sphere.

Explicitly, we have

$$\begin{aligned} \frac{1}{\omega_{d-1}\mu^{d-1}} \frac{dI}{d\mu} \quad = \quad & -\int_{-\frac{\mu}{2}}^{0} dp_x \int_0^{\sqrt{\mu^2-(p_x+\mu)^2}} dp_T \omega_{d-2} p_T^{d-2} J(p_x, p_T) \\ & + \int_{-\frac{\mu}{2}}^{\mu} dp_x \int_{\sqrt{\mu^2-p_x^2}}^{\infty} dp_T \omega_{d-2} p_T^{d-2} J(p_x, p_T) \\ & + \int_{\mu}^{\infty} dp_x \int_0^{\infty} dp_T \omega_{d-2} p_T^{d-2} J(p_x, p_T) \end{aligned} \tag{2.45}$$

where

$$J(p_x,p_T) = \frac{1}{\sqrt{\mu^2+m^2}\sqrt{p_x^2+p_T^2+m^2}\sqrt{(\mu+p_x)^2+p_T^2+m^2}} \cdot \frac{1}{(\sqrt{\mu^2+m^2}+\sqrt{p_x^2+p_T^2+m^2}+\sqrt{(\mu+p_x)^2+p_T^2+m^2})^2} .$$

We find that in 2 + 1 dimensions, the entanglement entropy decreases with μ as

$$I_2(\mu) \to \frac{2\pi}{3\mu} \tag{2.46}$$

when $\mu \gg m$, while in 3+1 dimensions, we have

$$I_3(\mu) \to 2\pi^2(\ln(4)-1)\mu$$

for $\mu \gg m$. We interpret the the $3+1$ dimensional result as saying that in this case the μ^3 growth in the number of degrees of freedom below scale μ overwhelms the $1/\mu$ falloff of the effective dimensionless coupling. These expressions are exact (and finite) as $m \to 0$. For 4+1 dimensions and higher, (2.45) diverges – we will discuss this divergence below.

2.4.3 ϕ^4 theory

Finally, consider the ϕ^4 field theory in 1+1 dimensions. From (2.31),

$$\begin{aligned} S(\mu)/V &= -\lambda^2\log(\lambda^2)\frac{1}{16(2\pi)^3}\int_{\{p_i\}_\mu}\prod d^dp_i \frac{\delta(p_1+\cdots+p_4)}{\omega_1\cdots\omega_4(\omega_1+\cdots+\omega_4)^2}+\mathcal{O}(\lambda^2) \\ &\equiv -\lambda^2\log(\lambda^2)\frac{1}{16(2\pi)^3}I(\mu) . \end{aligned}$$

Thus, we study $I(\mu)=\int_{\{p_i\}_\mu} dp_1dp_2dp_3dp_4\,\delta(p_1+p_2+p_3+p_4)\,J(p_1,p_2,p_3,p_4)$, where $J(p_1,p_2,p_3,p_4)^{-1}=\omega_1\omega_2\omega_3\omega_4(\omega_1+\omega_2+\omega_3+\omega_4)^2$. It is again more convenient to evaluate

$$\frac{1}{2}\frac{dI}{d\mu} = \left\{\int_\mu^\infty dp_1\int_\mu^{p_1}dp_2+\int_\mu^\infty dp_1\int_{-\frac{p_1+\mu}{2}}^{-\mu}dp_2\right\}J(p_1,p_2,\mu,-p_1-p_2-\mu)$$

$$- \int_{-\mu}^{-\frac{\mu}{3}} dp_1 \int_{p_1}^{-\frac{p_1+\mu}{2}} dp_2\, J(p_1, p_2, \mu, -p_1 - p_2 - \mu) \qquad \equiv \qquad \frac{1}{m^4} F(\mu/m)\ .$$

A numerical integration determines F, giving the final result

$$S(\mu)/V = -\lambda^2 \log(\lambda^2) \frac{1}{384\pi^3} \frac{1}{m^3} \int_0^{\mu/m} dx F(x)\ . \tag{2.47}$$

The function $F(x)$ is plotted in Fig. 2.2B. By analyzing (analytically) the behavior of F for large and small x, we find that the entropy $S(\mu)$ behaves as μ/m^4 for small μ and as

$$S \sim \frac{1}{\mu^3} \ln^2(\mu/m)$$

for large μ. As for the ϕ^3 theory, the decay at large μ is related to the fact that the ϕ^4 theory in 1+1 dimensions is free in the UV.

The leading perturbative contribution to the entanglement entropy $S(\mu)$ of ϕ^4 theory can be similarly evaluated in 2+1 dimensions. The integrals there are more difficult to evaluate numerically, but are convergent. For 3+1 and higher dimensions, the integral expression for the leading contribution to $S(\mu)$ in the ϕ^4 theory has a UV divergence, which we discuss further below.

2.4.4 General remarks

Massless limits: We found above that in two and higher space dimensions, the entanglement entropy $S(\mu)$ has a finite limit as we take the mass to zero. However, in 1+1 dimensions, the results for both ϕ^3 theory and ϕ^4 theory diverge in the massless limit. These divergences suggest that $S(\mu)$ is not an "infrared-safe" quantity for a massless scalar theory in 1+1 dimensions. However, the ratio $S(\mu)/S(\mu_0)$ has a finite limit if we hold μ and μ_0 fixed as we take m to zero. The result is

$$S(\mu)/S(\mu_0) = (\mu_0/\mu)^3 \tag{2.48}$$

Thus, while it may not be sensible to talk about $S(\mu)$ directly for $m = 0$ and infinite volume, the ratio for different scales appears to be well-defined even in $1 + 1$ dimensions.

General understanding of large μ behavior: The results above agree with the following heuristic derivation of the power law behavior of $S(\mu)$ for large μ. The behavior is influenced by two significant effects. First, the number of degrees of freedom per unit volume below a momentum scale μ grows like μ^d. All else being equal, we expect S to scale like the number of degrees of freedom (for example, it is extensive). However the interactions in a general field theory depend on the scale, and this scale dependence also contributes to the behavior of $S(\mu)$. The dimensionless effective coupling for a ϕ^n interaction at scale μ behaves as $1/\mu^{d+1-n(d-1)/2}$. Since S goes like λ^2 (plus logarithmic corrections), we can estimate that $S(\mu)$ should behave as

$$S \sim \mu^d \times \left(\frac{1}{\mu^{d+1-n(d-1)/2}} \right)^2 = \frac{1}{\mu^{d+2-n(d-1)}}$$

up to possible logarithmic corrections. This is consistent with our results. At a technical level this scaling arises in the integrals for entanglement entropy from two sources: (a) the measure factors (i.e. the density of modes), and (b) the energy denominators in the interaction terms. These are the same ingredients that affect the scaling of physical observables during renormalization.

Divergences: In various specific case considered above, we found that the leading perturbative contribution to $S(\mu)/V$ is finite in the limit where the IR and UV cutoffs are removed. However, for the ϕ^3 theory in 4+1 and higher dimensions or the ϕ^4 theory in 3+1 and higher dimensions, the integral expressions for the leading perturbative contribution to $S(\mu)$ diverge. The divergence is associated with the sum over states in (2.20), or in the sum or integral over the momenta in (2.30) or (2.31) respectively. The leading divergence comes from the sum over states where one momentum has magnitude less than μ while the rest have magnitudes greater than μ. The

divergence is proportional to a power (or logarithm) of the UV cutoff Λ.

Of course, ultraviolet divergences are commonplace in quantum field theory. Typically, they are associated with integrals over momenta that appear beyond the leading order in perturbative calculations, and are dealt with by expressing the results in terms of renormalized (physical) parameters rather than bare parameters. However, here the divergences appear in leading order perturbative results. Since the bare and renormalized parameters are the same at leading order in perturbation theory, the divergences will not be eliminated by expressing the results in terms of renormalized parameters.[34]

Such divergences in leading order expressions indicate a breakdown of perturbation theory for the specific quantity that is diverging. To see this, note first that similar divergences appear even in fermionic theories, for example fermions in 1+1 dimensions with a $(\bar{\psi}\psi)^2$ interaction (see Appendix A.2 for details). Furthermore, the divergence is present even at finite volume, since it is associated with the infinite number of high-momentum modes which are still present with an IR regulator. But for a theory of fermionic fields at finite volume, the Hilbert space associated with degrees of freedom with momentum below a scale μ is finite-dimensional. In this case, there is an upper bound $S(\mu) < \log(N)$, where N is the dimension of the Hilbert space. Now, consider the theory with a UV cutoff Λ. The leading perturbative expression for $S(\mu)$ will be finite for any finite Λ. But since this expression diverges as Λ is taken to infinity, there will be some finite Λ above which this leading contribution to $S(\mu)$ is larger than the bound $\log(N)$. Here, Λ is still finite, so $S(\mu)$ is clearly well-defined, and the correct result for $S(\mu)$ must certainly be less than $\log(N)$, so the only possibility is that the leading perturbative expression is not a good approximation to the correct result.

Our conclusion should not be particularly surprising: regardless of how small the coupling parameter of a theory is, there will always be quantities that cannot be computed in perturbation theory. Here, the breakdown of

[34] We do expect the standard divergences to appear in higher order perturbative corrections, even for quantities whose leading order result is finite. These divergences should be cured in the usual way by expressing results in terms of physical parameters, or by using renormalized perturbation theory with the appropriate counterterms.

perturbation theory seems to be associated with computing the entanglement entropy between a finite set of modes with the infinite set of degrees of freedom above scale μ. We will see below that in cases where perturbation theory breaks down for this quantity, it is still possible to perturbatively calculate less inclusive quantities, such as the mutual information between degrees of freedom associated with two finite regions of momentum space. We will also encounter a case where the exact result is know but perturbation theory fails. In cases where no divergence appears at leading order, the finite leading-order perturbative result should be reliable so long as subsequent terms in the perturbative expansion (after renormalization) are small compared to the leading terms.

aaaaaaaaaaaaaaaaaa

2.5 The extent of entanglement between scales

So far, we have considered the entanglement between modes in a field theory above and below some scale μ. In this section, we ask about the entanglement entropy associated with a single mode of the field theory, or the mutual information between two individual modes. A version of the former observable has been considered previously in the condensed matter literature (see e.g. [71]). We will first consider the entanglement entropies of bounded regions of momentum space. These sorts of observables are useful for two reasons: (a) they can be finite even when the entanglement entropy for the low-energy density matrix diverges, (b) they are a much more sensitive and clear probe of the extent of entanglement since they don't sum over the entire tower of UV modes.

2.5.1 An aggregate measure of the range of entanglement

The quantity $S(\mu)$ measures entanglement between the complete set of degrees of freedom below the scale μ and the complete set of degrees of freedom above the scale μ. Is this entanglement largely between modes just above and below the scale μ, or is the entanglement "long-range" in momentum

space?

One way to address this question is to consider the entanglement entropy for the annular region $\mu_1 \leq |p| \leq \mu_2$ in momentum space. If the entanglement is short-range, then for $\mu_2 \gg \mu_1$, the entanglement entropy $S([\mu_1, \mu_2]) \equiv S(\mu_1 \leq |p| \leq \mu_2)$ should be dominated by entanglement between modes just above and below the scales μ_1 and μ_2. In addition, these separate contributions to the entanglement entropy should be well measured by $S(\mu_2)$ and $S(\mu_1)$. Thus, for short-range entanglement we would expect

$$S([\mu_1, \mu_2]) \approx S(\mu_1) + S(\mu_2) \qquad\qquad \mu_2 \gg \mu_1 \,. \tag{2.49}$$

Alternatively, consider the mutual information between the degrees of freedom with $|p| \geq \mu_2$ and $|p| \leq \mu_1$:

$$I(\mu_1, \mu_2) = S(\mu_1) + S(\mu_2) - S([\mu_1, \mu_2]) \,. \tag{2.50}$$

For short-range momentum space entanglement we expect (2.49). Hence, when $\mu_2 \gg \mu_1$ we expect that $I(\mu_1, \mu_2) \approx 0$. The rate of falloff of $I(\mu_2, \mu_1)$ as μ_2/μ_1 increases from 1 is a characterization of the extent of entanglement.

In ϕ^4 theory the infinite volume expression for $I(\mu_1, \mu_2)$ is (using (2.22))

$$S([\mu_1, \mu_2])/V = -\lambda^2 \log(\lambda^2) \frac{1}{24} \int_* \prod_i \frac{d^d p_i}{2(2\pi)^d} \frac{(2\pi)^d \delta(p_1 + p_2 + p_3 + p_4)}{(\omega_1 + \omega_2 + \omega_3 + \omega_4)^2 \omega_1 \omega_2 \omega_3 \omega_4} + \mathcal{O}(\lambda^2) \tag{2.51}$$

where the asterisk indicates that we integrate over momenta such that at least one $|p|$ is in the range $[\mu_1, \mu_2]$ and at least one $|p|$ is not in this range. For simplicity, we specialize to $d = 1$ and take the mass $m = 1$. It is simplest to first evaluate the quantity

$$\frac{d^2 S}{d\mu_1 d\mu_2} \,. \tag{2.52}$$

We can see that the only contribution to this will be from regions of the integral above where one momentum is at μ_1 and another momentum is at

μ_2.

This is equal to the integral above with $p_1 = \pm\mu_2$, $p_2 = \pm\mu_1$, and the remaining $|p|$s either both inside or both outside the interval $[\mu_1, \mu_2]$. The distinct choices of momenta satisfying these constraints are

$$\begin{array}{cccc} p_1 = \mu_2 & p_2 = \mu_1 & p_3 \in (-\infty, -2\mu_2 - \mu_1] \cup [-2\mu_1 - \mu_2, -\frac{1}{2}(\mu_1 + \mu_2)] \\ p_1 = \mu_2 & p_2 = -\mu_1 & p_3 \in (-\infty, -2\mu_2 + \mu_1] \cup [-\mu_1, \frac{1}{2}(\mu_1 - \mu_2)] \end{array}$$

or momenta obtained from these via $p_i \to -p_i$, where in all cases, p_4 is determined by the delta function constraint. Thus, we have

$$\begin{aligned} \frac{1}{V}\frac{d^2S}{d\mu_1 d\mu_2} = -\lambda^2 \log(\lambda^2)\frac{1}{12}\frac{1}{16(2\pi)^3}\Big\{ & \int_{-\infty}^{-2\mu_2-\mu_1} dp\; J(\mu_2, \mu_1, p, -p-\mu_1-\mu_2) \\ & + \int_{-2\mu_1-\mu_2}^{-\frac{1}{2}(\mu_1+\mu_2)} dp\; J(\mu_2, \mu_1, p, -p-\mu_1-\mu_2) \\ & + \int_{-\infty}^{-2\mu_2+\mu_1} dp\; J(\mu_2, -\mu_1, p, -p+\mu_1-\mu_2) \\ & + \int_{-\mu_1}^{\frac{1}{2}(\mu_1-\mu_2)} dp\; J(\mu_2, -\mu_1, p, -p+\mu_1-\mu_2)\Big\} \end{aligned}$$

where

$$J(p_1, p_2, p_3, p_4) = \frac{1}{(\omega_1+\omega_2+\omega_3+\omega_4)^2\omega_1\omega_2\omega_3\omega_4}\ .$$

To determine $S(\mu_1, \mu_2)$ from this expression, we can use $S(\mu, \mu) = 0$, $S(0, \mu) = S(\mu)$, and $\frac{\partial S}{\partial \mu_2}(0, \mu) = \frac{dS}{d\mu}(\mu)$. From these, we have

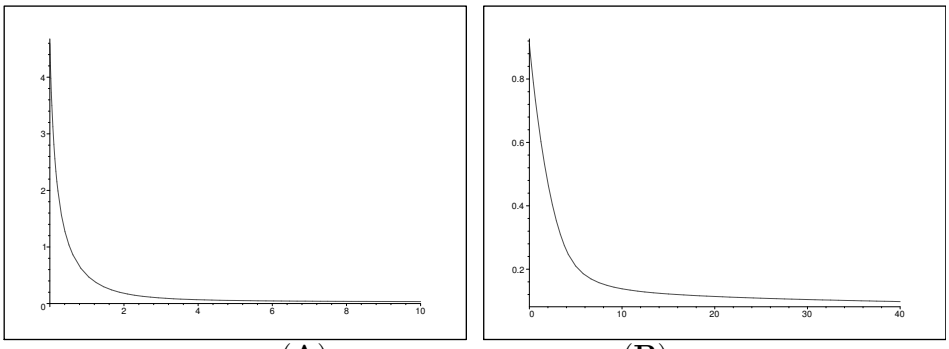
$$\frac{\partial S}{\partial \mu_2}(\mu_1, \mu_2) = \int_0^{\mu_1} d\tilde{\mu}_1 \frac{d^2S}{d\mu_1 d\mu_2}(\tilde{\mu}_1, \mu_2) + \frac{dS}{d\mu}(\mu_2) \tag{2.53}$$

and

$$S(\mu_1, \mu_2) = \int_{\mu_1}^{\mu_2} d\tilde{\mu}_2 \frac{dS}{d\mu_2}(\mu_1, \tilde{\mu}_2) = \int_{\mu_1}^{\mu_2} d\tilde{\mu}_2 \int_0^{\mu_1} d\tilde{\mu}_1 \frac{d^2S}{d\mu_1 d\mu_2}(\tilde{\mu}_1, \tilde{\mu}_2) + \int_{\mu_1}^{\mu_2} d\tilde{\mu}_2 \frac{dS}{d\mu}(\tilde{\mu}_2)\ . \tag{2.54}$$

Here, $S(\mu)$ is the quantity that we evaluated in previous sections.

To investigate whether the entropy $S(\mu_1, \mu_2)$ is dominated by entanglement between degrees of freedom close to μ_1 and μ_2, we can vary μ_2 and ask whether the variation of $S(\mu_1, \mu_2)$ is well approximated by the variation of



(**A**) (**B**)

Figure 2.3: Ratio of first and second terms in (2.53) vs $\mu = (\mu_2 - \mu_1)/m$ for (**A**) $\mu_1 = 1$ and (**B**) $\mu_1 = 4$. This is a measure of the range of entanglement in ϕ^4 theory in $1+1$ dimensions. We have taken the mass to be $m = 1$.

$S(\mu_2)$ (these variations would be equal if $S(\mu_1, \mu_2) = S(\mu_1) + S(\mu_2)$). From (2.53), the difference between the variations is equal to the first term on the right hand side, so we ask whether this term is small compared with the other term. In Fig. 2.3, we plot the ratio of these terms as a function of $\mu = (\mu_2 - \mu_1)/m$. The ratio declines as $1/\ln\mu$ increases and approaches a finite value as $(\mu_2 - \mu_1)/m \to 0$.

The slow rate of decline is surprising given that the ϕ^4 theory in 1+1 dimensions enjoys the property of decoupling. Note however, that the quantity we are computing integrates over all of the UV modes. Thus it is an aggregate measure of entanglement. A more refined way to ask about the range of entanglement in momentum space is to consider the mutual information between individual modes at two different momenta p and q as we do below. We will see that this mutual information falls off as a power law with $|q|$ when $|q| \gg |p|$.

2.5.2 Single mode entanglement

In this section, we calculate the entanglement entropy for a single mode with momentum $\vec{p}$. This measures the entanglement between a single mode and the rest of the field theory. The leading result for a ϕ^n scalar field theory

follows immediately from (2.30):

$$S(\vec{p}) = -\lambda^2 \log(\lambda^2) \sum_{\{p_2,\ldots,p_n\}} \frac{\delta_{p+p_2+\cdots+p_n}}{2^n L^{d(n-2)} \omega_p \omega_2 \cdots \omega_n (\omega_p + \cdots + \omega_n)^2} + \mathcal{O}(\lambda^2) \tag{2.55}$$

where the sum is over all distinct sets of $(n-1)$ momenta.[35] In the infinite volume limit, this gives

$$\begin{aligned} S(\vec{p}) &= -\lambda^2 \log(\lambda^2) \frac{1}{2^n (2\pi)^{d(n-2)}} \int_{\{p_2,\ldots,p_n\}} \prod_{i=2}^{n} d^d p_i \frac{\delta^d(p + p_2 + \cdots + p_n)}{\omega_p \omega_2 \cdots \omega_n (\omega_p + \cdots + \omega_n)^2} + \mathcal{O}(\lambda^2) \\ &\equiv s_1(|\vec{p}|) \ . \end{aligned}$$

By rotational invariance, the result is a function only of $|\vec{p}|$. All explicit factors of the volume have canceled out without dividing by volume on the left side.

A natural interpretation of this *finite* quantity is that it gives the entanglement entropy density for degrees of freedom in an infinitesimal range $d^d p$ around the momentum $\vec{p}$. The number of modes in the box $d^d p$ is proportional to spatial volume, so if the entanglement entropy for one mode has no explicit volume dependence, the entanglement entropy for the set of modes in the box should be proportional to volume. This entropy is also proportional to the momentum space volume $d^d p$ (if this is infinitesimal), so the entanglement entropy associated with an infinitesimal volume $d^d x$ in position space and volume $d^d p$ in momentum space takes the form

$$dS(\vec{p}) = \frac{d^d x \, d^d p}{(2\pi)^d} \ s_1(|\vec{p}|) \ . \tag{2.56}$$

It is interesting that the phase space volume appears naturally here.[36]

As an explicit example, we have plotted $s_1(p)$ for ϕ^3 theory in two, three, and four spacetime dimensions in Fig. 2.4. In the figure, the entropies are normalized by their value at $p = 0$. For $1+1$, $2+1$, and $3+1$ dimensions,

[35] For $\vec{p} = 0$, we have the further restriction that not all momenta are zero.

[36] Note that while this entropy is spatially extensive, it is not extensive in momentum space. That is, it is not true that $S(R)/V = \int_R d^d p f(p)$.

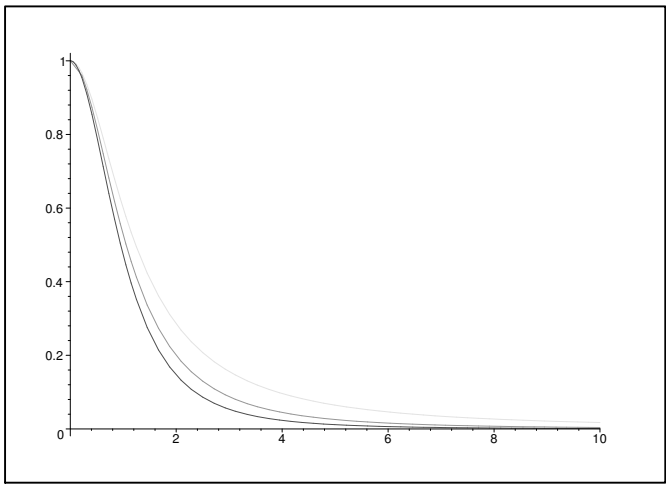


Figure 2.4: Single-mode entanglement entropy vs magnitude of mode momentum for ϕ^3 field theory in 1+1 (bottom), 2+1 (middle), and 3+1 (top) dimensions. The entropies are normalized by their values at $p = 0$.

the entanglement entropy decreases like $1/p^4$,$1/p^3$, and $1/p^2$ respectively. Thus we see that in this case the entanglement of a *single mode* with the rest of the theory declines as power-law of the momentum, even though we found above that the *integrated* entanglement between modes above and below scales μ_2 and μ_1 only declines logarithmically. The slow decay in the latter case is arising from the sum over modes.

2.5.3 Mutual information between individual modes

It is also interesting to investigate the mutual information between two specific field theory momenta. In the large volume limit, the natural quantity to consider is the mutual information between degrees of freedom in an infinitesimal range $d^d p$ around some momentum p and degrees of freedom in an infinitesimal range $d^d q$ around some momentum q. Starting from the basic formula (2.22), only contributions from the second term in curly brackets survive the large volume limit. These involve matrix elements between the vacuum state and states where one particle is excited in each of the regions $d^d p$ and $d^d q$, and the remaining particles lie outside these regions. The re-

sulting mutual information is proportional to $d^d p$ and $d^d q$ and to spatial volume, so we have:

$$I(\vec{p},\vec{q})/V = d^d p\, d^d q\, \mathcal{I}(\vec{p},\vec{q});. \tag{2.57}$$

For ϕ^n scalar field theory in $d+1$ dimensions, the leading contribution to $\mathcal{I}$ is

$$\mathcal{I}(\vec{p},\vec{q}) = -\lambda^2 \log(\lambda^2) \frac{1}{2^n (2\pi)^{d(n-1)}} \int_{\{p_3,\ldots,p_n\}} \prod_{i=3}^{n} d^d p_i \frac{\delta^d(p+q+p_3+\cdots+p_n)}{\omega_p \omega_q \omega_3 \cdots \omega_n (\omega_p+\cdots+\omega_n)^2} + \mathcal{O}(\lambda^2) \tag{2.58}$$

where the integral is over distinct sets of $n-2$ momenta. For ϕ^3 theory, this is

$$\mathcal{I}(\vec{p},\vec{q}) = -\lambda^2 \log(\lambda^2) \frac{1}{8(2\pi)^{2d}} \frac{1}{\omega_p \omega_q \omega_{p+q} (\omega_p+\omega_q+\omega_{p+q})^2} + \mathcal{O}(\lambda^2) \tag{2.59}$$

Thus, the mutual information is enhanced when $\vec{p}$, $\vec{q}$ or $(\vec{p}+\vec{q})$ are near zero, and for fixed p, the mutual information falls off as $1/|q|^4$ for large q. While this expression gives the formal leading order result in any number of dimensions, we will see below that it should only be trusted as an accurate approximation to the exact result for $d \leq 4$ space dimensions.

2.5.4 Convergence and validity of leading order expressions.

As for the entanglement entropy $S(\mu)$ considered in the previous section, the integrals in the leading order contributions to the mutual information and entanglement entropy of single modes can contain UV divergences. As we argued in Sec. 2.4.4, such divergences indicate a breakdown of perturbation theory for the quantity in question. In this subsection, we classify the scalar field theories for which the perturbative calculation of single-mode quantities $s_1(p)$ and $\mathcal{I}(p,q)$ gives sensible results.

We begin with the expression (2.58) for the single mode mutual information of ϕ^n scalar field theory in $d+1$ spacetime dimensions. Naively, this will converge (i.e. there are enough powers of momenta being integrated

over in the denominator) if

$$d < 1 + \frac{3}{n-3} \, . \tag{2.60}$$

Thus, we have convergence in any dimension for $n = 3$ for $d \leq 3$ for $n = 4$, for $d \leq 2$ for $n = 5$, and only for $d = 1$ for any higher n.

Since higher order interactions (i.e. interactions with more powers of the field) are more likely to lead to divergences, we should be concerned that such higher order interactions generated in the quantum effective action will produce divergences at higher orders in perturbation theory. In ϕ^3 theory we get an effective ϕ^n vertex at order λ^n from a one loop diagram. As a function of the external momenta, this scales like $1/p^{2n-d-1}$ as these momenta are taken large. The contribution to $I(p, q)$ from such a vertex will naively be convergent if

$$d < 5 + \frac{3}{n-1} \, . \tag{2.61}$$

This is satisfied for any n so long as $d \leq 5$, but leads to a divergence in 6 and higher space dimensions. Thus, it appears that $I(p, q)$ can be computed in perturbation theory for ϕ^3 theory in $d \leq 5$ (the same dimensions for which the theory is renormalizible).

For ϕ^4 theory, the effective action contains effective ϕ^{2n} interactions coming from one-loop diagrams at order λ^n. These scale with external momenta like $1/p^{2n-d-1}$. The contribution to $I(p, q)$ from such a vertex will naively be convergent if

$$d < 3 + \frac{1}{2n-1} \, ,$$

which is satisfied for any n as long as $d \leq 3$. Thus, it appears that $I(p, q)$ is a well defined quantity for ϕ^4 theory in $d \leq 3$ (again, the same dimensions for which the theory is renormalizible).

An almost identical analysis shows that the mutual information between degrees of freedom in any two finite region of momentum space converges whenever $I(p, q)$ converges. Note that it would be incorrect to suppose from the considerations above that the leading order $I(p, q)$ is necessarily well defined for every renormalizable theory. For example, according to

Theory	Dimensions where $I(A,B)$ converges	Dimensions where $S(A)$ converges
ϕ^3	$d \leq 5$	$d \leq 3$
ϕ^4	$d \leq 3$	$d \leq 2$
ϕ^5	$d \leq 2$	$d = 1$
$\phi^{n\geq 6}$	$d = 1$	$d = 1$

Table 2.1: Spatial dimensions where momentum space mutual information and entanglement entropy converge. The results apply for any bounded regions A and B in momentum space.

(2.61), the leading order $I(p,q)$ diverges for the renormalizible ϕ^6 theory in 3 dimensions.

We can also ask when the entanglement of a single mode (or a finite region of momentum space) with the rest of the field theory is well defined. For ϕ^n theory, we find convergence for

$$d < 1 + \frac{3}{n-2} \tag{2.62}$$

This result extends to entanglement entropy of any finite region of momentum space.

2.6 Entanglement in the Thirring model

In this section we wish to explore the entanglement structure of a 1+1 dimensional fermionic interacting field theory, known as the Thirring model. The reason is that powerful techniques have been developed to study such a theory. In fact this theory is known to be equivalent to a scalar field theory [11]. It is an example of a more general principle known as bosonization [18, 34], we will briefly review this concept and use it to obtain the exact ground state in the massless Thirring model, which we will then use to discuss the entanglement entropy of the theory.

2.6.1 The correspondence

The Lagrangian for the massless Thirring model is:

$$\mathcal{L}_F = -i\bar{\Psi}\gamma^\mu\partial_\mu\Psi + \frac{g}{2}J^\mu J_\mu, \tag{2.63}$$

where $J^\mu = \bar{\Psi}\gamma^\mu\Psi$. Here γ^μ are the usual gamma matrices in $1+1$ dimensions.

Bosonization gives us a correspondence between correlation functions in a free bosonic theory and a free fermionic theory [18, 34]. At the level of fields this correspondence can be written in terms of bilinears in the Fermi theory and scalars in the bosonic theory, in particular we have: $J^\mu \cong -\epsilon^{\mu\nu}\frac{1}{\sqrt{\pi}}\partial_\nu\phi$. Thus adding the term J^2 in the fermionic theory should be equivalent to adding the term $\frac{1}{\pi}(\partial\phi)^2$. Thus this interaction simply rescales the free boson action to:

$$\mathcal{L}_B = \frac{1}{2}\tilde{g}\left(\partial\phi\right)^2, \tag{2.64}$$

where $\tilde{g} = 1 + g/\pi$.

2.6.2 The ground state

Boson ground state

Now if we wish to find the exact ground state of the Thirring model what we need to do is find the exact ground state of the free bosonic model and "refermionize" it.

The free boson ground state $|\Omega\rangle$ can be written in path integral form by going to position basis:

$$\langle f(x)|\,\Omega\rangle = \int_{-\infty}^{\phi(0,x)=f(x)} D\phi\ \exp\left(-\tilde{g}\int\frac{1}{2}\left(\partial\phi\right)^2\right). \tag{2.65}$$

$$\tag{2.66}$$

Change the path integration variables from $\phi \to \phi_s + \delta\phi$ where ϕ_s is a

solution of the equations of motion with boundary condition $\phi_s(0,x) = f(x)$. Then using $\partial^2\phi_s = 0$ and integrating the action by parts with $\delta\phi(0,x) = 0$ it can be seen that the path integral factorizes into a piece that depends on $f(x)$ and one that does not.

$$\langle f(x)|\,\Omega\rangle = \int_{-\infty}^{\phi(0,x)=f(x)} D\phi\ \exp\left(-\tilde{g}\int \frac{1}{2}\left(\partial\phi\right)^2\right), \tag{2.67}$$

$$= \int_{-\infty}^{\delta\phi(0,x)=0} D\delta\phi\ \exp\left(-\tilde{g}\left(\int \frac{1}{2}\left(\partial\phi_s + \partial\delta\phi\right)^2\right)\right), \tag{2.68}$$

$$= \int_{-\infty}^{\delta\phi(0,x)=0} D\delta\phi\ \exp\left(-\tilde{g}\left(\int \frac{1}{2}\left(\partial\delta\phi\right)^2 + \int \frac{1}{2}\left(\partial\phi_s\right)^2\right)\right), \tag{2.69}$$

$$= \int_{-\infty}^{\delta\phi(0,x)=0} D\delta\phi\ \exp\left(-\tilde{g}\int \frac{1}{2}\left(\partial\delta\phi\right)^2\right)\exp\left(\tilde{g}\int dx\frac{1}{2}\left(f(x)\partial_t\phi_s(0,x)\right)\right), \tag{2.70}$$

$$= C\exp\left(\tilde{g}\int dx\frac{1}{2}\left(f(x)\partial_t\phi_s(0,x)\right)\right) \tag{2.71}$$

So we just need to find ϕ_s. This is easy to do once we have the Green's function (satisfies the delta function equation and vanishes on the boundary) since we have:

$$\phi_s(t,x) = \int dt'dx'\phi_s(t',x')\partial'^2 G(t,x;t',x') = \int' \left(\partial'_i(\phi'_s\partial'_i G) - \partial'_i\phi'_s\partial'_i G\right), \tag{2.72}$$

$$= \int'^{boundary} dA_i\left(\phi'_s\partial'_i G\right) - \int' \left(-\partial'^2\phi'_s G + \partial'_i(\partial'_i\phi'_s G)\right), \tag{2.73}$$

$$= \int dx'\phi_s(t'=0,x')\partial_{t'}G(t',x';t,x)|_{t'=0} \tag{2.74}$$

where the plus sign in the last line comes from the fact that dA_i is the outward normal, which points in the $+t$ direction. In general we only know the solution to the translation invariant equation $\partial^2 g(t-t',x-x') = \delta(t-t',x-x')$ and we use this to satisfy the boundary condition for $G(t,x;t',x')$. De-

manding G vanish at the boundary can be achieved by setting $G(t,x;t',x') = -g(t+t',x-x') + g(t-t',x-x')$ (this satisfies the Green's equation as well since t and t' must always be negative, thus we only get one delta function when we take the Laplacian).

Thus we can swap derivatives by t,x with negative derivatives by t',x', and since G satisfies Laplace's equation (on the boundary the δ function does not contribute) we can swap second derivatives by t,t' by negative second derivatives by x,x'. Then our vacuum state Eq. (2.71) becomes:

$$\ln\langle f(x)|\Omega\rangle = \tilde{g}\int dx\frac{1}{2}\left(f(x)\partial_t\phi_s(0,x)\right) \tag{2.75}$$

$$= \tilde{g}\int dxdx'\frac{1}{2}\left(f(x)f(x')\partial_t\partial_{t'}G(t,t';x,x')|_{t=0,t'=0}\right) \tag{2.76}$$

$$= \tilde{g}\int dxdx'\frac{1}{2}\left(f(x)f(x')\partial_t\partial_{t'}\left(g(t-t';x-x') - g(t+t',x-x')\right)|_{t=0,t'=0}\right) \tag{2.77}$$

$$= \tilde{g}\int dxdx'\frac{1}{2}\left(f(x)f(x')\partial_t^2\left(-g(t-t';x-x') - g(t+t',x-x')\right)|_{t=0,t'=0}\right) \tag{2.78}$$

$$= \tilde{g}\int dxdx'\frac{1}{2}\left(f(x)f(x')\partial_x^2 2g(0;x-x')\right) \tag{2.79}$$

$$= -\tilde{g}\int dxdx'\left(\partial_x f(x)\partial_{x'}f(x')g(0;x-x')\right) \tag{2.80}$$

For flat $D=2$ space the Green's function is just a logarithm of the distance [37]. Correctly normalized

$$g(t-t',x-x') = -\frac{1}{2\pi}\ln\sqrt{(t-t')^2+(x-x')^2}. \tag{2.81}$$

2.6.3 Thirring ground state

Using the correspondence between bosons and fermions mentioned at the beginning we can now go back to the fermion variables via $J^\mu \cong -\epsilon^{\mu\nu}\frac{1}{\sqrt{\pi}}\partial_\nu\phi$. Thus $-\partial_x f(x) = -\partial_x\phi(x) = \sqrt{\pi}J^0 = \sqrt{\pi}\rho(x) = \sqrt{\pi}\Psi^\dagger(x)\Psi(x)$.

[37] The electrostatic potential in 2D is $\sim \log|\mathbf{r}|$.

So using these relations in Eq. (2.80) with (2.81) we finally have that the ground state in fermionic variables is

$$\Omega[\Psi,\Psi^\dagger] = \exp\left(-\tilde{g}\int dxdx'\left(\Psi^\dagger(x)\Psi(x)\pi g(0,x-x')\Psi^\dagger(x')\Psi(x')\right)\right) \tag{2.82}$$

$$= \exp\left(\frac{\tilde{g}}{2}\int dxdx'\left(\Psi^\dagger(x)\Psi(x)\ln|x-x'|\Psi^\dagger(x')\Psi(x')\right)\right) \tag{2.83}$$

Two interpretations of this ground state will be explored:

- $\Omega[\Psi,\Psi^\dagger]$ gives a probability of Grassman configurations of bilinears, to be interpreted in the path integral sense over the grassman fields Ψ.
- $\Omega[\Psi,\Psi^\dagger] = \Omega[\rho(x)]$ gives the probability of the *bosonic* configuration $\rho(x) = \Psi^\dagger(x)\Psi(x)$. The probability has meaning on its own, and the path integral should be taken over the bosonic fields $\rho(x)$.

We will investigate both of these possibilities and the extent to which they differ.

Note that for the bose kinetic term to have the correct sign we have $\tilde{g} > 0$. Then from Eq. (2.83) configurations with two well separated lumps are more probable than two lumps side by side.

2.6.4 Checks

In this section we check that the ground state obtained is reasonable. The first quantized ground states that we will find from (2.83) have been explored in [36, 37].

Particles in $\mathbb{R}^2$

To check that 2.83 describes the ground state let us check that, in first quantization, $\tilde{g} = 1$ does indeed produce the ground state for the free theory. To do this go to fixed particle number subspace. Then $J^0 = \Psi^\dagger\Psi = \rho$ and for n

particles we have $\rho = \sum_{i=1}^{n} \delta(x - x_i)$. We can think about this as integrating all configurations in (2.83) with the constraint $\delta\left(\Psi^\dagger(x)\Psi(x) - \sum \delta(x - x_i)\right)$. This gives

$$\Omega_n = \prod_{i<j} |x_i - x_j|^{\tilde{g}}. \tag{2.84}$$

Particles at the same position do not appear. This would have been true even with a UV regularization of the Green's function due to the Grassman statistics, implementing the Pauli exclusion principle. In the bosonic interpretation this is nothing but the removal of the self energy.

This state has come out with the wrong statistics. This is what we should expect though since in first quantized form one must impose the correct statistics by hand. Adding the proper statistics we have:

$$\Omega_n = \prod_{i<j} |x_i - x_j|^{\tilde{g}} \mathrm{sgn}(x_i - x_j). \tag{2.85}$$

So now we search for a first quantized Hamiltonian for which this wave function is an eigenfunction. Computing the kinetic term we find

$$\sum_i -\frac{\partial^2}{\partial x_i^2}\Omega_n = -\sum_{i<j} 2\frac{\tilde{g}(\tilde{g}-1)}{(x_i - x_j)^2}\Omega_n. \tag{2.86}$$

Thus choosing the $1/r^2$ potential $V(x_i - x_j) = \frac{1}{(x_i - x_j)^2}$ we have that Ω_n is an eigenfunction of

$$H = p^2 + 2\tilde{g}(\tilde{g}-1)\sum_{i<j} V(x_i - x_j), \tag{2.87}$$

with energy $E = 0$. Since H is positive for $\tilde{g} \geq 1$, Ω_n is the ground state. Thus we now see that for $\tilde{g} = 1$, Ω_n becomes the ground state for the free Hamiltonian.

Particles in $\mathbb{R} \times S^1$

To investigate the theory on the cylinder one simply makes the conformal transformation $z = x - it \to \exp iz$. Since the free bosonic theory is conformally invariant we simply have $G(x-y) \to G(\exp(ix) - \exp(iy))$ and so the ground state becomes:

$$\Omega^{S^1}[\Psi, \Psi^\dagger] = \exp\left(\frac{\tilde{g}}{2}\int dxdy\left(\Psi^\dagger(x)\Psi(x)\ln|\exp(ix) - \exp(iy)|\Psi^\dagger(y)\Psi(y)\right)\right). \tag{2.88}$$

Then going to the n particle subspace again we have

$$\Omega_n^{S^1} = \prod_{i<j} |\exp(ix_i) - \exp(ix_j)|^{\tilde{g}}, \tag{2.89}$$

This product is nothing but a Vandermonde determinant, and so for $\tilde{g} = 1$ we have:

$$\Omega_n^{S^1} = |\det(\exp(ik_i x_j))|, \tag{2.90}$$

$$\to \exp(i\varphi)\det(\exp(ik_i x_j)). \tag{2.91}$$

with $k_i = 0, 1, .., n-1$, which is the quantization we would expect for modes on a circle. In going from line 1 to line 2 we have imposed the correct Fermi statistics, and noted that in so doing there is a phase ambiguity. This phase can depend on the coordinates, so long as it doesn't mess up the statistics. In fact choosing $\varphi = -p\sum_{i=1}^n x_i$ will shift the momenta k_i by $-p$ [38]. With

[38] This is easily seen by writing

$$\det(\exp(ik_i x_j)) = \sum_\sigma \mathrm{sgn}(\sigma)\exp(i\sum_j k_j x_{\sigma_j}) \tag{2.92}$$

where σ is a given permutation. Then the shift on the momenta follows from writing $\varphi = -p\sum_{i=1}^n x_i = -p\sum_{i=1}^n x_{\sigma_i}$.

such a choice the free Hamiltonian acting on this state gives:

$$-\sum_{i=1}^{n}\frac{\partial^2}{\partial^2 x_i}\Omega_n^{S^1}=\sum_{i=1}^{n}(k_i-p)^2\Omega_n^{S^1}. \tag{2.93}$$

Thus $\Omega_n^{S^1}$ is a free eigenstate, and it is in fact the ground state since the proper choice of phase will give it minimal energy. The proper choice is of course the center of mass frame, achieved when $p=(N-1)/2$.

2.6.5 Scaling

In what follows we will use some simple scaling arguments to find the entanglement entropy for the massless Thirring ground state at any coupling. The result is the same whether we use the bosonic or Grassman interpretation of (2.83).

Position space I

Consider how an interval of length L is entangled with its complement. We split up the fields $\Psi(x)=\Psi(x)^{<}+\Psi(x)^{>}$ into pieces inside ($<$) and outside ($>$) L. Then the reduced density matrix is:

$$\rho_L(\{\Psi_1^{<},\Psi_1^{\dagger<}\},\{\Psi_2^{<},\Psi_2^{\dagger<}\})= \tag{2.94}$$

$$\int D\Psi(x)^{>\dagger}D\Psi(x)^{>}\ \Omega[\{\Psi_1^{<},\Psi_1^{\dagger<}\},\{\Psi^{>},\Psi^{\dagger>}\}]\Omega[\{\Psi_2^{<},\Psi_2^{\dagger<}\},\{\Psi^{>},\Psi^{\dagger>}\}]. \tag{2.95}$$

Then

$$\text{Tr }\rho_L^n=\int\prod_{i=1}^{n}D\Psi_i(x)^{>\dagger}D\Psi_i(x)^{>}\rho_L(1,2)\rho_L(2,3)...\rho_L(n,1). \tag{2.96}$$

To give finite results this should be regulated in the UV by a lattice spacing a. The measure splits up into degrees of freedom living on each site. Let us also put the system in a box of size B ($B>>L$).

Now consider the effect of local field rescalings (note we are not making a conformal transformation, as we are leaving the position fixed). Under $\Psi^> \to \sqrt{k}\Psi^>$ we have $\rho_L \to k^{(B-L)/a}\rho_L^k$ (where we have indicated that the new density matrix has some dependence on k). Under $\Psi^< \to \sqrt{k}\Psi^<$ we have $\text{Tr } \rho_L^n \to k^{(nL)/a} \text{ Tr } (\rho_L^k)^n$. Thus under both rescalings $\Psi \to \sqrt{k}\Psi$, we simply have

$$\text{Tr } \rho_L^n = k^{nB/a} \text{ Tr } \left(\rho_L^k\right)^n, \tag{2.97}$$

where B/a is simply the total number of sites. Now using Eq. (1.41) the overall factor disappears and we have

$$\text{Tr } \rho_L \ln \rho_L = \text{ Tr } \rho_L^k \ln \rho_L^k. \tag{2.98}$$

If we choose the arbitrary parameter k to be $1/\tilde{g}^{1/2}$ then the RHS of this last equation becomes independent of $\tilde{g}$. This means that the entanglement entropy for the Thirring model with arbitrary coupling g is the same as the free theory with $g = 0$ ($\tilde{g} = 1$). The entanglement entropy for the free theory has been well studied and the result is (for $B \to \infty$)

$$S = S_{\text{free}} = \frac{c}{6} \ln \frac{L}{a}. \tag{2.99}$$

Position Space II

Another way to get this same result is as follows. Birrell and Davies [7] have shown that the conformal anomaly in curved space is independent of the coupling so that

$$< T_\mu^\mu >= -\frac{1}{24\pi} R. \tag{2.100}$$

Thus the Thirring model is conformal invariant and so we must have

$$S = \frac{c(g)}{6} \ln \frac{L}{a}, \tag{2.101}$$

where we have noted that the only possible way that the dimensionless cou-

pling g can now enter is through the central charge. We know, however, that the central charge counts some physical degrees of freedom in the system. Thus starting from the free theory and adiabatically turning on the coupling, the physical degrees of freedom cannot continuously change, and so we are left with the same result as before:

$$S = \frac{c}{6} \ln \frac{L}{a}. \tag{2.102}$$

Momentum space

The calculation of momentum space entanglement entropy is basically the same as in position space. Consider the entanglement between modes in an interval of length μ in momentum space and its complement. The Fourier transform of the ground state (2.83) is denoted: [39]

$$\Omega[\psi_k, \psi_k^\dagger] = \exp\left(\frac{\tilde{g}}{2} \int \frac{dp}{2\pi} G(p) | \int \frac{dk}{2\pi} \psi_{k+p}^\dagger \psi_k |^2\right). \tag{2.105}$$

Again splitting fields into those with support inside the μ interval and those outside, $\psi = \psi^< + \psi^>$ the reduced density matrix for modes inside is:

$$\rho_\mu(\{\psi_{1k}^<, \psi_{1k}^{\dagger <}\}, \{\psi_{2k}^<, \psi_{2k}^{\dagger <}\}) = \int D\psi_k^{>\dagger} D\psi_k^> \, \Omega[\{\psi_{1k}^<, \psi_{1k}^{\dagger <}\}, \{\psi_k^>, \psi_k^{\dagger >}\}] \Omega[\{\psi_{2k}^<, \psi_{2k}^{\dagger <}\}, \{\psi_k^>, \psi_k^{\dagger >}\}] \tag{2.106}$$

Impose a UV cutoff in momenta denoted Λ and put the system in a box of size V ($\Lambda >> 1/V$).

Now under $\psi^> \to \sqrt{k}\psi^>$ we have $\rho_\mu \to k^{(\Lambda-\mu)V/\pi} \rho_\mu^k$. Under $\psi^< \to \sqrt{k}\psi^<$ we have Tr $\rho_\mu^n \to k^{(n\mu V)/\pi}$ Tr $(\rho_\mu^k)^n$. Thus under both rescalings

[39] In what follows the Fourier transform requires regularization. In the IR this can be accomplished by adding a mass term. For small mass one uses

$$\ln(m\Delta_{xy}) \approx K_0(m\Delta_{xy}) \tag{2.103}$$

$$= \int dp \frac{1}{\sqrt{p^2 + m^2}} e^{ip\Delta_{xy}}, \tag{2.104}$$

K_0 being the modified Bessel function of the second kind.

$\psi \to \sqrt{k}\psi$, we simply have

$$\text{Tr } \rho_\mu^n = k^{n\Lambda V/\pi} \text{ Tr } \left(\rho_\mu^k\right)^n \tag{2.107}$$

Again using Eq. (1.41) the overall factor disappears and we have

$$\text{Tr } \rho_\mu \ln \rho_\mu = \text{ Tr } \rho_\mu^k \ln \rho_\mu^k. \tag{2.108}$$

So again if we choose k to be $1/\tilde{g}^{1/2}$ the entanglement entropy at any coupling is the same as in the free theory, which must vanish. Thus the entanglement between modes in the massless interacting Thirring model must vanish identically.

2.6.6 Direct calculation

Grassman way

In this section we will directly calculate the momentum-space entanglement of the ground state. We put the system in a box of length L with anti-periodic boundary conditions for the fermions. We will consider for simplicity the single mode entanglement.

We regulate the Green's function in the IR by adding a small mass term. Then since our definition of the entanglement is independent of normalization, we multiply by the infinite constant [40]:

$$\Omega \to \det{}^{-1/2}(G_p)\Omega \tag{2.109}$$

$$= \int Du_q \exp\left(-\frac{1}{L}\sum_q G_q|u_q|^2\right) \exp\left(\frac{\tilde{g}}{2L^3}\sum_p G_p|\sum_k \psi^\dagger_{k+p}\psi_k|^2\right) \tag{2.110}$$

$$= \int Du_q \exp\left(-\frac{1}{L}\sum_q G_q|u_q|^2 - \frac{\sqrt{2\tilde{g}}}{L^2}\sum_{k,m}\psi^\dagger_k G_{k-m}u_{k-m}\psi_m\right) \tag{2.111}$$

[40] The mass also removes the zero mode in the determinant, $\det{}^{-1} G_p = \prod_p \sqrt{p^2+m^2} = \prod_n \sqrt{(2\pi/L)^2\ n^2 + m^2}$.

where in going from line 2 to 3 we have made the shift $u_q \to u_q + \sqrt{\tilde{g}/2L^2}\sum_k \psi^\dagger_{k+q}\psi_k$ (k is half-integer moded and q is integer).

Then one finds for the reduced density matrix:

$$\rho(\psi^<_L,\psi^<_R) = \int D\psi^{\dagger>}_k D\psi^>_k \Omega[\psi^<_L,\psi^>]\Omega[\psi^<_R,\psi^>] \tag{2.112}$$

$$= \int Du_S \ \det\left(\sum_S M^{>>}_S\right) \exp\left(-\frac{1}{L}\sum_{q,S} G_q |\,(u_S)_q\,|^2\right) \exp\left(-\frac{\sqrt{2\tilde{g}}}{L^2}\psi^{\dagger<}_S \mathbb{M}^{<<}_{SS'}\psi^<_{S'}\right) \tag{2.113}$$

$$\begin{cases} \sum_S f_S = f_L + f_R \\ Du_S = Du_L Du_R \\ (M_S)_{k,m} = G_{k-m}\,(u_S)_{k-m} \\ \mathbb{M}^{<<}_{SS'} = \begin{pmatrix} \mathbb{M}^{<<}_{LL} & \mathbb{M}^{<<}_{LR} \\ \mathbb{M}^{<<}_{RL} & \mathbb{M}^{<<}_{RR} \end{pmatrix} \\ \qquad\quad = \begin{pmatrix} M^{<<}_L + M^{<>}_L \left(\sum_S M^{>>}_S\right)^{-1} M^{><}_L & M^{<>}_L \left(\sum_S M^{>>}_S\right)^{-1} M^{><}_R \\ M^{<>}_R \left(\sum_S M^{>>}_S\right)^{-1} M^{><}_L & M^{<<}_R + M^{<>}_R \left(\sum_S M^{>>}_S\right)^{-1} M^{><}_R \end{pmatrix} \end{cases} \tag{2.114}$$

Now we wish to compute Tr ρ^n. However, this is a fermionic path integral which can require non trivial boundary conditions, as in the standard case of the thermal partition function where the Fermi fields are anti-periodic. The boundary conditions in position space have been investigated in [10, 25] where it was found that there is an extra minus sign coming in between each cut. This comes from considering the way spinors rotate. In momentum space we will consider both periodic and anti-periodic boundary conditions, i.e. $\psi^i_R = e^{i\pi\alpha}\psi^{i+1}_L$, periodic and anti-periodic being recovered for $\alpha = 0, 1$. We will find that the result is independent of our choice of boundary conditions. Note that we cannot consider $\psi^{i\dagger}_R = e^{i\pi\alpha}\psi^{i+1}_L$ since these fields do not transform in the same way.

Using these boundary conditions we arrive at:

$$\text{Tr } \rho^n = \int \prod_i^n \left[Du_S^i \ \exp\left(-\frac{1}{L} \sum_{q,S} G_q |\left(u_S^i\right)_q|^2 \right) \det\left(\sum_S \left(M^i\right)_S^{>>} \right) \right] \det_n \left(\mathbb{M}\left(u_S^j\right) \right), \tag{2.115}$$

$$\equiv \int \prod_i \mathcal{D}U^i \ \det_n \left(\mathbb{M}\left(u_S^j\right) \right), \tag{2.116}$$

where

$$\det_n \left(\mathbb{M}\left(u_S^i\right) \right) = \det \begin{pmatrix} \mathbb{M}_{RR}^n + \mathbb{M}_{LL}^1 & e^{i\pi\alpha}\mathbb{M}_{LR}^1 & 0 & \cdots & 0 & e^{-2i\pi\alpha}\mathbb{M}_{RL}^n \\ e^{-i\pi\alpha}\mathbb{M}_{RL}^1 & \mathbb{M}_{RR}^1 + \mathbb{M}_{LL}^2 & e^{i\pi\alpha}\mathbb{M}_{LR}^2 & 0 & \cdots & 0 \\ 0 & e^{-i\pi\alpha}\mathbb{M}_{RL}^2 & \mathbb{M}_{RR}^2 + \mathbb{M}_{LL}^3 & \ddots & & \\ \vdots & 0 & \ddots & \ddots & & \\ \vdots & \vdots & & & & \\ 0 & & & & & \\ e^{2i\pi\alpha}\mathbb{M}_{LR}^n & 0 & \cdots & & & \mathbb{M}_{RR}^{n-1} + \mathbb{M}_{LL}^n \end{pmatrix}, \tag{2.117}$$

and where we have defined

$$\left\{ \mathbb{M}_{SS'}^i \equiv \mathbb{M}_{SS'}^{<<}\left(u_L^i, u_R^i\right). \right. \tag{2.118}$$

Notice the extra factor of $e^{2i\pi\alpha}$ that comes in between the first and last entry, due to the fermion statistics.

Now consider the parity of $d(u_L^i, u_R^i) \equiv \det\left(\sum_S \left(M^i\right)_S^{>>}\right)$ in Eq. (2.115). With a cutoff $\Lambda = (2\pi/L)N_{max}$ such that $|k| \leq \Lambda$ then

$$\dim\ \left(M^i\right)_S^{>>} = 2\dim(\Lambda) - \dim(\mu), \tag{2.119}$$

$$= 2N_{max} - n_\mu, \tag{2.120}$$

with μ denoting the length of the interval we are entangling with the rest of the system. This implies that

$$d(-u_L^i, -u_R^i) = (-1)^{n_\mu}\, d(u_L^i, u_R^i). \tag{2.121}$$

We are considering the single mode entanglement and so the determinant is odd, $d(-u_L^i, -u_R^i) = -d(u_L^i, u_R^i)$.

It then follows that

$$\text{Tr}\ \rho^n = \int \prod_i \mathcal{D}U^i\ \det{}_n\left(\mathbb{M}\left(u_S^j\right)\right), \tag{2.122}$$

$$= \int \prod_i \mathcal{D}U^i \left(2\prod_i \mathbb{M}_{RR}^i + \left[e^{i(n+1)\alpha\pi}\prod_i \mathbb{M}_{LR}^i + c.c.\right]\right), \tag{2.123}$$

$$= 2\left(\int \mathcal{D}U^j \mathbb{M}_{RR}^j\right)^n + \left[e^{i(n+1)\alpha\pi}\left(\int \mathcal{D}U^j \mathbb{M}_{LR}^j\right)^n + c.c.\right], \tag{2.124}$$

$$= 2a^n + \left[e^{i(n+1)\alpha\pi} b^n + c.c.\right], \tag{2.125}$$

which gives for the entropy

$$S = \ln\left(a + \cos(2\pi\alpha)b\right) - \frac{a\ln a + \cos(2\pi\alpha) b\ln b - (1+\alpha)\pi\sin(2\pi\alpha)b}{a + \cos(2\pi\alpha)b} + \ln 2. \tag{2.126}$$

This actually only depends on the ratio $r = b/a$, and for the usual fermion boundary conditions where $\alpha = 0, 1$ we find:

$$S = \ln(1+r) - \frac{r\ln r}{1+r} + \ln 2. \tag{2.127}$$

Thus we find S is bounded between $\ln 2$ (for $r \to \infty$ or $r \to 0$) and $\ln 4$

(for $r = 1$), which is the maximum entropy for a single mode of a relativistic fermion.

Bosonic way

In this approach we treat $\rho(x)$ as the fundamental field variable. Whether ρ is a density matrix or fermion density should be clear by its argument. Then

$$\Omega[\rho(x)] = \exp\left(\frac{\tilde{g}}{2}\int dxdx' \left(\rho(x)\ln|x-x'|\rho(x')\right)\right). \tag{2.128}$$

This is diagonal in k-space

$$\Omega[\rho(k)] = \exp\left(\frac{\tilde{g}}{2}\int \frac{dk}{2\pi} G(k)\rho(k)\rho^*(k)\right), \tag{2.129}$$

so that the integration over large modes k simply produces an overall normalization, which can be dropped. The reduced density matrix is

$$\rho[\rho_L^<(k),\rho_R^<(k)] = \exp\left(\frac{\tilde{g}}{2}\int \frac{dk}{2\pi}\begin{pmatrix}\rho_L^<(k) & \rho_R^<(k)\end{pmatrix}\begin{pmatrix}G(k) & 0\\ 0 & G(k)\end{pmatrix}\begin{pmatrix}\rho_L^<(-k)\\ \rho_R^<(-k)\end{pmatrix}\right). \tag{2.130}$$

Being diagonal, the traces in the replica trick simply factorize

$$Z_n = (Z_1)^n, \tag{2.131}$$

and so we simply have

$$S = 0. \tag{2.132}$$

This is the result we argued for in the scaling section. These considerations suggest that the Thirring ground state 2.83 makes sense in terms of density profiles rather than Grassman configurations.

2.7 Comments

We have obtained a number of results for the scaling of entanglement entropies and mutual informations with the upper bound on a momentum interval. At a technical level, these results all follow from the density of modes (measure) in the integrals over momenta, and the energy denominators in the interactions. These are the same ingredients that lead to the decoupling property of local quantum field theories. Indeed, decoupling is usually understood simply as the power law suppression of higher dimension operators in a low energy effective theory. This suppression means that high momentum degrees of freedom have weak effects on the dynamics of low-momentum degrees of freedom other than renormalizing the interaction strengths and wavefunctions. Our study of entanglement between degrees of freedom with different momenta, and the resulting entanglement entropies and mutual informations, refines this understanding of the influence between momentum scales.

In more detail, "decoupling" between UV and IR physics implies that starting from a generic action $S^{\Lambda}(g_I)$ at scale Λ, that depends on an infinite number of parameters g_I, the Wilsonian effective action at a much lower scale μ will be very close to some action $S^{\mu}_{W}(g_i)$ in a family parameterized by a small number of physical parameters g_i. In other words, the operation of integrating out degrees of freedom to successively lower scales results in a flow in the space of effective actions that converges to a low dimensional subspace at scales $\mu \ll \Lambda$. Now, according to (2.5), our effective action S^{μ}_{W} at scale μ completely determines the reduced density matrix $\rho(\mu)$ for the degrees of freedom with $|p| < \mu$. Thus we conclude that for the ground state of a generic field theory defined at scale Λ, the reduced density matrix for the degrees of freedom below some much lower scale μ will be very close to some family of density matrices $\rho(\mu, g_i)$ that depend on a small number of physical parameters g_i. Consequently, knowing the state of the low-momemtum degrees of freedom tells us relatively little about the details of the state at much higher scales.

The paucity of information about UV physics contained in the low-

momentum density matrix should be reflected in some of the measures of quantum information we have discussed. Specifically, it seems likely that there is a connection between the decoupling behavior of field theories and the power-law fall off in mutual information observed in Sec. 2.5. It would be interesting to make this connection precise.

Relation to AdS/CFT: In the context of gauge-theory / gravity duality (the AdS/CFT correspondence) [82], there is now evidence that certain measures of entanglement in quantum field theory carry geometrical information about the dual spacetime. For field theories with a weakly curved dual gravity description, Ryu and Takayanagi have proposed 1.86 that the entanglement entropy for a spatial region A is proportional to the area of the minimal surface $\tilde{A}$ in the bulk space whose boundary coincides with the boundary of A,

$$S(A) = \frac{Area(\tilde{A})}{4G_N} \ .$$

Given the holographic interpretation of position-space entanglement entropy, it is natural to ask whether the momentum-space quantities considered in this chapter have some simple dual geometrical interpretation for field theories with gravity duals. As an example, the quantity $S(\mu)$ measures the entanglement between degrees of freedom above and below the scale μ. Since energy/momentum scale in holographic field theories corresponds to radial position in the dual geometry, we might guess that $S(\mu)/V$ is related to the area (per unit field theory volume) of a surface separating the IR region $r < r(\mu)$ of the dual geometry from the UV region $r > r(\mu)$. For the dual geometry to a translation-invariant field theory state, this area function is a well-defined observable.[41] However, we currently have no way to check whether this or some similar observable corresponds to momentum-space entanglement entropy, since we cannot calculate this entropy for any strongly coupled field theory with a gravity dual.[42]

[41] If the spatial part of the dual metric is $dr^2 + f(r)dx^2$, the area of the surface at radius $r(\mu)$ per unit field theory volume is proportional to a power of $f(r(\mu))$.

[42] $S(\mu)$ will probably not always correspond in a simple way to the specific area observable mentioned, since that area would be well-defined even for gravity duals of 0+1

Relation to DMRG and MERA: Here we have explored various aspects of entanglement in quantum field theory and the connection to renormalization theory. In the condensed matter literature, the ideas of entanglement and renormalization have come together previously in various schemes for approximating the ground state of many-body systems [64, 65]. While the focus and details of that work are rather different from the present discussion, it may be useful to briefly review those ideas here.

Consider a quantum many-body system described by some lattice of degrees of freedom, for which the Hilbert space decomposes as a tensor product of Hilbert spaces for the individual sites. The dimension of the full Hilbert space is d^N where d is the dimension of the individual Hilbert spaces and N is the number of sites. A general state (and in particular, the exact ground state of the system for a given Hamiltonian) can be represented exactly by a tensor $T^{i_1 \cdots i_N}$ that gives the coefficient of the basis state $|i_1\rangle \otimes \cdots \otimes |i_N\rangle$.

A general numerical determination of the ground state is impractical due to the large number d^N of independent coefficients. For certain systems, usually in 1+1 dimensions, an efficient variational approach to approximating the ground state is to consider tensors T that can be decomposed into contractions of lower-rank tensors. For example, the "Matrix Product State" (MPS) decomposition corresponds to

$$T^{i_1 \cdots i_N} = (M_1)^{i_1}_{a_1 a_2} (M_2)^{i_2}_{a_2 a_3} \cdots (M_N)^{i_N}_{a_N a_1} \ .$$

In practice, one uses this decomposition as a variational ansatz, varying the individual matrices M^i to arrive at the best approximation to the ground state. If the dimension of the matrices M^i is large enough, any tensor T can be represented in this way, so the variational method gives an exact result. However, for a wide class of systems, it has been found that the ground state can be well approximated by matrices of much lower dimension. In this case, the matrix product ansatz represents a truncation of the Hilbert

dimensional field theories, for which there is no way to divide up the degrees of freedom by spatial momentum, and therefore no way to define $S(\mu)$. Of course such low-dimensional gauge/gravity dualities (e.g. $\mathrm{AdS}_2/\mathrm{CFT}_1$) also have many other special features [62, 63]).

space to a subspace of lower dimension, and in cases where it is effective, the true ground state is close to the ground state in this subspace.

It turns out that the success of this method is related to the entanglement properties of the ground state. The optimal method of truncating to a lower-dimensional Hilbert space is to retain as much of the entanglement entropy for the various subsystems (blocks of sites) as possible.[43] The procedure works most efficiently (i.e. for smallest matrices M) when there is limited entanglement between the subsystems corresponding to blocks of sites. For systems with a highly entangled ground state, the method is much less efficient.

Another approach that is more successful in cases with long-range entanglement is the "Multiscale Entanglement Renormalization Ansatz" (MERA [65]). In this approach the tensor T is represented by an iterative procedure. The tensor is first written in terms of a "disentangled" tensor $\tilde{T}$ using unitary matrices U:

$$T_{(n)}^{i_1\cdots i_{2N}} = (U_1^{(n)})_{j_1 j_2}^{i_1 i_2} \cdots (U_N^{(n)})_{j_{2N-1} j_{2N}}^{i_{2N-1} i_{2N}} \tilde{T}_{(n)}^{j_1\cdots j_{2N}}$$

and then $\tilde{T}$ is represented in terms of a lower rank tensor using "projectors" P:

$$\tilde{T}_{(n)}^{j_1\cdots j_{2N}} = (P_1^{(n)})_{I_1}^{j_1 j_2} \cdots (P_N^{(n)})_{I_N}^{j_{2N-1} j_{2N}} T_{(n+1)}^{I_1\cdots I_N} \ .$$

The latter step can be understood as a "coarse-graining" of the system, though the dimension of the index space I is not necessarily the same as that of the original index space i. The original tensor $T_{(1)}$ is thus represented by the individual matrices $(U_i^{(n)})_{i'j'}^{ij}$ and $(P_i^{(n)})_I^{ij}$ which are the variational parameters used to approximate the ground state.

The introduction of the U matrices is motivated by the observation that coarse graining works most efficiently when there is little entanglement between the adjacent blocks. The unitary matrices U can remove short-range

[43] This idea arose first in the "Density Matrix Renormalization Group" (DMRG) [64] an iterative renormalization procedure on the *state* of the system that truncates the Hilbert space in each step while retaining as much entanglement entropy as possible. The DMRG is now understood to give results equivalent to this MPS variational method.

entanglement between adjacent blocks before coarse graining. In this way, the matrices $U_{(n)}$ encode the entanglement between sites at the nth level, which corresponds in the original picture to blocks of 2^n sites. Thus, in the MERA representation of a ground state the unitary matrices U encode entanglement at different scales. This information is certainly related to the scale-dependent entanglement entropies considered in this chapter, though the MERA entanglements would seem to be more closely related to position space entanglement. Also, the original MERA applies only to discrete systems, though an extension to continuum quantum field theories has been recently proposed in [66]. An interesting connection between MERA and the AdS/CFT proposal above has been given in [67].

Chapter 3

Momentum-space entanglement for interacting fermions at finite density

3.1 Introduction

The physics of finite density systems of interacting fermions plays a crucial role in our understanding of a wide variety of condensed matter systems, from ordinary materials to nuclear matter in a neutron star. To understand even crude macroscopic properties of these systems, quantum mechanics (e.g. the Pauli Exclusion Principle) is essential. Nevertheless, our theoretical investigations often focus on classical observables such as thermodynamic quantities, correlation functions, and response functions, since these quantities are simpler to access via experiment. In this chapter, we instead introduce and investigate some intrinsically quantum observables in simple examples of finite-density fermion systems.

One of the key features that distinguishes quantum systems from classical systems is the possibility of entanglement between different degrees of freedom. This entanglement can be quantified: given any subset A of degrees of freedom, the von Neumann entropy $S_A = -\operatorname{tr}(\rho_A \log \rho_A)$ of the density matrix $\rho_A = \operatorname{tr}_{\bar{A}} |\Psi\rangle\langle\Psi|$ provides a measure of the entanglement between A and the rest of the system $\bar{A}$. This is known as the entanglement entropy of the subsystem A (for a review, see [68]). Entanglement entropy and related observables have been studied extensively in quantum field theory and many-body systems over the past several years (see, for example

[69–71, 73]), but these studies typically choose A to be the subset of degrees of freedom inside a particular spatial region. In this work, we study the entanglement entropy for even simpler subsystems: we take A a subset of *momentum space*, focusing on the simplest possible subset consisting of a single field theory mode (i.e. a single allowed momentum). In the context of spin systems, similar momentum-space entanglement observables have been studied for example in [72].

Momentum-space entanglement entropy in quantum field theory systems was investigated in Chapter 2. There, it was emphasized that entanglement entropy in momentum space vanishes in the ground state of non-interacting systems and remains finite in the continuum limit, in contrast to the position-space entanglement entropy which is non-zero for non-interacting systems and diverges in the continuum limit. The work [74] showed further that in the presence of weak interactions, the momentum-space entanglement entropy can often be computed in perturbation theory, as we review in Sec. 2. The goal of this chapter is to carry out such perturbative calculations of the entanglement entropy for the simplest possible interacting finite-density fermion systems, a gas of free non-relativistic or relativistic fermions in one dimension perturbed by a four-fermion (two-body) interaction.

We focus on two quantities in particular: the single-mode entanglement entropy $S(p)$, and the mutual information $I(p,q) = S(p) + S(q) - S(p,q)$ that measures entanglement and correlations between two individual modes at distinct momenta. In the non-relativistic case, for either lattice fermions (Sec. 2) or continuum fermions (Sec. 3), we show that the leading perturbative expression for $S(p)$ diverges logarithmically in $|p_F - p|$ as p approaches the Fermi momentum from above or below. The logarithmic divergence also indicates a breakdown of perturbation theory when the momentum is within e^{-c/λ^2} of the Fermi momentum, where c is some λ-independent constant. For continuum fermions, we also calculate the leading perturbative contribution to the mutual information between any two modes, both in the non-relativistic case and (in Sec. 5) for Dirac fermions with a $(\bar{\psi}\psi)^2$ interaction. This quantity shows discontinuities when either of the momenta cross

the Fermi surface and is largest when both momenta are near the Fermi point.

The renormalization group picture of such interacting fermion systems (see [75] for a review) suggests that the low-energy physics of the systems we consider should be described by a scale-invariant Luttinger liquid system. Luttinger liquids correspond to stable renormalization group (RG) fixed-points, so the low-energy physics should be largely insensitive to the details of physics for modes far from the Fermi surface. As a check of this, we show that the behavior of the entanglement entropy for modes near the Fermi-surface is the same in a theory with a cutoff $||k| - k_F| < \Lambda$. As discussed in [74], there may be a direct connection between the behavior of systems under renormalization group flows and the momentum-dependence of entanglement observable; investigating this further is an interesting question for future work.

The results in this chapter represent a preliminary exploration of momentum-space entanglement observables in the simplest interacting fermion systems. The fact that our results show a striking momentum dependence for the single-mode entanglement with a sharp peak at the Fermi point gives new insight into the entanglement structure of the ground state away from the free limit and suggests that perhaps single-mode entanglement can be used more generally as a new signature of the Fermi surface for interacting systems. The results motivate the study of this observable for more strongly interacting systems and for systems in higher dimensions. As a particular application, it may be interesting to study the the entanglement between modes in trapped cold-atom systems. In this case, the experimental system is such that the interaction strengths can be varied in a controlled way, and the occupation of particular modes can be observed experimentally. Thus, it may be possible to experimentally investigate observables similar to those studied in this chapter.[44]

[44] We thank Kirk Madison for discussions on this point.

3.2 Momentum-space entanglement in perturbation theory

In this section, we review the general result found in Chap. 2 for entanglement entropy at leading order in perturbation theory and its application to the calculation of various measures of entanglement for subsystems corresponding to subsets of modes in momentum space. Although this chapter is concerned with a perturbative calculation of mode entanglement we emphasize that the observable can be defined via a path integral for general field theories [74] and in particular does not assume that the theory under consideration includes particle or quasiparticle states.

General result for entanglement entropy in perturbation theory

Consider any quantum system with Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ and Hamiltonian $H = H_A \otimes \mathbb{1} + \mathbb{1} \otimes H_B + \lambda H_{AB}$. Letting $|n\rangle$ and $|N\rangle$ be energy eigenstates of H_A and H_B respectively, a state $|n\rangle \otimes |N\rangle$ (an energy eigenstate of the $\lambda = 0$ Hamiltonian, e.g. the vacuum state) has no entanglement between the subsystems. Turning on the interaction, the perturbed eigenstate may be calculated using ordinary quantum-mechanical perturbation theory. From this, we can compute the density matrix for subsystem A and the associated entanglement entropy S_A. As shown in ([74]), the leading order perturbative expression in the non-degenerate case is [45]

$$S_A = -\lambda^2 \log(\lambda^2) \sum_{n' \neq n, N' \neq N} \frac{|\langle n', N'|H_{AB}|n, N\rangle|^2}{(E_n + \tilde{E}_N - E_{n'} - \tilde{E}_{N'})^2} + \mathcal{O}(\lambda^2)\,. \qquad (3.1)$$

This involves a sum over matrix elements of the interaction Hamiltonian between the original state and states for which both subsystems A and B have been changed. [46] In Appendix B.1, we present a simple application of

[45] Note that this perturbative result is perfectly valid for systems with degeneracy, so long as the state about which we are perturbing is *itself* non-degenerate.

[46] In cases where the coupling λ is dimensionful there is a corresponding dimensionful quantity entering the log which serves to make it dimensionless. The details of this quantity affect the result only at $O(\lambda^2)$ so we do not write it explicitly.

this result, and as a demonstration that this leading perturbative expression is reliable in a case where the exact answer is known.

Entanglement entropy for a region of momentum space in quantum field theory

For quantum field theory, we can start at finite volume so that the the Hilbert space has a discrete Fock-space decomposition $\mathcal{H} = \otimes_p \mathcal{H}_p$ and the unperturbed Hamiltonian is a sum of terms

$$H_0 = \sum_p E_{p,\alpha} a^\dagger_{p,\alpha} a_{p,\alpha} \ ,$$

each acting on a single factor of the tensor product. Here, α labels the species of particle if there is more than one, and $E_{p,\alpha}$ is the energy of a particle of species α with momentum p. This may include the contribution of a chemical potential added to give a ground state with finite density

$$|\Psi_i\rangle = \prod_i a^\dagger_{p_i,\alpha_i} |0\rangle \ . \tag{3.2}$$

We can take A to be some subset of the allowed momenta for the theory, i.e. all modes with a particular set of allowed wavelengths, and consider the entanglement entropy of the modes in region A for the state (3.2). Taking the large volume limit with the region A of momentum space fixed, we find [74] that the formula (3.1) gives an entanglement entropy for modes in the region A that is extensive (i.e. proportional to spatial volume), with

$$S_A/V = -\lambda^2 \log(\lambda^2) \sum_f \int^* \prod \frac{d^d p_a}{(2\pi)^d} (2\pi)^d \delta(p_f - p_i) \frac{|\mathcal{M}_{fi}|^2}{(E_f - E_i)^2} + \mathcal{O}(\lambda^2) \, . \tag{3.3}$$

Here, the matrix element $\mathcal{M}_{fi}$ is defined by

$$\langle \Psi_f | H_I | \Psi_i \rangle = (2\pi)^d \delta(p_f - p_i) \mathcal{M}_{fi}(p_1, \ldots, p_N) \tag{3.4}$$

where $|\Psi_f\rangle$ are occupation number basis elements of the form $a^\dagger_{p_1,\alpha_1} \cdots a^\dagger_{p_n,\alpha_n} |0\rangle$.

The sum and integral are over the possible states $|\Psi_f\rangle$ appearing in the matrix element. Specifically, the sum is over the possible number of particles of each type present in the state $|\Psi_f\rangle$, while the integral is over the momenta p_a of the particles that have been added/removed from the initial state $|\Psi_i\rangle$ to produce $|\Psi_f\rangle$, with the constraint that we have added/removed at least one particle with momentum in the region A and at least one particle with momentum in the complementary region $\bar{A}$.

Single-mode entanglement entropy in quantum field theory

In the case where A corresponds to a single mode, the entanglement entropy is finite and volume-independent in the large volume limit (since we are no longer keeping the momentum-space volume of the the region A fixed in the limit). We find

$$S(p) = -\lambda^2 \log(\lambda^2) \sum_f \int^* \prod \frac{d^d p_a}{(2\pi)^d} (2\pi)^d \delta(p_f - p_i) \frac{|\mathcal{M}_{fi}|^2}{(E_f - E_i)^2} + \mathcal{O}(\lambda^2)\,. \tag{3.5}$$

where now in the sum/integrals over the basis state $|\Psi_f\rangle$ the constraint is that the occupation number of the mode with momentum p is different than for $|\Psi_i\rangle$, and the occupation number of at least one mode with some other momentum has changed.

For the special case of spinless fermions, the Hilbert space associated with a single mode is only two-dimensional, and this allows us to give an explicit result for the order λ^2 terms in the entanglement entropy in terms of the order $\lambda^2 \log(\lambda^2)$ piece. From equation (18) in [74], we see that for

$$S(p) = -\lambda^2 \log(\lambda^2) a + \mathcal{O}(\lambda^2),$$

we must have

$$S(p) = -\lambda^2 \log(\lambda^2) a + \lambda^2 a(1 - \log(a)) + \mathcal{O}(\lambda^3),$$

Below, we will write explicitly only the leading $\mathcal{O}(\lambda^2 \log(\lambda^2))$ terms.

Mutual information between modes in quantum field theory

Finally, we will consider the mutual information between two modes with momentum p and q. Letting A_p and A_q represent subsets of momentum space corresponding to infinitesimal volumes $d^d p$ and $d^d q$ about momenta p and q, we find that the mutual information $I(A_p, A_q) \equiv S(A_p) + S(A_q) - S(A_p \cup A_q)$ is proportional to volume in the large volume limit, and also proportional to $d^d p$ and $d^d q$. If we define $\mathcal{I}(p, q)$ by

$$I(A_p, A_q)/V = \frac{d^d p}{(2\pi)^d} \frac{d^d q}{(2\pi)^d} \mathcal{I}(p, q) \,, \tag{3.6}$$

then we find

$$\mathcal{I}(p, q) = -\lambda^2 \log(\lambda^2) \sum_f \int^* \prod \frac{d^d p_a}{(2\pi)^d} (2\pi)^d \delta(p_f - p_i) \frac{|\mathcal{M}_{fi}|^2}{(E_f - E_i)^2} + \mathcal{O}(\lambda^2) \,, \tag{3.7}$$

where now the state $|\Psi_f\rangle$ in the matrix element is required to differ from the state $|\Psi_i\rangle$ in the occupation numbers of modes p and q and at least one other mode.

3.3 Entanglement entropy for lattice fermions with nearest neighbor interactions

We now study the entanglement between modes in several models of fermions at finite density, starting with a systems of spinless fermions on a lattice in one spatial dimension with nearest neighbor interactions.

We choose a Hamiltonian[47]

$$H = -\frac{1}{2} \sum_j (\psi^\dagger_{j+1} \psi_j + \psi^\dagger_j \psi_{j+1}) + \lambda \sum_j (\psi^\dagger_j \psi_j - \frac{1}{2})(\psi^\dagger_{j+1} \psi_{j+1} - \frac{1}{2})$$

[47] Here, we have rescaled H to be dimensionless. To restore the physical dimensions, we can multiply by $1/(a^2 m)$, where a is the lattice spacing and m is the effective particle mass, so that excitations with momentum $p \ll 1/a$ about the unfilled state have energy $p^2/(2m)$.

for which the ground state has half-filling and an exact particle-hole symmetry [75]. We define the momentum-space modes ψ_p by

$$\psi_p = \sqrt{a}\sum_j \psi_j e^{ijpa} \qquad\qquad \psi_j = \sqrt{a}\int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} \frac{dp}{2\pi}\, \psi_p e^{-ijpa}$$

such that

$$\{\psi_p, \psi_q^\dagger\} = (2\pi)\, \delta_{\frac{2\pi}{a}}(p-q)\ ,$$

a delta function with period $2\pi/a$, where a is the lattice spacing.

The Hamiltonian becomes

$$H = \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} \frac{dp}{2\pi}\, \psi_p^\dagger \psi_p(-\cos(pa)) + \frac{\lambda}{4} - \lambda \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} \frac{dp}{2\pi}\, \psi_p^\dagger \psi_p + \lambda H_I \tag{3.8}$$

where

$$H_I = -\frac{a}{8\pi^3}\int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} dP dQ dp dq \delta(P+Q-p-q)\psi_P^\dagger \psi_Q^\dagger \psi_p \psi_q e^{i(Q-q)a}\ . \tag{3.9}$$

The third term in (3.8) represents an adjustment to the chemical potential such that the state remains at half-filling in the presence of the interaction.

For $\lambda = 0$, the mode energy $-\cos(pa)$ is negative for $|p| < \pi/(2a)$, so the ground state is

$$|\Psi_i\rangle = \prod_{|p|<\frac{\pi}{2a}} \psi_p^\dagger |0\rangle\ .$$

We can now calculate the entanglement entropy of a single mode with momentum k/a using (3.5). In this case, all states $|\Psi_f\rangle$ for which the matrix element (3.4) is nonzero have the same number of particles as $|\Psi_i\rangle$, with two particles removed inside the Fermi surface (at momenta p and q), and two particles added outside the Fermi surface (at momenta P and Q). Denoting such a state by $|\Psi_i; P, Q, p, q\rangle$, we have

$$\langle \Psi_i; P, Q, p, q | H_I | \Psi_i \rangle = -a(2\pi)\delta(P+Q-p-q)(e^{i(Q-q)a} - e^{i(Q-p)a} - e^{i(P-q)a} + e^{i(P-p)a})$$

so that the squared matrix element appearing in (3.5) is

$$|\mathcal{M}(P,Q,p,q)|^2 = a^2|(e^{iQa}-e^{iPa})(e^{iqa}-e^{ipa})|^2 \ .$$

When the momentum k/a is inside the Fermi surface, we can take $p = k/a$ and then integrate over all q inside the Fermi surface and (P,Q) outside the Fermi surface. Performing the Q integral using the delta function sets $Q = k/a + q - P$ in the integrand and the remaining integral is over all possible q inside the Fermi surface and P outside the Fermi surface such that $Q = k/a + q - P$ is also outside the Fermi surface. The result is that the leading perturbative contribution to the entanglement entropy is

$$S_{in}(k) = -\frac{1}{4\pi^2}\lambda^2\log(\lambda^2)\int_{R_1} dqdP \frac{4(\cos(P-k)-\cos(P-q))^2}{(\cos(q)+\cos(k)-\cos(P)-\cos(k+q-P))^2},$$

where the region of integration is

$$R_1 = \{-k \le q \le \frac{\pi}{2}; -\pi+\frac{k+q}{2} \le P \le -\frac{\pi}{2}\} \cup \{-\frac{\pi}{2} \le q \le -k; \frac{\pi}{2} \le P \le \frac{k+q}{2}+\pi\} \ ,$$

and we have absorbed a factor of a into the integration variables.

Similarly, when the momentum k/a is outside the Fermi surface, we can take $P = k/a$, and integrate over Q outside the Fermi surface and (p,q) inside the Fermi surface. The result is

$$S_{out}(k) = -\frac{1}{4\pi^2}\lambda^2\log(\lambda^2)\int_{R_2} dQdp \frac{4(\cos(p-k)-\cos(p-Q))^2}{(\cos(Q)+\cos(k)-\cos(p)-\cos(k+Q-p))^2},$$

where

$$\begin{aligned} R_2 \ = \ & \{\frac{\pi}{2} \le Q \le \pi; -\frac{\pi}{2} \le p \le \frac{Q+k}{2} - \pi\} \cup \{-\pi \le Q \le -k; -\frac{\pi}{2} \le p \le \frac{k+Q}{2}\} \\ & \cup \{-k \le Q \le -\frac{\pi}{2}; \frac{k+Q}{2} \le p \le \frac{\pi}{2}\} \ . \end{aligned}$$

It is straightforward to show that $S_{out}(k) = S_{in}(\pi/2-k)$; thus, the entanglement entropy is exactly symmetric about the Fermi surface, a consequence of particle-hole symmetry.

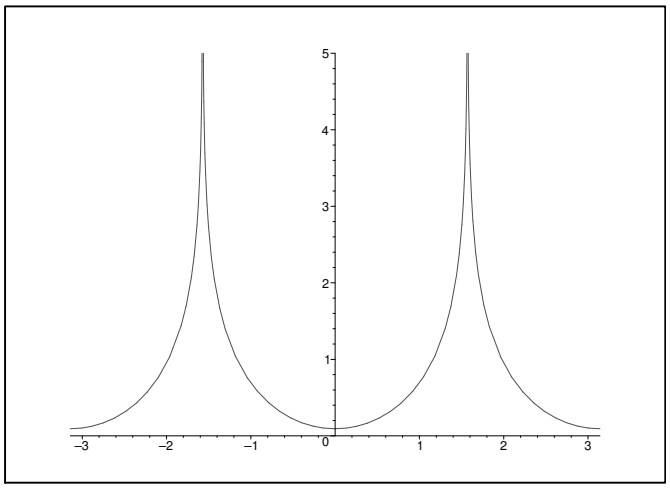


Figure 3.1: Leading perturbative contribution to single mode entanglement entropy $S(k)$ as a function of mode momentum for lattice fermions at half filling. This diverges logarithmically at the Fermi points $k = \pm\pi/2$.

The function $S(k)$ is plotted in Fig. 3.1. We see that this leading perturbative expression diverges at the Fermi momentum; the divergence is logarithmic in $|k - k_F|$ and remains if we include a cutoff $|k - k_F| < \epsilon$ restricting to momenta near the Fermi surface. In this case, we find

$$S(k) = -\frac{1}{4\pi^2}\lambda^2 \log(\lambda^2) \log\left(\frac{\epsilon}{|k - k_F|}\right) + C(\epsilon) + \mathcal{O}(k - k_F),$$

where C is a momentum-independent constant of order ϵ^2.

Since the Hilbert space for a single mode is two-dimensional, the exact mode entanglement entropy is bounded by $\log(2)$. Thus, the divergence in our leading perturbative expression indicates a breakdown in perturbation theory when the momentum is taken too close to the Fermi surface. Specifically, we expect that the perturbative result is reliable only if it is much less than one. This requires that $|k - k_F| \gg e^{-1/\lambda^2}$.

In the exponentially small region $|k - k_F| \lesssim e^{-1/\lambda^2}$ non-perturbative effects are important. One expects on general grounds that an exact calculation would smooth out the singularity in this region, giving a finite entanglement entropy. Outside this exponentially tiny region, our results

are reliable and demonstrate that modes become sharply entangled with the rest of the system as the mode momentum approaches the Fermi surface.

3.4 Entanglement entropy for continuum non-relativistic fermions

We now consider the continuum limit of the previous model, obtained by taking the lattice spacing to zero and adjusting the chemical potential so that states up to some fixed momentum (independent of a) remain occupied. Restoring the overall factor of $1/(a^2 m)$ in the Hamiltonian, we can rewrite the interaction (3.11) as

$$H_I = -\frac{1}{32\pi^3 am}\int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} dPdQdpdq\delta(P+Q-p-q)\psi_P^\dagger\psi_Q^\dagger\psi_p\psi_q(e^{iQa}-e^{iPa})(e^{iqa}-e^{ipa})\ . \tag{3.10}$$

In the limit $a \to 0$, this gives

$$H_I = \frac{a}{32\pi^3 m}\int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} dPdQdpdq\delta(P+Q-p-q)\psi_P^\dagger\psi_Q^\dagger\psi_p\psi_q(Q-P)(q-p)\ . \tag{3.11}$$

Rescaling $\lambda \to \lambda/(am)$ so that the Hamiltonian is independent of a in the limit, we finally obtain

$$H = H_0 + \lambda H_I$$

where

$$H_0 = \int_{-\infty}^{\infty}\frac{dp}{2\pi}\ \psi_p^\dagger\psi_p(\frac{p^2}{2m}-\mu) \tag{3.12}$$

and

$$H_I = \frac{1}{32\pi^3 m^2}\int_{-\infty}^{\infty} dPdQdpdq\delta(P+Q-p-q)\psi_P^\dagger\psi_Q^\dagger\psi_p\psi_q(Q-P)(q-p)\ . \tag{3.13}$$

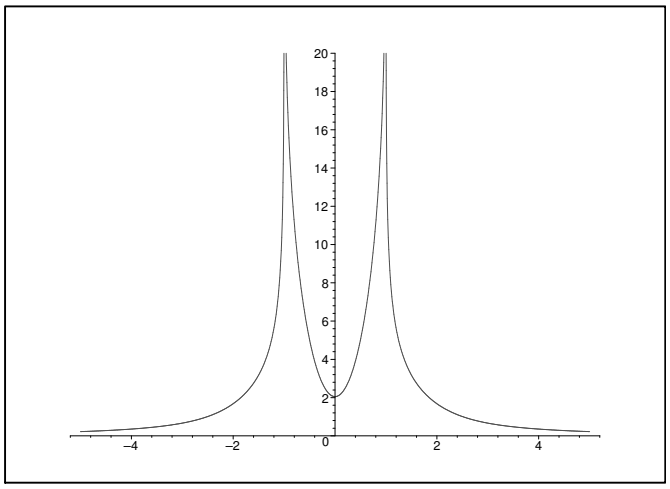


Figure 3.2: Leading perturbative contribution to single mode entanglement entropy $S(p_F k)$ as a function of mode momentum (as a fraction of the Fermi momentum) for weakly interacting continuum non-relativistic fermions. This diverges logarithmically at the Fermi points.

We now follow the same steps as in the previous section to obtain results for the single mode entanglement entropy. For $0 < k < 1$, we find

$$\frac{S(p_F k)}{\lambda^2 \log(\lambda^2)} = -\frac{p_F^2}{4\pi^2 m^2} \left\{ \int_{-k}^{1} dq \int_{-\infty}^{-1} dP + \int_{-1}^{-k} dq \int_{1}^{\infty} dP \right\} \frac{(k-q)^2(2P-k-q)^2}{(P-k)^2(P-q)^2} \, .$$

For $1 < k < 3$, we find

$$\frac{S(p_F k)}{\lambda^2 \log(\lambda^2)} = -\frac{p_F^2}{4\pi^2 m^2} \left\{ \int_{-k}^{-1} dQ \int_{\frac{k+Q}{2}}^{1} dp + \int_{-2-k}^{-k} dQ \int_{-1}^{\frac{k+Q}{2}} dp \right\} \frac{(k-Q)^2(2p-k-Q)^2}{(p-k)^2(p-Q)^2} \, .$$

For $k > 3$, we have

$$\frac{S(p_F k)}{\lambda^2 \log(\lambda^2)} = -\frac{p_F^2}{4\pi^2 m^2} \left\{ \int_{-k}^{2-k} dQ \int_{\frac{k+Q}{2}}^{1} dp + \int_{-2-k}^{-k} dQ \int_{-1}^{\frac{k+Q}{2}} dp \right\} \frac{(k-Q)^2(2p-k-Q)^2}{(p-k)^2(p-Q)^2} \, .$$

In each case, we have rescaled the integration variables by a factor of p_F (the Fermi momentum) to make them dimensionless. All of these integrals

may be evaluated analytically to obtain

$$S(p_F k) = -\frac{p_F^2}{4\pi^2 m^2}\lambda^2 \log(\lambda^2) f(k)$$

where

$$f(k) = \begin{cases} -(k+1)^2\log(1-k) - (1-k)^2\log(1+k) + k^2\left(\frac{14}{3} + 2\log(2)\right) + \frac{2}{3} + 2\log(2) \\ \qquad\qquad 0 < k < 1 \\ -(k+1)^2\log(k-1) + (k+1)^2\left(\frac{7}{3} + \log(2)\right) - \frac{34}{3}(k+1) + \frac{28}{3} - \frac{8}{3(k+1)} \\ \qquad\qquad 1 < k < 3, \\ \frac{16}{3(k^2-1)} \qquad\qquad 3 < k \end{cases} . \tag{3.14}$$

The results for $k < 0$ are obtained using $S(-k) = S(k)$.

The single-mode entanglement entropy $S(p_F k)$ is plotted in figure 3.2 . It diverges logarithmically at $k = \pm 1$ with the behavior on either side described by

$$f(k) = -4\log|1-k| + \frac{16}{3}\mathrm{sgn}(1-k) + 4\log(2) + \mathcal{O}(1-k).$$

At $k = 3$, the function $f(k)$ and its first, second, and third derivatives are all continuous, with a discontinuity appearing only in the fourth derivative.

3.4.1 Mutual information between modes

We can also look at the entanglement structure in more detail by calculating the mutual information between individual modes with momenta $p_F k$ and $p_F l$. Specifically, we calculate the function $\mathcal{I}(p_F k, p_F l)$ defined in (3.6).

The mutual information $\mathcal{I}(p_F k, p_F l)$ satisfies

$$\mathcal{I}(p_F k, p_F l) = \mathcal{I}(p_F l, p_F k) = \mathcal{I}(-p_F k, -p_F l) \ ,$$

so we can restrict to the region $\{k > 0, |l| < k\}$ and find $\mathcal{I}$ for the other values using the symmetries. For each choice of k and l, the integral in (3.7) is over distinct pairs of momenta such that together with the momenta $p_F k$

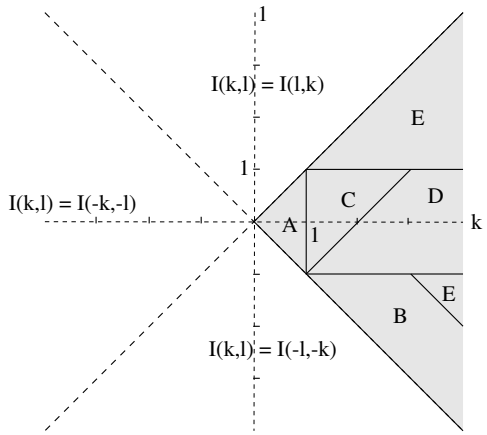


Figure 3.3: Regions of (k, l) space with different behaviors for $\mathcal{I}(k, l)$. Results for the unshaded regions may be obtained from the results for the shaded regions using the indicated symmetries of $\mathcal{I}$.

and $p_F l$, we have two momenta inside the Fermi surface and two momenta outside the Fermi surface (otherwise the matrix element in (3.7) vanishes). Performing one integral using the delta function, we are left with a single integral in each case. For the various regions depicted in figure 3.3, we find the following results:

- Region A: $\{0 < k < 1, |l| < k\}$

$$\begin{aligned}\mathcal{I}(p_F k, p_F l) &= -\lambda^2 \log(\lambda^2) \frac{p_F}{m^2} \int_{-\infty}^{-1} \frac{dP}{2\pi} \frac{(k-l)^2(2P-k-l)^2}{(P-k)^2(P-l)^2} \\ &= -\lambda^2 \log(\lambda^2) \frac{p_F}{2\pi m^2} \left\{ 2(k-l) \log\left(\frac{1+k}{1+l}\right) + \frac{(k-l)^2(k+l+2)}{(k+1)(l+1)} \right\}\end{aligned}$$

- Region B: $\{l < -1, -l < k < 2 - l\}$

$$\begin{aligned}\mathcal{I}(p_F k, p_F l) &= -\lambda^2 \log(\lambda^2) \frac{p_F}{m^2} \int_{\frac{k+l}{2}}^{1} \frac{dp}{2\pi} \frac{(k-l)^2(2p-k-l)^2}{(p-k)^2(p-l)^2} \\ &= -\lambda^2 \log(\lambda^2) \frac{p_F}{2\pi m^2} \left\{ 2(l-k) \log\left(\frac{1-l}{k-1}\right) + \frac{(k-l)^2(k+l-2)}{(k-1)(l-1)} \right\}\end{aligned}$$

- Region C: $\{-1 < l < 1, 1 < k < 2 + l\}$

$$\begin{aligned}\mathcal{I}(p_F k, p_F l) &= -\lambda^2 \log(\lambda^2) \frac{p_F}{m^2} \int_{-1}^{k-1-l} \frac{dp}{2\pi} \frac{(p-l)^2(2k-p-l)^2}{(k-p)^2(k-l)^2} \\ &= -\lambda^2 \log(\lambda^2) \frac{p_F}{2\pi m^2} \left\{ \frac{(k-l)^3}{(k+1)(l+1)} + \frac{1}{3} \frac{(k+1)^3 - (l+1)^3}{(k-l)^2} - 2(k-l) \right\}\end{aligned}$$

- Region D: $\{-1 < l < 1, k > 2 + l\}$

$$\begin{aligned}\mathcal{I}(p_F k, p_F l) &= -\lambda^2 \log(\lambda^2) \frac{p_F}{m^2} \int_{-1}^{1} \frac{dp}{2\pi} \frac{(p-l)^2(2k-p-l)^2}{(k-p)^2(k-l)^2} \\ &= -\lambda^2 \log(\lambda^2) \frac{p_F}{2\pi m^2} \left\{ \frac{1}{3} \frac{(k+1)^3 - (k-1)^3}{(k-l)^2} + \frac{2(k-l)^2}{k^2-1} - 4 \right\}\end{aligned}$$

Region E: $\{k > 1, l > 1\} \cup \{k > 3, 2 - k < l < -1\}$

$$\mathcal{I}(p_F k, p_F l) = 0 \, .$$

The function $\mathcal{I}(p_F k, p_F l)$ is plotted in figure 3.4, with plots for specific values of k given in figure 3.5. We see that this leading-order contribution to $\mathcal{I}$ is generally discontinuous as one momentum crosses $\pm p_F$ and diverges when both momenta approach one of the Fermi points, unless the two momenta are equal.

3.5 Entanglement entropy for relativistic fermions

In this section, we consider Dirac fermions with a four-fermion interaction, described by the action

$$S = \int d^2x \left\{ i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi - \tilde{\lambda}\bar{\psi}\psi\bar{\psi}\psi \right\} \, .$$

This reduces exactly to the non-relativistic model in the previous section (with $\lambda = \tilde{\lambda}/2$) if we take chemical potential $\mu = m + \tilde{\mu}$ with $\tilde{\mu} \ll m$ and consider observables related to energy scales small compared to m so that

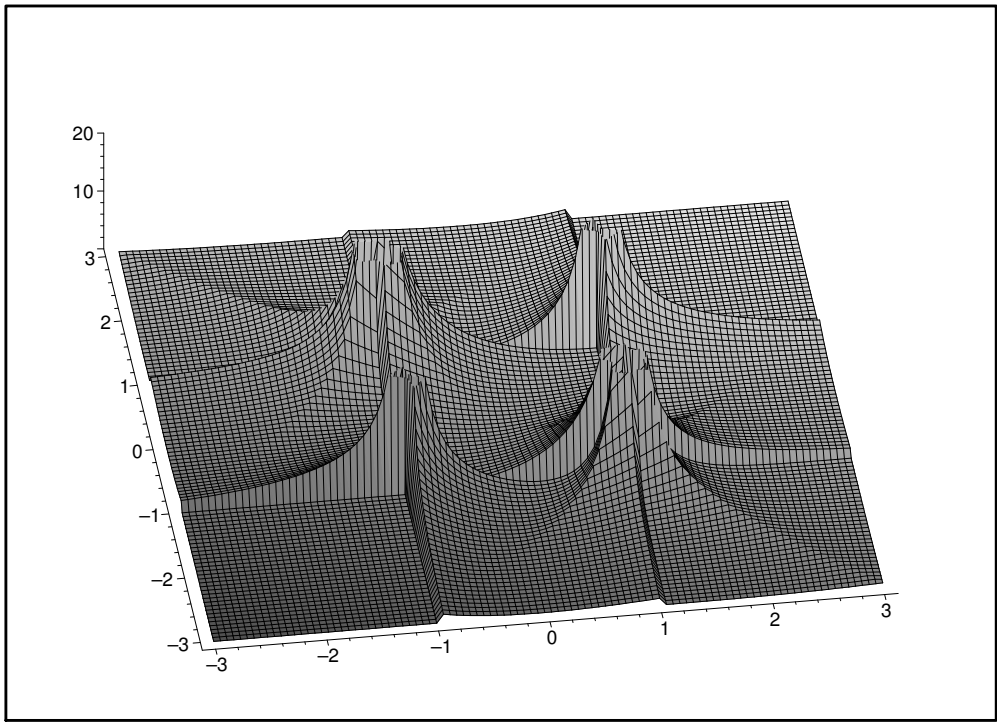


Figure 3.4: Mutual information $\mathcal{I}(p_F k, p_F l)$ between individual modes for non-relativistic fermions with Fermi momentum p_F.

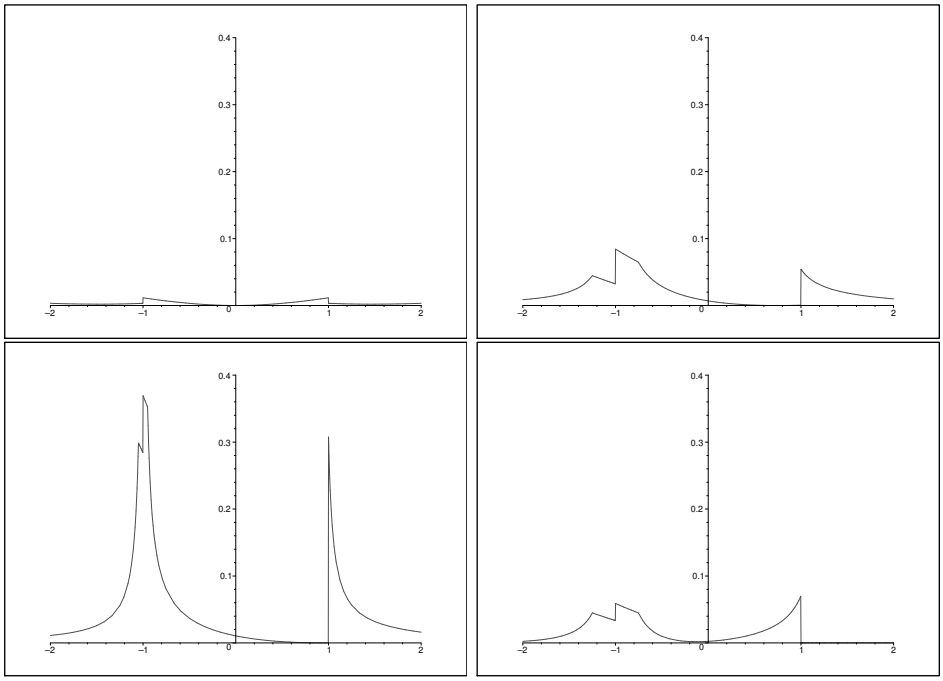


Figure 3.5: Mutual information $\mathcal{I}(p_F k, p_F l)$ vs l for non-relativistic fermions with $k = 0$ (top left), $k = 0.75$ (top right), $k = 0.95$ (bottom left), and $k = 1.25$ (bottom right). The overall scale for $\mathcal{I}$ is arbitrary

the particle number is fixed and antiparticles decouple.

Here, we consider general values of the chemical potential μ and the corresponding ground state $|\mu\rangle$ for which all particle states with energy less than μ are occupied (and none of the antiparticle states are occupied).

The field may be expanded as usual in terms of creation and annihilation operators as

$$\psi(x) = \int \frac{dk}{2\pi} \frac{1}{\sqrt{2\omega_k}} \left(a_k \, u_k \, e^{-ik\cdot x} + b_k^\dagger \, v_k \, e^{ik\cdot x} \right) . \tag{3.15}$$

where a_k and b_k respectively annihilate a particle and antiparticle with momentum k.

For the calculation of any of the entanglement observables in Sec. 3.2, the non-vanishing matrix elements are those between the ground state and states obtained by

1. adding two particles and two antiparticles, corresponding to the operator combinations $a^\dagger b^\dagger a^\dagger b^\dagger$ in the expansion of $\int dx \, (\bar{\psi}\psi)^2$;

2. adding two particles and removing two particles, corresponding to $a^\dagger a a^\dagger a$; or

3. adding two particles, removing one particle and adding an antiparticle, corresponding to $a^\dagger a a^\dagger b^\dagger$ and $a^\dagger b^\dagger a^\dagger a$.

Particles can only be added outside the Fermi surface and can only be removed inside the surface; antiparticles can be added anywhere, but cannot be removed without annihilating the ground state.

Using (A.1), we can write the relevant terms in the interaction Hamiltonian in terms of the momentum-space creation and annihilation operators as

$$\begin{aligned}
H_1 &= \int \left\{ \prod \frac{dk_i}{2\pi\sqrt{2\omega_i}} \right\} \bar{u}(k_1) v(k_2) \bar{u}(k_3) v(k_4) a_{k_1}^\dagger b_{k_2}^\dagger a_{k_3}^\dagger b_{k_4}^\dagger (2\pi)\delta(k_1 + k_2 + k_3 + k_4) \\
H_2 &= 2 \int \left\{ \prod \frac{dk_i}{2\pi\sqrt{2\omega_i}} \right\} \bar{u}(k_1) v(k_2) \bar{u}(k_3) u(k_4) a_{k_1}^\dagger b_{k_2}^\dagger a_{k_3}^\dagger a_{k_4} (2\pi)\delta(k_1 + k_2 + k_3 - k_4) \\
H_3 &= \int \left\{ \prod \frac{dk_i}{2\pi\sqrt{2\omega_i}} \right\} \bar{u}(k_1) u(k_2) \bar{u}(k_3) u(k_4) a_{k_1}^\dagger a_{k_3}^\dagger a_{k_2} a_{k_4} (2\pi)\delta(k_1 + k_3 - k_2 - k_4)
\end{aligned}$$

Using these, we can calculate the matrix elements appearing in the calculation of entanglement observables. As an example, consider the matrix element of H_1 between the ground state $|\mu\rangle$ and a state $|\mu, p_1p_2; p_3p_4\rangle$ where two particles with momenta p_1 and p_2 and two antiparticles with momenta p_3 and p_4 have been added to the ground state. We find

$$\langle \mu, p_1p_3; p_2p_4|H_1|\mu\rangle = \frac{(2\pi)\delta(\sum p_i)}{4\sqrt{\omega_1\omega_2\omega_3\omega_4}} 2(\bar{u}(p_1)v(p_2)\bar{u}(p_3)v(p_4) - \bar{u}(p_1)v(p_4)\bar{u}(p_3)v(p_2)) \tag{3.16}$$

so that from (3.4), we get

$$\begin{aligned} \frac{|\mathcal{M}_{p_1p_3;p_2p_4}|^2}{(\Delta E)^2} &= \frac{1}{4\omega_1\omega_2\omega_3\omega_4} \frac{|\bar{u}(p_1)v(p_2)\bar{u}(p_3)v(p_4) - \bar{u}(p_1)v(p_4)\bar{u}(p_3)v(p_2)|^2}{(\omega_1+\omega_2+\omega_3+\omega_4)^2} \\ &= \frac{(p_1\cdot p_3 - m^2)(p_2\cdot p_4 - m^2)}{\omega_1\omega_2\omega_3\omega_4(\omega_1+\omega_2+\omega_3+\omega_4)^2} \\ &\equiv J_1(p_1,p_2,p_3,p_4) \end{aligned}$$

Similarly, we can calculate the matrix element of H_2 between the ground state and the state $|\mu, p_1p_3\bar{p}_2; p_4\rangle$ where two particles and an antiparticle have been added with momenta p_1, p_3 and p_4 respectively, and a particle with momentum p_2 has been removed. We find

$$\begin{aligned} \frac{|\mathcal{M}_{p_1p_3\bar{p}_2;p_4}|^2}{(\Delta E)^2} &= \frac{1}{4\omega_1\omega_2\omega_3\omega_4} \frac{|\bar{u}(p_1)v(p_4)\bar{u}(p_3)u(p_2) - \bar{u}(p_3)v(p_4)\bar{u}(p_1)u(p_2)|^2}{(\omega_1+\omega_3-\omega_2+\omega_4)^2} \\ &= \frac{(p_1\cdot p_3 - m^2)(p_2\cdot p_4 + m^2)}{\omega_1\omega_2\omega_3\omega_4(\omega_1+\omega_3-\omega_2+\omega_4)^2} \\ &\equiv J_2(p_1,p_2,p_3,p_4)\ . \end{aligned}$$

Finally, we have

$$\begin{aligned} \frac{|\mathcal{M}_{p_1p_3\bar{p}_2\bar{p}_4}|^2}{(\Delta E)^2} &= \frac{1}{4\omega_1\omega_2\omega_3\omega_4} \frac{|\bar{u}(p_1)u(p_4)\bar{u}(p_3)u(p_2) - \bar{u}(p_3)u(p_4)\bar{u}(p_1)u(p_2)|^2}{(\omega_1+\omega_3-\omega_2-\omega_4)^2} \\ &= \frac{(p_1\cdot p_3 - m^2)(p_2\cdot p_4 - m^2)}{\omega_1\omega_2\omega_3\omega_4(\omega_1+\omega_3-\omega_2-\omega_4)^2} \\ &\equiv J_3(p_1,p_2,p_3,p_4) \end{aligned}$$

for the matrix element of H_3 between the ground state and the state $|\mu, p_1p_3\bar{p}_2\bar{p}_4\rangle$

where particles have been added with momenta p_1 and p_3 and particles with momenta p_2 and p_4 have been removed.

Single-mode entanglement

To calculate the entanglement entropy for a single mode with momentum p, we use the expression (3.5), taking the sum over the three types of final states discussed above. However, we find that the integrals in the terms involving H_1 and H_2 diverge. Thus, as noted in Chapter 2 for the case without chemical potential, the leading perturbative expression for the single-mode entanglement entropy in this model is ill-defined. Since the exact answer is necessarily less than $\log(2)$, this must indicate a breakdown of perturbation theory, as discussed in more detail in section 2.4.4. Thus, for this model, we focus on the mutual information between modes, which can be computed in perturbation theory.

Mutual information between modes

To calculate the mutual information between modes, we use (3.7). In each case, $\mathcal{I}(p,q)$ is calculated using matrix elements for which the occupation number of the particle modes with momenta p and q have been changed relative to the ground state and for which occupation numbers for two other particles or antiparticles (with momenta P and Q) have been changed.

When p and q are both inside the Fermi surface, we have:

$$\mathcal{I}(p,q) \;=\; -\lambda^2 \log(\lambda^2)\frac{1}{2\pi}\left[\frac{1}{2}\int_> dP \int_> dQ\, \delta(P+Q-p-q) J_3(P,p,Q,q)\right],$$

where $\int_>$ and $\int_<$ indicates that the integration variable ranges outside and inside the Fermi surface, respectively.

When p is inside, and q is outside:

$$\begin{aligned}\mathcal{I}(p,q) \;=\; -\lambda^2 \log(\lambda^2)\frac{1}{2\pi}\Bigg[&\int_> dP \int dQ\, \delta(P+Q-p+q) J_2(q,p,P,Q) \\ &+ \int_> dP \int_< dQ\, \delta(P-Q-p+q) J_3(q,p,P,Q)\Bigg].\end{aligned}$$

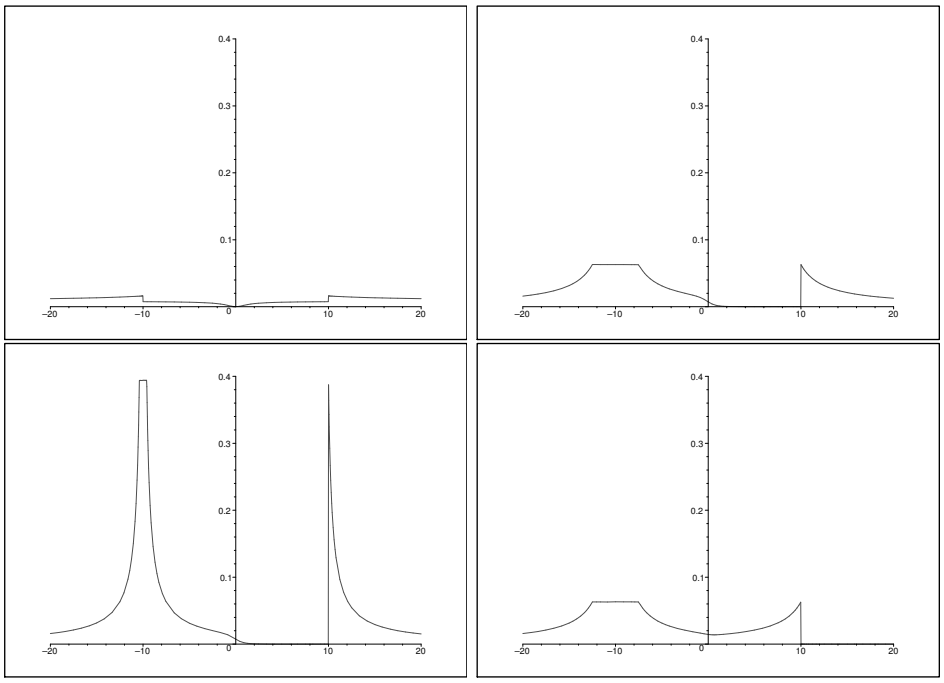


Figure 3.6: Mutual information $\mathcal{I}(p,q)$ vs q for relativistic fermions with $\mu = 10m$ and $p = 0$ (top left), $p = 7.5$ (top right), $p = 9.5$ (bottom left), and $p = 12.5$ (bottom right). Momenta are given in units of m, so the Fermi points are at ± 10. The overall scale for $\mathcal{I}$ is arbitrary.

This also covers the case when p is outside and q is inside, since $I(p,q) = I(q,p)$.

Finally, when both p and q are outside:

$$\begin{aligned}\mathcal{I}(p,q) &= -\lambda^2 \log(\lambda^2)\frac{1}{2\pi}\left[\frac{1}{2}\int dP \int dQ\, \delta(P+Q+p+q) J_1(p,P,q,Q)\right. \\ &\qquad + \int_< dP \int dQ\, \delta(-P+Q+p+q) J_2(p,P,q,Q) \\ &\qquad \left. + \frac{1}{2}\int_< dP \int_< dQ\, \delta(-P-Q+p+q) J_3(p,P,q,Q)\right].\end{aligned}$$

All these integrals are straightforward to perform numerically. We find behaviour qualitatively similar to the non-relativistic case, with discontinuities at the Fermi momenta and the largest mutual information when both

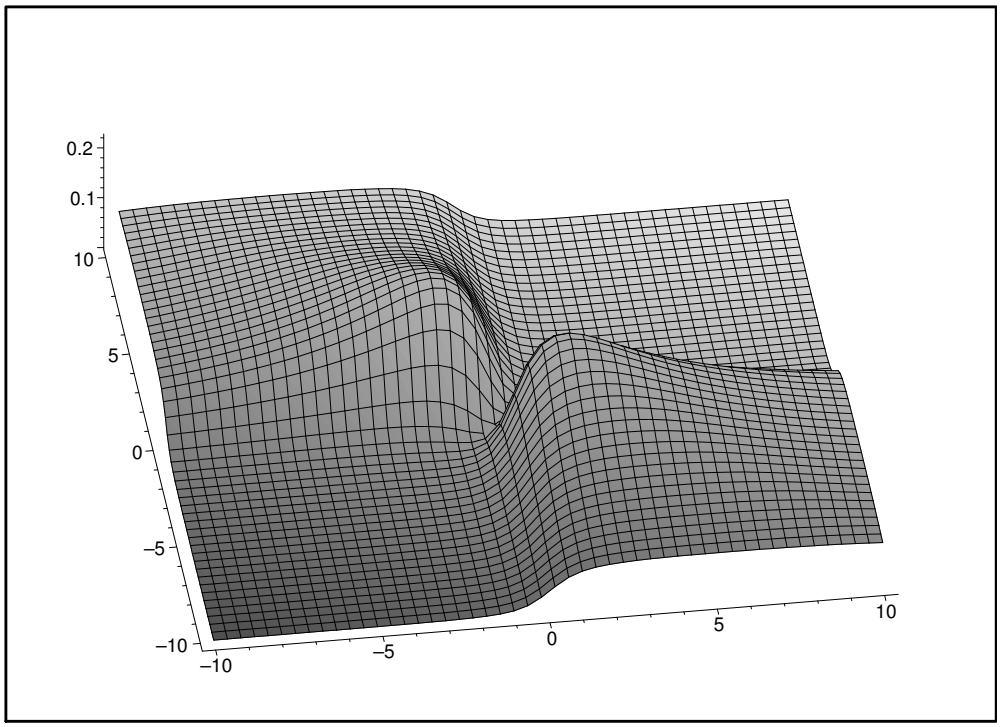


Figure 3.7: Mutual information $\mathcal{I}(p,q)$ as a function of p and q for relativistic fermions with $\mu = 0$. Momenta are given in units of m.

modes are close to the Fermi point. As an example, in figure 3.6 the mutual information $\mathcal{I}(p,q)$ at $\mu = 10m$ is plotted as a function of q for several fixed values of p. These may be compared with the non-relativistic results in figure 3.5.

For comparison, we plot in Fig. 3.7 the mutual information $\mathcal{I}(p,q)$ vs q for several values of p in the case with zero chemical potential, where the unperturbed ground state is the Fock-space vacuum. Here, we find smooth behavior with the largest mutual information between pairs of momenta of opposite sign and momenta of order the mass scale. The mutual information falls off as $1/q$ for fixed p.

3.6 Discussion

In this chapter we have investigated the entanglement between modes in both finite density and relativistic weakly interacting fermi systems. In particular the entanglement entropy $S(p)$ is an observable which measures the degree of entanglement between a degree of freedom of a given wavelength (corresponding to p) and the rest of the system. It should be stressed that while a correlation function in momentum space also measures classical correlations, $S(p)$ is a measure of purely quantum mechanical correlations.

At finite density we found a sharp amplification in quantum correlations as we approach the Fermi surface: degrees of freedom near the fermi surface are the most strongly entangled with the rest of the system. This is in line with the intuition that fermi surface physics dominates the low energy behaviour.

This interesting structure for the quantum correlations in the perturbative regime motivates the study of these observables in more strongly coupled theories, where the Luttinger model is sure to be important.

Chapter 4

Gravitational Dynamics From Entanglement "Thermodynamics"

4.1 Introduction

Since the first connections between gravity and thermodynamics were realized in the study of black hole physics [76–78], various attempts have been made to derive Einstein's equations from the thermodynamics of some underlying degrees of freedom, starting with Jacobson's intriguing paper [79] (see also [80, 81]). With the AdS/CFT correspondence [82, 83], the underlying degrees of freedom for certain theories of gravity with AdS asymptotics have been explicitly identified as the degrees of freedom of a conformal field theory. It is thus interesting to ask whether the Einstein's equations in the gravitational theory can be derived from some thermodynamic relations for the CFT degrees of freedom.

In this note, following [84–88] we demonstrate that at least to linear order in perturbations around pure AdS, Einstein's equations do follow from a relation $dE = dS$ closely related to the First Law of Thermodynamics, but where the entropy S is the *entanglement entropy* of a spatial region in the field theory, and E is a certain energy associated with this region. A key point is that dS and dE can be defined and the relation $dS = dE$ shown to hold for *arbitrary* perturbations around the vacuum state; thus, the relation is more general than the ordinary first law which applies only in situations of thermodynamic equilibrium.

The specific relation we employ, which we write as

$$\delta S_A = \delta E_A^{hyp} \tag{4.1}$$

was derived recently by Blanco, Casini, Hung, and Myers in [88]. Here A represents a ball-shaped spatial region, δS_A represents the change in entanglement entropy of the region A relative to the vacuum state, and δE_A^{hyp} represents the "hyperbolic" energy of the perturbed state in the region A, the expectation value of an operator which maps to the Hamiltonian of the CFT on hyperbolic space times time under a conformal transformation that takes the domain of dependence of the region A to $H^d \times$ time. We review the derivation of this relation in the next section.

For holographic conformal field theories, each side of (4.1) has an interpretation in the dual gravity theory. Assuming that the perturbed state $|\Psi\rangle$ corresponds to some weakly-curved classical spacetime, the entanglement entropy S_A may be calculated (at the leading order in the $1/N$ to which we work) via the Ryu-Takayanagi proposal [84] and its covariant generalization [85] as the area of an extremal surface in the bulk, as we review in Sec. 4.3.1. In Sec. 4.3.2, we recall that the energy δE_A can be calculated from the asymptotic behavior of the metric. Thus, the field theory relation $\delta S_A = \delta E_A^{hyp}$ translates to a constraint on the dual geometry.

In section 4, we show that this constraint is precisely that the bulk metric corresponding to $|\Psi\rangle$ must satisfy Einstein's equations to linear order in the perturbation around pure AdS (the geometry corresponding to the CFT vacuum state). That solutions of Einstein's equations satisfy $\delta S_A = \delta E_A^{hyp}$ has already been shown in [88] (see also the related earlier work [87, 89, 90]). For completeness, we provide an alternate demonstration of this in section 4.2. In section 4.3, we go the other direction, showing that any perturbation to pure AdS satisfying $\delta S_A = \delta E_A^{hyp}$ must satisfy Einstein's equations. This requires more than simply reversing the arguments of section 4.2 (or of [88]). In particular, demanding that $\delta S_A = \delta E_A^{hyp}$ for all ball-shaped spatial regions A in a particular Lorentz frame only places mild constraints on the metric, determining the combination $H_{xx} + H_{yy}$ in terms of the other

components. It is only when we demand that $\delta S_A = \delta E_A^{hyp}$ in an arbitrary Lorentz frame (i.e. for ball-shaped regions on arbitrary spatial slices) that the full set of linearized Einstein's equations is implied.

In Appendix C.1, we give an alternative proof that Einstein's equations imply $\delta S_A = \delta E_A^{hyp}$ that is perhaps more straightforward, but assumes that the metric is analytic.

We conclude in section 5 with a discussion.

4.2 Entropy-energy relation

In this section, we review the relation $\delta S_A = \delta E_A^{hyp}$, derived by Blanco, Casini, Hung, and Myers in [88] as a special case of an inequality that follows from the positivity of relative entropy.

General expression for variation of the entanglement entropy

Consider a CFT on $R^{d,1}$ in some state $|\Psi\rangle$. Choosing a spatial region A, define ρ_A to be the reduced density matrix associated with this region for the state $|\Psi\rangle$,

$$\rho_A = \mathrm{tr}_{\bar{A}} |\Psi\rangle\langle\Psi| \, .$$

From this, we can define the modular Hamiltonian H_A by

$$\rho_A = e^{-H_A} \, .$$

For general states, this modular Hamiltonian is not related to the usual Hamiltonian, and cannot be written as the integral of a local density. We now consider an arbitrary variation of the state $|\Psi\rangle$. The change in entanglement entropy S_A for the region A is given by

$$\begin{aligned} \delta S_A &= \delta(-\mathrm{tr}(\rho_A \log \rho_A)) \\ &= -\mathrm{tr}(\delta\rho_A \log \rho_A) \\ &= \mathrm{tr}(\delta\rho_A H_A) \\ &= \delta\langle H_A \rangle \end{aligned}$$

where we have used the fact that $\mathrm{tr}(\delta\rho_A) = 0$, a consequence of assuming that the density matrix has a fixed normalization. In the last line, H_A is the original modular Hamiltonian associated with the density matrix ρ_A for the original state. Thus, we have the general relation

$$\delta S_A = \delta\langle H_A\rangle \ , \tag{4.2}$$

valid in any spatial region A for arbitrary perturbations of an arbitrary state.

"Thermodynamic" relation for perturbations around the vacuum state

We now specialize to the case where $|\Psi\rangle$ is the vacuum state, and the region A is a ball of radius R. In this case, the domain of dependence of the ball-shaped region[48] can be mapped by a conformal transformation to hyperbolic space times time. As shown in [86], such a transformation maps the vacuum density matrix for the region A to the thermal density matrix $e^{-\beta H_{hyp}}$ for the hyperbolic space theory, where the temperature is related to the hyperbolic space curvature radius R_H by $\beta = 2\pi R_{hyp}$. In this case H_{hyp} is the integral of the local operator T^{00}_{hyp} over hyperbolic space. Mapping back to the ball-shaped region of Minkowski space, it follows [86] that the modular Hamiltonian can be written as

$$H_A^{vac} = 2\pi \int_A d^d x \frac{R^2 - r^2}{2R} T^{00}$$

where T^{00} is the energy density operator for the CFT and r is a radial coordinate centered at the center of the ball.

In this case, we have

$$\delta\langle H_A\rangle = 2\pi \int_A d^d x \frac{R^2 - r^2}{2R} \delta T^{00} \equiv \delta E_A^{hyp} \ , \tag{4.3}$$

i.e. the variation in the expectation value of the vacuum modular Hamilto-

[48] The domain of dependence of A is the set of points p for which all inextensible causal curves passing through p also pass through A.

nian H_A^{vac} under a small perturbation away from the vacuum state is equal to the change in the "hyperbolic" energy of the region. Thus, the general relation (4.2) gives

$$\delta S_A = \delta E_A^{hyp} \,, \tag{4.4}$$

reminiscent of the First Law of Thermodynamics. We emphasize however that the entanglement entropy S_A can be defined for any state, in contrast to the usual thermodynamic entropy which applies to equilibrium states. Thus, (4.4) represents a much more general result.

4.3 Gravitational implications of $dS = dE$ in holographic theories

Let us now consider the case of a holographic conformal field theory on Minkowski space, whose states correspond to asymptotically AdS spacetimes in some quantum theory of gravity. In this case, each side of the relation $\delta S_A = \delta E_A^{hyp}$ has a straightforward gravitational interpretation. As we review below, the left side may be calculated using the Ryu-Takayanagi proposal [84, 85], while the right side can be calculated from the asymptotic form of the metric. The equality of these quantities represents a constraint on the gravitational dynamics implied by the dual field theory. In the next section, we show that this constraint is precisely equivalent to Einstein's equations linearized about AdS.

4.3.1 Gravitational calculation of dS

According to the Ryu-Takayanagi proposal [84] and its covariant generalization [85], the entanglement entropy S_A for a state with a geometrical gravity dual is proportional[49] to the area of the extremal co-dimension two surface $\tilde{A}$ in the bulk whose boundary coincides with the boundary of the region A

[49] Here, we are working to leading order in $1/N$. See the discussion section for comments on $1/N$ corrections.

on the AdS boundary,

$$S_A = \frac{\text{Area}(\tilde{\text{A}})}{4\text{G}_\text{N}} \ .$$

The surface $\tilde{A}$ is an extremum of the area functional

$$A(G, X_{ext}) = \int d^d\sigma\sqrt{g}$$

where

$$g = \det(g_{ab}) = \det(G_{\mu\nu})\frac{dX^\mu}{d\sigma^a}\frac{dX^\nu}{d\sigma^b} \ .$$

Starting from pure AdS, with metric[50]

$$ds^2 = G^0_{\mu\nu}dx^\mu dx^\nu = \frac{1}{z^2}(-dt^2 + dz^2 + d\vec{x}^2) \tag{4.5}$$

the extremal surface ending on the spatial boundary sphere of radius R is described by the spacetime surface

$$\vec{x}^2 + z^2 = R^2 \ . \tag{4.6}$$

We now consider a small variation

$$G_{\mu\nu} = G^0_{\mu\nu} + \delta G_{\mu\nu} \ . \tag{4.7}$$

In this case, the extremal surface changes, and the new area is

$$A(G_0 + \delta G, X^0_{ext} + \delta X)$$

where the variation δX will be of order δG. Since the original surface was extremal, we have

$$A(G_0, X^0_{ext} + \delta X) = A(G_0, X^0_{ext}) + \mathcal{O}(\delta X^2) \ .$$

Thus, the variation of the surface gives rise to changes in the area that start at order δG^2. To find the order δG variation of the area, we need only

[50] Throughout this chapter, we set the AdS radius to one.

evaluate

$$A(G_0 + \delta G, X^0_{ext}) - A(G, X^0_{ext})$$

expanded to linear order in δG. We find that

$$\delta A = \int d^d\sigma \frac{1}{2}\sqrt{g_0} g_0^{ab} \delta g_{ab} \ , \tag{4.8}$$

where we have used lower-case letters to represent pullbacks to the extremal surface. Thus, for field theory state $|\Psi\rangle$ close to the vacuum state with dual geometry described by (4.7), the change in the entanglement entropy for region A relative to the vacuum state is given by an integral of the metric perturbation over the original extremal surface $\tilde{A}$. Using the explicit metric (4.5) and parameterizing the extremal surface (4.6) by the boundary coordinates x^i, we have finally that

$$\delta S = \frac{R}{8G_N} \int d^d x (\delta_{ij} - \frac{1}{R^2} x_i x_j) H_{ij} \ . \tag{4.9}$$

4.3.2 Gravitational calculation of dE

General asymptotically AdS spacetimes with a Minkowski space boundary geometry may by described using Fefferman-Graham coordinates by a metric

$$ds^2 = \frac{1}{z^2}(dz^2 + dx_\mu dx^\mu + z^d H_{\mu\nu}(x,z) dx^\mu dx^\nu) \ . \tag{4.10}$$

where pure AdS, dual to the CFT vacuum, corresponds to $H_{\mu\nu} = 0$. With this parametrization, the expectation value $t_{\mu\nu}$ of the field theory stress-energy tensor is simply related to the asymptotic metric by [91, 92]

$$t_{\mu\nu}(x) = \frac{d+1}{16\pi G_N} H_{\mu\nu}(z = 0, x) \ .$$

Thus, we may write the change in the hyperbolic energy (4.3) relative to the vacuum state as

$$\delta E^{hyp}_A = \frac{d+1}{16 G_N} \int_A d^d x \frac{R^2 - r^2}{R} \delta H_{00}(0, x) \ . \tag{4.11}$$

This is an integral of the boundary value of H over the region A.

4.4 Derivation of linearized Einstein's equations from $dE = dS$

We are now ready to demonstrate that using the holographic dictionary reviewed in the previous section, the CFT relation $\delta S_A = \delta E_A^{hyp}$ is equivalent to the constraint that the metric corresponding to the perturbed CFT satisfies Einstein's equations to linear order. For clarity, we focus on the case of 2+1 dimensional conformal field theories, corresponding to gravitational theories with four non-compact dimensions. However, the result can also be proven for general higher-dimensional theories.

Using the results (4.9) and (4.11), the CFT relation $\delta S_A = \delta E_A^{hyp}$ implies that for a disk of any radius R centered at any point (x_0, y_0) on the boundary, the integral

$$\begin{aligned}\delta\hat{S} \;=\; \int_{D_R} dxdy \Big\{ & H_{xx}(\sqrt{R^2-x^2-y^2}, t, x+x_0, y+y_0)(R^2-x^2) \\ & + H_{yy}(\sqrt{R^2-x^2-y^2}, t, x+x_0, y+y_0)(R^2-y^2) \\ & - 2H_{xy}(\sqrt{R^2-x^2-y^2}, t, x+x_0, y+y_0)xy \Big\} \end{aligned} \tag{4.12}$$

over the bulk extremal surface must equal the integral

$$\delta\hat{E} = \frac{3}{2}\int_{D_R} dxdy(R^2-x^2-y^2)H_{tt}(0, t, x+x_0, y+y_0) \tag{4.13}$$

over the $z = 0$ surface, where we have absorbed a factor of $1/8G_N R$ to define $\delta\hat{S}(R, x_0, y_0)$ and $\delta\hat{E}(R, x_0, y_0)$ (we drop the hats from now on). We will now show that this equality is true for all disks *in all Lorentz frames* if and only if the bulk metric satisfies Einstein's equations to linear order in H. As shown in [88], these are equivalent to the set of equations

$$H_\alpha{}^\alpha = 0 \qquad \partial_\mu H^{\mu\nu} = 0 \qquad \frac{1}{z^4}\partial_z\left\{z^4\partial_z H_{\mu\nu}\right\} + \partial^2 H_{\mu\nu} = 0 \tag{4.14}$$

that arise by plugging the Fefferman-Graham form of the metric (4.10) into

the zz, $z\mu$, and $\mu\nu$ components of Einstein's equations

$$W_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R - 3 g_{\mu\nu} = 0 \ ,$$

respectively and using the fact that H is regular at $z = 0$. In (4.14), the last equation is equivalent to saying that each component of $z^3 H$ must satisfy the Laplace equation on the AdS background.

4.4.1 Proof that $\delta S = \delta E$ for solutions of Einstein's equations

We begin by showing that solutions of the linearized Einstein's equations obey the equality $\delta S = \delta E$. This has already been checked in Sec. 3.1 of [88] by demonstrating the result for a complete basis of solutions to the equations (4.14). In this section, we offer an alternative proof that does not require using an explicit basis of solutions. A third proof that is perhaps more straightforward but assumes a series expansion of H is given in Appendix C.1.

Using the equations (4.14), we have:

$$\begin{aligned}
& \partial_t^2 H_{tt} = \partial_t^2 (H_{xx} + H_{yy}) \\
\Rightarrow \quad & \partial_t (\partial_x H_{xt} + \partial_y H_{yt}) = \partial_t^2 (H_{xx} + H_{yy}) \\
\Rightarrow \quad & \partial_x^2 H_{xx} + \partial_y^2 H_{yy} + 2 \partial_x \partial_y H_{xy} = \partial_t^2 (H_{xx} + H_{yy}) \\
\Rightarrow \quad & \partial_x^2 H_{xx} + \partial_y^2 H_{yy} + 2 \partial_x \partial_y H_{xy} = (\partial_x^2 + \partial_y^2)(H_{xx} + H_{yy}) + \tfrac{1}{z^4} \partial_z (z^4 \partial_z (H_{xx} + H_{yy})) \\
\Rightarrow \quad & 2 \partial_x \partial_y H_{xy} = \partial_y^2 H_{xx} + \partial_x^2 H_{yy} + \tfrac{1}{z^4} \partial_z (z^4 \partial_z (H_{xx} + H_{yy}))
\end{aligned} \tag{4.15}$$

We would like to use the last equation to eliminate H_{xy} from (4.12). However, we have H_{xy} rather than $\partial_x \partial_y H_{xy}$ in (4.12). To make progress, we begin by differentiating δS by x_0 and y_0 (the coordinates of the center of the boundary disk). This gives

$$\begin{aligned}
\partial_{x_0} \partial_{y_0} \delta S \quad = \quad & \int_{D_R} dx dy \Big\{ \partial_x \partial_y H_{xx} (\sqrt{R^2 - x^2 - y^2}, t, x + x_0, y + y_0)(R^2 - x^2) \\
& + \partial_x \partial_y H_{yy} (\sqrt{R^2 - x^2 - y^2}, t, x + x_0, y + y_0)(R^2 - y^2)
\end{aligned}$$

$$-2\partial_x\partial_y H_{xy}(\sqrt{R^2-x^2-y^2},t,x+x_0,y+y_0)xy\Big\} \tag{4.16}$$

Now, using (4.15), we have

$$\begin{aligned}\partial_{x_0}\partial_{y_0}\delta S &= \int_{D_R} dxdy\,\big\{\partial_x\partial_y H_{xx}(R^2-x^2)+\partial_x\partial_y H_{yy}(R^2-y^2)\\ &\qquad -xy\left(\partial_y^2 H_{xx}+\partial_x^2 H_{yy}+\frac{1}{z^4}\partial_z(z^4\partial_z(H_{xx}+H_{yy}))\right)\Big\}\end{aligned} \tag{4.17}$$

It is straightforward to check that this expression is equal to the integral over the extremal surface of an exact form dA, where A is defined for all (x,y,z,t) as

$$\begin{aligned}A &= \left(-xz\partial_z H_{xx}-3xH_{xx}+z^2\partial_x H_{yy}\right)dx\\ &\quad +\left(z^2\partial_y H_{xx}-yz\partial_z H_{yy}-3yH_{yy}\right)dy\\ &\quad +\left(-yz\partial_y H_{xx}-xz\partial_x H_{yy}\right)dz\ .\end{aligned} \tag{4.18}$$

By Stokes theorem, this equals the integral of A over the boundary of the extremal surface, so we have

$$\begin{aligned}\partial_{x_0}\partial_{y_0}\delta S &= \int_{\partial D_R} A\\ &= -3\int_{\partial D_R}(xH_{xx}dx+yH_{yy}dy)\end{aligned}$$

where we have used the fact that all other terms in A vanish for $z=0$.

Similarly, we find that $\partial_{x_0}\partial_{y_0}\delta E$ may be written as

$$\begin{aligned}\partial_{x_0}\partial_{y_0}\delta E &= \frac{3}{2}\int_{D_R} dxdy\partial_{x_0}\partial_{y_0}H_{tt}(0,t,x+x_0,y+y_0)(R^2-x^2-y^2)\\ &= \frac{3}{2}\int_{D_R} dxdy\partial_x\partial_y(H_{xx}(0,t,x+x_0,y+y_0)+H_{yy}(0,t,x+x_0,y+y_0))(R^2-x^2-y^2)\\ &= \frac{3}{2}\int_{D_R} d\hat{A}\ ,\end{aligned}$$

where we can choose

$$\hat{A}=\left(-2xH_{xx}+(R^2-x^2-y^2)\partial_x H_{yy}\right)dx+\left(-2yH_{yy}+(R^2-x^2-y^2)\partial_y H_{xx}\right)dy\ .$$

Again, using Stokes theorem, this reduces to the integral of $(3/2)\hat{A}$ over the boundary, so

$$\begin{aligned}\partial_{x_0}\partial_{y_0}\delta E &= \frac{3}{2}\int_{\partial D_R}\hat{A} \\ &= -3\int_{\partial D_R}(xH_{xx}dx + yH_{yy}dy) \\ &= \partial_{x_0}\partial_{y_0}\delta S\end{aligned}$$

We conclude that for any H satisfying Einstein's equations,

$$\delta S(x_0, y_0, R; H) - \delta E(x_0, y_0, R; H) = C_x(x_0, R; H) + C_y(y_0, R; H) \ ,$$

where C_x and C_y are some functionals linear in H that do not depend on y_0 or x_0 respectively. Now, consider the class of functions H that vanish for sufficiently large $x_0^2 + y_0^2$ at the time $t = 0$ where we evaluate δS and δE. In this case, fixing any x_0 and taking $y_0 \to \infty$ or fixing any y_0 and taking $x_0 \to \infty$, the left side must vanish. For this to be true on the right side, both C_x and C_y must be constant (as functions of x_0 and y_0), with $C_x + C_y = 0$. Thus, the right side vanishes for any H that vanishes as $x_0^2 + y_0^2 \to \infty$. But more general H can be written as linear combinations of such functions, and since the right side is a linear functional in H, it must vanish for all H. This completes the argument that $\delta S_A = \delta E_A^{hyp}$ for solutions of Einstein's equations.

4.4.2 Proof that $\delta S = \delta E$ implies the linearized Einstein's equations

In this section, we go the other direction to show that the relation $\delta S = \delta E$ implies that the metric satisfies Einstein's equations to linear order, i.e. that the equivalence of (4.12) and (4.13) implies the relations (4.14).

Given the boundary stress tensor $t_{\mu\nu}$, let $H_{\mu\nu}^{EE}$ be the corresponding metric perturbation that follows from Einstein's equations, i.e. the solution of (4.14) satisfying $H_{\mu\nu}^{EE}(0, t, x, y) = (16\pi G_N/3)t_{\mu\nu}$. We will show that there is no other H with these boundary conditions for which $\delta S = \delta E$ in

all frames of reference.

Suppose there were another H for which $\delta S = \delta E$ for all disk shaped regions in all Lorentz frames. Then the difference $\Delta = H - H^{EE}$ must satisfy

$$\Delta_{\mu\nu}(z = 0, t, x, y) = 0 \ , \tag{4.19}$$

and

$$\begin{aligned} 0 \ = \ \int_{D_R} dxdy \Big\{ & \Delta_{xx}(\sqrt{R^2 - x^2 - y^2}, x + x_0, y + y_0)(1 - \frac{x^2}{R^2}) \\ & + \Delta_{yy}(\sqrt{R^2 - x^2 - y^2}, x + x_0, y + y_0)(1 - \frac{y^2}{R^2}) \\ & - 2\Delta_{xy}(\sqrt{R^2 - x^2 - y^2}, x + x_0, y + y_0)\frac{xy}{R^2} \Big\} \end{aligned} \tag{4.20}$$

for arbitrary R, x_0, and y_0, and in an arbitrary Lorentz frame.

Let us first see the consequences of demanding this result in a fixed frame. To begin, we note that (4.20) may be expanded in powers of R using the basic integral

$$\int_{D_R} dxdy(R^2 - x^2 - y^2)^{\frac{n}{2}} x^{2m_x} y^{2m_y} = R^{n+2m_x+2m_y+2} I_{n,m_x,m_y} \ ,$$

where

$$I_{n,m_x,m_y} = \frac{\Gamma(m_x + \frac{1}{2})\Gamma(m_y + \frac{1}{2})\Gamma(\frac{n}{2} + 1)}{\Gamma(\frac{n}{2} + m_x + m_y + 2)} \ . \tag{4.21}$$

Defining

$$\Delta_{\mu\nu}(z, x, y) = \sum_{n=0}^{\infty} z^n \Delta^{(n)}_{\mu\nu}(x, y) \tag{4.22}$$

we find that (4.20) becomes[51]

$$\begin{aligned} 0 = \sum R^{n+2m_x+2m_y+2}\Big\{ & \tfrac{1}{(2m_x)!(2m_y)!}\partial_x^{2m_x}\partial_y^{2m_y}\Delta_{xx}^{(n)}(t,x_0,y_0)(I_{n,m_x,m_y} - I_{n,m_x+1,m_y}) \\ + & \tfrac{1}{(2m_x)!(2m_y)!}\partial_x^{2m_x}\partial_y^{2m_y}\Delta_{yy}^{(n)}(t,x_0,y_0)(I_{n,m_x,m_y} - I_{n,m_x,m_y+1}) \\ -2 & R^2\tfrac{1}{(2m_x+1)!(2m_y+1)!}\partial_x^{2m_x+1}\partial_y^{2m_y+1}\Delta_{xy}^{(n)}(t,x_0,y_0)I_{n,m_x+1,m_y+1}\Big\} \end{aligned} \tag{4.23}$$

The vanishing of the terms at order R^{N+2} implies that

$$\begin{aligned} \Delta_{xx}^{(N)}(t,x_0,y_0) + \Delta_{yy}^{(N)}(t,x_0,y_0) = & \sum_{(m_x,m_y)\neq(0,0)} C_{xx}^{N,m_x,m_y}\partial_x^{2m_x}\partial_y^{2m_y}\Delta_{xx}^{(N-2m_x-2m_y)} \\ & + C_{yy}^{N,m_x,m_y}\partial_x^{2m_x}\partial_y^{2m_y}\Delta_{yy}^{(N-2m_x-2m_y)} \\ & + C_{xy}^{N,m_x,m_y}\partial_x^{2m_x-1}\partial_y^{2m_y-1}\Delta_{xy}^{(N-2m_x-2m_y)}, \end{aligned}$$

where the C coefficients can be read off from (4.23). As examples, the first few equations give

$$\begin{aligned} \Delta_{xx}^{(0)}(t,x_0,y_0) + \Delta_{yy}^{(0)}(t,x_0,y_0) &= 0 \\ \Delta_{xx}^{(1)}(t,x_0,y_0) + \Delta_{yy}^{(1)}(t,x_0,y_0) &= 0 \\ \Delta_{xx}^{(2)}(t,x_0,y_0) + \Delta_{yy}^{(2)}(t,x_0,y_0) &= -\frac{1}{4}(\partial_y^2\Delta_{xx}^{(0)}(t,x_0,y_0) + \partial_x^2\Delta_{yy}^{(0)}(t,x_0,y_0)) \\ &\quad -\frac{3}{20}(\partial_x^2\Delta_{xx}^{(0)}(t,x_0,y_0) + \partial_y^2\Delta_{yy}^{(0)}(t,x_0,y_0)) \\ &\quad +\frac{1}{5}\partial_x\partial_y\Delta_{xy}^{(0)}(t,x_0,y_0) \\ \Delta_{xx}^{(3)}(t,x_0,y_0) + \Delta_{yy}^{(3)}(t,x_0,y_0) &= -\frac{1}{6}(\partial_y^2\Delta_{xx}^{(1)}(t,x_0,y_0) + \partial_x^2\Delta_{yy}^{(1)}(t,x_0,y_0)) \\ &\quad -\frac{1}{6}(\partial_x^2\Delta_{xx}^{(1)}(t,x_0,y_0) + \partial_y^2\Delta_{yy}^{(1)}(t,x_0,y_0)) \\ &\quad +\frac{1}{9}\partial_x\partial_y\Delta_{xy}^{(1)}(t,x_0,y_0) \end{aligned} \tag{4.24}$$

We see that this set of equations completely determines the combination $\Delta_{xx} + \Delta_{yy}$ at each order in z in terms of the lower order terms in the expansion of Δ. However, apart from the constraint (4.19) on the boundary behavior (equivalent to $\Delta_{\mu\nu}^{(0)} = 0$), the remaining elements of $\Delta_{\mu\nu}$ are com-

[51] Here, we are assuming that the function Δ is analytic. It would be useful to find a derivation of our result that holds more generally.

pletely unconstrained.

To constrain $\Delta_{\mu\nu}$ further, we need to use the requirement that the relation (4.20) should hold in an arbitrary Lorentz frame. Thus, for each choice of reference frame, we will have equations analogous to (4.24). Specifically, consider a general boost

$$\Lambda = \begin{pmatrix} \gamma & \gamma\beta_x & \gamma\beta_y \\ \gamma\beta_x & 1+\beta_x^2\frac{\gamma^2}{\gamma+1} & \beta_x\beta_y\frac{\gamma^2}{\gamma+1} \\ \gamma\beta_y & \beta_x\beta_y\frac{\gamma^2}{\gamma+1} & 1+\beta_y^2\frac{\gamma^2}{\gamma+1} \end{pmatrix}$$

In the equations for a general frame of reference obtained by such a boost, the left sides in (4.24) will be replaced by

$$\Lambda_x{}^\mu\Lambda_x{}^\nu\Delta_{\mu\nu} + \Lambda_y{}^\mu\Lambda_y{}^\nu\Delta_{\mu\nu} \ .$$

Up to an overall constant factor, this gives

$$\Delta_{ii} + 2\beta_i\Delta_{it} + \beta^2(\Delta_{tt} - \frac{1}{2}\Delta_{ii}) + (\beta_i\beta_j - \frac{1}{2}\delta_{ij}\beta^2)\Delta_{ij} \ .$$

Now, consider the general version of the second equation in (4.24) (the first equation already holds by (4.19)). This requires the vanishing of

$$\Delta^{(1)}_{ii} + 2\beta_i\Delta^{(1)}_{it} + \beta^2(\Delta^{(1)}_{tt} - \frac{1}{2}\Delta^{(1)}_{ii}) + (\beta_i\beta_j - \frac{1}{2}\delta_{ij}\beta^2)\Delta^{(1)}_{ij} \ .$$

For a fixed x_0 and y_0, this is a polynomial in β_i that must vanish for all values of β_i. Thus, the polynomial must be identically zero. At order β^0, this gives

$$\Delta^{(1)}_{ii}(t, x_0, y_0) = 0$$

as we had before. At order β, we get

$$\Delta^{(1)}_{it}(t, x_0, y_0) = 0 \ .$$

At order β^2, this gives

$$\Delta^{(1)}_{tt}(t, x_0, y_0) = \frac{1}{2}\Delta^{(1)}_{ii}(t, x_0, y_0) = 0$$

and

$$\Delta^{(1)}_{ij}(t, x_0, y_0) - \frac{1}{2}\Delta_{ij}\Delta^{(1)}_{kk}(t, x_0, y_0) = 0 \ .$$

Thus, we have $\Delta^{(1)}_{\mu\nu} = 0$. We can now continue to analyze the remaining equations in (4.24) in turn. Supposing that we have shown $\Delta^{(k)}_{\mu\nu} = 0$ for $k < n$, the general version of the nth equation in (4.24) requires the vanishing of

$$\Delta^{(n)}_{ii} + 2\beta_i\Delta^{(n)}_{it} + \beta^2(\Delta^{(n)}_{tt} - \frac{1}{2}\Delta^{(n)}_{ii}) + (\beta_i\beta_j - \frac{1}{2}\delta_{ij}\beta^2)\Delta^{(n)}_{ij} \ ,$$

since the right hand side in (4.24) will be zero. Repeating the analysis above, we conclude that $\Delta^{(n)}_{\mu\nu} = 0$. By induction, this holds for all n, so we have shown that $\Delta_{\mu\nu} = 0$, completing the proof.

4.5 Discussion

In this chapter, we have seen that to linear order in perturbations about the vacuum state, the emergence of gravitational dynamics in the theory dual to a holographic CFT is directly related to a general relation satisfied by CFT entanglement entropies on ball-shaped regions. This relation is closely related to the First Law of Thermodynamics, but is more general since it applies to arbitrary perturbations of the state rather than perturbations for which the system remains in thermal equilibrium.

While the CFT relation (1) is an exact equivalence, we have made use of this relation only at the leading order in $1/N$ where the entanglement entropy maps over to the extremal surface area. This corresponds to working in the classical limit in the bulk. According to [93], $1/N$ corrections to the CFT entanglement entropy correspond to bulk quantum corrections including the entropy of entanglement of bulk quantum fields across the extremal surface. It will be interesting to understand the implications of the CFT relation (1) beyond the classical level in the bulk, but we leave this for

future work.

The derivations in Sec. 4.4 were written specifically for the case of four-dimensional gravity. However, the proof given in [88] that Einstein's equations imply $\delta S = \delta E$, and our method of proof in section 4.2 that $\delta S = \delta E$ implies the linearized Einstein's equations work for general dimensions.[52]

The linearized Einstein's equations we derived are for the metric components in the field theory directions and radial direction of the bulk. Any additional fields in the gravitational theory, including metric components in any compactified directions, are not constrained by the CFT relation we have considered. At linear order, the equations for these fields decouple from the linearized Einstein's equations for the metric in the non-compact directions. Thus, we can say that the universal relation $\delta S = \delta E$ is equivalent to the universal sector of the linearized bulk equations.

Our results do not imply that all holographic theories are dual to gravitational theories whose metric perturbations satisfy Einstein's equations. In this chapter, we assumed that entanglement entropies are related to areas via the usual Ryu-Takayanagi formula, and that the stress-energy tensor in the dual field theory is related to the asymptotic form of the metric. In more general theories, the entanglement entropy may correspond to a more complicated functional of the bulk geometry and the relation between the stress tensor and asymptotic metric may be modified. In these cases, we expect that the bulk equations will be different, for example involving α' corrections with higher-derivative terms. However, it may be possible following the methods in this chapter to derive the linearized version of these more general equations given a particular choice for the holographic entanglement entropy formula and the holographic formula for the stress tensor.

It will be interesting to see whether the first non-linear corrections to Einstein's equations in the bulk are equivalent to some simple property of entanglement entropies.

[52] Specifically, Eq. (20) becomes $0 = \int_{D_R} d^d x (\Delta_{ii} - x_i x_j / R^2 \Delta_{ij})$; expanding this in powers of R using the generalization of Eq. (21) yields at each order in R an equation that relates $\Delta_{ii}(n)$ to quantities calculated from Δ at lower orders in n. The steps in the proof are as before.

Finally, we comment on the relation to the work of Jacobson [79], which partly motivated our investigations. Jacobson realized that Einstein's equations could be derived from the assumption that the energy flux through a part of any bulk Rindler horizon gives rise to a proportional local change in area of this horizon. Interpreting the area as an entropy, such a relation looks like the first law of thermodynamics. However, in Jacobson's work, it was not clear why areas of segments of an arbitrary bulk Rindler horizon (not necessarily associated with any black hole) should correspond to an entropy, so the origin of the thermodynamic relation remained mysterious.

In our case, the "thermodynamic relation" $dS = dE$ is an exact quantum relation (i.e. not really thermodynamics) derived to hold for the underlying fundamental degrees of freedom associated with our gravitational system. Thus, while our final result (in contrast to Jacobson's work) applies so far only at the linearized level, the starting point is well understood. In detail, the bulk interpretation of our $dS = dE$ relation is somewhat different that Jacobson's starting point (the bulk surfaces/horizons we deal with are global rather than local and the energy has a different interpretation), but the two relations were similar enough to motivate the question of whether Einstein's equations could be derived from the first law of [88].

Chapter 5

Outlook

There are many interesting questions which are left open in the work of this thesis. In fact one of the questions asked in the introduction by way of motivation was left unanswered: *does momentum-space entanglement entropy correspond to some simple geometric quantity in the gravity theory?* Because of the UV/IR connection in AdS/CFT which relates a radial coordinate r to a CFT energy scale μ it is tempting to speculate that the the entanglement entropy between modes above and below some scale μ is related to the area of a gravitational radial slice at $r(\mu)$, i.e.

$$S(\mu) \stackrel{?}{=} \frac{Area(r(\mu))}{4G_N}. \tag{5.1}$$

We have been confined mainly to results in perturbation theory, where our expressions are valid for small coupling. To answer this question we would require a better understanding of the strong coupling behaviour, where the gravity side is classical. There is, however, a consistency check: the area of a radial slice should be proportional to the field theory volume and we do indeed find that the momentum-space entanglement entropy scales with the volume in our calculations. [53]

We have shown how to obtain a reduced density matrix for low momentum modes and it would be interesting to investigate how quantities in field theories compare when computed via low momentum effective actions versus the usual low energy Wilsonian effective actions which show up in the renormalization group. Also, as mentioned earlier, there is a possibility that the types of quantities we have been discussing can be measured in the lab

[53] This is related to the fact that the number of modes below a scale μ to entangle is proportional to the volume of the system, and so this result should be true outside of perturbation theory.

providing a possible experimental validation of the results presented.

We have been able to give one specific answer to the question: *how entanglement in a CFT can affect the gravitational structure of its dual?* Changing the amount of energy in a ball in the CFT incurs a change in the amount of entanglement in that region, and in the dual theory this is precisely tantamount to the Einstein equations. This was carried to linear order in the metric, however, and it would be interesting to investigate what non-linear contributions to Einstein's equations correspond to in the CFT, and if in particular these are related to some simple properties of entanglement.

It would be unbearable to end this thesis without pointing out an amusing irony which it entails. Einstein's great achievement was his theory of General Relativity. This classical theory stood in stark contrast to quantum mechanics, which Einstein did not accept as a complete theory of nature, and in particular pointed to entanglement as one of its quintessentially bizarre and counterintuitive features. As we have seen though, holography (AdS/CFT specifically) indicates to us that General Relativity is the way it is, and in particular Einstein's equations are obeyed, precisely due to the quantum mechanical phenomenon of entanglement. Thus an anachronistic student of holography, finding themself in one of Einstein's lectures, may have dared to raise their hand and say: *Einstein, be not afraid of action at a distance, for your very theory depends on its existence!*

Bibliography

[1] O. Aharony, S. S. Gubser, J. Maldacena, H. Ooguri, and Y. Oz. Large N field theories, string theory and gravity. , 323:183–386, January 2000, hep-th/9905111.

[2] Alain Aspect, Jean Dalibard, and Gérard Roger. Experimental test of bell's inequalities using time- varying analyzers. *Phys. Rev. Lett.*, 49:1804–1807, Dec 1982.

[3] V. Balasubramanian, M. B. McDermott, and M. Van Raamsdonk. Momentum-space entanglement and renormalization in quantum field theory. , 86(4):045014, August 2012, 1108.3568.

[4] J.S. Bell. On the einstein podolsky rosen paradox. *Physics (Long Island City, N.Y.)*, 1:195–200, 1964.

[5] C. H. Bennett and G. Brassard. in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India* (IEEE, New York, 1984), pp 175-179.

[6] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, Mar 1993.

[7] N. D. Birrell and P. C. W. Davies. Massless Thirring model in curved space - Thermal states and conformal anomaly. , 18:4408–4421, December 1978.

[8] Luca Bombelli, Rabinder K. Koul, Joohan Lee, and Rafael D. Sorkin.

Quantum source of entropy for black holes. *Phys. Rev. D*, 34:373–383, Jul 1986.

[9] P. Calabrese and J. Cardy. Entanglement entropy and quantum field theory. *Journal of Statistical Mechanics: Theory and Experiment*, 6:2, June 2004, hep-th/0405152.

[10] H. Casini, C. D. Fosco, and M. Huerta. Entanglement and alpha entropies for a massive Dirac field in two dimensions. *Journal of Statistical Mechanics: Theory and Experiment*, 7:7, July 2005, arXiv:cond-mat/0505563.

[11] Sidney Coleman. Quantum sine-gordon equation as the massive thirring model. *Phys. Rev. D*, 11:2088–2097, Apr 1975.

[12] B. Czech, J. L. Karczmarek, F. Nogueira, and M. Van Raamsdonk. The gravity dual of a density matrix. *Classical and Quantum Gravity*, 29(15):155009, August 2012, 1204.1330.

[13] C. M. Dawson and M. A. Nielsen. Frustration, interaction strength, and ground-state entanglement in complex quantum systems. , 69(5):052316, May 2004, quant-ph/0401061.

[14] E. D'Hoker and D. Z. Freedman. Supersymmetric Gauge Theories and the AdS/CFT Correspondence. *ArXiv High Energy Physics - Theory e-prints*, January 2002, hep-th/0201253.

[15] W. Dür, L. Hartmann, M. Hein, M. Lewenstein, and H.-J. Briegel. Entanglement in Spin Chains and Lattices with Long-Range Ising-Type Interactions. *Physical Review Letters*, 94(9):097203, March 2005, quant-ph/0407075.

[16] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, May 1935.

[17] D. V. Fursaev and S. N. Solodukhin. Description of the Riemannian geometry in the presence of conical defects. , 52:2133–2143, August 1995, hep-th/9501127.

[18] T. Giamarchi. *Quantum Physics in One Dimension*. International Series of Monographs on Physics. Clarendon, 2004.

[19] Berry Groisman, Sandu Popescu, and Andreas Winter. Quantum, classical, and total amount of correlations in a quantum state. *Phys. Rev. A*, 72:032317, Sep 2005.

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

[21] M. Headrick and T. Takayanagi. Holographic proof of the strong subadditivity of entanglement entropy. , 76(10):106013, November 2007, 0704.3719.

[22] C. Holzhey, F. Larsen, and F. Wilczek. Geometric and renormalized entropy in conformal field theory. *Nuclear Physics B*, 424:443–467, August 1994, hep-th/9403108.

[23] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Reviews of Modern Physics*, 81:865–942, April 2009, quant-ph/0702225.

[24] T.-C. L. Hsu, M. B. McDermott, and M. Van Raamsdonk. Momentum-space entanglement for interacting fermions at finite density. *Journal of High Energy Physics*, 11:121, November 2013, 1210.0054.

[25] D. Kabat. Black hole entropy and entropy of entanglement. *Nuclear Physics B*, 453:281–299, February 1995, arXiv:hep-th/9503016.

[26] N. Lashkari, M. B. McDermott, and M. Van Raamsdonk. Gravitational Dynamics From Entanglement "Thermodynamics". *ArXiv e-prints*, August 2013, 1308.3716.

[27] A. Lewkowycz and J. Maldacena. Generalized gravitational entropy. *Journal of High Energy Physics*, 8:90, August 2013, 1304.4926.

[28] D. Marolf. Holography without strings? *Classical and Quantum Gravity*, 31(1):015008, January 2014, 1308.1977.

[29] J. McGreevy. Holographic duality with a view toward many-body physics. *ArXiv e-prints*, September 2009, 0909.0518.

[30] J. Polchinski. Introduction to Gauge/Gravity Duality. *ArXiv e-prints*, October 2010, 1010.6134.

[31] A. V. Ramallo. Introduction to the AdS/CFT correspondence. *ArXiv e-prints*, October 2013, 1310.4319.

[32] S. Ryu and T. Takayanagi. Holographic Derivation of Entanglement Entropy from the anti de Sitter Space/Conformal Field Theory Correspondence. *Physical Review Letters*, 96(18):181602, May 2006, hep-th/0603001.

[33] M. Srednicki. Entropy and area. *Physical Review Letters*, 71:666–669, August 1993, hep-th/9303048.

[34] M. Stone. *Bosonization*. Bosonization. World Scientific, 1994.

[35] L. Susskind. The world as a hologram. *Journal of Mathematical Physics*, 36:6377–6396, November 1995, hep-th/9409089.

[36] B. Sutherland. Quantum Many-Body Problem in One Dimension: Ground State. *Journal of Mathematical Physics*, 12:246–250, February 1971.

[37] Bill Sutherland. Exact results for a quantum many-body problem in one dimension. *Phys. Rev. A*, 4:2019–2021, Nov 1971.

[38] B. Swingle. Mutual information and the structure of entanglement in quantum field theory. *ArXiv e-prints*, October 2010, 1010.4038.

[39] G. 't Hooft. Dimensional Reduction in Quantum Gravity. *ArXiv General Relativity and Quantum Cosmology e-prints*, October 1993, gr-qc/9310026.

[40] W. G. Unruh. Notes on black-hole evaporation. *Phys. Rev. D*, 14:870–892, Aug 1976.

[41] M. Van Raamsdonk. Comments on quantum gravity and entanglement. *ArXiv e-prints*, July 2009, 0907.2939.

[42] V. Vedral. High-temperature macroscopic entanglement. *New Journal of Physics*, 6:102, August 2004, quant-ph/0405102.

[43] Frank Verstraete, Markus Popp, and J Ignacio Cirac. Entanglement versus correlations in spin systems. *Physical review letters*, 92(2):27901, 2004.

[44] Reinhard F. Werner. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, Oct 1989.

[45] E. Witten. Anti-de Sitter space and holography. *Advances in Theoretical and Mathematical Physics*, 2:253, 1998, hep-th/9802150.

[46] Nielsen, Michael A. and Chuang, Isaac L., "Quantum computation and quantum information," Cambridge University Press (2000)

[47] K. G. Wilson, "Renormalization group and critical phenomena. 1. Renormalization group and the Kadanoff scaling picture," Phys. Rev. **B4**, 3174-3183 (1971).
K. G. Wilson, "Renormalization group and critical phenomena. 2. Phase space cell analysis of critical behavior," Phys. Rev. **B4**, 3184-3205 (1971).
K. G. Wilson, J. B. Kogut, "The Renormalization group and the epsilon expansion," Phys. Rept. **12**, 75-200 (1974).

[48] S. Minwalla, M. Van Raamsdonk, N. Seiberg, "Noncommutative perturbative dynamics," JHEP **0002**, 020 (2000). [hep-th/9912072].

[49] P. Calabrese, J. L. Cardy, "Entanglement entropy and quantum field theory," J. Stat. Mech. **0406**, P06002 (2004). [hep-th/0405152]

[50] H. Casini, M. Huerta, "A Finite entanglement entropy and the c-theorem," Phys. Lett. **B600**, 142-150 (2004). [hep-th/0405111]

[51] F. Lombardo, F. D. Mazzitelli, "Coarse graining and decoherence in quantum field theory," Phys. Rev. **D53**, 2001-2011 (1996). [hep-th/9508052].

[52] Y. Shi, "Entanglement in Relativistic Quantum Field Theory," Phys. Rev. D **70** (2004) 105001 [arXiv:hep-th/0408062].

[53] D. Mazur, J. S. Heyl, "Characterizing entanglement entropy produced by nonlinear scalar interactions during inflation," Phys. Rev. **D80**, 023523 (2009). [arXiv:0810.0521 [gr-qc]]

[54] A. Altland and B. Simons, "Condensed matter field theory," Cambridge, UK: Univ. Pr. (2006) 624 p

[55] M.M. Wolf, F. Verstraete, M.B. Hastings, J.I. Cirac, "Area laws in quantum systems: mutual information and correlations," Phys. Rev. Lett. 100, 070502 (2008), [arXiv:0704.3906[quant-ph]]

[56] J. M. Maldacena, "The large N limit of superconformal field theories and supergravity," Adv. Theor. Math. Phys. **2**, 231 (1998) [Int. J. Theor. Phys. **38**, 1113 (1999)] [arXiv:hep-th/9711200].

[57] S. Ryu and T. Takayanagi, "Holographic derivation of entanglement entropy from AdS/CFT," Phys. Rev. Lett. **96**, 181602 (2006) [arXiv:hep-th/0603001].

[58] S. Ryu and T. Takayanagi, "Aspects of Holographic Entanglement Entropy," JHEP **0608**, 045 (2006) [arXiv:hep-th/0605073].

[59] T. Nishioka, S. Ryu and T. Takayanagi, "Holographic Entanglement Entropy: An Overview," J. Phys. A **42**, 504008 (2009) [arXiv:0905.0932 [hep-th]].

[60] M. Headrick, "Entanglement Renyi entropies in holographic theories," Phys. Rev. D **82**, 126010 (2010) [arXiv:1006.0047 [hep-th]].

[61] P. Hayden, M. Headrick, A. Maloney, "Holographic Mutual Information is Monogamous," [arXiv:1107.2940 [hep-th]].

[62] A. Strominger, "AdS(2) quantum gravity and string theory," JHEP **9901**, 007 (1999) [arXiv:hep-th/9809027].

[63] V. Balasubramanian, J. de Boer, M. M. Sheikh-Jabbari and J. Simon, "What is a chiral 2d CFT? And what does it have to do with extremal black holes?," JHEP **1002**, 017 (2010) [arXiv:0906.3272 [hep-th]].

[64] S. R.White, Phys. Rev. Lett. 69, 2863 (1992), Phys. Rev. B 48, 10345 (1993).
U. Schollwoeck, Rev. Mod. Phys. 77, 259 (2005), cond-mat/0409292

[65] G. Vidal, "Entanglement renormalization," cond-mat/0512165, Phys. Rev. Lett. 99, 220405 (2007)
Guifre Vidal, "Entanglement Renormalization: an introduction," arXiv:0912.1651

[66] J. Haegeman, T. J. Osborne, H. Verschelde and F. Verstraete, "Entanglement renormalization for quantum fields," arXiv:1102.5524 [hep-th].

[67] B. Swingle, "Entanglement Renormalization and Holography," arXiv:0905.1317 [cond-mat.str-el].

[68] Nielsen, Michael A. and Chuang, Isaac L., "Quantum computation and quantum information," Cambridge University Press (2000)

[69] P. Calabrese, J. L. Cardy, "Entanglement entropy and quantum field theory," J. Stat. Mech. **0406**, P06002 (2004). [hep-th/0405152]

[70] H. Casini, M. Huerta, "A Finite entanglement entropy and the c-theorem," Phys. Lett. **B600**, 142-150 (2004). [hep-th/0405111]

[71] Y. Shi, "Entanglement in Relativistic Quantum Field Theory," Phys. Rev. D **70** (2004) 105001 [arXiv:hep-th/0408062].

[72] Ronny Thomale, D. P. Arovas, B. Andrei Bernevig, "Non-local order in gapless systems: Entanglement Spectrum in Spin Chains," Phys. Rev. Lett. **105**, 116805 (2010)

[73] A. Kitaev and J. Preskill, "Topological entanglement entropy," Phys. Rev. Lett. **96**, 110404 (2006) [hep-th/0510092].

[74] V. Balasubramanian, M. B. McDermott and M. Van Raamsdonk, "Momentum-space entanglement and renormalization in quantum field theory," Phys. Rev. D **86**, 045014 (2012) [arXiv:1108.3568 [hep-th]].

[75] R. Shankar, "Renormalization group approach to interacting fermions," Rev. Mod. Phys. **66**, 129 (1994).

[76] J. D. Bekenstein, "Black holes and entropy," Phys. Rev. D **7**, 2333 (1973).

[77] J. M. Bardeen, B. Carter and S. W. Hawking, "The Four laws of black hole mechanics," Commun. Math. Phys. **31**, 161 (1973).

[78] S. W. Hawking, "Particle Creation by Black Holes," Commun. Math. Phys. **43**, 199 (1975) [Erratum-ibid. **46**, 206 (1976)].

[79] T. Jacobson, "Thermodynamics of space-time: The Einstein equation of state," Phys. Rev. Lett. **75**, 1260 (1995) [gr-qc/9504004].

[80] T. Padmanabhan, "Thermodynamical Aspects of Gravity: New insights," Rept. Prog. Phys. **73**, 046901 (2010) [arXiv:0911.5004 [gr-qc]].

[81] E. P. Verlinde, "On the Origin of Gravity and the Laws of Newton," JHEP **1104**, 029 (2011) [arXiv:1001.0785 [hep-th]].

[82] J. M. Maldacena, "The large N limit of superconformal field theories and supergravity," Adv. Theor. Math. Phys. **2**, 231 (1998) [Int. J. Theor. Phys. **38**, 1113 (1999)] [arXiv:hep-th/9711200].

[83] O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri and Y. Oz, "Large N field theories, string theory and gravity," Phys. Rept. **323**, 183 (2000) [arXiv:hep-th/9905111].

[84] S. Ryu and T. Takayanagi, "Holographic derivation of entanglement entropy from AdS/CFT," Phys. Rev. Lett. **96**, 181602 (2006) [hep-th/0603001].

[85] V. E. Hubeny, M. Rangamani and T. Takayanagi, "A Covariant holographic entanglement entropy proposal," JHEP **0707**, 062 (2007) [arXiv:0705.0016 [hep-th]].

[86] H. Casini, M. Huerta and R. C. Myers, "Towards a derivation of holographic entanglement entropy," JHEP **1105**, 036 (2011) [arXiv:1102.0440 [hep-th]].

[87] M. Nozaki, T. Numasawa, A. Prudenziati and T. Takayanagi, "Dynamics of Entanglement Entropy from Einstein Equation," arXiv:1304.7100 [hep-th].

[88] D. D. Blanco, H. Casini, L. -Y. Hung and R. C. Myers, "Relative Entropy and Holography," arXiv:1305.3182 [hep-th].

[89] D. Allahbakhshi, M. Alishahiha and A. Naseh, "Entanglement Thermodynamics," arXiv:1305.2728 [hep-th].

[90] G. Wong, I. Klich, L. A. Pando Zayas and D. Vaman, "Entanglement Temperature and Entanglement Entropy of Excited States," arXiv:1305.3291 [hep-th].

[91] R. C. Myers, "Stress tensors and Casimir energies in the AdS / CFT correspondence," Phys. Rev. D **60**, 046002 (1999) [hep-th/9903203].

[92] S. de Haro, S. N. Solodukhin and K. Skenderis, "Holographic reconstruction of space-time and renormalization in the AdS / CFT correspondence," Commun. Math. Phys. **217**, 595 (2001) [hep-th/0002230].

[93] T. Faulkner, A. Lewkowycz and J. Maldacena, "Quantum corrections to holographic entanglement entropy," JHEP **1311**, 074 (2013) [arXiv:1307.2892 [hep-th]].

Appendix A

A.1 Momentum-space entanglement and correlators

Starting with the general expression (2.20) for the leading order perturbative contribution to entanglement entropy, we can now specialize to the case of quantum field theory. We find that

$$\begin{aligned}
S(P) &= -\lambda^2 \log(\lambda^2) \sum_{n\neq 0, N\neq 0} \frac{|\langle n, N|H_{AB}|0,0\rangle|^2}{(E_0+\tilde{E}_0-E_n-\tilde{E}_N)^2} + \mathcal{O}(\lambda^2) \\
&= -\lambda^2 \log(\lambda^2) \sum_{n\neq 0, N\neq 0} \int_0^\infty d\tau\tau \langle 0,0|H_I|n,N\rangle e^{(E_{0,0}-E_{n,N})\tau} \langle n,N|H_I|0,0\rangle + \mathcal{O}(\lambda^2) \\
&= -\lambda^2 \log(\lambda^2) \sum_{n\neq 0, N\neq 0} \int_0^\infty d\tau\tau \langle 0,0|e^{H_0\tau} H_I e^{-H_0\tau}|n,N\rangle\langle n,N|H_I|0,0\rangle + \mathcal{O}(\lambda^2) \\
&= -\lambda^2 \log(\lambda^2) \int_0^\infty d\tau\tau \langle 0,0|e^{H_0\tau} H_I e^{-H_0\tau}\Pi_A H_I|0,0\rangle + \mathcal{O}(\lambda^2) \\
&= -\lambda^2 \log(\lambda^2) \int_0^\infty d\tau\tau \langle H_I(-i\tau)\Pi_A H_I(0)\rangle + \mathcal{O}(\lambda^2) \\
&= -\lambda^2 \log(\lambda^2) \int_0^\infty d\tau\tau \int d^3x d^3y \langle \mathcal{H}_I(-i\tau,x)\Pi_A \mathcal{H}_I(0,y)\rangle + \mathcal{O}(\lambda^2) \\
&= -V\lambda^2 \log(\lambda^2) \int_0^\infty d\tau\tau \int d^3x \langle \mathcal{H}_I(-i\tau,x)\Pi_A \mathcal{H}_I(0,0)\rangle + \mathcal{O}(\lambda^2)
\end{aligned}$$

Here, we use the standard definition of time-dependent operators in the "interaction picture":

$$H_I(t) \equiv e^{iH_0 t} H_I e^{-iH_0 t} \ .$$

The operator Π is projects to intermediate states with at least one particle having momentum in the subset P and at least one particle having momentum in the complementary subset of momenta.

The factor of volume in the last line comes from the y integral in the previous line, which is trivial since the correlator in that line can depend only on the combination $x - y$. The entropy per unit volume $S(P)/V$ will have a finite limit, so that $S(P)$ is an extensive quantity.

A.2 Entanglement entropy in a fermionic system

Here we calculate the entanglement entropy in a fermionic theory with a $(\bar{\psi}\psi)^2$ interaction. Consider for definiteness the renormalizable theory in 1+1 dimensions. The fermion fields are expanded as

$$\psi(x) = \sum_p \frac{1}{L^{\frac{1}{2}}} \frac{1}{\sqrt{2\omega_p}} \left(a_p u(p) e^{-ipx} + b_p^\dagger v(p) e^{ipx} \right). \tag{A.1}$$

As a straightforward application of (2.20) the entanglement entropy is

$$S_\mu = -\lambda^2 \log(\lambda^2) \sum_t {\sum_p}^* \frac{\left| \langle \{p,t\}_1, ..., \{p,t\}_4 \left| \left(\bar{\psi}\psi\right)^2 \right| 0 \rangle \right|^2}{\left(\omega_1 + \omega_2 + \omega_3 + \omega_4\right)^2} + O(\lambda^2), \tag{A.2}$$

where t indicates the type of fermion (i.e. particle or antiparticle). The star indicates that the sum over momenta is restricted to the set where at least one momentum is above and at least one momentum is below the scale μ. Substituting the expansion (A.1) into (A.2)

$$S_\mu = -\lambda^2 \log(\lambda^2) \frac{6 \cdot 4}{2^4 L^2} {\sum_p}^* \frac{\delta_{\sum_i p_i}}{\left(\sum_i \omega_{p_i}\right)^2 \prod_i \omega_{p_i}} \left| \bar{u}(p_1) v(p_2) \bar{u}(p_3) v(p_4) - \bar{u}(p_1) v(p_4) \bar{u}(p_3) v(p_2) \right|^2, \tag{A.3}$$

where 6 is the number of ways or choosing 2 particles and 2 antiparticles. Using 1+1 dimensional spinor and gamma matrix identities, and passing to the infinite volume limit we are left with

$$S_\mu / L = -\lambda^2 \log(\lambda^2) \frac{6}{(2\pi)^3} \int^* dp_1 ... dp_4 \delta(\sum_i p_i) \frac{(p_1 \cdot p_3 - m^2)(p_2 \cdot p_4 - m^2)}{\left(\sum_i \omega_{p_i}\right)^2 \prod_i \omega_{p_i}}.$$

In the region where three momenta are taken to be large this integral diverges linearly.

Appendix B

B.1 Two qubit system example

In order to elucidate the essential features of the type of entanglement we consider in this thesis, this section will be concerned with detailing a simple and explicit example. As a simple model for a system with multiple fermionic degrees of freedom, consider a system of two spins with Hamiltonian (units are chosen so that $\hbar = 1$):

$$H_f = \frac{\sigma_z^{(1)}}{2} \otimes \mathbb{1} + \mathbb{1} \otimes \frac{\sigma_z^{(2)}}{2}, \tag{B.1}$$

Our notation here for spin up is $|u>$ and for spin down it is $|d>$. These are defined by $\sigma_z|u>=|u>$ and $\sigma_z|d>=-|d>$.

The ground state has both spins pointing down, $|0>=|d,d>$ with energy -1. This state is separable and there is no entanglement between the spins.

Now imagine adding some interaction which generates entanglement,

$$H_{int} = \sigma_x^{(1)} \otimes \sigma_x^{(2)}, \tag{B.2}$$

so that

$$H = H_f + \lambda H_{int}. \tag{B.3}$$

We first consider $|\lambda| << 1$ so that we may get a perturbative solution for the ground state. One can easily see that to $O(\lambda^3)$ this state is simply given by $|s>=|d,d>-\frac{\lambda}{2}|u,u>$ since we have [54]

[54] In our notation we have $\sigma_x|d>=|u>$ and $\sigma_x|u>=|d>$.

$$H|s>=H_f|s>+\lambda H_{int}|s>, \tag{B.4}$$

$$=-|d,d>-\frac{\lambda}{2}|u,u>+\lambda|u,u>-\lambda^2|d,d>, \tag{B.5}$$

$$=-(1+\frac{\lambda^2}{2})|s>+O(\lambda^3). \tag{B.6}$$

Thus up to $O(\lambda^3)$ the ground state is proportional to $|s>$ with energy $-(1+\lambda^2/2)$. The correctly normalized ground state is

$$|\Omega>=\frac{1}{\sqrt{1+\frac{\lambda^2}{4}}}\left(|d,d>-\frac{\lambda}{2}|u,u>\right)+O(\lambda^3). \tag{B.7}$$

The corresponding density matrix is then

$$\rho=|\Omega><\Omega|, \tag{B.8}$$

$$=\left(1-\frac{\lambda^2}{4}\right)|d,d><d,d|-\frac{\lambda}{2}\left(|u,u><d,d|+|d,d><u,u|\right)+\frac{\lambda^2}{4}|u,u><u,u|+O(\lambda^3). \tag{B.9}$$

Note that to this order $\mathrm{Tr}\ \rho=1$ as it should.

The reduced density matrix for the first spin (or the second) is obtained by tracing out the second (or the first),

$$\rho^{(1)}=\mathrm{Tr}^{(2)}\rho, \tag{B.10}$$

$$=\left(1-\frac{\lambda^2}{4}\right)|d><d|+\frac{\lambda^2}{4}|u><u|+O(\lambda^3). \tag{B.11}$$

Thus the density matrix is diagonal to this order and the entanglement entropy is easily calculated:

$$S = -\ \mathrm{Tr}\ \rho^{(1)} \log \rho^{(1)}, \tag{B.12}$$

$$= -\left(1 - \frac{\lambda^2}{4}\right) \log\left(1 - \frac{\lambda^2}{4}\right) - \frac{\lambda^2}{4} \log \frac{\lambda^2}{4} + O(\lambda^3), \tag{B.13}$$

$$= -\frac{1}{4}\lambda^2 \log \lambda^2 + \frac{1 + 2\log 2}{4}\lambda^2 + O(\lambda^3 \log \lambda^3). \tag{B.14}$$

Now let us calculate the lowest order term in the entanglement entropy of the ground state by using our general formula Eq. (3.1)

$$S = -\lambda^2 \log(\lambda^2) \sum_{n \neq 0, N \neq 0} \frac{|\langle n, N | H_{int} | 0 \rangle|^2}{\left(E_0 + \tilde{E}_0 - E_n - \tilde{E}_N\right)^2} + O(\lambda^2). \tag{B.15}$$

Let us quickly recall what is meant in this formula. Consider a Hilbert space which decomposes into the tensor product of two factors $\mathcal{H} = \mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$. Then Eq. (B.15) applies to any Hamiltonian which is the sum of free and interacting pieces $H = H_f + \lambda H_{int}$, with the free Hamiltonian being given by a sum acting on each factor separately, i.e.

$$H_f = H^{(1)} \otimes \mathbb{1} + \mathbb{1} \otimes H^{(2)}, \tag{B.16}$$

Then the states $|n>$ and $|N>$ are the energy eigenstates of $H^{(1)}$ and $H^{(2)}$ respectively, and the energies E_n and $\tilde{E}_N$ their corresponding eigenvalues.

Now specializing to our qubit system (Eqs. (B.1) and (B.2)) the free ground state was $|d, d>$, and thus this state corresponds to $|0, 0>$ in Eq. (B.15). Then from the free spin hamiltonian Eq. (B.1) we have $E_d = \tilde{E}_d = -1/2$ and $E_u = \tilde{E}_u = +1/2$. Since the sum in Eq. (B.15) exludes $n = d$ and $N = d$ it reduces to a single term:

$$S = -\lambda^2 \log(\lambda^2) \frac{\left| \left\langle u,u \right| \left(\sigma_x^{(1)} \otimes \sigma_x^{(2)} \right) \left| d,d \right\rangle \right|^2}{\left(E_d + \tilde{E}_d - E_u - \tilde{E}_u \right)^2} + O(\lambda^2), \tag{B.17}$$

$$= -\frac{1}{4}\lambda^2 \log(\lambda^2) + O(\lambda^2), \tag{B.18}$$

in agreement with what we found in Eq. (B.14).

Now of course this system is exactly solvable and so we can check these results.

The eigenvalues of $H = H_f + \lambda H_{int}$ are found to be $\sqrt{1+\lambda^2}, \lambda, -\lambda, -\sqrt{1+\lambda^2}$. Thus for any finite λ, $-\sqrt{1+\lambda^2}$ is the lowest energy with corresponding eigenvector $|d,d> -\frac{\lambda}{1+\sqrt{1+\lambda^2}}|u,u>$ as one can easily check. Then the exact ground state of this model is simply,

$$|G> = \frac{\left(1+\sqrt{1+\lambda^2}\right)^{1/2}}{2^{1/2}\left(1+\lambda^2\right)^{1/4}} \left(|d,d> -\frac{\lambda}{1+\sqrt{1+\lambda^2}}|u,u> \right). \tag{B.19}$$

Then one finds for the reduced density matrix

$$\rho^{(1)} = \text{Tr}\ ^{(2)}\rho, \tag{B.20}$$

$$= \frac{1+\sqrt{1+\lambda^2}}{2\left(1+\lambda^2\right)^{1/2}} \left(|d><d| + \frac{\lambda^2}{\left(1+\sqrt{1+\lambda^2}\right)^2}|u><u| \right), \tag{B.21}$$

$$= \left(\frac{a+1}{2a}\right)|d><d| + \left(\frac{a-1}{2a}\right)|u><u|. \tag{B.22}$$

where we have defined the constant $a(\lambda) \equiv \sqrt{1+\lambda^2}$. Then one arrives at the following simple result for the entanglement entropy between the two qubits,

$$S = -\frac{a+1}{2a}\log\frac{a+1}{2a} - \frac{a-1}{2a}\log\frac{a-1}{2a}. \tag{B.23}$$

Thus as we expect, the system is unentangled at $\lambda = 0$ ($a = 1$) and becomes maximally entangled as $\lambda \to \infty$ ($a \to \infty$). Expanding (B.23) to $O(\lambda^3 \log \lambda^3)$ we have

$$S = -\frac{1}{4}\lambda^2 \log \lambda^2 + \frac{1 + \log 4}{4}\lambda^2 + O(\lambda^3 \log \lambda^3). \tag{B.24}$$

This is in agreement with Eq. (B.14), and with Eq. (B.18) obtained from the general formula (B.15).

Appendix C

C.1 Alternative derivation of linearized Einstein's equations from $\delta E = \delta S$

In this appendix, we offer an alternative proof that solutions of Einstein's equations satisfy $\delta S_A = \delta E_A^{hyp}$. This proof replaces $\delta S_A = \delta E_A^{hyp}$ with the infinite set of relations obtained by matching the terms in the power series expansion of this relation in R, the radius of the disk A, as we did in section 4.2.

C.1.1 Expansion of $\delta E = \delta S$ in powers of R

To begin, we expand both (4.12) and (4.13) in powers of R. Defining

$$H_{\mu\nu}(z,x,y) = \sum_{n=0}^{\infty} z^n H_{\mu\nu}^{(n)}(x,y) \tag{C.1}$$

we have

$$\delta E = \frac{3}{2} \sum_{m_x,m_y=0} R^{2+2m_x+2m_y} I_{2,m_x,m_y} \partial_x^{2m_x} \partial_y^{2m_y} H_{tt}^{(0)}(t,x_0,y_0) \tag{C.2}$$

while

$$\begin{aligned} \delta S = \sum R^{n+2m_x+2m_y+2} \Big\{ & \quad \tfrac{1}{(2m_x)!(2m_y)!} \partial_x^{2m_x} \partial_y^{2m_y} H_{xx}^{(n)}(t,x_0,y_0)(I_{n,m_x,m_y} - I_{n,m_x+1,m_y}) \\ & + \quad \tfrac{1}{(2m_x)!(2m_y)!} \partial_x^{2m_x} \partial_y^{2m_y} H_{yy}^{(n)}(t,x_0,y_0)(I_{n,m_x,m_y} - I_{n,m_x,m_y+1}) \\ & -2 \quad R^2 \tfrac{1}{(2m_x+1)!(2m_y+1)!} \partial_x^{2m_x+1} \partial_y^{2m_y+1} H_{xy}^{(n)}(t,x_0,y_0) I_{n,m_x+1,m_y+1} \Big\} \end{aligned} \tag{C.3}$$

where I was defined in (4.21).

C.1.2 Checking that solutions of Einstein's equations satisfy $\delta S = \delta E$

Using these expansions, it is straightforward to verify that any solution of the linearized Einstein's equations (4.14) satisfies $\delta E = \delta S$, as was done originally in [88] and by another alternative approach in section 4.

Using the expansion (C.1), the equations (4.14) become

$$
\begin{aligned}
H^{(n)}_{tt} &= H^{(n)}_{xx} + H^{(n)}_{yy} && \text{(C.4)}\\
\partial_t H^{(n)}_{tt} &= \partial_x H^{(n)}_{tx} + \partial_y H^{(n)}_{ty} && \text{(C.5)}\\
\partial_t H^{(n)}_{tx} &= \partial_x H^{(n)}_{xx} + \partial_y H^{(n)}_{xy} && \text{(C.6)}\\
\partial_t H^{(n)}_{ty} &= \partial_x H^{(n)}_{xy} + \partial_y H^{(n)}_{yy} && \text{(C.7)}\\
H^{(n)}_{\mu\nu} &= \frac{1}{n(n+3)}(\partial_t^2 - \partial_x^2 - \partial_y^2) H^{(n-2)}_{\mu\nu} \qquad n \geq 2 && \text{(C.8)}\\
H^{(1)}_{\mu\nu} &= 0\,. && \text{(C.9)}
\end{aligned}
$$

Starting with (C.4) and then using (C.5), (C.6), (C.7), and finally (C.8), we find:

$$
\begin{aligned}
& \partial_t^2 H^{(n)}_{tt} = \partial_t^2 (H^{(n)}_{xx} + H^{(n)}_{yy}) \\
\Rightarrow\ & \partial_t (\partial_x H^{(n)}_{xt} + \partial_y H^{(n)}_{yt}) = \partial_t^2 (H^{(n)}_{xx} + H^{(n)}_{yy}) \\
\Rightarrow\ & \partial_x^2 H^{(n)}_{xx} + \partial_y^2 H^{(n)}_{yy} + 2\partial_x\partial_y H^{(n)}_{xy} = \partial_t^2 (H^{(n)}_{xx} + H^{(n)}_{yy}) \\
\Rightarrow\ & \partial_x^2 H^{(n)}_{xx} + \partial_y^2 H^{(n)}_{yy} + 2\partial_x\partial_y H^{(n)}_{xy} = \partial_t^2 (H^{(n)}_{xx} + H^{(n)}_{yy}) \\
\Rightarrow\ & \partial_x^2 H^{(n)}_{xx} + \partial_y^2 H^{(n)}_{yy} + 2\partial_x\partial_y H^{(n)}_{xy} = (\partial_x^2 + \partial_y^2)(H^{(n)}_{xx} + H^{(n)}_{yy}) + (n+2)(n+5)(H^{(n+2)}_{xx} + H^{(n+2)}_{yy}) \\
\Rightarrow\ & 2\partial_x\partial_y H^{(n)}_{xy} = \partial_y^2 H^{(n)}_{xx} + \partial_x^2 H^{(n)}_{yy} + (n+2)(n+5)(H^{(n+2)}_{xx} + H^{(n+2)}_{yy})
\end{aligned}
$$

Using this last equation, we can eliminate $H^{(n)}_{xy}$ from (C.3). This gives

$$
\begin{aligned}
\delta S &= \sum R^{n+2m_x+2m_y+2} \Big\{ \frac{1}{(2m_x)!(2m_y)!} \partial_x^{2m_x} \partial_y^{2m_y} H^{(n)}_{xx}(t, x_0, y_0) C^{xx}_{n,m_x,m_y} \\
&\quad + \frac{1}{(2m_x)!(2m_y)!} \partial_x^{2m_x} \partial_y^{2m_y} H^{(n)}_{yy}(t, x_0, y_0) C^{yy}_{n,m_x,m_y} \Big\} \qquad \text{(C.10)}
\end{aligned}
$$

where for $n \geq 2$ we have

$$\begin{aligned} C^{xx}_{n,m_x,m_y} &= I_{n,m_x,m_y} - I_{n,m_x+1,m_y} - \frac{2m_y}{2m_x+1} I_{n,m_x+1,m_y} - \frac{n(n+3)}{(2m_x+1)(2m_y+1)} I_{n-2,m_x+1,m_y+1} \\ &= 0 \\ C^{yy}_{n,m_x,m_y} &= I_{n,m_x,m_y} - I_{n,m_x,m_y+1} - \frac{2m_x}{2m_y+1} I_{n,m_x,m_y+1} - \frac{n(n+3)}{(2m_x+1)(2m_y+1)} I_{n-2,m_x+1,m_y+1} \\ &= 0 \end{aligned}$$

while for $n = 1$ and $n = 0$, we have

$$\begin{aligned} C^{xx}_{1,m_x,m_y} &= I_{1,m_x,m_y} - I_{1,m_x+1,m_y} - \frac{2m_y}{2m_x+1} I_{1,m_x+1,m_y} = \frac{4}{3} I_{3,m_x,m_y} \\ C^{yy}_{1,m_x,m_y} &= I_{1,m_x,m_y} - I_{1,m_x,m_y+1} - \frac{2m_x}{2m_y+1} I_{1,m_x,m_y+1} = \frac{4}{3} I_{3,m_x,m_y} \end{aligned}$$

and

$$\begin{aligned} C^{xx}_{0,m_x,m_y} &= I_{0,m_x,m_y} - I_{0,m_x+1,m_y} - \frac{2m_y}{2m_x+1} I_{0,m_x+1,m_y} = \frac{3}{2} I_{2,m_x,m_y} \\ C^{yy}_{0,m_x,m_y} &= I_{0,m_x,m_y} - I_{0,m_x,m_y+1} - \frac{2m_x}{2m_y+1} I_{0,m_x,m_y+1} = \frac{3}{2} I_{2,m_x,m_y} \ . \end{aligned}$$

In each case, we have made simplifications using the definition (4.21) of I. Using these results together with (C.9), we find that (C.10) simplifies to

$$\begin{aligned} \delta S &= \sum R^{2m_x+2m_y+2} \frac{1}{(2m_x)!(2m_y)!} \partial_x^{2m_x} \partial_y^{2m_y} (H^{(0)}_{xx}(t,x_0,y_0) + H^{(0)}_{yy}(t,x_0,y_0))(\frac{3}{2} I_{2,m_x,m_y}) \\ &= \frac{3}{2} \sum R^{2m_x+2m_y+2} \frac{1}{(2m_x)!(2m_y)!} \partial_x^{2m_x} \partial_y^{2m_y} H^{(0)}_{tt}(t,x_0,y_0) I_{2,m_x,m_y} \\ &= \delta E \end{aligned}$$

Thus, we have verified that $\delta S = \delta E$ for linearized solutions of Einstein's equations, providing an alternate argument to the one in [88].