Localising Information in Bandlimited Quantum Field Theory

by

Jason Pye

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Applied Mathematics

Waterloo, Ontario, Canada, 2015

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

A deeply-held belief in fundamental physics is that interactions must happen locally. Despite the great success of local models, there are many indications that locality must be abandoned in a theory of quantum gravity. For example, this nonlocality may appear as a minimum length or an ultraviolet cutoff at the Planck scale.

But without a full theory of quantum gravity, can we model the effect that this nonlocality may have on sub-Planckian physics? In this work, we study a model for an ultraviolet cutoff based on Shannon's sampling theorem of classical information theory. We apply the sampling theorem to quantum field theory, which results in the ability to represent fields on a lattice without breaking Euclidean symmetries. After building up the bandlimited quantum field theory, we study the locality and entanglement of its degrees of freedom. We find that each degree of freedom occupies an incompressible volume of space, on the order of the Planck scale in size.

Acknowledgements

I would first like to thank my supervisor Prof. Achim Kempf and collaborator Dr. William Donnelly for their invaluable guidance, mentorship, and assistance throughout the course of this project.

I would like to thank my colleagues Robert Jonsson, Daniel Hümmer, David Layden, Dr. Eduardo Martin-Martinez, Mikhail Panine, and Guillaume Verdon-Akzam with whom I have had countless thoughtful discussions over the course of my time in the Physics of Information lab.

I would also like to thank the University of Waterloo, the Ontario Ministry of Training, Colleges and Universities (MTCU), and the National Sciences and Engineering Research Council (NSERC) for financial support during the course of this work through the Ontario Graduate Scholarship (OGS) program of MTCU and the Canada Graduate Scholarship (CGS-M) program of NSERC.

Table of Contents

List of Figures

List of Symbols

$\mathcal{H}$	a Hilbert space
$\|\psi\rangle$	an element (occasionally referred to as a vector) in a Hilbert space
$\langle\psi\|$	a dual vector
$\|\psi+\phi\rangle \equiv \|\psi\rangle + \|\phi\rangle$	equivalent notations for addition
$\|A\psi\rangle \equiv A\,\|\psi\rangle$	equivalent notations for the action of an operator
$\langle\cdot\|\cdot\rangle$	inner product, conjugate-linear in the first argument
$\mathcal{B}(\mathcal{H})$	the space of bounded linear functions acting on $\mathcal{H}$
$\mathbb{1}$	the identity operator
$D(A)$	the domain of an operator A
$R(A)$	the range of an operator A
$\oplus$	direct sum
$\otimes$	tensor product
$\langle\cdot\rangle$	ensemble average or expectation value

Chapter 1

Introduction

One of the most fundamental principles in modern physics is the principle of locality. Simply put, the principle of locality states that an event cannot instantaneously influence events at distant places. This principle is implemented by requiring that, at least locally, nothing can travel through spacetime faster than the speed of light. Mathematically, this means that physics should be expressed in terms of fields, leading to the subjects of classical and quantum field theory. The principle of locality is one of the most cherished because of its connection with the principle of causality.

However, locality also causes problems, most notably in quantum field theory. In order to achieve locality, one must be able to probe arbitrarily small distance scales. This in turn requires there to be arbitrarily small wavelengths, and it is known that these small wavelengths cause ultraviolet divergences in calculations of physical quantities in quantum field theory. Also, simply from the Heisenberg uncertainty principle,

$$\Delta x \Delta p \geq 1/2 \tag{1.1}$$

(where $\hbar = c = k_B = 1$ throughout), we see that if we were to probe arbitrarily small distance scales, this would cause an enormous uncertainty in the momentum. Since the momentum contributes to the energy-momentum tensor in the Einstein field equations, this large uncertainty in momentum would cause a large uncertainty in curvature. The large uncertainty in curvature would then render it impossible to perform the local measurements in position. [61]

Generally, there are many indications that when gravity is made dynamical, locality is no longer possible. One indication is that if there is a fundamental minimum length close to the Planck length ℓ_P, then the Bekenstein-Hawking entropy would agree with

the entanglement entropy of quantum fields outside of the horizon. We will return to discuss black holes in Section 5.2. Recent work in Refs.[31, 23] also suggests that the gauge-invariant observables one must use in quantum gravity are fundamentally nonlocal. Therefore, it is not clear how to even define locality in the context of gravity. These are few among many reasons to believe that there is a fundamental minimum length in nature [82, 29, 54, 87, 43].

If the fundamental degrees of freedom of nature are nonlocal, then how can we characterise this nonlocality? In this work, we consider a model based on Shannon's sampling theory from classical communication theory. The sampling theorem allows us to represent bandlimited functions on any one of a family of lattices. We apply the sampling theorem to quantum field theory and find that the degrees of freedom are nonlocal. We probe the nonlocality of the degrees of freedom using entanglement entropy of subsystems of the field. Entanglement entropy is a natural choice for studying nonlocality because it can be used to measure statistical independence of degrees of freedom and it seems to be sensitive to Planck-scale physics.

Chapters 2 through 5 review background material for our study. The contributions of this thesis are mainly confined to Chapter 6, where we present a bandlimited quantum field theory model, calculate and determine the scaling behaviour of the entanglement entropy in this model, and use the entropy to localise degrees of freedom of the field to incompressible Planck-scale volumes of space.

Chapter 2

Theory of Self-Adjoint Extensions

The purpose of Chapter 2 will be to provide an overview of self-adjoint extensions of symmetric operators. The purpose of this review is to provide the background at the heart of the functional-analytic approach to sampling theory, which will be presented in Chapter 3. The development of the theory of self-adjoint extensions in this chapter will be in the light of its role in quantum theory, as it provides a natural and illustrative arena for discussion of the issues at hand.

2.1 Basic Definitions and Theorems

This section will present the basic definitions and state the relevant theorems required to engage in the discussion of self-adjoint extensions. The culmination of this section will be with the statement of the spectral theorem and the definition of the functional calculus for self-adjoint operators. First we must review some definitions.

Definition 1 (Adjoint) *Consider a (densely defined) linear operator A acting on a Hilbert space $\mathcal{H}$. Now, for a given $|\phi\rangle \in \mathcal{H}$, suppose there exists a vector $|\chi\rangle \in \mathcal{H}$ such that*

$$\langle \phi | A\psi \rangle = \langle \chi | \psi \rangle \tag{2.1}$$

for all $|\psi\rangle \in \mathcal{H}$. Then the **adjoint** *of A, denoted $A^\dagger$, is defined by the map $A^\dagger |\phi\rangle := |\chi\rangle$. Thus, one often writes $A^\dagger$ as the unique operator satisfying*

$$\langle A^\dagger \phi | \psi \rangle = \langle \phi | A\psi \rangle, \tag{2.2}$$

where $|\psi\rangle \in D(A)$ *and* $|\phi\rangle \in D(A^\dagger)$ *(which is defined as the set of vectors in* $\mathcal{H}$ *which satisfy the above equality).*

Note that the assumption that the operator is densely defined (i.e., the domain of A, denoted $D(A)$, is such that the closure $\overline{D(A)} = \mathcal{H}$) is required so that the operator $A^\dagger$ is well-defined. For consider omitting this assumption. Suppose $\overline{D(A)} \neq \mathcal{H}$, then there is a non-zero element $|\xi\rangle \in \overline{D(A)}^\perp$. Now, as in the definition of the adjoint of A, suppose for a given $|\phi\rangle \in \mathcal{H}$ there is a $|\chi\rangle \in \mathcal{H}$ such that

$$\langle \phi | A\psi \rangle = \langle \chi | \psi \rangle \tag{2.3}$$

so that $A^\dagger |\phi\rangle = |\chi\rangle$. However, since $|\xi\rangle \in \overline{D(A)}^\perp$ and $|\psi\rangle \in D(A)$, then

$$\langle \phi | A\psi \rangle = \langle \chi | \psi \rangle = \langle \chi + \xi | \psi \rangle \tag{2.4}$$

so that we have both $A^\dagger |\phi\rangle = |\chi\rangle$ and $A^\dagger |\phi\rangle = |\chi\rangle + |\xi\rangle$. Therefore, $A^\dagger$ is not well-defined if the domain of A is not dense in $\mathcal{H}$. [55]

Definition 2 (Symmetric operator) *A (densely defined) linear operator* A *acting on a Hilbert space* $\mathcal{H}$ *is* **symmetric** iff

$$\langle \phi | A\psi \rangle = \langle A\phi | \psi \rangle \tag{2.5}$$

for all $|\psi\rangle, |\phi\rangle \in D(A)$.

Definition 3 (Self-Adjoint operator) *A (densely defined) linear operator* A *acting on a Hilbert space* $\mathcal{H}$ *is* **self-adjoint** *iff* A *is symmetric and* $D(A) = D(A^\dagger)$.

Note that by the definition of a symmetric operator A,

$$\langle \phi | A\psi \rangle = \langle A\phi | \psi \rangle, \tag{2.6}$$

which means that $A^\dagger$ exists and $A^\dagger |\phi\rangle = A |\phi\rangle$ for all $|\phi\rangle \in D(A)$. Therefore, one always has $D(A^\dagger) \supset D(A)$ for a symmetric operator A. However, it is possible that $D(A^\dagger) \neq D(A)$, in which case A is symmetric but not self-adjoint.

Definition 4 (Spectrum [35]) *The* **spectrum** *of a linear operator* A*, denoted* $\sigma(A)$*, is the set of* $\lambda \in \mathbb{C}$ *for which* $(A - \lambda \mathbb{1})$ *does not have a bounded inverse.*

The set of eigenvalues of an operator A lie in its spectrum. Recall that $\lambda \in \mathbb{C}$ is an eigenvalue of A if there is some $|\psi\rangle \in \mathcal{H}$ such that $A|\psi\rangle = \lambda|\psi\rangle$. Therefore, $(A-\lambda\mathbb{1})|\psi\rangle = 0$, which means that $(A-\lambda\mathbb{1})$ has a kernel and thus is not invertible. Thus, $\lambda \in \sigma(A)$.

An important property of symmetric operators is that their spectrum is real. Note that for $z := x + iy \in \mathbb{C}$, $x, y \in \mathbb{R}$, we have for a symmetric operator A

$$\begin{aligned} \langle (A-z\mathbb{1})\psi | (A-z\mathbb{1})\psi\rangle &= \langle (A-x\mathbb{1})\psi | (A-x\mathbb{1})\psi\rangle - iy\langle (A-x\mathbb{1})\psi|\psi\rangle \\ &\quad + iy\langle\psi|(A-x\mathbb{1})\psi\rangle + y^2\langle\psi|\psi\rangle \\ \|(A-z\mathbb{1})|\psi\rangle\|^2 &= \|(A-x\mathbb{1})|\psi\rangle\|^2 + y^2\||\psi\rangle\|^2 \\ &\geq y^2\||\psi\rangle\|^2. \end{aligned} \tag{2.7}$$

Where the symmetry property of A was used to cancel the middle terms in the first line. Thus, we see that if $y \neq 0$, then $\|(A-z\mathbb{1})|\psi\rangle\|$ is bounded from below. It follows that $(A-z\mathbb{1})^{-1}$ exists and is bounded, and so $\sigma(A) \subset \mathbb{R}$. [55].

The key importance of the property of self-adjointness lies in the following theorem:

Theorem 1 (Spectral theorem for Self-Adjoint operators [35]) *Let A be a self-adjoint linear operator acting on a Hilbert space $\mathcal{H}$. Let $\mathcal{B}(\mathcal{H})$ be the space of bounded linear operators acting on $\mathcal{H}$. Then there is a unique projection-valued measure $\mu_A : \sigma(A) \to \mathcal{B}(\mathcal{H})$ such that*

$$A = \int_{\sigma(A)} \lambda \, d\mu_A(\lambda). \tag{2.8}$$

Notice that this theorem holds for self-adjoint operators but not for more general symmetric operators.

The spectral decomposition property of self-adjoint operators allows for the following definition.

Definition 5 (Functional Calculus [35]) *For any measurable function f on $\sigma(A)$, where A is a self-adjoint operator, define the operator $f(A)$ by*

$$f(A) := \int_{\sigma(A)} f(\lambda) \, d\mu_A(\lambda). \tag{2.9}$$

Note that if f is bijective, $f(A)$ simply amounts to a relabelling of the eigenvalues.

2.2 The Role of Self-Adjointness in Quantum Theory

In this section we will illustrate the importance of the technical distinction between symmetric and self-adjoint operators by discussing its role in quantum theory. Surprisingly, the importance of this distinction can be illustrated using the simple and well-known infinite potential well example encountered in introductory quantum mechanics courses. The functional analysis of this example is also similar to that of the sampling theorem which will be encountered in Section 3.2. First, however, we will take a step back and recall the general role of self-adjointness in quantum theory.

2.2.1 The Axioms

The axioms of quantum theory (for pure states) may be organised as follows:

Axiom 1: The state of each physical degree of freedom is described by a normalised element $|\psi\rangle$ in a Hilbert space $\mathcal{H}$.

Axiom 2: Systems consisting of multiple degrees of freedom are described by elements in the tensor product of the Hilbert spaces corresponding to the individual degrees of freedom, e.g., $\otimes_i \mathcal{H}_i$ (where $\mathcal{H}_i$ is the Hilbert space associated with the degree of freedom labelled by i).

Axiom 3: Physical observables are described by self-adjoint linear operators acting on the Hilbert space associated with the quantum system. Possible measurement outcomes of an observable are labelled by the eigenvalues of the associated linear operator. The normalised eigenvectors of the operator represent the states associated with each measurement outcome. The probability associated with a measurement outcome is the squared modulus of the overlap of the state of the system at the time of measurement with the eigenvector corresponding to the measurement outcome. The post-measurement state is the eigenvector corresponding to the measurement outcome.

Axiom 4: The state of a quantum system evolves in time (if no measurement is performed) according to a one-parameter group of unitaries

$$|\psi(t)\rangle = U(t)\,|\psi(0)\rangle \tag{2.10}$$

where the unitaries are given by

$$U(t) := e^{-itH} \tag{2.11}$$

and where H is the Hamiltonian of the system.

These axioms only provide a mathematical framework within which one describes quantum systems; it is the art of modelling a quantum system where one identifies the independent variables of the system as well as an appropriate Hilbert space model to describe them.

The analysis of finite-dimensional quantum systems in principle only involves the use of elementary linear algebra. However, infinite-dimensional quantum systems require the use of more general functional-analytic techniques. There are often subtleties in the analysis of operators on infinite-dimensional Hilbert spaces which do not arise in finite-dimensional Hilbert spaces. For example, the particular subtlety alluded to above that will be the focus of the remainder of this chapter is the distinction between symmetric and self-adjoint operators.

It is important that the physical observables be represented by self-adjoint operators. It is natural for the operator to at least be symmetric since measurement outcomes (eigenvalues of the corresponding operator) should be represented by real numbers. Note, however, that it is not necessary for an operator to be symmetric to have a real spectrum, even in finite dimensions. For consider the operator, A, represented by the matrix

$$A \leftrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \tag{2.12}$$

acting on the Hilbert space $\mathbb{C}^2$. Clearly, for any nonzero $\lambda \in \mathbb{C} \setminus \{0, 1\}$,

$$(A - \lambda \mathbb{1})^{-1} \leftrightarrow \begin{bmatrix} \frac{-1}{\lambda - 1} & \frac{-1}{\lambda(\lambda-1)} \\ 0 & \frac{-1}{\lambda} \end{bmatrix} \tag{2.13}$$

exists and is bounded. Thus, $\sigma(A) = \{0, 1\} \subset \mathbb{R}$, however A is clearly not symmetric.

The true importance of the self-adjointness of the operators representing physical observables lies in the fact that they have a unique spectral decomposition (from the spectral theorem) and thus a well-defined functional calculus.

First, the spectral decomposition is useful for characterising measurement outcomes (Axiom 3). This is because once the spectral decomposition of the self-adjoint operator, A, corresponding to an observable is known, then the possible measurement outcomes of this observable are the Borel subsets $E \subset \sigma(A)$ and the probability of this outcome for a state $|\psi\rangle$ is

$$p(E) = \langle\psi| \mu_A(E) |\psi\rangle , \tag{2.14}$$

with post-measurement state

$$|\psi'\rangle = \frac{\mu_A(E)\,|\psi\rangle}{\sqrt{\langle\psi|\,\mu_A(E)\,|\psi\rangle}}. \tag{2.15}$$

Note that due to the functional calculus of A, for any measurable, bijective, and real function f, $f(A)$ is simply a relabelling of the measurement outcomes [21, 34].

There is also a more general type of measurement corresponding to positive-operator valued measures rather than projection valued measures. These positive-operator valued measures along with a labelling of the measurement outcomes can be used to give a non-unique decomposition of a symmetric (not self-adjoint) operator [66] representing a more general type of "unsharp" observable [8]. Here we will restrict our attention to the more conventional notion of observables, represented by self-adjoint operators.

Second, the functional calculus of the Hamiltonian for the system guarantees the existence, uniqueness, and unitarity of the time evolution operator (Axiom 4) [58, 33]. In the special case where the Hamiltonian is time-independent, we also obtain the useful result:

Theorem 2 (Stone's theorem [35]) *Let $\{U(t)\}_t$ be a strongly continuous one-parameter group of unitaries acting on a Hilbert space $\mathcal{H}$. Then the infinitesimal generator,*

$$H\,|\psi\rangle := \lim_{t\to 0^+} \frac{U(t)\,|\psi\rangle - |\psi\rangle}{-it} \tag{2.16}$$

is densely defined and self-adjoint, and

$$U(t) = e^{-itH} \tag{2.17}$$

for all t.

In Axiom 4 as well as in the above theorem, the exponential function $H \mapsto e^{-itH}$ should be understood in terms of the functional calculus of H.

Typically, the Hamiltonian of a system is built from more fundamental observables. In the Dirac quantisation programme [21], the core observables are those associated with the canonical coordinates and momenta of the system. This is because in classical physics, one typically represents a degree of freedom of a system as a real-valued quantity representing some observable property of the system. Then all other properties of the system (e.g., the energy) are represented as functions of the degrees of freedom and their corresponding conjugate momenta. Therefore, in the quantum theory, if a self-adjoint Hamiltonian is to

be constructed from the operators representing the quantised degrees of freedom and their conjugate momenta, it is natural to require that all of these operators are also self-adjoint. That is to say it is natural to require that the generators of the algebra of observables (the core quantum degrees of freedom) are represented by self-adjoint operators in order for the free evolution of the system to be unitary.

2.2.2 Example: Infinite Potential Well

The previous section highlighted the theoretical importance of the requirement that quantum mechanical operators representing physical observables are self-adjoint and not merely symmetric. Often it is the case in physics where such technical distinctions can be overlooked for most practical purposes. However, in this section we will present a simple situation which may lead to confusion if the distinction between symmetric and self-adjoint is not recognised.

This example in this section is based on [13]. Consider a particle in an infinite potential well of length L. That is, the Hamiltonian is

$$H = \frac{1}{2m}P^2 + V \tag{2.18}$$

with

$$V(x) = \begin{cases} 0 & \text{if } x \in [0, L] \\ \infty & \text{if } x \notin [0, L] \end{cases} \tag{2.19}$$

where $V(x)$ is the Schrödinger representation of the operator V in the x-basis. The physical wavefunctions $\psi(x)$ which describe this system are those in $L_2[0, L]$ with Dirichlet boundary conditions: $\psi(0) = \psi(L) = 0$. Thus, the domain of H is:

$$\begin{aligned} D(H) = \{|\psi\rangle : \psi(x) \in L_2[0, L], (H\psi)(x) \in L_2[0, L], \\ \text{and } \psi(0) = \psi(L) = 0\} \end{aligned} \tag{2.20}$$

Now, suppose we are given the physical wavefunction

$$\psi(x) = Cx(L - x) \tag{2.21}$$

(where $C \neq 0$ is simply a normalisation constant) and we are asked to calculate the uncertainty in the energy

$$\Delta E^2 = \langle H^2 \rangle - \langle H \rangle^2 \tag{2.22}$$

The second term is simply calculated from

$$\begin{aligned}\langle H\rangle &= \langle\psi|H\psi\rangle \\ &= \frac{-1}{2m}\int_0^L dx\ \psi^*(x)\frac{d^2}{dx^2}\psi(x) \\ &= \frac{|C|^2L^3}{6m}\end{aligned} \tag{2.23}$$

At first, it may seem that one could calculate the first term in different ways, for example, one could calculate either $\langle H^2\rangle_\psi = \langle H\psi|H\psi\rangle$ or $\langle H^2\rangle_\psi = \langle H^2\psi|\psi\rangle$ in the position basis of the Schrödinger representation. However, it is simple to show that these are not equivalent:

$$\begin{aligned}\langle H\psi|H\psi\rangle &= \frac{1}{4m^2}\int_0^L dx\ \left(\frac{d^2}{dx^2}\psi^*(x)\right)\left(\frac{d^2}{dx^2}\psi(x)\right) \\ &= \frac{|C|^2L}{m^2}\end{aligned} \tag{2.24}$$

and

$$\begin{aligned}\langle H^2\psi|\psi\rangle &= \frac{1}{4m^2}\int_0^L dx\ \left(\frac{d^4}{dx^4}\psi^*(x)\right)\psi(x) \\ &= 0\end{aligned} \tag{2.25}$$

What went wrong? Well, clearly H is symmetric, since for $|\phi\rangle, |\psi\rangle \in D(H)$,

$$\begin{aligned}\langle\phi|H\psi\rangle &= \frac{-1}{2m}\int_0^L dx\ \phi^*(x)\left(\frac{d^2}{dx^2}\psi(x)\right) \\ &= \frac{-1}{2m}\phi^*(x)\frac{d}{dx}\psi(x)\Big|_{x=0}^{L} - \frac{-1}{2m}\phi^*(x)\psi(x)\Big|_{x=0}^{L} + \frac{-1}{2m}\int_0^L dx\ \left(\frac{d^2}{dx^2}\phi^*(x)\right)\psi(x) \\ &= \langle H\phi|\psi\rangle\end{aligned} \tag{2.26}$$

where the boundary terms vanish since $|\phi\rangle, |\psi\rangle \in D(H)$ obey Dirichlet boundary conditions on the interval $[0, L]$. Therefore $\langle\phi|H\psi\rangle = \langle H\phi|\psi\rangle$ for all $|\psi\rangle, |\phi\rangle \in D(H)$. However, in order to have

$$\langle H\psi|H\psi\rangle = \langle H^2\psi|\psi\rangle \tag{2.27}$$

we must have both $|\psi\rangle \in D(H)$ and $|H\psi\rangle \in D(H)$. In the above example, this is clearly not the case since $(H\psi)(x) = C/m$ is not in $D(H)$ because the constant function $C/m(\neq 0)$ does not obey Dirichlet boundary conditions.

This issue arises because H is not a self-adjoint operator. Furthermore, the momentum operator P is also not a self-adjoint operator. First, we will consider only the momentum operator before returning to discuss the Hamiltonian operator. The domain of the momentum operator is

$$\begin{aligned} D(P) = \{|\psi\rangle : \psi(x) \in L_2[0,L], (P\psi)(x) \in L_2[0,L], \\ \text{and } \psi(0) = \psi(L) = 0\} \end{aligned} \tag{2.28}$$

Now let us calculate the adjoint of P. Note that for $\psi \in D(P)$,

$$\begin{aligned} \langle\phi|P\psi\rangle &= \int_0^L dx\, \phi^*(x) \left(-i\frac{d}{dx}\psi(x)\right) \\ &= -i\phi^*(x)\psi(x)|_{x=0}^L + \int_0^L dx\, \left(-i\frac{d}{dx}\phi(x)\right)^* \psi(x) \\ &= \int_0^L dx\, \left(-i\frac{d}{dx}\phi(x)\right)^* \psi(x) \\ &= \langle P^\dagger\phi|\psi\rangle \end{aligned} \tag{2.29}$$

The boundary term vanishes since $|\psi\rangle \in D(P)$ obeys Dirichlet boundary conditions. Clearly, if $|\phi\rangle \in D(P)$ then $P^\dagger|\phi\rangle = P|\phi\rangle$. However, in the above calculation we did not have to impose Dirichlet boundary conditions on $|\phi\rangle$. Therefore the operator $P^\dagger$ is defined on a larger set of vectors, namely

$$D(P^\dagger) = \{|\psi\rangle : \psi(x) \in L_2[0,L] \text{ and } (P^\dagger\psi)(x) \in L_2[0,L]\} \tag{2.30}$$

Thus we see that P is symmetric but not self-adjoint, since $D(P^\dagger) \neq D(P)$.

Despite the fact that the momentum and Hamiltonian operators are not self-adjoint, it is possible to find an extension of their respective domains on which they are self-adjoint. Given this *s*elf-adjoint extension of the Hamiltonian, one can use the functional calculus of H to unambiguously calculate $\langle H^2\rangle_\psi$ as

$$\langle H^2\rangle_\psi = \int_{\sigma(H)} E^2 d\langle\psi|\mu_H(E)\psi\rangle \tag{2.31}$$

(Note that the projection $\mu_H(E)$ is self-adjoint.)

The next section will describe the theory behind self-adjoint extensions of symmetric operators which will resolve the issues presented in our example. Although one can construct self-adjoint extensions of H independently of the self-adjoint extensions of P, for convenience we will focus on the construction of the self-adjoint extensions of P and then define $H = P^2/2m$ and $H^2 = P^4/4m^2$ in terms of the functional calculus of the self-adjoint extensions of P.

2.3 Self-Adjoint Extensions

This section will provide an overview of the construction of self-adjoint extensions of symmetric operators. By an **extension** of an operator A, we mean that $\tilde{A}$ is an extension of A if $D(A) \subset D(\tilde{A})$ and $\tilde{A} = A$ on $D(A)$ [35]. The key insight which provides a systematic method for constructing self-adjoint extensions is that symmetric extensions of a symmetric operator are in one-to-one correspondence with isometric extensions of a particular isometric operator related to A, called the Cayley transform of A (which will be defined momentarily). There is a natural way of constructing these isometric extensions, which in turn provides a natural method for constructing the required extensions of the symmetric operator A.

First, a couple of definitions:

Definition 6 (Isometric operator [3]) *An operator U acting on a Hilbert space $\mathcal{H}$ is called* **isometric** *if*

$$\langle U\phi | U\psi \rangle = \langle \phi | \psi \rangle \tag{2.32}$$

for all $|\phi\rangle, |\psi\rangle \in D(U) \subset \mathcal{H}$.

Here we assume U maps $\mathcal{H}$ to itself, however clearly one can generalise this definition for a map U between two different Hilbert spaces.

Definition 7 (Unitary operator [3]) *An operator U acting on a Hilbert space $\mathcal{H}$ is called* **unitary** *if it is isometric, $D(U) = \mathcal{H}$, and $R(U) = \mathcal{H}$.*

If U is unitary, it can be shown that its inverse exists and is equal to its adjoint, i.e., $U^{-1} = U^\dagger$ [3].

Recall that the spectrum $\sigma(A)$ of a symmetric operator A is real, since for $z := x+iy \in \mathbb{C}$ $(x, y \in \mathbb{R})$,

$$\begin{aligned} \| (A - z\mathbb{1}) |\psi\rangle \| &\geq |y| \| |\psi\rangle \| - \| (A - x\mathbb{1}) |\psi\rangle \| \\ &\geq |y| \| |\psi\rangle \| \end{aligned} \tag{2.33}$$

Therefore, if $y \neq 0$, then $(A - z\mathbb{1})^{-1}$ exists and is bounded, and so z is in the **resolvent set** of A, denoted $\rho(A)$ [35]. The spectrum is $\sigma(A) = \mathbb{C} \setminus \rho(A)$. The above calculation shows that the upper-half complex plane, $\mathrm{Im}(z) > 0$, and the lower-half complex plane, $\mathrm{Im}(z) < 0$, are connected subsets of $\rho(A)$ (for symmetric A).

Now one can compare this to connected subsets of the resolvent set of an isometric operator U. This will motivate the form of a natural map between symmetric and isometric operators (the Cayley transform). Note that for an isometric operator U, $\xi \in \mathbb{C}$, and $|\xi| > 1$ [3]

$$\begin{aligned} \|(U - \xi\mathbb{1})\,|\psi\rangle\,\| &\geq |\xi|\|\,|\psi\rangle\,\| - \|U\,|\psi\rangle\,\| \\ &= (|\xi| - 1)\|\,|\psi\rangle\,\| \end{aligned} \tag{2.34}$$

In the case $|\xi| < 1$, [3]

$$\begin{aligned} \|(U - \xi\mathbb{1})\,|\psi\rangle\,\| &\geq \|U\,|\psi\rangle\,\| - |\xi|\|\,|\psi\rangle\,\| \\ &= (1 - |\xi|)\|\,|\psi\rangle\,\| \end{aligned} \tag{2.35}$$

Thus we see that for $|\xi| \neq 1$, $(U - \xi\mathbb{1})^{-1}$ exists and is bounded. Therefore, the spectrum of U, $\sigma(U)$, is a subset of the unit circle in $\mathbb{C}$ and the sets of $\xi \in \mathbb{C}$ inside the unit circle, $|\xi| < 1$, and outside the unit circle, $|\xi| > 1$, respectively form connected subsets of the resolvent set $\rho(U)$.

Consider the following map $f : \mathbb{C} \to \mathbb{C}$, defined by

$$f : z \mapsto \frac{z - i}{z + i} \tag{2.36}$$

and inverse map

$$f^{-1} : z \mapsto -i\frac{z + 1}{z - 1} \tag{2.37}$$

The transformation f of $\mathbb{C}$ maps the real line to the unit circle, and the upper- and lower-half planes to the interior and exterior of the unit circle, respectively. Therefore, the transformation naturally maps the spectrum and resolvent set of a symmetric operator A to the spectrum and resolvent set (respectively) of some corresponding isometric operator U. This motivates the following definition:

Definition 8 (Cayley Transform [3, 35]) *The* **Cayley transform** *of a symmetric operator A is an isometric operator $U_A : R(A + i\mathbb{1}) \to R(A - i\mathbb{1})$ (where $R(A)$ denotes the range of an arbitrary operator A) given by*

$$U_A = (A - i\mathbb{1})(A + i\mathbb{1})^{-1} \tag{2.38}$$

The inverse transformation is

$$A = -i(U_A + \mathbb{1})(U_A - \mathbb{1})^{-1} \tag{2.39}$$

In this definition, note that $(A+i\mathbb{1})^{-1}$ exists and is bounded since $-i \in \rho(A)$. It can also be shown that $(U_A - \mathbb{1})^{-1}$ exists and is bounded [3]. Indeed, U_A is isometric, since one can write $U_A = (A - i\mathbb{1})(A + i\mathbb{1})^{-1}$ as

$$\begin{aligned} U_A \left|f\right\rangle &= (A - i\mathbb{1}) \left|h\right\rangle \\ (A + i\mathbb{1}) \left|h\right\rangle &= \left|f\right\rangle \end{aligned} \tag{2.40}$$

Then for any $\left|f\right\rangle, \left|g\right\rangle \in R(A+i\mathbb{1})$ where $\left|f\right\rangle = (A+i\mathbb{1})\left|h\right\rangle$ and $\left|g\right\rangle = (A+i\mathbb{1})\left|k\right\rangle$, we have

$$\begin{aligned} \langle f|g\rangle &= \langle (A+i\mathbb{1})h|(A+i\mathbb{1})k\rangle \\ &= \langle Ah|Ak\rangle + i\langle Ah|k\rangle - i\langle h|Ak\rangle + \langle h|k\rangle \\ &= \langle Ah|Ak\rangle + \langle h|k\rangle \end{aligned} \tag{2.41}$$

and

$$\begin{aligned} \langle U_A f|U_A g\rangle &= \langle (A-i\mathbb{1})h|(A-i\mathbb{1})k\rangle \\ &= \langle Ah|Ak\rangle + i\langle h|Ak\rangle - i\langle Ah|k\rangle + \langle h|k\rangle \\ &= \langle Ah|Ak\rangle + \langle h|k\rangle \end{aligned} \tag{2.42}$$

where in both equations we have used the fact that A is symmetric. By comparing the last line of each equation, we see that

$$\langle U_A f|U_A g\rangle = \langle f|g\rangle \tag{2.43}$$

thus U_A is isometric.

Before proceeding, the concept of the *deficiency indices* of a symmetric operator A will prove to be useful.

Definition 9 (Deficiency [3]) *The* **deficiency subspaces** *of a symmetric operator A are defined to be*

$$\begin{aligned} N_+ &:= R(A+i\mathbb{1})^\perp \\ N_- &:= R(A-i\mathbb{1})^\perp \end{aligned}$$

The **deficiency indices** *are the respective dimensions of the deficiency subspaces:*

$$\begin{aligned} n_+ &:= \dim(N_+) \\ n_- &:= \dim(N_-) \end{aligned}$$

Note that since the Cayley transform of A, U_A, maps $R(A+i\mathbb{1})$ to $R(A-i\mathbb{1})$, we also have

$$N_+ = D(U_A)^\perp$$
$$N_- = R(U_A)^\perp$$

The following theorems guarantee that symmetric extensions of A are in one-to-one correspondence with the isometric extensions of U_A, and that A is self-adjoint when U_A is unitary. Then the Neumann formulas will be presented, which will provide compact representation for construction symmetric extensions of A.

Theorem 3 *Let A, $\tilde{A}$ be two symmetric operators and U_A, $U_{\tilde{A}}$ be their respective Cayley transforms. Then $\tilde{A}$ is an extension of A iff $U_{\tilde{A}}$ is an extension of U_A. [3]*

Theorem 4 *A symmetric operator is maximal iff one of its deficiency indices is zero. A symmetric operator is self-adjoint iff both of its deficiency indices are zero. [3]*

Theorem 5 *Let A be a symmetric operator with domain $D(A)$ and deficiency subspaces N_+ and N_-. Then the domain of its adjoint, $A^\dagger$, is [3]*

$$D(A^\dagger) = D(A) \oplus N_+ \oplus N_- \tag{2.44}$$

The above theorems give a method for constructing self-adjoint extensions of a symmetric operaotor A. Recall that $N_+ = D(U_A)^\perp$ and $N_- = R(U_A)^\perp$, and consider extending U_A by adjoining an arbitrary isometric operator $U' : N_+ \to N_-$ so that the extension $\tilde{U}_A$ is given by

$$\tilde{U}_A |\psi\rangle := \begin{cases} U_A |\psi\rangle & \text{for } |\psi\rangle \in D(U_A) \\ U' |\psi\rangle & \text{for } |\psi\rangle \in N_+ \end{cases} \tag{2.45}$$

or compactly written as $\tilde{U}_A = U_A \oplus U'$. Clearly $\tilde{U}_A$ is an isometric extension of U_A. Furthermore, if the dimensions of N_+ and N_- are equal, i.e., if $n_+ = n_-$, and U' maps N_+ onto N_-, then the deficiency indices of $\tilde{U}_A$ are both zero and therefore this operator is unitary [3]. A self-adjoint extension for the corresponding operator A is given by the inverse Cayley transform of $\tilde{U}_A$. If the deficiency indices are not equal, it is not possible to construct a self-adjoint extension of the operator A, but only a maximally symmetric extension [3].

Note that a self-adjoint extension of the symmetric operator A is not unique, since the isometric operator $U' : N_+ \to N_-$ remains arbitrary. In general, one obtains an entire

family of self-adjoint extensions parametrised by $U(n)$ (where $n := n_+ = n_-$), i.e., one obtains a self-adjoint extension of A for each element of the group $U(n)$.

Let us now examine the form of the self-adjoint extension $\tilde{A}$ given by the inverse Cayley transform $\tilde{A} := -i(\tilde{U}_A + \mathbb{1})(\tilde{U}_A - \mathbb{1})^{-1}$. For present purposes, this transformation will be written as

$$\begin{aligned} \tilde{A}\,|f\rangle &= -i(\tilde{U}_A + \mathbb{1})\,|h\rangle \\ (\tilde{U}_A - \mathbb{1})\,|h\rangle &= |f\rangle \end{aligned} \tag{2.46}$$

By construction, the domain of $\tilde{U}_A$ is $D(\tilde{U}_A) = D(U_A) \oplus N_+$. From the above transformation, we see that

$$\begin{aligned} D(\tilde{A}) &= R(\tilde{U}_A - \mathbb{1}) \\ &= (\tilde{U}_A - \mathbb{1})D(\tilde{U}_A) \\ &= [(U_A - \mathbb{1}) \oplus (U' - \mathbb{1})]\,[D(U_A) \oplus N_+] \\ &= (U_A - \mathbb{1})D(U_A) \oplus (U' - \mathbb{1})N_+ \\ &= D(A) \oplus (U' - \mathbb{1})N_+ \end{aligned} \tag{2.47}$$

Therefore, the domain of a self-adjoint extension of A consists of vectors of the form

$$|\psi\rangle = |\psi_0\rangle + (U' - \mathbb{1})\,|h\rangle \tag{2.48}$$

where $|\psi_0\rangle \in D(A), |h\rangle \in N_+ = R(A + i\mathbb{1})^\perp$. The action of $\tilde{A}$ is given by

$$\tilde{A}\,|\psi\rangle = A\,|\psi_0\rangle - i(U' + \mathbb{1})\,|h\rangle \tag{2.49}$$

These last two equations are called the **Neumann formulas** [3].

Therefore, to construct a self-adjoint extension of a symmetric operator A, first one must determine the deficiency subspaces $N_\pm := R(A \pm i\mathbb{1})^\perp$. Then, if $n = n_+ = n_-$, find $U' : N_+ \to N_-$ which is unique modulo $U(n)$ and the self-adjoint extension $\tilde{A}$ of A is determined by the Neumann formulas. To illustrate these ideas, we shall conclude this chapter by returning to the infinite potential well example from the previous section.

2.3.1 Example: Infinite Potential Well (Resolution)

Recall that the issue which arose when calculating the second moment of the Hamiltonian was caused by the fact that the momentum operator P on the interval $[0, L]$ is merely

symmetric and not self-adjoint. Constructing a self-adjoint extension for the momentum operator will allow this issue to be resolved by using the spectral decomposition and functional calculus of the self-adjoint extension of the momentum operator. Recall that in the above example, the domain of P was defined to be

$$\begin{aligned} D(P) = \{|\psi\rangle : \psi(x) \in L_2[0,L],\, (P\psi)(x) \in L_2[0,L], \\ \text{and } \psi(0) = \psi(L) = 0\} \end{aligned} \tag{2.50}$$

We begin by examining the deficiency subspaces of the momentum operator, defined as $N_\pm = R(P \pm i\mathbb{1})^\perp$. Therefore, for $|\phi_\pm\rangle \in N_\pm$ and any $|\psi\rangle \in D(P)$, we have

$$\begin{aligned} 0 &= \langle\phi_\pm|(P \pm i\mathbb{1})\psi\rangle \\ &= \int_0^L dx\, \phi_\pm^*(x) \left(-i\frac{d}{dx}\psi(x)\right) \pm i\int_0^L dx\, \phi^*(x)\psi(x) \\ &= \int_0^L dx\, i\left[\left(\frac{d}{dx} \pm 1\right)\phi_\pm(x)\right]^* \psi(x) \end{aligned} \tag{2.51}$$

where the boundary term in the integration by parts vanished since $|\psi\rangle \in D(P)$. Thus we have $|\phi_\pm\rangle \in N_\pm$ if

$$\begin{aligned} \frac{d}{dx}\phi_\pm(x) &= \mp\phi_\pm(x) \\ \phi_\pm(x) &= ce^{\mp x} \end{aligned} \tag{2.52}$$

for some constant c (arbitrary since $e^{\mp x} \in L_2[0,L]$). The spaces $N_\pm$ respectively consist of multiples of $e^{\mp}$, and the deficiency indices are $n_\pm = 1$. Therefore, in order to construct the self-adjoint extensions of P, we can define an isometric operator parametrised by $U(1)$ as $U_\alpha : N_+ \to N_-$ by $U_\alpha : e^{-x} \mapsto e^{i\alpha - L}e^x$ (where the e^{-L} is introduced so that U_α is isometric). The corresponding momentum operator will be written P_α with domain given by the first Neumann formula,

$$D(P_\alpha) = \{\psi_0(x) + c(e^{i\alpha-L}e^x - e^{-x}) : |\psi_0\rangle \in D(P), c \in \mathbb{C}\} \tag{2.53}$$

and action given by the second Neumann formula

$$\begin{aligned} \psi(x) &:= \psi_0(x) + c(e^{i\alpha-L}e^x - e^{-x}) \\ (P_\alpha\psi)(x) &= -i\frac{d}{dx}\psi_0(x) - ic(e^{i\alpha-L}e^x + e^{-x}) \end{aligned} \tag{2.54}$$

Note that $(P_\alpha\psi)(x)$ can simply be expressed as

$$(P_\alpha\psi)(x) = -i\frac{d}{dx}\psi(x) \tag{2.55}$$

Furthermore, note that $\psi(0) = c(e^{i\alpha-L} - 1)$ and $\psi(L) = c(e^{i\alpha} - e^{-L})$. It is also easy to check that

$$\left|\frac{\psi(L)}{\psi(0)}\right| = \left|\frac{e^{i\alpha} - e^{-L}}{e^{i\alpha-L} - 1}\right| = 1 \tag{2.56}$$

thus $\psi(0)$ and $\psi(L)$ are simply related by a phase. This shows that the operator P_α is formally represented by $-id/dx$ with domain consisting of $L_2[0, L]$ functions with periodic boundary conditions up to a phase. Thus, after a reparametrisation, the self-adjoint extensions of P are P_α with

$$\begin{aligned} D(P_\alpha) = \{|\psi\rangle : \psi(x) \in L_2[0, L], (P_\alpha\psi)(x) \in L_2[0, L], \\ \text{and } \psi(L) = e^{i\alpha}\psi(0)\} \end{aligned} \tag{2.57}$$

Now that we have the self-adjoint operator P_α, it is simple to calculate its eigenvectors and eigenvalues:

$$\begin{aligned} P_\alpha |\phi_n^{(\alpha)}\rangle &= k_n^{(\alpha)} |\phi_n^{(\alpha)}\rangle \\ \phi_n^{(\alpha)}(x) &= \frac{1}{\sqrt{L}} e^{ik_n^{(\alpha)}x} \\ k_n^{(\alpha)} &= \frac{2\pi n + \alpha}{L}, \quad n \in \mathbb{Z} \end{aligned} \tag{2.58}$$

Therefore we obtain the spectral decomposition

$$P_\alpha = \sum_{n\in\mathbb{Z}} k_n^{(\alpha)} |\phi_n^{(\alpha)}\rangle \langle\phi_n^{(\alpha)}| \tag{2.59}$$

Now returning to the example from the previous section, after making a choice of self-adjoint extension one can calculate expectation values for $\langle H_\alpha^2\rangle_\psi = \langle P_\alpha^4\rangle/(4m^2)$. For example, one could choose anti-periodic boundary conditions ($\alpha = \pi$), so that for the state $\psi(x) = Cx(L - x)$ we have

$$\begin{aligned} \langle\phi_n^{(\pi)}|\psi\rangle &= \int_0^L dx\, \frac{1}{\sqrt{L}} e^{-ik_n^{(\pi)}x} Cx(L - x) \\ &= \frac{-4iC}{k^3\sqrt{L}} \end{aligned} \tag{2.60}$$

Then

$$\begin{aligned}\langle H_\pi^2\rangle_\psi &= \frac{1}{4m^2}\sum_{n\in\mathbb{Z}}(k_n^{(\pi)})^4|\langle\phi_n^{(\pi)}|\psi\rangle|^2 \\ &= \frac{4|C|^2L}{\pi^2m^2}\sum_{n\in\mathbb{Z}}\frac{1}{(2n+1)^2} \\ &= \frac{|C|^2L}{m^2}\end{aligned} \tag{2.61}$$

which is the result obtained previously for $\langle H\psi|H\psi\rangle$.

Note that since the spectrum and the projections of P_α depend on the particular choice of self-adjoint extension, the expectation values will also generally depend on the choice of self-adjoint extension. Therefore, one must provide additional information (e.g., from experiment) in order to motivate the choice of self-adjoint extension [13].

Chapter 3

Sampling Theory

Chapter 3 will review Shannon's sampling theorem from classical information theory. It will be examined in the light of the functional analysis outlined in Chapter 2 in terms of the family of self-adjoint extensions of a position (or multiplication) operator.

3.1 Shannon's Sampling Theorem

Shannon's sampling theorem is a central result of classical information theory. It is intriguing for information theorists because it provides an equivalence between continuous and discrete representations of information. It is also of great practical value; it has found numerous applications in communications engineering and signals processing.

Before discussing further, we shall state the theorem:

Theorem 6 (Shannon sampling theorem) *Let f be a* **bandlimited function** *with* **bandlimit** Ω*, i.e., f is a function whose Fourier transform has support on the interval $(-\Omega, \Omega)$. Then f can be reconstructed from its values on the lattice $x_n = n\pi/\Omega$ $(n \in \mathbb{Z})$ via the reconstruction formula:*

$$f(x) = \sum_{n \in \mathbb{Z}} \operatorname{sinc}[(x - x_n)\Omega] f(x_n) \tag{3.1}$$

where $\operatorname{sinc}(x) := \frac{\sin(x)}{x}$.

This theorem states that if a bandlimited function (with bandlimit Ω) is known on the lattice $\{x_n\}_{n\in\mathbb{Z}}$ then it is also known at all points $x \in \mathbb{R}$.

The sampling theorem also holds with the same reconstruction formula for any shifted lattice $x_n \mapsto x_n^{(\alpha)} = (2\pi n - \alpha)/(2\Omega)$. Further generalisations of the theorem will be mentioned below.

3.2 Functional Analytic Approach

In this section, we shall discuss the functional analytic approach to deriving the sampling theorem. This will be accomplished by examining the self-adjoint extensions of a particular operator, namely the position operator (in quantum mechanical language), also referred to as the multiplication operator (in functional analysis language). Here we will adopt the quantum mechanical language since one may view the results of this section as a "first-quantised" version of the sampling theorem, which we will later "second-quantise" in Section 6.1.

We will denote the Fourier transform of a function ψ by $\tilde{\psi}$. Consider the Hilbert space of square-integrable bandlimited functions with bandlimit Ω, which we shall denote $B(\Omega) := \{\psi(x) \in L_2(\mathbb{R}) : \tilde{\psi}(k) = 0 \text{ for } |k| \geq \Omega\}$, and a position operator X acting on this space. In Fourier space (momentum basis) the position operator is represented by id/dk. The domain of X is

$$D(X) = \{|\psi\rangle : \tilde{\psi}(k), (X\tilde{\psi})(k) \in L_2[-\Omega, \Omega] \text{ and } \tilde{\psi}(-\Omega) = \tilde{\psi}(\Omega) = 0\} \tag{3.2}$$

Notice that this problem is dual to the infinite potential well problem studied in Chapter 2, where the position and momentum operators have switched roles. Thus we can construct the self-adjoint extensions of the position operator analogously to the construction of the self-adjoint extensions of the momentum operator in the infinite potential well problem.

Recall that the deficiency subspaces are defined as $N_\pm := R(X \pm i\mathbb{1})$. Let $|\phi_\pm\rangle \in N_\pm$ and $|\psi\rangle \in D(X)$, then

$$\begin{aligned} 0 &= \langle\phi_\pm|(X \pm i\mathbb{1})\psi\rangle \\ &= \int_{-\Omega}^{\Omega} dk\, (-i) \left[\left(\frac{d}{dk} \mp 1\right) \tilde{\phi}_\pm(k)\right]^* \tilde{\psi}(k) \end{aligned} \tag{3.3}$$

Therefore the deficiency subspaces are

$$N_\pm = \text{span}_{\mathbb{C}}(e^{\pm k}) \tag{3.4}$$

and the deficiency indices are $n_\pm = 1$. Thus the isometric operators $U_\alpha : N_+ \to N_-$ generating the self-adjoint extensions of the position operator are parametrised as

$$U_\alpha : e^k \mapsto e^{i\alpha} e^{-k} \tag{3.5}$$

The resulting Neumann formulas are:

$$\tilde{\psi}(k) = \tilde{\psi}_0(k) + c(e^{i\alpha} e^{-k} - e^k) \tag{3.6}$$

$$(X_\alpha \tilde{\psi})(k) = i\frac{d}{dk}\tilde{\psi}_0(k) - ic(e^{i\alpha} e^{-k} + e^k) \tag{3.7}$$

for $|\psi_0\rangle \in D(X)$, $c \in \mathbb{C}$. Note that the action of X can simply be written as

$$(X\tilde{\psi})(k) = i\frac{d}{dk}\tilde{\psi}(k) \tag{3.8}$$

In a similar manner to the infinite potential well example, since $\tilde{\psi}(-\Omega) = c(e^{i\alpha} e^{\Omega} - e^{-\Omega})$ and $\tilde{\psi}(\Omega) = c(e^{i\alpha} e^{-\Omega} - e^{\Omega})$ then

$$\left| \frac{\tilde{\psi}(\Omega)}{\tilde{\psi}(-\Omega)} \right| = \left| \frac{e^{i\alpha} e^{-\Omega} - e^{\Omega}}{e^{i\alpha} e^{\Omega} - e^{-\Omega}} \right| = 1 \tag{3.9}$$

Therefore, we can reparametrise the self-adjoint extensions so that for $\alpha \in [0, 2\pi)$,

$$(X_\alpha \tilde{\psi})(k) = i\frac{d}{dk}\tilde{\psi}(k) \tag{3.10}$$

with

$$D(X_\alpha) = \{|\psi\rangle : \tilde{\psi}(k), (X_\alpha\tilde{\psi})(k) \in L_2[-\Omega, \Omega] \text{ and } \tilde{\psi}(\Omega) = e^{i\alpha}\tilde{\psi}(-\Omega)\} \tag{3.11}$$

The eigenvectors and eigenvalues of the operator X_α are easily calculated to be

$$X_\alpha |x_n^{(\alpha)}\rangle = x_n^{(\alpha)} |x_n^{(\alpha)}\rangle \tag{3.12}$$

$$\langle k|x_n^{(\alpha)}\rangle = \frac{1}{\sqrt{2\Omega}} e^{-ikx_n^{(\alpha)}} \tag{3.13}$$

$$x_n^{(\alpha)} = \frac{2\pi n - \alpha}{2\Omega} \tag{3.14}$$

Notice that the spectrum of X_α is discrete and describes a lattice in position space. Furthermore, the eigenvectors of each X_α form an orthogonal basis for the space of square-integrable bandlimited functions (with bandlimit Ω) and admit a resolution of identity

$$\sum_{n\in\mathbb{Z}} |x_n^{(\alpha)}\rangle \langle x_n^{(\alpha)}| = \mathbb{1} \tag{3.15}$$

Since the eigenvectors on the corresponding lattice span the Hilbert space $B(\Omega)$, any function in this Hilbert space can be uniquely represented on this lattice. This is the essence of the sampling theorem.

It remains to derive the reconstruction formula. Note that although the eigenvectors of a particular self-adjoint extension X_α form an orthogonal basis for $B(\Omega)$, the eigenvectors from two different self-adjoint extensions are not orthogonal, since

$$\begin{aligned}\langle x_n^{(\alpha)}|x_{n'}^{(\alpha')}\rangle &= \int_{-\Omega}^{\Omega} dk\, \langle x_n^{(\alpha)}|k\rangle\langle k|x_{n'}^{(\alpha')}\rangle \\ &= \int_{-\Omega}^{\Omega} \frac{dk}{2\Omega}\, e^{ik(x_n^{(\alpha)}-x_{n'}^{(\alpha')})} \\ &= \mathrm{sinc}[(x_n^{(\alpha)}-x_{n'}^{(\alpha')})\Omega] \end{aligned} \tag{3.16}$$

If $\alpha' = \alpha$, then

$$\mathrm{sinc}[(x_n^{(\alpha)}-x_{n'}^{(\alpha')})\Omega] = \delta_{nn'} \tag{3.17}$$

However, if $\alpha' \neq \alpha$, then $|x_n^{(\alpha)}\rangle, |x_{n'}^{(\alpha')}\rangle$ will generally have a nontrivial overlap.

Above we noted that the spectrum of a particular self-adjoint extension X_α is a lattice in position space $\{x_n^{(\alpha)} = (2\pi n - \alpha)/(2\Omega)\}_{n\in\mathbb{Z}}$. The parameter $\alpha \in [0, 2\pi)$ which parametrises the self-adjoint extensions corresponds to a shift in the lattice. The distance between two adjacent points on the same lattice is $\Delta x = \pi/\Omega$ (a uniform density of points). Note that the union of the spectra of the entire family of self-adjoint extensions (parametrised by $\alpha \in [0, 2\pi)$) provides a covering of $\mathbb{R}$. Therefore, if we correspondingly take the union of the eigenvectors of the family of self-adjoint extensions, we will obtain a continuum basis $\{|x\rangle\}_{x\in\mathbb{R}}$ where $|x\rangle$ is identified with $|x_n^{(\alpha)}\rangle$ when $x = x_n^{(\alpha)}$. Using the continuum basis and the resolution of identity for the eigenvectors of a single self-adjoint extension, one readily obtains the reconstruction formula for an arbitrary $|\psi\rangle \in B(\Omega)$:

$$\begin{aligned}\langle x|\psi\rangle &= \sum_{n\in\mathbb{Z}} \langle x|x_n^{(\alpha)}\rangle\langle x_n^{(\alpha)}|\psi\rangle \\ \psi(x) &= \sum_{n\in\mathbb{Z}} \mathrm{sinc}[(x-x_n^{(\alpha)})\Omega]\psi(x_n^{(\alpha)}) \end{aligned} \tag{3.18}$$

A note on terminology: a lattice such as $\{x_n^{(\alpha)}\}_{n\in\mathbb{Z}}$ from which one can reconstruct a function in $B(\Omega)$ is called a **sampling lattice**. A kernel G used to perform the reconstruction $\psi(x) = \sum_n G(x, x_n)\psi(x_n)$ is in general called a **reconstruction kernel**.

We further note that in the continuum basis one obtains an overcomplete resolution of identity,

$$\frac{\Omega}{\pi}\int_{\mathbb{R}} dx\ |x\rangle\langle x| = \mathbb{1} \tag{3.19}$$

which can be checked on a basis:

$$\begin{aligned}
\langle x_n^{(\alpha)}|\left(\frac{\Omega}{\pi}\int_{\mathbb{R}} dx\ |x\rangle\langle x|\right)|x_m^{(\alpha)}\rangle &= \frac{\Omega}{\pi}\int_{\mathbb{R}} dx\int_{-\Omega}^{\Omega} dk\int_{-\Omega}^{\Omega} dk'\ \langle x_n^{(\alpha)}|k\rangle\langle k|x\rangle\langle x|k'\rangle\langle k'|x_m^{(\alpha)}\rangle \\
&= \frac{\Omega}{\pi}\int_{\mathbb{R}} dx\int_{-\Omega}^{\Omega} dk\int_{-\Omega}^{\Omega} dk'\ \frac{1}{(2\Omega)^2}e^{ik(x_n^{(\alpha)}-x)}e^{ik'(x-x_m^{(\alpha)})} \\
&= \int_{-\Omega}^{\Omega}\frac{dk}{2\Omega}\ e^{ik(x_n^{(\alpha)}-x_m^{(\alpha)})} \\
&= \delta_{nm}
\end{aligned} \tag{3.20}$$

Using this overcomplete resolution of identity, one can obtain a reproducing kernel for $B(\Omega)$:

$$\begin{aligned}
\langle x|\psi\rangle &= \frac{\Omega}{\pi}\int_{\mathbb{R}} dx'\ \langle x|x'\rangle\langle x'|\psi\rangle \\
\psi(x) &= \int_{\mathbb{R}} dx'\ K(x,x')\psi(x')
\end{aligned} \tag{3.21}$$

where $K(x,x') := (\Omega/\pi)\,\mathrm{sinc}[(x-x')\Omega]$.

Since the reproducing kernel is simply a continuous representation of the evaluation map, one can consider restricting the range of the reproducing kernel to one of the sampling lattices $\{x_n^{(\alpha)}\}_{n\in\mathbb{Z}}$. One can then think of the reproducing kernel and reconstruction kernel as inverses of one another, where the reproducing kernel maps from the continuous to the discrete representation of a bandlimited function,

$$K : \psi(x) \mapsto \psi(x_n^{(\alpha)}) = \int_{\mathbb{R}} dx\ K(x_n^{(\alpha)},x)\psi(x) \tag{3.22}$$

and the reconstruction kernel maps from the discrete to the continuous representation,

$$G : \psi(x_n^{(\alpha)}) \mapsto \psi(x) = \sum_{n\in\mathbb{Z}} G(x,x_n^{(\alpha)})\psi(x_n^{(\alpha)}) \tag{3.23}$$

Here we see explicitly that the sampling theorem provides an equivalence between continuous and discrete representations of information.

3.3 Generalised Sampling

Above we saw that a function bandlimited by Ω can be completely recovered solely from knowledge of the values it takes on any one of a family of lattices $\{x_n^{(\alpha)} = (2\pi n - \alpha)/(2\Omega)\}_{n\in\mathbb{Z}}$ (with $\alpha \in [0, 2\pi)$). The fact that the set of eigenvectors of any one of the self-adjoint extensions provides a basis for $B(\Omega)$ means that we can choose any one of the lattices parametrised by α on which to represent a bandlimited function. However, it turns out the set of lattices from which one can reconstruct a function in $B(\Omega)$ is much more general. There is a condition that the density of points on a lattice must obey in order to be a sampling lattice. Before presenting this condition, we must first define a certain notion of density.

Definition 10 (Beurling density [40, 15]) *The* **Beurling density** *of a set of real points* $\{x_n\}_n$ *is*

$$D(x_n) := \lim_{R\to\infty} \frac{|\{x_n : x_n \in [-R, R]\}|}{2R} \tag{3.24}$$

where $|S|$ *denotes the cardinality of the set* S*.*

The following theorem is built on the works of Nyquist [65] and Beurling [15], but culminated in the work of Landau [56]:

Theorem 7 (Minimum sampling density [56]) *The set* $\{x_n\}_{x\in\mathbb{Z}}$ *is a sampling lattice for* $B(\Omega)$ *iff*

$$D(x_n) \geq \frac{\Omega}{\pi} \tag{3.25}$$

Roughly, this theorem states that for a lattice to be a sampling lattice for $B(\Omega)$, the average density of points (more precisely, the density D) must be at least Ω/π. The minimum sampling rate Ω/π will be referred to as the **Nyquist density or rate**, and the maximum spacing π/Ω as the **Nyquist spacing**. Notice that the lattices considered in the previous section have a uniform spacing equal to the Nyquist spacing. However, the condition $D(x_n) \geq \Omega/\pi$ is unfortunately not sufficient for **stable sampling**, which is when a sampling lattice $\{x_n\}_n$ also satisfies the condition $\|\psi(x)\|_{L_2} \leq C\|\psi(x_n)\|_{\ell_2}$ for $|\psi\rangle \in B(\Omega)$ and some constant C [40].

Although the above theorem provides a condition for which sampling and reconstruction of bandlimited functions is possible, it is generally very difficult to derive a reconstruction formula for a given set of samples. Efforts to understand generalisations of Shannon's

reconstruction formula and sampling theorem has produced a rich field of research [40, 41, 47, 79, 7, 89, 18], including generalising the functional analytic approach to sampling theory as presented above [36, 37, 38]. In the functional analytic approach to sampling theory [36, 37, 38], the non-Nyquist spaced sample points are associated with eigenvectors in the function space from different self-adjoint extensions of the position operator. Although these vectors are not orthogonal, they can still be used to span the function space (thus reconstruction is still possible). Non-Nyquist sampling is useful for reconstructing functions which are localised to a particular region or which have local high-frequency oscillations (e.g. superoscillations [50, 2, 9, 25, 90]). This allows one to sample at a high rate only in these localised regions and at a lower rate in other regions, thus introducing the idea of a spatially-varying Nyquist rate. The amount of sample data necessary to represent these functions is reduced in comparison with the case of Nyquist sampling where one requires a relatively high uniform sampling rate. However, the price of non-Nyquist sampling is that the reconstruction formula becomes more sensitive to noise in the samples.

For our purposes below we shall not require any reconstruction formula for a non-uniform sampling lattice, only the fact that the Beurling density of the sampling lattice must exceed the Nyquist rate.

Chapter 4

Quantum Field Theory

Chapter 4 will present an overview of the concepts of quantum field theory that will be required in later chapters. We will restrict our attention only to free scalar quantum fields on nondynamical classical spacetimes. First we will briefly mention how fields are related to the principle of locality, which motivates the use of fields to model physical systems. Then we shall review classical scalar field theory before proceeding to canonical quantisation as well as path integral quantisation. These last three sections will largely be based on [48, 62, 68] and to some extent on [30].

4.1 Why fields?

The glib answer to why one should describe nature using fields is that it works. For example, the quantum theory of fields has been very successful in developing the standard model of particle physics, the theory of inflation in cosmology, as well as in condensed matter physics. A more complete answer is that fields are used to implement the principle of locality [34, 84, 68]. The field concept was introduced to physics in order to provide an understanding of the electromagnetic and gravitational forces in terms of mechanisms. The electric and gravitational forces between two static and electrically and/or massively charged particles were first described by the Coulomb force law and Newtonian gravitational force law (respectively):

$$\vec{F}_e = \frac{q_1 q_2}{4\pi\epsilon_0}\frac{\vec{r}_1 - \vec{r}_2}{|\vec{r}_1 - \vec{r}_2|^3}, \quad \text{and} \quad \vec{F}_m = -Gm_1m_2\frac{\vec{r}_1 - \vec{r}_2}{|\vec{r}_1 - \vec{r}_2|^3} \tag{4.1}$$

Instead of specifying the charge of the first particle a-priori, as a technical tool one could describe the influence of particle 2 on a test particle at position $\vec{r}_1$ by introducing the electric field $\vec{E}(\vec{r}_1) := \vec{F}_e/q_1$ and the gravitational field $\vec{g}(\vec{r}_1) := \vec{F}_m/m_1$.

The idea behind the principle of locality is to take these fields as fundamental, so that movements of the electric or massive charges cause local perturbations in the electric or gravitational fields which propagate to influence other electric or massive charges at distant points [34]. Explicitly, the propagation of these influences are described by Maxwell's equations for electromagnetism,

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right) A^\mu(t, \vec{x}) = \mu_0 j^\mu(t, \vec{x}) \tag{4.2}$$

(in flat spacetime and in the Lorenz gauge, $\partial^\nu A_\nu = 0$), and originally Poisson's equation for gravitation,

$$\nabla^2 \Phi(t, \vec{x}) = 4\pi G \rho_m(t, \vec{x}) \tag{4.3}$$

which was later corrected in the form of Einstein's field equations,

$$R_{\mu\nu} - \tfrac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{4.4}$$

Perturbations in the metric around flat space, $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, propagate (up to first order in $h_{\mu\nu}$) according to

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right) \overline{h}_{\mu\nu}(t, \vec{x}) = 16\pi G T_{\mu\nu}(t, \vec{x}) \tag{4.5}$$

(in the analog of the Lorenz gauge, $\partial^\nu \overline{h}_{\mu\nu} = 0$), where $\overline{h}_{\mu\nu} := h_{\mu\nu} - \frac{1}{2} h \eta_{\mu\nu}$ [81]. In both equations (4.2) and (4.5), we see that influences from the charge and matter distributions $j^\mu(t, \vec{x})$ and $T_{\mu\nu}(t, \vec{x})$ propagate via waves travelling at the speed of light $c = 1$. In this way, the existence of the electromagnetic and gravitational fields provide a local mechanism for the electric and gravitational forces. That is, an object can only affect other objects locally, and its influence can propagate no faster than the speed of light c (in the background, or unperturbed, flat spacetime).

Though the idea of local mechanisms provides a satisfying physical picture, the true importance of the principle of locality is due to its connection to the principle of causality. This link came to light with special relativity. Consider two events, A and B, which occur at the same time (so that $\Delta t := t_B - t_A = 0$) and at a distance $\Delta x := x_B - x_A > 0$ from one another in the coordinate frame (t, x) (let us simply consider $1+1$ dimensions). Then consider an observer moving with a velocity $v > 0$, so that this observer sees the temporal

and spatial differences between the events as

$$\Delta t' := t'_B - t'_A = \gamma(0 - v\Delta x) < 0 \tag{4.6}$$
$$\Delta x' := x'_B - x'_A = \gamma(\Delta x - 0) \tag{4.7}$$

Therefore, we see that if event A could influence event B instantaneously in the reference frame (t, x) (and so A is a cause with an effect B), then there are observers who would see the event A occur after the event B. This would imply that events could occur before their causes, and so causality is violated. This situation is generally considered to be unphysical, and is remedied by the requirement that influences cannot propagate faster than the speed of light. Therefore, it is natural to choose fields to describe physical systems, with equations of motion such as (4.2) and (4.5).

Naively it seems that it should be unphysical for any system to contain a degree of freedom which occupies a finite volume. Intuitively, fields describe physical systems with a degree of freedom at every point in space, but each degree of freedom individually does not describe a finite volume of space. (Note that one could also consider relativistic point particles which have a degree of freedom at a single point in space.) If a degree of freedom were to describe a finite volume, a change in the state of this degree of freedom would instantaneously have an influence within the entire volume, thus violating causality. Of course determining whether fundamental physics can be described with degrees of freedom occupying finite volumes is still an open problem, and as discussed in the introduction, preliminary studies in quantum gravity tend to indicate that there should be some form of nonlocality.

Although in classical physics one could also describe relativistic point particles without difficulty, this seems not to be the case in quantum theory. Consider a free quantum particle with an initial wavefunction $|\psi\rangle = |x_0\rangle$. Suppose the particle has a relativistic energy expression $H = \sqrt{p^2 + m^2}$. Then consider the matrix elements of the evolution

$$U(x, t; x_0, 0) = \langle x | e^{-it\sqrt{p^2+m^2}} | x_0 \rangle \tag{4.8}$$

It is relatively straightforward to show that well outside the light-cone $(x - x_0)^2 \gg t^2$, the amplitude corresponding to a transition from x_0 to x in time t is nonzero:

$$U(x, t; x_0, 0) \sim e^{-m\sqrt{(x-x_0)^2-t^2}} \tag{4.9}$$

Therefore we get a violation of causality [68]. We will see below that quantum fields do not suffer from this problem, thus they can be used to model propagation of information without violating causality, whereas relativistic point particles cannot.

In quantum theory, one typically implements locality by ensuring that observables, $\{O_A\}$ and $\{O_B\}$, located at spacelike-separated spacetime points A and B (respectively), commute with one another [34]:

$$[O_A, O_B] = 0 \tag{4.10}$$

Consider a state ρ and two positive-operator valued measures (POVMs) $\{E_\mu\}_\mu$ at A and $\{F_\nu\}$ at B. Suppose each of these POVM elements have respective Kraus representation elements (see, e.g., [64]) $\{P_{\mu m}\}_m$ and $\{Q_{\nu n}\}_n$ (where each $P_{\mu m}$ and $Q_{\nu n}$ are in the respective algebras generated by the observables $\{O_A\}$ and $\{O_B\}$, so that $[P_{\mu m}, Q_{\nu n}] = 0$). Now suppose B is at a later time than A, (but A and B are still spacelike-separated), and that measurements are performed at A and B using the above POVMs. Then the probability of a particular outcome at B, without knowledge of the outcome of the measurement at A, is given by [67]

$$\begin{aligned} p_\nu &= \sum_\mu \mathrm{tr}_A \left(\sum_{m,n} Q_{\nu n} P_{\mu m} \rho P_{\mu m}^\dagger Q_{\nu n}^\dagger \right) \\ &= \sum_\mu \mathrm{tr}_A \left(\sum_{m,n} P_{\mu m} Q_{\nu n} \rho Q_{\nu n}^\dagger P_{\mu m}^\dagger \right) \\ &= \mathrm{tr}_A \left(\sum_{\mu,m} P_{\mu m}^\dagger P_{\mu m} \sum_n Q_{\nu n} \rho Q_{\nu n}^\dagger \right) \\ &= \mathrm{tr}_A \left(\sum_n Q_{\nu n} \rho Q_{\nu n}^\dagger \right) \end{aligned} \tag{4.11}$$

where we have used the fact that $\{E_\mu\}_\mu$ is a POVM, so that $\sum_\mu E_\mu = \sum_{\mu,m} P_{\mu m}^\dagger P_{\mu m} = \mathbb{1}$. The result of the above equation is the same as one would obtain if no measurement at A was performed. Thus, we see explicitly that measurements performed at spacelike-separated spacetime points cannot affect one another [67]. Therefore, in quantum field theory, forcing observables at spacelike separated spacetime points to commute with each other will ensure causality is not violated.

4.2 Classical Scalar Field Theory

4.2.1 Equations of Motion

A model for a physical system is often provided by the statement of an action, as it provides a unifying framework for building physical models. Here we will consider classical, free, scalar, real fields on a classical background (nondynamical) spacetime with metric g. These fields will then be quantised in Sections 4.3 and 4.4. We will choose a metric signature where timelike vectors have positive norm and spacelike vectors have negative norm (this is a common choice in particle physics so that energy is positive). The Klein-Gordon action for real scalar fields in $n+1$ dimensions (where the 0^{th} coordinate is temporal) is [62, 68]

$$\begin{aligned} I[\phi] &= \int_{\mathcal{M}} d^{n+1}x\, \mathcal{L}(\phi(x), \partial_\mu \phi(x)) \\ &= \frac{1}{2}\int_{\mathcal{M}} d^{n+1}x \sqrt{-g(x)}\, \left(g^{\alpha\beta}(x)\partial_\alpha\phi(x)\partial_\beta\phi(x) - m^2\phi^2(x)\right) \end{aligned} \tag{4.12}$$

where $g^{\alpha\beta}(x)$ denotes the components of the inverse of g at x in the x-coordinate basis. The Euler-Lagrange equations provide the equations of motion for the system, in this case:

$$\frac{1}{\sqrt{-g(x)}}\partial_\alpha\left(\sqrt{-g(x)}g^{\alpha\beta}(x)\partial_\beta\phi(x)\right) + m^2\phi(x) = 0 \tag{4.13}$$

In Minkowski spacetime $g_{\mu\nu}(x) = \eta_{\mu\nu}$ this becomes

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\phi(t,\vec{x}) + m^2\phi(t,\vec{x}) = 0 \tag{4.14}$$

The Klein-Gordon field $\phi(x) = \phi(t,\vec{x})$ can be intuitively thought of as representing a system composed of degrees of freedom located at every point in space which evolve according to the above equations of motion. Notice that in flat space the equations of motion have a similar form to (4.2) and (4.5), except for the mass term, which modifies the speed of wave propagation according to the dispersion relation $\omega_{\vec{k}}^2 = \vec{k}^2 + m^2$.

The canonical quantisation procedure is built upon the Hamiltonian formalism, thus we shall proceed to derive the canonical momenta, Hamiltonian, and Poisson brackets for the above system. The canonical momentum associated with the degree of freedom $\phi(x)$ is given by

$$\pi(x) = \frac{\partial\mathcal{L}}{\partial(\partial_0\phi(x))} = \sqrt{-g}g^{0\alpha}\partial_\alpha\phi(x) \tag{4.15}$$

The Hamiltonian is given by the Legendre transform of the Lagrangian

$$\begin{aligned} H &:= \int_\Sigma d^n\vec{x}\,(\pi(x)\partial_0\phi(x) - \mathcal{L}) \\ &= \int_\Sigma d^n\vec{x}\sqrt{-g}\left[\frac{1}{g^{00}}\left(\frac{1}{\sqrt{-g}}\pi - g^{0i}\partial_i\phi\right)^2 - g^{ij}\partial_i\phi\partial_j\phi + m^2\phi^2\right] \end{aligned} \tag{4.16}$$

where the integral is over the surface Σ orthogonal to the time coordinate, and latin indices represent spatial directions.

The change of a functional $f[\phi,\pi]$ along the time coordinate $t \equiv x^0$ can be determined from

$$\frac{df[\phi,\pi]}{dt} = \{f[\phi,\pi], H[\phi,\pi]\}_{PB} \tag{4.17}$$

where

$$\{f[\phi,\pi], g[\phi,\pi]\}_{PB} := \int_\Sigma d^n\vec{x}\left(\frac{\delta f[\phi,\pi]}{\delta\phi(x)}\frac{\delta g[\phi,\pi]}{\delta\pi(x)} - \frac{\delta f[\phi,\pi]}{\delta\pi(x)}\frac{\delta g[\phi,\pi]}{\delta\phi(x)}\right) \tag{4.18}$$

In this sense, the Hamiltonian H generates translations along t.

The equations of motion for the degrees of freedom $\phi(x)$ and their canonical momenta $\pi(x)$ are

$$\begin{aligned} \dot{\phi}(x) &= \frac{\delta H[\phi,\pi]}{\delta\pi(x)} \\ &= \frac{1}{g^{00}}\left(\frac{1}{\sqrt{-g}}\pi - g^{0i}\partial_i\phi\right) \end{aligned} \tag{4.19}$$

and

$$\begin{aligned} \dot{\pi}(x) &= -\frac{\delta H[\phi,\pi]}{\delta\phi(x)} \\ &= -\partial_i\left[\sqrt{-g}\frac{g^{0i}}{g^{00}}\left(\frac{1}{\sqrt{-g}}\pi - g^{0j}\partial_j\phi\right)\right] - \partial_i\left[\sqrt{-g}g^{ij}\partial_j\phi\right] - \sqrt{-g}m^2\phi \end{aligned} \tag{4.20}$$

It is straightforward to show that these equations of motion are equivalent to those obtained in the Lagrangian picture.

Note that the action is invariant under any coordinate transformation $x^\alpha \mapsto \tilde{x}^\alpha$. However, if the time coordinate is changed under this transformation, one obtains a different

Hamiltonian, generating translations along $\tilde{t} \equiv \tilde{x}^0$ instead of $t \equiv x^0$. This is because the new momentum variables are now

$$\tilde{\pi}(\tilde{x}) = \frac{\partial \mathcal{L}}{\partial(\tilde{\partial}_0 \phi(\tilde{x}))} = \sqrt{-\tilde{g}}\tilde{g}^{0\alpha}\phi(\tilde{x}) \tag{4.21}$$

Therefore the Hamiltonian is determined from

$$\tilde{H} := \int_{\tilde{\Sigma}} d^n \vec{\tilde{x}} \left(\tilde{\pi}\tilde{\partial}_0 \phi - \mathcal{L} \right) \tag{4.22}$$

The equations of motion for $\phi, \tilde{\pi}$ and the Poisson bracket structure follow from a similar procedure as above.

4.2.2 A note on Poisson Brackets

The functionals one typically considers, such as those above, are functionals of only ϕ and π. Therefore, using the easily verified linearity, antisymmetry, and derivation properties of the Poisson bracket,

- Linearity: $\{f + g, h\} = \{f, h\} + \{g, h\}$
- Antisymmetry: $\{f, g\} = -\{g, f\}$
- Derivation: $\{fg, h\} = f\{g, h\} + \{f, h\}g$

one can determine the Poisson bracket between sufficiently well-behaved functionals without explicitly performing the functional derivatives, but solely from knowing $\{\phi(x), \phi(x')\}$, $\{\phi(x), \pi(x')\}$, and $\{\pi(x), \pi(x')\}$. In the previous section, the equal-time brackets between fundamental degrees of freedom are given by:

$$\{\phi(t, \vec{x}), \phi(t, \vec{x'})\} = 0 \tag{4.23}$$
$$\{\phi(t, \vec{x}), \pi(t, \vec{x'})\} = \delta^{(n)}(\vec{x} - \vec{x'}) \tag{4.24}$$
$$\{\pi(t, \vec{x}), \pi(t, \vec{x'})\} = 0 \tag{4.25}$$

(The nonequal-time brackets are determined from the equal-time brackets and solutions to the equations of motion for ϕ and π.) Thus one could consider, for example, the Poisson

bracket between the functionals

$$f[\phi,\pi] = \int_\Sigma d^n\vec{x}\,\phi^2(t,\vec{x}) \tag{4.26}$$

$$g[\phi,\pi] = \int_\Sigma d^n\vec{x}\,\pi(t,\vec{x}) \tag{4.27}$$

which is:

$$\begin{aligned}
\{f,g\} &= \{\int_\Sigma d^n\vec{x}\,\phi^2(t,\vec{x}), \int_\Sigma d^n\vec{x'}\,\pi(t,\vec{x'})\} \\
&= \int_{\Sigma^2} d^n\vec{x}d^n\vec{x'}\,\{\phi^2(t,\vec{x}),\pi(t,\vec{x'})\} \\
&= \int_{\Sigma^2} d^n\vec{x}d^n\vec{x'}\,\Big(\phi(t,\vec{x})\{\phi(t,\vec{x}),\pi(t,\vec{x'})\} + \{\phi(t,\vec{x}),\pi(t,\vec{x'})\}\phi(t,\vec{x})\Big) \\
&= \int_{\Sigma^2} d^n\vec{x}d^n\vec{x'}\,\Big(\phi(t,\vec{x})\delta^{(n)}(\vec{x}-\vec{x'}) + \delta^{(n)}(\vec{x}-\vec{x'})\phi(t,\vec{x})\Big) \\
&= \int_\Sigma d^n\vec{x}\,2\phi(t,\vec{x})
\end{aligned} \tag{4.28}$$

The fact that the bracket between any two functionals in phase space is determined from the bracket between the degrees of freedom is important because in canonical quantisation the brackets between the basic phase space variables are often thought of as fundamental. Then the dynamics of any function of phase space can be determined from these basic brackets.

However, the fundamental brackets are not always the canonical ones (4.23). If the dynamics of the system do not occur in the entire phase space, but on some smaller space defined by a set of constraints on the phase space variables, then brackets between phase space variables can differ from the canonical brackets. The general procedure for deriving the brackets from a set of constraints was outlined by Dirac in [22]. We will not reproduce Dirac's method for dealing with constraints, but in Chapter 6 we will encounter a situation where the Poisson brackets are no longer canonical due to constraints imposed on the phase space variables. Dirac's procedure for determining the correct brackets will not be needed in Chapter 6, but we simply note that the fact that the brackets are not canonical does not conflict with the ideas in the classical field theory presented in this section nor the canonical quantisation in the next section.

4.3 Canonical Quantisation

In this section we will simply consider Minkowski spacetime, since the canonical quantisation on other backgrounds will not be needed. The quantisation program as outlined by Dirac [21] dictates that quantisation of a system is achieved by promoting the phase space coordinates representing the classical degrees of freedom to self-adjoint operators (i.e., observables) whose action is represented on some Hilbert space. Since the Klein-Gordon field intuitively represents a degree of freedom at every point in space, this would suggest the following promotion:

$$\phi(\vec{x}) \mapsto \hat{\phi}(\vec{x}) \tag{4.29}$$
$$\pi(\vec{x}) \mapsto \hat{\pi}(\vec{x}). \tag{4.30}$$

Technically, the quantity $\hat{\phi}(\vec{x})$ is an operator-valued distribution. In order to obtain an operator, one must perform some regularisation scheme, such as smearing the operator by integrating it over some smooth function [34].

After replacing the phase space variables with operators, one can represent the Poisson bracket with commutators:

$$\{\cdot,\cdot\}_{PB} \mapsto -i[\cdot,\cdot]. \tag{4.31}$$

Since the commutator has the same properties as the Poisson bracket as outlined in Subsection 4.2.2, the quantum theory inherits much of the kinematical structure of the classical theory outlined in Subsection 4.2.2. In particular, we mentioned that once the Poisson brackets between the fundamental degrees of freedom are specified, one can then work out the Poisson bracket between any two functionals which depend only on these degrees of freedom. This is also true in the quantum theory, where the Poisson brackets between the fundamental degrees of freedom become the equal-time canonical commutation relations:

$$\left[\phi(\vec{x}), \pi(\vec{x'})\right] = i\delta^{(n)}(\vec{x} - \vec{x'}), \tag{4.32}$$
$$\left[\phi(\vec{x}), \phi(\vec{x'})\right] = 0, \tag{4.33}$$
$$\left[\pi(\vec{x}), \pi(\vec{x'})\right] = 0. \tag{4.34}$$
$$\tag{4.35}$$

The inherited kinematical structure also implies that the equations of motion for the quantum degrees of freedom will take the same form as the classical equations of motion. Therefore, for simplicity, one can first solve the classical equations of motion and promote

the solutions to these equations to self-adjoint operators. Hereafter we will drop the ˆ from the operators. For Minkowski space, the equations of motion can be written as:

$$\left(\frac{\partial^2}{\partial t^2} - \nabla^2\right)\phi(t,\vec{x}) + m^2\phi(t,\vec{x}) = 0, \tag{4.36}$$

$$\pi(t,\vec{x}) = \dot{\phi}(t,\vec{x}). \tag{4.37}$$

The solutions of the equation of motion for ϕ are plane waves

$$e^{\pm i(\omega_k t - \vec{k}\cdot\vec{x})}, \tag{4.38}$$

where $\omega_k := \sqrt{\vec{k}^2 + m^2}$. Then we can write a general solution for ϕ as a linear combination:

$$\phi(t,\vec{x}) = \int \frac{d\vec{k}}{(2\pi)^n} \frac{1}{\sqrt{2\omega_k}} \left(a_{\vec{k}} e^{-i(\omega_k t - \vec{k}\cdot\vec{x})} + a^*_{\vec{k}} e^{i(\omega_k t - \vec{k}\cdot\vec{x})}\right), \tag{4.39}$$

where the $1/\sqrt{2\omega_k}$ is simply included as a normalisation factor. One can then promote the coefficients $a_{\vec{k}}$ and $a^*_{\vec{k}}$ to operators in order to perform the quantisation.

Note that if we impose the commutation relations

$$\left[a_{\vec{k}}, a^\dagger_{\vec{k}'}\right] = \delta^{(n)}(\vec{k} - \vec{k}'), \tag{4.40}$$

$$\left[a_{\vec{k}}, a_{\vec{k}'}\right] = 0, \text{ and} \tag{4.41}$$

$$\left[a^\dagger_{\vec{k}}, a^\dagger_{\vec{k}'}\right] = 0, \tag{4.42}$$

then this will ensure that ϕ and π obey canonical commutation relations. If we substitute the expression for ϕ and the corresponding expression for π into the Klein-Gordon Hamiltonian, we obtain

$$H = \int \frac{d\vec{k}}{(2\pi)^n}\, \omega_k \left(a^\dagger_{\vec{k}} a_{\vec{k}} + \frac{1}{2}\right). \tag{4.43}$$

Here we see explicitly that each mode $\vec{k}$ simply has the structure of a quantum harmonic oscillator. Therefore, for each mode we can construct a Fock basis,

$$|n\rangle_{\vec{k}} := \frac{\left(a^\dagger_{\vec{k}}\right)^n}{\sqrt{n!}} |0\rangle_{\vec{k}}\,, \tag{4.44}$$

where the vacuum state is the vector such that

$$a_{\vec{k}} |0\rangle_{\vec{k}} = 0. \tag{4.45}$$

Note that the operator $a^\dagger_{\vec{k}}$ creates a state with energy $\omega_k = \sqrt{\vec{k}^2 + m^2}$. This is the energy of a particle with mass m propagating with momentum $\vec{k}$. Furthermore, the state $|n\rangle_{\vec{k}}$ is a state with energy $n\omega_k$ (i.e., this is the energy of n particles all with momentum $\vec{k}$). Therefore, the state $a^\dagger_{\vec{k}}|0\rangle_{\vec{k}}$ is typically referred to as a one-particle state of momentum $\vec{k}$, and correspondingly $|n\rangle_{\vec{k}}$ is an n-particle state of momentum $\vec{k}$.

We can construct the global ground state of the system by

$$|0\rangle = \bigotimes_{\vec{k}} |0\rangle_{\vec{k}}\,, \tag{4.46}$$

and we can act upon this state with creation and annihilation operators of any mode to create a wide variety of n-particle states. The total Hilbert space of the system is taken to be the direct sum of all n-particle states.

In Minkowski spacetime, the fields take a particularly simple form which allows for a relatively simple quantisation procedure. On other background spacetimes this is typically very difficult unless the metric takes a very simple form (such as FLRW spacetime). For general spacetimes where canonical quantisation is not feasible, it is often still possible to perform certain calculations using path integral techniques.

4.4 Path Integral Quantisation

The path integral quantisation procedure provides an alternative to the canonical quantisation procedure. Path integrals deal with matrix elements of density matrices and evolution operators rather than with the operators themselves. Path integrals have proven useful in providing very powerful techniques for calculating certain quantities in quantum field theory, such as scattering matrix elements for particle collision experiments. Furthermore, path integral techniques do not always require one to solve the equations of motion in order to perform calculations, as opposed to the canonical quantisation procedure. This is particularly useful when studying quantum fields on general spacetimes where solutions to the equations of motion are not known. Though powerful, one disadvantage of path integral techniques is that for some calculations they do not provide much physical insight.

We will begin by demonstrating a derivation of the path integral for a single degree of freedom, then we shall examine the path integral for real Klein-Gordon fields.

4.4.1 Path integral for single degree of freedom

Here we will simply provide a brief outline for deriving the path integral for a single degree of freedom. Many of the details overlooked here can be found in Refs.[68, 62].

We begin with Schrödinger's equation,

$$i\partial_t \left|\psi(t)\right\rangle = H(t) \left|\psi(t)\right\rangle , \tag{4.47}$$

which has a formal solution

$$\left|\psi(t)\right\rangle = \mathcal{T} e^{-i\int_0^t dt' H(t')} \left|\psi(0)\right\rangle , \tag{4.48}$$

where $\mathcal{T}$ is the time-ordering operator. We can represent this time evolution in a basis $|q\rangle$ as:

$$\psi(q,t) = \int dq\, K(q,t;q',0)\psi(q',0), \tag{4.49}$$

where $K(q,t;q',0) := \langle q| \mathcal{T} e^{-i\int_0^t dt' H(t')} |q'\rangle$ is the propagator for the system.

Evidently, one would like to be able to calculate the matrix elements of the propagator. We will perform this calculation for the Hamiltonian

$$H = \frac{1}{2m}p^2 + V(q). \tag{4.50}$$

Note that the propagators have a convolution property:

$$K(q,t;q',t') = \int dq''\, K(q,t;q'',t'')K(q'',t'';q',t') \tag{4.51}$$

for $t' < t'' < t$. Consider discretising time into small steps Δt, then for the time interval $[t_i, t_{i+1}]$,

$$\begin{aligned} K(q_{i+1},t_{i+1};q_i,t_i) &= \langle q_{i+1}| e^{-i\Delta t\, \hat{H}} |q_i\rangle \\ &\approx \langle q_{i+1}| \left(\mathbb{1} - i\Delta t \left(\frac{1}{2m}\hat{p}^2 + \hat{V}(q)\right)\right) |q_i\rangle \\ &= \int \frac{dp}{2\pi} \left(1 - i\Delta t\, H(q,p)\right) e^{ip(q_{i+1}-q_i)} \\ &\approx \int \frac{dp}{2\pi}\, e^{ip(q_{i+1}-q_i) - i\Delta t H(q,p)}. \end{aligned} \tag{4.52}$$

Now, using the composition property of the propagators, and taking the time-ordering operator $\mathcal{T}$ into account, it is possible to show that

$$K(q,t;q',t') = \int \left(\prod_i \frac{dq_i dp_i}{2\pi} \right) \, e^{i\Delta t \sum_i (p_i(q_{i+1}-q_i)/(\Delta t) - H(q_i,p_i))}. \tag{4.53}$$

Since H is quadratic in p, then one can explicitly perform the Gaussian integrals and absorb the corresponding factor into the measure. Then, taking a continuum limit, one can obtain the path integral,

$$K(q,t;q',t') = \int_{q''(t')=q'}^{q''(t)=q} \mathcal{D}q'' \, e^{iI[q'']}, \tag{4.54}$$

where $I[q]$ is the action functional. Recall that the discretised version of the integral was an integration over all possible values of the degree of freedom q at each time step. Therefore, the path integral can be thought of as the integral over all possible paths between q' at t' and q at t. The matrix elements of K are parametrised by the boundary conditions given at t' and t.

4.4.2 Path integral for Klein-Gordon field

In analogy with the procedure for the single degree of freedom, it is also possible to write a path integral to determine the matrix elements for the evolution operator of the Klein-Gordon field:

$$\langle \phi | \, e^{-i(t-t')H} \, | \phi' \rangle = \int_{\phi''(t',\vec{x})=\phi'(\vec{x})}^{\phi''(t,\vec{x})=\phi(\vec{x})} \mathcal{D}\phi'' \, e^{iI[\phi'']}, \tag{4.55}$$

where $I[\phi]$ is the Klein-Gordon action. Notice that the matrix elements are also parametrised by boundary conditions given at t' and t, except now the boundary conditions are not just real numbers but field configurations on the spacelike hypersurface t' and t (respectively).

One convenient use of the path integral is to find the matrix elements for the ground state. First, we perform a Wick rotation $\tau = it$. The action is transformed into a Euclidean action, $iI[\phi] \overset{\tau=it}{\longmapsto} -I_E[\phi]$, and the path integral becomes

$$\langle \phi | \, e^{-(\tau-\tau')H} \, | \phi' \rangle = \int_{\phi''(\tau',\vec{x})=\phi'(\vec{x})}^{\phi''(\tau,\vec{x})=\phi(\vec{x})} \mathcal{D}\phi'' \, e^{-I_E[\phi'']}. \tag{4.56}$$

Note that if we let $\tau' \to -\infty$, then

$$e^{\tau' H} \, | \phi' \rangle = | 0 \rangle \, . \tag{4.57}$$

And if we let $\tau \to 0$, then we find

$$\langle \phi | 0 \rangle = \int^{\phi''(0^-,\vec{x})=\phi(\vec{x})} \mathcal{D}\phi''\, e^{-I_E[\phi'']}. \tag{4.58}$$

Similarly, one can perform $\tau \to +\infty$ and $\tau' \to 0$ to find

$$\langle 0 | \phi' \rangle = \int_{\phi''(0^+,\vec{x})=\phi'(\vec{x})} \mathcal{D}\phi''\, e^{-I_E[\phi'']}. \tag{4.59}$$

Putting these together, we can express the matrix elements of the ground state as:

$$\langle \phi | 0 \rangle \langle 0 | \phi' \rangle = \int_{\phi''(0^+,\vec{x})=\phi'(\vec{x})}^{\phi''(0^-,\vec{x})=\phi(\vec{x})} \mathcal{D}\phi''\, e^{-I_E[\phi'']}. \tag{4.60}$$

Chapter 5

Entanglement Entropy

The purpose of Chapter 5 is to describe methods for calculating the entanglement entropy of quantum fields. One method is based on the Hamiltonian formalism (Section 5.3), and a second on the path integral formalism (Section 5.4). Although only the Hamiltonian method will be used in Chapter 6, the path integral method is useful for illustrating the general structure and scaling behaviour of the entanglement entropy.

5.1 Information, Entropy, and Entanglement

Suppose our collaborator conducts N coin flips and we want to know what the outcomes were. How much information must we obtain from our collaborator in order to gain complete knowledge of the set of outcomes? Or, more concretely, how many bits must we obtain? Of course, assuming the coin flips are done identically and independently, we need N bits. For instance, we could determine the outcome of the N coin flips by simply asking for each individual coin flip whether the outcome was heads. Alternatively, we could determine the outcome of the N coin flips by asking whether the number of heads occurring in the N coin flips was even, then ask whether the number of heads occurring in the first $N-1$ coin flips was even, ..., whether the number of heads occurring in the first 2 coin flips was even, and finally whether the first coin flip was heads. One can easily become convinced for this simple example that no matter how the questioning is arranged, we must obtain the answer to at least N binary questions in order to guarantee that we determine the outcome of the N coin flips. (Note that it is not necessary to ask N binary questions, since if we were to simply guess the outcome of the experiment, then for some instances of the experiment we will get the correct answer on the first guess. The above statement

is in regards to the fewest number of binary questions that need to be answered in order to *guarantee* that we will know the outcome for any instantiation of the experiment; it is possible that there are other algorithms which perform better some of the time (i.e., for some instances of the experiment).) One must have N binary questions answered because the number of possible outcomes is 2^N, and every time we ask a binary question we eliminate half of the state space as possible outcomes (provided there is no redundancy in the questions we ask). Thus, after asking N binary questions, we have divided the space of possible states by 2 each time to leave $2^N/2^N = 1$ possible outcome which is consistent with the answers we have been given. This illustrates the idea that there are N bits of **uncertainty** about the outcome or N bits of **information** that we can obtain by gaining knowledge of the outcome.

Suppose instead of N coin flips, the experiment was N rolls of a 4-sided die. One could then ask N questions with four possible answers to determine the outcome of the N rolls. However, if the experimenter insisted on us asking only binary questions, we could ask two binary questions to determine the state of each roll (for example, if the outcomes are labelled $\{1, 2, 3, 4\}$, we could ask whether it was even and whether it was less than 3). Therefore, we would need to obtain $2N$ bits of information in order to determine the outcome. The generalisation of these kinds of experiments is that if there are W possible outcomes (i.e., the state space contains W elements) each equally probable, then the number of bits one requires or the number of binary questions one needs to ask in order to determine the state is

$$H = \log_2 W \tag{5.1}$$

One would like to generalise beyond these simple cases to deal with situations where the outcomes are not equally probable. In his foundational work on information theory, Shannon [71] argued that for a discrete set of outcomes Ω with each outcome $i \in \Omega$ occurring with probability p_i, any function $H(\{p_i\}_{i\in\Omega})$ representing the uncertainty about the outcome should satisfy the following properties:

1. H should be continuous in the p_i.

2. If the outcomes are equally probable, $p_i = 1/|\Omega|$, then H should be a monotonically increasing function of $|\Omega|$ (the number of possible outcomes). This is because with a larger number of equally probable outcomes there is more uncertainty.

3. If a question can be broken down into two successive questions, then the information obtained from an answer to this question should be the weighted sum of the infor-

mation retrieved from the two successive questions. This is a generalisation of the example above where we broke down a question with four possible outcomes into two binary questions. The example Shannon gives [71] is for three outcomes with probabilities $p_1 = 1/2, p_2 = 1/3, p_3 = 1/6$. This can be broken down into two questions: i) whether outcome 1 occurred, with $p_1 = 1/2, p_{2\cup 3} = 1/2$, and ii) whether outcome 2 occurred, with $p_{2\cap\neg 1} = 2/3, p_{3\cap\neg 1} = 1/3$ (where $\neg$ denotes logical negation, so that in the σ-algebra of Ω, $\neg 1 \equiv \{1\}^c := \Omega \setminus \{1\}$). Then for this example, the condition Shannon wishes to assert is

$$H\left(\tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{6}\right) = H\left(\tfrac{1}{2}, \tfrac{1}{2}\right) + \tfrac{1}{2} H\left(\tfrac{2}{3}, \tfrac{1}{3}\right) \tag{5.2}$$

where the coefficient of $\frac{1}{2}$ is included because one must only ask the second question half of the time, since $p_{2\cup 3} = 1/2$.

Shannon [71] then shows that these conditions uniquely determine H up to a constant:

$$H = -K \sum_i p_i \log p_i \tag{5.3}$$

The constant K simply determines the choice of units for H. For information measured in bits, the natural choice is $K = 1$ and a base 2 logarithm. Therefore, the information obtained from determining the outcome of a probabilistic event with a discrete state space Ω and probabilities $\{p_i\}_{i\in\Omega}$, called the **entropy**, is

$$H(\{p_i\}_{i\in\Omega}) := -\sum_{i\in\Omega} p_i \log p_i \tag{5.4}$$

A natural generalisation of the entropy to a continuous state space with probability density function $p(x)$ is

$$H(x) := -\int dx\, p(x) \log p(x) \tag{5.5}$$

This logarithm is typically taken to be the natural logarithm. The generalisation of the continuous entropy to multiple random variables is simply obtained by replacing $p(x)$ with the joint probability distribution for the multiple variables. For example, consider an n-dimensional Gaussian distribution with zero mean and covariance matrix Σ,

$$p(\vec{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det \Sigma}} e^{-\frac{1}{2}\vec{x}^T \Sigma^{-1} \vec{x}} \tag{5.6}$$

Then the entropy of this distribution is

$$\begin{aligned} H(x) &= -\int_{\mathbb{R}^n} d\vec{x}\, p(\vec{x}) \log p(\vec{x}) \\ &= \int_{\mathbb{R}^n} d\vec{x}\, p(\vec{x}) \left(\tfrac{1}{2}\vec{x}^T\Sigma^{-1}\vec{x} + \tfrac{1}{2}\log[(2\pi)^n \det\Sigma]\right) \\ &= \tfrac{1}{2}n + \tfrac{1}{2}\log[2\pi \det\Sigma] \end{aligned} \tag{5.7}$$

Consider the case $n = 1$ with variance σ^2, then

$$H(x) = \tfrac{1}{2}\log[2\pi e\sigma^2] \tag{5.8}$$

We see that for broad distributions, i.e., for large σ^2, there is more information contained in the distribution. Note however, that it is possible to choose σ^2 so that $2\pi e\sigma^2 < 1$, which causes the entropy to be negative! This does not happen in the discrete case since every $0 \leq p_i \leq 1$, so that $\log p_i < 0$ and equation (5.4) is always positive. Continuous probability density functions, however, are not constrained to take values between 0 and 1. Indeed, for very peaked distributions the values of the probability distribution can become very large near the peak. This is what happens for small σ^2 in this example.

But how can we make sense of negative information? As explained in [71], there is an ambiguity in the definition of the entropy for continuous probability distributions due to a choice of "coordinate system". Consider a change of variables $x^i \mapsto y^j$, then

$$\begin{aligned} H(y) &= -\int \prod_j dy^j\, p(\vec{y}) \log p(\vec{y}) \\ &= -\int \det\left(\frac{\partial y^j}{\partial x^i}\right) \prod_i dx^i\ \det\left(\frac{\partial x^i}{\partial y^j}\right) p(\vec{x}) \log\left[\det\left(\frac{\partial x^i}{\partial y^j}\right) p(\vec{x})\right] \\ &= H(x) - \int \prod_i dx^i\, p(\vec{x}) \log\det\left(\frac{\partial x^i}{\partial y^j}\right) \end{aligned} \tag{5.9}$$

Therefore, we see that the entropy depends on the choice of variables, so one must specify a preferred choice of coordinates based on external criteria in order to determine the entropy unambiguously. The negative information obtained in the above example is simply due to a poor choice of coordinates. We could always find some set of coordinates such that the second term in the above equation shifts the entropy to take on positive values for a given range of σ^2 that we are interested in.

Consider a linear change of variables given by $\vec{y} = A\vec{x}$ [71] so that $\det\left(\frac{\partial y^j}{\partial x^i}\right) = \det A$, then

$$H(y) = H(x) + \log\det A \tag{5.10}$$

Note that if $\det A < 1$, then since $\prod_j dy^j = \det A \prod_i dx^i$, we see that the measure $\prod_j dy^j$ in the y coordinates gives a smaller value to the same region of the state space Ω than does $\prod_i dx^i$ in the x coordinates. Thus the change of coordinates causes distributions to appear more peaked in the y coordinate system, which is reflected in the fact that the entropy decreases: $H(y) < H(x)$.

Although central to information theory, the concept of entropy had its origin in statistical mechanics. However, entropy in statistical mechanics is used in the same manner and is interpreted in the same way as in information theory. If we are given some probability distribution ρ over microstates in phase space for some set of degrees of freedom $\{q_i, p_i\}_i$, then the entropy is simply the information needed to determine the microstate:

$$\begin{aligned} S(\rho) &:= -\int \prod_i \frac{dq^i dp^i}{2\pi}\, \rho(\vec{q},\vec{p}) \log\rho(\vec{q},\vec{p}) \\ &= -\langle \log\rho \rangle \end{aligned} \tag{5.11}$$

For example, consider a Boltzmann distribution,

$$\rho(\vec{q},\vec{p}) = \frac{1}{Z} e^{-\beta H(\vec{q},\vec{p})} \tag{5.12}$$

then the entropy is

$$\begin{aligned} S(\rho) &= -\int \frac{d\vec{q}d\vec{p}}{(2\pi)^N}\, \rho(\vec{q},\vec{p}) \log\rho(\vec{q},\vec{p}) \\ &= \beta\langle H\rangle + \log Z \end{aligned} \tag{5.13}$$

However, since

$$\langle H\rangle = -\frac{\partial \log Z}{\partial\beta} \tag{5.14}$$

then the entropy can be expressed compactly as

$$S = \left(1 - \beta\frac{\partial}{\partial\beta}\right)\log Z \tag{5.15}$$

If the Hamiltonian for the distribution is quadratic in the phase space variables,

$$H(\vec{q},\vec{p}) = \tfrac{1}{2}\vec{p}^T M\vec{p} + \tfrac{1}{2}\vec{q}^T K\vec{q} \tag{5.16}$$

then the partition function is,

$$\begin{aligned} Z &= \int \frac{d\vec{q}d\vec{p}}{(2\pi)^N} e^{-\beta H(\vec{q},\vec{p})} \\ &= \det(\beta M)^{-1/2} \det(\beta K)^{-1/2} \end{aligned} \tag{5.17}$$

and the entropy is

$$\begin{aligned} S &= \left(1 - \beta\frac{\partial}{\partial\beta}\right) \log Z \\ &= N - \tfrac{1}{2} \log\left[\det(\beta M) \det(\beta K)\right] \end{aligned} \tag{5.18}$$

where N is the number of degrees of freedom. Note that S is roughly an extensive quantity since the first term is simply N, however the second term may introduce nonlinearities in the scaling behaviour due to coupling between the degrees of freedom.

Since the above probability distribution is for continuous random variables, there is an ambiguity in the definition of the entropy due to an ambiguity of the choice of coordinates in phase space. However, one cannot make arbitrary changes of coordinates in phase space, but only those which preserve its symplectic structure. This is enforced by requiring that any change of coordinates must preserve the Poisson bracket. I.e., for a transformation $(q_i, p_j) \mapsto (q'_i, p'_j)$ we require

$$\{\cdot,\cdot\} = \partial_{q_i} \otimes \partial_{p_i} - \partial_{p_i} \otimes \partial_{q_i} \mapsto \{\cdot,\cdot\}' = \partial'_{q_i} \otimes \partial'_{p_i} - \partial'_{p_i} \otimes \partial'_{q_i} \tag{5.19}$$

Because the Poisson bracket is preserved under a canonical transformation, then its inverse, the symplectic form

$$\omega = dp_i \wedge dq_i \tag{5.20}$$

is also preserved. Note that the Poisson bracket and symplectic form are inverses of one another in the same sense that $g^{\mu\nu}\partial_\mu \otimes \partial_\nu$ is the inverse of a metric $g = g_{\mu\nu} dx^\mu \otimes dx^\nu$ in Riemannian geometry. Since the symplectic form is preserved under canonical transformations, then so is the symplectic volume form,

$$\begin{aligned} \omega^{\wedge N} &:= \omega \wedge \omega \wedge \cdots \wedge \omega \\ &= \bigwedge_{i=1}^{N} dp_i \wedge dq_i \end{aligned} \tag{5.21}$$

which we use to define a measure for phase space:

$$\int \frac{\omega^{\wedge N}}{(2\pi)^N} = \int \prod_{i=1}^{N} \frac{dp_i dq_i}{2\pi} \tag{5.22}$$

The factor of $1/(2\pi)^N$ is to ensure any calculated quantities agree with the quantum mechanical results in the high temperature limit. Therefore, the measure factor ambiguity in the definition of the entropy is resolved if we allow for only canonical transformations of the phase space coordinates.

For example, consider a linear change of phase space coordinates $q_i \mapsto q_i' := A_{ij}q_j$ and $p_i \mapsto p_i' := B_{ij}p_j$. Then,

$$\begin{aligned} \{q_i', p_j'\} &= A_{ik}B_{jl}\{q_k, p_l\} \quad (\text{since } \{\cdot,\cdot\} \text{ is bilinear}) \\ &= A_{ik}B_{jl}\delta_{kl} \\ &= A_{ik}B_{jk} \end{aligned} \tag{5.23}$$

Therefore, in order to have $\{q_i', p_j'\} = \delta_{ij}$, we must have $B^T = A^{-1}$. Notice that the entropy for the quadratic Hamiltonian calculated above is invariant under this transformation, since

$$\begin{aligned} S = N - \tfrac{1}{2}\log\left[\det(\beta M)\det(\beta K)\right] &\mapsto N - \tfrac{1}{2}\log\left[\det(\beta B^{-T}MB^{-1})\det(\beta A^{-T}KA^{-1})\right] \\ &= N - \tfrac{1}{2}\log\left[\det(B)^{-2}\det(\beta M)\det(B)^2\det(\beta K)\right] \\ &= N - \tfrac{1}{2}\log\left[\det(\beta M)\det(\beta K)\right] \\ &= S \end{aligned} \tag{5.24}$$

where we have used the fact that $B^T = A^{-1}$.

In quantum mechanics, one deals with density matrices rather than probability distributions. Correspondingly, one can define a notion of entropy for density matrices, called the **von Neumann entropy**:

$$S(\rho) := -\operatorname{tr}(\rho\log\rho) \tag{5.25}$$

for a density matrix ρ. The operator $\log\rho$ is defined in the sense of the functional calculus of the self-adjoint operator ρ. In the case where ρ is diagonal, thus a classical statistical ensemble, the von Neumann entropy reduces to the classical entropy as defined above.

For example, given a thermal state,

$$\rho = \frac{1}{Z}\sum_{n=0}^{\infty} e^{-\beta\omega(n+1/2)}\left|n\right\rangle\left\langle n\right| = \frac{1}{\langle n\rangle + 1}\sum_{n=0}^{\infty}\left(\frac{\langle n\rangle}{\langle n\rangle + 1}\right)^n \left|n\right\rangle\left\langle n\right|, \tag{5.26}$$

where

$$Z = \frac{e^{-\beta\omega/2}}{1-e^{-\beta\omega}}, \qquad \langle n\rangle = \frac{1}{e^{\beta\omega}-1}, \tag{5.27}$$

the von Neumann entropy is

$$\begin{aligned} S(\rho) &= -\operatorname{tr}(\rho \log \rho) \\ &= \beta\langle H\rangle + \log Z \\ &= \left(1 - \beta\frac{\partial}{\partial\beta}\right)\log Z \\ &= (\langle n\rangle + 1)\log(\langle n\rangle + 1) - \langle n\rangle \log\langle n\rangle. \end{aligned} \tag{5.28}$$

The von Neumann entropy is commonly used as a measure of entanglement in a bipartite system. Given two systems A, B, with respective Hilbert spaces $\mathcal{H}_A$, $\mathcal{H}_B$, and a pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$, one can define the reduced density matrix on $\mathcal{H}_A$ by performing a partial trace on system B:

$$\rho_A := \operatorname{tr}_B(|\psi\rangle\langle\psi|). \tag{5.29}$$

The entanglement entropy is defined as the von Neumann entropy of this reduced density matrix,

$$S(\rho_A) := -\operatorname{tr}(\rho_A \log \rho_A). \tag{5.30}$$

One can equivalently trace out the system A to get an entanglement entropy $S(\rho_B)$, however one can show via Schmidt decomposition that for pure states $S(\rho_A) = S(\rho_B)$. The Schmidt decomposition of a pure state $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ is a unique representation $|\psi\rangle = \sum_i \lambda_i |i\rangle_A |i\rangle_B$ where $\{|i\rangle_A\}_i$ and $\{|i\rangle_B\}$ are orthonormal bases for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. If one of the two systems has a larger dimension than the other, the smaller basis can be appended with a set of zero vectors, since the number of nonzero λ_i's will be at most the dimension of the smaller Hilbert space (thus, the formula makes sense). The Schmidt decomposition for a state

$$|\psi\rangle = \sum_{j,k} a_{jk} |e_j\rangle_A |f_k\rangle_B , \tag{5.31}$$

with $\{|e_j\rangle_A\}$ and $\{|f_k\rangle_B\}$ respective orthonormal bases for $\mathcal{H}_A$ and $\mathcal{H}_B$, can be obtained by performing a singular value decomposition of the matrix A whose elements are a_{jk}. That is, if

$$a_{jk} = \sum_i u_{ji}\lambda_i v_{ik}, \tag{5.32}$$

where u_{ji} and v_{ik} are elements of unitary matrices, then

$$\begin{aligned} |\psi\rangle &= \sum_{j,k} a_{jk} |e_j\rangle_A |f_k\rangle_B \\ &= \sum_i \lambda_i \left(\sum_j u_{ji} |e_j\rangle_A \right) \left(\sum_k v_{ik} |f_k\rangle_B \right) \\ &= \sum_i \lambda_i |i\rangle_A |i\rangle_B , \end{aligned} \tag{5.33}$$

where the orthonormal bases $\{|i\rangle_A\}$ and $\{|i\rangle_B\}$ are defined by the quantities in brackets in the second line. It is then easy to see that

$$\rho_A = \sum_i |\lambda_i|^2 |i\rangle_A \langle i|_A \tag{5.34}$$

and

$$\rho_B = \sum_i |\lambda_i|^2 |i\rangle_B \langle i|_B , \tag{5.35}$$

so that

$$S(\rho_A) = S(\rho_B) = -\sum_i |\lambda_i|^2 \log |\lambda_i|^2. \tag{5.36}$$

The number of nonzero Schmidt coefficients, λ_i, is called the Schmidt rank of the state. The Schmidt rank is 1 if and only if the state is separable [64], i.e., if we can write $|\psi\rangle = |\psi_A\rangle |\psi_B\rangle$, in which case the entanglement entropy is zero.

The simplest example of a nonseparable, or entangled, state is a Bell state such as: $|\psi\rangle = \frac{1}{2}(|00\rangle + |11\rangle) \in \mathcal{H}_A \otimes \mathcal{H}_B$. For this state,

$$\rho_A = \tfrac{1}{2} |0\rangle \langle 0| + \tfrac{1}{2} |1\rangle \langle 1| , \tag{5.37}$$

so that $S(\rho_A) = 1$ (assuming the log in the definition of $S(\rho_A)$ is base 2). Thus we see that the two subsystems share 1 bit of entanglement in this state.

An important property of the von Neumann entropy is that it is preserved under unitary evolution. For consider an evolution, $\rho \mapsto U\rho U^\dagger$, then

$$\begin{aligned} S(U\rho U^\dagger) &= -\operatorname{tr}(U\rho U^\dagger \log[U\rho U^\dagger]) \\ &= -\operatorname{tr}(U^\dagger U \rho \log \rho) \\ &= S(\rho). \end{aligned} \tag{5.38}$$

This demonstrates that the information (as measured by the von Neumann entropy) of a system is preserved under unitary time evolution.

As two systems interact with one another, they tend to become entangled. Thus, for an open system interacting with the environment, as time passes the system becomes entangled with the environment and the entanglement entropy between the system and environment tends to increase. This can be thought of as the system losing information to the environment as they interact. More precisely, it is the amount of information contained in the environment which needs to be acquired in order to know the state of the system with no uncertainty. The fact that the entanglement entropy tends to increase implies that the state of the system with the environment traced out will undergo non-unitary evolution (otherwise the entanglement entropy would not change).

5.2 Entropy of Black Holes

The previous section was a simple overview of the ideas of information, entropy, and entanglement as they pertain to information theory, statistical mechanics, and quantum information theory in general. In this section we will an overview of one of the major reasons (at least historically) for studying entanglement entropy of quantum fields in particular, namely, in an attempt to understand the entropy of black holes [12].

Interest in black holes began with the observations of Bekenstein and Hawking that a black hole of mass M should radiate with a temperature

$$T = \frac{1}{8\pi GM}, \tag{5.39}$$

and have a corresponding entropy proportional to the area of the horizon,

$$S_{BH} = \frac{A}{4G}. \tag{5.40}$$

Despite having an explicit formula for the entropy of a black hole, the microscopic origins of this entropy are not well understood. It was first proposed in [12] (see also [76]) that the black hole entropy may originate in the entanglement entropy of quantum fields on the Schwarzschild spacetime. Recall that in position space, the Hamiltonian for a Klein-Gordon field represents a system a coupled harmonic oscillators, each residing at a point in space. Thus, the ground state of the Hamiltonian will generally be entangled. The proposal that the black hole entropy may be entanglement entropy is due to the observation that if

the quantum field is in a pure state and is split into two subsystems consisting of the field in some region R and the complementary region $\overline{R}$, then the entanglement entropy $S(\rho_R)$ scales with the area of the boundary between R and $\overline{R}$ (called the entangling surface). Thus the proposal would be that if one traces out the field behind the black hole horizon, then the entanglement entropy between the field behind the horizon and the field outside the horizon would account for the entropy of the black hole.

The scaling of the entanglement entropy with the area of the boundary is commonly referred to as the area law. Intuitively, one can think of the entanglement entropy as arising due to correlations between subsystems R and $\overline{R}$ near the entangling surface. This is because degrees of freedom near the entangling surface locally interact with degrees of freedom on the other side of the surface. As we mentioned above, interaction between two degrees of freedom cause them to become entangled, thus one would need information about one of the degrees of freedom in order to eliminate uncertainty regarding the state of the other. The entanglement entropy of the state of one of these degrees of freedom is a measure of this uncertainty. The area law intuitively arises from the observation that most of this uncertainty will be generated by the degrees of freedom near the entangling surface.

Though we will explore the area law throughout the remainder of this work, one heuristic argument for the area law begins with the fact that the entanglement entropy for a pure state is symmetric between the two subsystems, i.e., $S(\rho_R) = S(\rho_{\overline{R}})$. The entanglement entropy will depend on the geometry of the region R, but it cannot be proportional to the volume because if we decrease the size of the region R, the size of the region $\overline{R}$ does not decrease, but we must have $S(\rho_R) = S(\rho_{\overline{R}})$ [76]. Thus, the entanglement entropy can only depend on geometry shared by R and $\overline{R}$, namely the boundary between them.

A more concrete connection can be seen by considering the Unruh effect. Here we will derive the Unruh effect using a formal trick in the path integral, as presented in, e.g., Ref.[77]. There are also techniques for deriving the Unruh effect based on canonical quantisation and detector models which are more physically insightful (see, e.g., Ref.[19] for a review). Nevertheless, this path integral trick will suffice for our purposes. Consider a scalar field on 1+1-dimensional Minkowski space, with metric

$$g = -dt \otimes dt + dx \otimes dx. \tag{5.41}$$

Recall from Section 4.4 that if we Wick-rotate to Euclidean space with imaginary time $\tau := it$ and metric

$$g = d\tau \otimes d\tau + dx \otimes dx \tag{5.42}$$

then the ground state of the field can be written as the partition function:

$$\langle\phi'|\left(|0\rangle\langle 0|\right)|\phi''\rangle = \frac{1}{Z}\int_{\phi(x,0^+)=\phi''(x)}^{\phi(x,0^-)=\phi'(x)} \mathcal{D}\phi\, e^{-S_E[\phi]}. \tag{5.43}$$

Now we proceed to trace out the field in the left-half plane: $x < 0$. Recall that the matrix elements of the density matrix are parametrised by field configurations on either side of $\tau = 0$. Therefore, the partial trace over the left-half plane can be accomplished by matching the boundary conditions in the left-half plane, and then integrating over all possible boundary conditions. That is, if we write the fields as

$$\phi(x,t) = \begin{cases} \phi_+(x,t) & \text{if } x \geq 0 \\ \phi_-(x,t) & \text{if } x < 0 \end{cases}, \tag{5.44}$$

then the reduced density matrix is

$$\rho_{\text{red}}(\phi_- \oplus \phi_+, \phi_- \oplus \phi'_+) = \int \mathcal{D}\phi_- \langle\phi_-|\langle\phi_+|\,\rho\,|\phi_-\rangle\,|\phi'_+\rangle\,. \tag{5.45}$$

This integration amounts to removing the boundary conditions in Eq. (5.43) to obtain:

$$\langle\phi'_+|\,\rho_{\text{red}}\,|\phi''_+\rangle = \frac{1}{Z_{\text{red}}}\int_{\phi(x,0^+)=\phi''_+(x),x\geq 0}^{\phi(x,0^-)=\phi'_+(x),x\geq 0} \mathcal{D}\phi\, e^{-S_E[\phi]}. \tag{5.46}$$

At this point, we perform a change of variables $t = r\sinh(\eta)$ and $x = r\cosh(\eta)$ in the original 1+1-dimensional Minkowski space to get the Rindler spacetime metric:

$$g = -r^2 d\eta \otimes d\eta + dr \otimes dr. \tag{5.47}$$

Note that the trajectory of a uniformly accelerated observer (as a function of the observer's proper time ξ) with acceleration a is: $t(\xi) = a^{-1}\sinh(a\xi)$ and $x(\xi) = a^{-1}\cosh(a\xi)$. Therefore, we see that in Rindler coordinates, the observer's trajectory is given by $r(\xi) = a^{-1}$ and $\eta(\xi) = a\xi$.

Now we do a Wick rotation in Rindler spacetime: $\theta = i\eta$, so that

$$g = r^2 d\theta \otimes d\theta + dr \otimes dr, \tag{5.48}$$

is simply the Euclidean space in polar coordinates. Recall that in Section 4.2 we mentioned that after a change of coordinates, one can obtain a different Hamiltonian which generates translations along a different time coordinate. After changing coordinates to Rindler

spacetime, we see that we obtain a different Hamiltonian, H_R, which generates translations along η. Therefore the time evolution operator is $\exp(-i(\eta-\eta')H_R)$. After Wick rotation, this becomes $\exp(-(\theta-\theta')H_R)$. We can write the matrix elements of this operator using the path integral

$$\langle\phi'|\, e^{-(\theta-\theta')H_R}\,|\phi''\rangle = \int_{\phi(\theta',r)=\phi''(r)}^{\phi(\theta,r)=\phi'(r)} \mathcal{D}\phi\, e^{-S_E[\phi]}. \tag{5.49}$$

Now notice that if we take $\theta' = 0$ and $\theta = 2\pi$, the right-hand side is the same as the right-hand side of Eq. (5.46) in polar coordinates. Therefore, we can immediately write down

$$\rho_{\text{red}} = \frac{1}{Z} e^{-2\pi H_R}. \tag{5.50}$$

This is simply a thermal distribution with inverse temperature $\beta = 2\pi$. Recall that for an accelerated observer, the Rindler coordinate η is related to the proper time of the observer, ξ by $\eta(\xi) = a\xi$. Therefore, the Wick-rotated angle $2\pi = \Delta\theta = i\Delta\eta = ia\Delta\xi$ will appear to the accelerated observer as $2\pi/a$. So the temperature seen by the accelerated observer will be the Unruh temperature,

$$T_U = \beta^{-1} = \frac{a}{2\pi}. \tag{5.51}$$

Therefore, an accelerated observer will see a thermal bath of particles with Unruh temperature T_U.

By the equivalence principle, one should also see similar effects in the presence of gravity. Indeed, for a Schwarzschild black hole, for a distant observer the surface gravity of the black hole, $\kappa = 1/4GM$, plays the role of the acceleration a in the Rindler case. Therefore, this observer will experience a temperature,

$$T_{BH} = \frac{\kappa}{2\pi} = \frac{1}{8\pi GM}, \tag{5.52}$$

which is exactly the Hawking temperature. Although the quantity we have calculated is the temperature for the *field*, to a distant observer it will appear as though the black hole is an object which is radiating at a temperature $T = 1/8\pi GM$. If indeed we consider the black hole as a thermodynamic object at thermal equilibrium, then we can employ the first law of thermodynamics,

$$dE = TdS, \tag{5.53}$$

to obtain an entropy for the black hole. Identifying the black hole energy with its mass, $E = M$, we find

$$dM = \frac{1}{8\pi GM} dS, \tag{5.54}$$

or

$$S = 4\pi GM^2. \tag{5.55}$$

Since the radius of the black hole is related to its mass by $R = 2GM$,

$$S = \frac{A}{4G}, \tag{5.56}$$

which is the Bekenstein-Hawking formula for the black hole entropy [77]. We see that this formula arises by considering the black hole as an object whose thermality is due to the effects of a quantum field on the Schwarzschild spacetime. If we wanted to calculate the entropy of the field itself, we would have to use one of the above formulas for the von Neumann entropy of the reduced density matrix ρ_{red}:

$$\begin{aligned} S(\rho_{\text{red}}) &= \beta\langle H\rangle + \log Z \\ &= \left(1 - \beta\frac{\partial}{\partial\beta}\right)\log Z. \end{aligned} \tag{5.57}$$

For this purpose, we would need to obtain the partition function Z for the reduced density matrix. The calculation of such partition functions is difficult; it is the goal of Section 5.4 to demonstrate a technique for doing this.

A major obstacle in understanding the connection between black hole entropy and entanglement entropy is the fact that the entanglement entropy of quantum fields is ultraviolet-divergent. The generality of this statement will be demonstrated in Section 5.4. In order to obtain a finite entropy, one can introduce an ultraviolet cutoff ϵ in the form of a minimum length. Then to leading order, the entropy will be of the form

$$S \sim \frac{A}{\epsilon^2}. \tag{5.58}$$

If we take this minimum length to be the order of the Planck length, $\epsilon = \ell_P := 2\sqrt{G}$, then the leading term of the entropy will take the form of the Bekenstein-Hawking entropy. Although we referred to this as an obstacle, it is generally considered a feature of the entanglement entropy because if the black hole entropy is indeed the entanglement entropy of fields on the Schwarzschild spacetime, then it may be providing an insight into the nature of Planck-scale physics. Namely, it would be indicating that there may be some fundamental minimum length on the order of the Planck length ℓ_P.

How the entropy of black holes is related to the entanglement entropy of quantum fields is still an unsolved mystery, but provides sufficient motivation for understanding properties of the entanglement entropy of quantum fields in general.

5.3 Entanglement Entropy in Quantum Field Theory: Phase Space method

In this section we will outline a method for calculating the entropy of Gaussian states of a system of quantum harmonic oscillators. This method will be used in Chapter 6 to calculate the entanglement entropy of subregions of a quantum field. The first version of this method was performed in [12], and was later simplified by [5, 1] (see also [20]). The procedure is most easily demonstrated by representing the density matrix of the system using Wigner functions, therefore we will first provide a brief overview of their definition and properties.

5.3.1 Wigner functions

Consider a quantum system described by degrees of freedom $\{q_i\}_i$ and conjugate momenta $\{p_i\}_i$ obeying the commutation relations:

$$[q_i, p_j] = i\delta_{ij}\mathbb{1}, \text{ and} \tag{5.59}$$
$$[q_i, q_j] = [p_i, p_j] = 0. \tag{5.60}$$

The Hilbert space for this quantum system will be written as $\mathcal{H} = \bigotimes_i \mathcal{H}_i$, and operators with subscript i will indicate that the operator acts on the subsystem $\mathcal{H}_i$. One can express the canonical variables in terms of creation and annihilation operators $a_i^{\pm} = \frac{1}{2}(q_i \pm \frac{p_i}{i})$ with commutation relations $\left[a_i^-, a_j^+\right] = \delta_{ij}$ and $\left[a_i^{\pm}, a_j^{\pm}\right] = 0$. Then one can define the displacement operator

$$D_i(\alpha) := e^{\alpha a_i^+ - \alpha^* a_i^-} \tag{5.61}$$

for $\alpha \in \mathbb{C}$, which generates coherent states:

$$|\alpha\rangle_i := D_i(\alpha)\,|0\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle_i \,. \tag{5.62}$$

The set of coherent states forms an overcomplete resolution of identity:

$$\frac{1}{\pi}\int d^2\alpha \;|\alpha\rangle\langle\alpha| = \mathbb{1}, \tag{5.63}$$

where $d^2\alpha := d\mathrm{Re}(\alpha) d\mathrm{Im}(\alpha)$.

For a state ρ on a single subsystem, one can define the corresponding Wigner function as

$$W(\alpha) := \int \frac{d^2\beta}{\pi^2} \ \mathrm{tr}\left[\rho D(\beta)\right] e^{\beta^*\alpha - \beta\alpha^*}. \tag{5.64}$$

It follows from the fact that

$$\int \frac{d^2\beta}{\pi^2} \, e^{\beta^*\alpha - \beta\alpha^*} = \delta^{(2)}(\alpha), \tag{5.65}$$

that we have

$$\int d^2\alpha \, W(\alpha) = \mathrm{tr}\, \rho = 1. \tag{5.66}$$

For example, for the ground state $\rho = |0\rangle \langle 0|$, the corresponding Wigner function is

$$W(\alpha) = \frac{2}{\pi} e^{-2|\alpha|^2}. \tag{5.67}$$

For a thermal state given by Eq. (5.26), it is tedious though straightforward to show that the corresponding Wigner function is

$$W(\alpha) = \frac{1}{\pi(\langle n \rangle + 1/2)} e^{\frac{-|\alpha|^2}{\langle n \rangle + 1/2}}. \tag{5.68}$$

For a system with multiple degrees of freedom, one can generalise the definition of the Wigner function to

$$W(\vec{\alpha}) := \int \prod_i \frac{d^2\beta_i}{\pi^2} \ \mathrm{tr}\left[\rho \otimes_i D_i(\beta_i)\right] e^{\sum_i (\beta_i^* \alpha_i - \beta_i \alpha_i^*)}. \tag{5.69}$$

It will be convenient for our purposes to re-express the Wigner functions in terms of $Q_i := \mathrm{Re}(\alpha_i)$ and $P_i := \mathrm{Im}(\alpha_i)$, which represent coordinates in a phase space whose vector elements can be written $\vec{X} := (Q_1, P_1, Q_2, P_2, \dots)$.

Notice that if we are given a product state $\rho = \otimes_i \rho_i$, then the Wigner function will have the form:

$$W(\vec{X}) = \prod_i W_i(X_i). \tag{5.70}$$

Also, tracing over a subsystem i corresponds to marginalising over the phase space coordinates $X_i = (Q_i, P_i)$ of the Wigner function.

5.3.2 Entanglement Entropy formula

We will now proceed to derive a formula describing how to calculate the entropy of a Gaussian state of a system of quantum harmonic oscillators. Gaussian states are quantum states which can be completely characterised by the averages and correlators of the canonical variables. In phase space, their Wigner functions take the form [1]:

$$W(\vec{X}) = \frac{1}{\pi\sqrt{\det \Sigma}} e^{-\frac{1}{2}\vec{X}^T \Sigma^{-1} \vec{X}}, \tag{5.71}$$

where Σ is the matrix of correlators. If we express the quantum mechanical operators in a vector form $\vec{x} := (q_1, p_1, q_2, p_2, \dots)$, then the matrix elements of Σ are given by

$$\Sigma_{ij} = \tfrac{1}{2}\langle x_i x_j + x_j x_i \rangle. \tag{5.72}$$

Here we are assuming that the expectation values of these operators are zero, i.e., $\langle x_i \rangle = 0$, as this will be sufficient for our purposes.

Since the commutation relations between the canonical variables are c-numbers, we can also define the matrix of commutators as:

$$[x_i, x_j] =: i\Lambda_{ij}\mathbb{1}. \tag{5.73}$$

We now employ a theorem due to Williamson [85, 86], which states that any positive-definite matrix can be put into a diagonal form by a symplectic transformation, i.e., one that preserves the symplectic form (thus the matrix of commutators Λ). That is, Williamson's theorem states that there is a transformation S, such that

$$S\Lambda S^T = \Lambda, \tag{5.74}$$

and

$$S\Sigma S^T = \bigoplus_j \begin{bmatrix} d_j^{(1)} & 0 \\ 0 & d_j^{(2)} \end{bmatrix}. \tag{5.75}$$

It also turns out that in each subspace indexed by j, the diagonal entries are equal: $d_j := d_j^{(1)} = d_j^{(2)}$. Therefore, if we write $\vec{Y} := S\vec{X}$, we can express the Wigner function in these coordinates as:

$$W(\vec{Y}) = \prod_j \frac{1}{\pi d_j} e^{-\frac{\|\vec{Y}_j\|^2}{d_j}}. \tag{5.76}$$

Notice that this is simply a product of Wigner functions corresponding to thermal states, where in each subspace we identify $d_j \leftrightarrow \langle n \rangle + 1/2$.

Therefore, if we are given a Gaussian state ρ, we can represent it as a tensor product of thermal states, $\rho = \otimes_j \rho_j$, where

$$\rho_j := \frac{1}{d_j + 1/2} \sum_{n=0}^{\infty} \left(\frac{d_j - 1/2}{d_j + 1/2} \right)^n |n\rangle \langle n| . \tag{5.77}$$

Since $S(\rho_1 \otimes \rho_2) = S(\rho_1) + S(\rho_2)$, using the formula for the entropy of a thermal state (Eq. (5.28)) we can directly write down the von Neumann entropy of ρ as:

$$S(\rho) = \sum_j f(d_j), \tag{5.78}$$

where

$$f(d_j) := (d_j + 1/2) \log(d_j + 1/2) - (d_j - 1/2) \log(d_j - 1/2). \tag{5.79}$$

We see that the entropy of ρ can be determined entirely in terms of the diagonal elements of Σ obtained from a symplectic transformation S.

This procedure however can be simplified, since finding a symplectic transformation to diagonalise Σ may not be straightforward. Notice that since the matrix of commutators Λ can be written

$$\Lambda = \bigoplus_i \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \tag{5.80}$$

then we can write

$$\begin{aligned} \Lambda^{-1} S \Sigma S^T &= (S \Lambda S^T)^{-1} S \Sigma S^T \\ &= S^{-T} \Lambda^{-1} \Sigma S^T \\ &= \bigoplus_j \begin{bmatrix} 0 & -d_j \\ d_j & 0 \end{bmatrix}. \end{aligned} \tag{5.81}$$

Note that the eigenvalues of this matrix are $\pm i d_j$. Furthermore, since the eigenvalues of $S^{-T} \Lambda^{-1} \Sigma S^T$ are the same as the eigenvalues of $\Lambda^{-1} \Sigma$, one can simply determine the d_j's as the positive imaginary parts of the eigenvalues of $\Lambda^{-1} \Sigma$. These eigenvalues are called the symplectic eigenvalues of the matrix Σ. This simplifies the process of finding the values of d_j, since calculating eigenvalues is often simpler than performing a symplectic diagonalisation.

Therefore, the method is simply to determine the matrix of commutators Λ and the matrix of correlators Σ, then calculate the eigenvalues of the matrix $\Lambda^{-1}\Sigma$, which come in pairs $\pm id_j$, and then use Eqs. (5.78) and (5.79) to calculate the entropy.

In order to calculate the entanglement entropy of a subsystem of a Gaussian state, recall that a partial trace over a subsystem corresponds to marginalising over the corresponding phase space variables in the Wigner function. Because marginalising over variables of a Gaussian distribution, such as the Wigner function (5.71), again results in a Gaussian distribution, one can use this method for calculating the entropy of the marginalised distribution by restricting the matrices Λ and Σ to the subspaces of the remaining phase space variables. In this way, one can also use this method to calculate entanglement entropies of Gaussian states.

5.4 Entanglement Entropy in Quantum Field Theory: Path Integral method

In this section we will present a method of calculating the entanglement entropy for the ground state of a field based on path integrals. Although out attention will be restricted to simple cases, this method can be extended to other fields as well as general spacetimes. The goal of this section is to identify the general scaling behaviour of the entanglement entropy.

The replica trick shown in Subsection 5.4.1 was first developed in [14], however much of this section is based on the reviews [16, 73]. Subsection 5.4.2 on functional determinants and heat kernels is based on Ref.[62] (see also [32, 80, 28]).

5.4.1 The Replica Trick

First, we define a quantity called the **Renyí entropy** of a density matrix ρ:

$$S_\alpha(\rho) := \frac{\log(\operatorname{tr}\rho^\alpha)}{1-\alpha}. \tag{5.82}$$

In the limit $\alpha \to 1$, this simply becomes the von Neumann entropy:

$$\begin{aligned}\lim_{\alpha\to 1} S_\alpha(\rho) &= \lim_{\alpha\to 0} \frac{\log(\operatorname{tr}\rho^\alpha)}{1-\alpha} \\ &= \lim_{\alpha\to 1} \frac{\frac{1}{\operatorname{tr}\rho^\alpha}\operatorname{tr}(\rho^\alpha\log(\rho^\alpha))}{-1} \\ &= -\operatorname{tr}(\rho\log\rho). \end{aligned} \tag{5.83}$$

Now, recall from Section 4.4 that the matrix elements of the ground state of a field at the $t = 0$ surface can be written as the following Euclidean space path integral:

$$\rho(\phi', \phi'') = \int_{\phi(x,0^+)=\phi''(x)}^{\phi(x,0^-)=\phi'(x)} \mathcal{D}\phi\, e^{-S_E}. \tag{5.84}$$

Note that the matrix elements of ρ are parametrised by boundary conditions placed on either side of the $t = 0$ surface.

If we wish to trace out a subregion Σ of the $t = 0$ surface, then similar to above, we can split the field configurations into the direct sum $\phi = \phi_\Sigma \oplus \phi_{\Sigma^c}$, or

$$\phi(x) = \begin{cases} \phi_\Sigma(x) & \text{if } x \in \Sigma \\ \phi_{\Sigma^c}(x) & \text{if } x \notin \Sigma \end{cases}. \tag{5.85}$$

Therefore, we can split a Hilbert space vector representing a field configuration as $|\phi\rangle = |\phi_\Sigma\rangle \otimes |\phi_{\Sigma^c}\rangle$. This can be understood in roughly the following way. Intuitively, quantum fields represent quantum systems which have a degree of freedom in every point in space. Therefore, formally, we can think of the field operator at every point, $\hat{\phi}(x)$, as an operator with an associated Hilbert space $\mathcal{H}_x$ and eigenvectors $|\phi(x)\rangle \in \mathcal{H}_x$ such that

$$\hat{\phi}(x)\,|\phi(x)\rangle = \phi(x)\,|\phi(x)\rangle\,. \tag{5.86}$$

Then one can think of the entire field as acting on a Hilbert space $\mathcal{H} = \otimes_{x\in\mathbb{R}}\mathcal{H}_x$ with eigenvectors $|\phi\rangle = \otimes_{x\in\mathbb{R}}|\phi(x)\rangle$. Therefore, one can think of the decomposition $|\phi\rangle = |\phi_\Sigma\rangle \otimes |\phi_{\Sigma^c}\rangle$ simply as factoring the Hilbert space into a bipartite system $\mathcal{H} = (\otimes_{x\in\Sigma}\mathcal{H}_x) \otimes (\otimes_{x\notin\Sigma}\mathcal{H}_x)$. Again, this is only formal since the field operator $\hat{\phi}(x)$ is not really an operator, but an operator-valued distribution. In order to obtain an operator, one must smear the distribution by integrating over some smooth function [34]. Nevertheless, this intuition may be helpful in understanding the formal manipulations in this section.

Then the reduced density matrix can be expressed as

$$\rho_{\rm red}(\phi'_\Sigma, \phi''_\Sigma) = \int \mathcal{D}\phi_{\Sigma^c} \ \langle \phi'_\Sigma | \otimes \langle \phi_{\Sigma^c} | \, \rho \, | \phi''_\Sigma \rangle \otimes | \phi_{\Sigma^c} \rangle \, . \tag{5.87}$$

Since the matrix elements of ρ are parametrised by boundary conditions on either side of $t = 0$, the matrix element in the above integral corresponds to matching the boundary conditions in the region Σ^c by setting them both to ϕ_{Σ^c}, and using the boundary condition ϕ''_Σ on Σ just above the surface (at $t = 0^+$) and ϕ'_Σ on Σ just below the surface (at $t = 0^-$). The reduced density matrix is then obtained by integrating over all possible boundary conditions in the region Σ^c. Thus the elements of the reduced density matrix can be obtained from the original Euclidean path integral by alleviating the boundary condition in the region Σ^c. That is,

$$\rho_{red}(\phi'_\Sigma, \phi''_\Sigma) = \int_{\phi(x,0^+)=\phi''_\Sigma(x), x\in\Sigma}^{\phi(x,0^-)=\phi'_\Sigma(x), x\in\Sigma} \mathcal{D}\phi \, e^{-S_E}. \tag{5.88}$$

In the following, unless otherwise stated we shall simply reuse the symbol ρ to represent the reduced density matrix $\rho_{\rm red}$.

The idea behind the replica trick begins with the observation that it is much easier to calculate an integer power of ρ than it is to calculate $\log \rho$. For example, for

$$\rho^2(\phi, \phi') = \int \mathcal{D}\phi'' \ \rho(\phi, \phi'')\rho(\phi'', \phi'), \tag{5.89}$$

the element $\rho(\phi, \phi'')$ is determined from performing a path integral by integrating over field configurations over the entire manifold $\mathcal{M}$ with the boundary conditions ϕ on Σ at $t = 0^-$ and ϕ'' on Σ at $t = 0^+$. Similarly, $\rho(\phi'', \phi')$ is also determined by integrating field configurations over the entire manifold $\mathcal{M}$ with appropriate boundary conditions. Then, the product $\rho(\phi, \phi'')\rho(\phi'', \phi')$ is an integration of field configurations over two copies of the manifold $\mathcal{M}$ whose field configurations are identified on Σ at $t = 0^+$ on the first manifold and Σ at $t = 0^-$ on the second manifold. Then the integration in Eq. (5.89) is an integration over field configurations on the identified surface. Therefore, one can view the path integral defining the matrix elements of $\rho^2(\phi, \phi')$ as an integrating over field configurations on a two-sheeted manifold, $\mathcal{M}_2$, where the $(t = 0^+, \Sigma)$ region of the first sheet is identified with the $(t = 0^-, \Sigma)$ region of the second sheet.

This method can be extended iteratively to define the matrix elements of an integer power n of the density matrix: $\rho^n(\phi, \phi'')$. One obtains n copies or "replicas" of the manifold $\mathcal{M}$ to get an n-sheeted manifold, $\mathcal{M}_n$, where the $(t = 0^+, \Sigma)$ region of the i^{th} sheet is

identified with the $(t = 0^-, \Sigma)$ region of the $(i+1)^{st}$ sheet. The trace $\mathrm{tr}(\rho^n)$ can then be obtained by identifying the $(t = 0^+, \Sigma)$ region of the n^{th} sheet is identified with the $(t = 0^-, \Sigma)$ region of the 1^{st} sheet in a similar manner. Then this trace can be represented as a partition function $Z(n)$ corresponding to the path integral

$$\mathrm{tr}(\rho^n) = Z(n) := \int \mathcal{D}\phi \, e^{-S_E}, \tag{5.90}$$

where the integral in the action is now taken over the n-sheeted manifold $\mathcal{M}_n$. Note that in the case $n = 1$, if ρ is normalised, then $Z(1) = 1$.

Once this partition function has been calculated, one can calculate the Renyí entropy:

$$S_n(\rho) = \frac{\log(\mathrm{tr}\,\rho^n)}{1-n} = \frac{\log Z(n)}{1-n}. \tag{5.91}$$

If one can find an analytic continuation of $Z(n)$ to non-integer n, then one can determine the von Neumann entropy (thus the entanglement entropy of the region Σ) via the relation:

$$\begin{aligned} S(\rho) &= \lim_{n\to 1} S_n(\rho) \\ &= \lim_{n\to 1} \frac{\log Z(n)}{1-n} \\ &= -\left.\frac{\partial}{\partial n}\right|_{n=1} \log Z(n). \end{aligned} \tag{5.92}$$

If we define a parameter $\beta := 2\pi n$ and add $0 = \log Z(1) = \log Z(n)|_{n=0}$, then we can rewrite the entropy in a suggestive form similar to the equation for the thermal entropy:

$$S(\rho) = \left(1 - \beta\frac{\partial}{\partial\beta}\right)\Big|_{\beta=2\pi} \log Z(\beta). \tag{5.93}$$

There is still the issue of performing an explicit calculation of $Z(n)$ and determining an analytic continuation to $Z(\beta)$. A strategy often employed for calculating partition functions is to recast the problem as the problem of calculating a functional determinant, which is in turn rewritten as the problem of calculating a heat kernel. There have been many studies devoted toward calculating heat kernel coefficients for manifolds with conical singularities (which also perform the required analytic continuation). For our purposes, we will mainly be interested in the general behaviour of these coefficients, and the implications this has for the relationship between the geometry of Σ and the corresponding entanglement entropy S.

In Subsection 5.4.2, we will briefly outline how to reformulate the calculation of a partition function as a calculation of heat kernel coefficients. Then in Subsection 5.4.3 we will examine the general structure of the entanglement entropy. Finally, we will end this section with explicit results obtained for the entanglement entropy using this method.

5.4.2 Functional Determinants and Heat Kernels

For simplicity, we will restrict our attention in this section to free, massless, scalar fields on a (compact) manifold $\mathcal{M}$. Consider the Euclidean action:

$$I_E[\phi] = \frac{1}{2}\int d^n x\sqrt{g}\,\phi(x)(-\Delta)\phi(x), \tag{5.94}$$

where

$$\Delta\phi = \frac{1}{\sqrt{g}}\partial_\mu\left(\sqrt{g}g^{\mu\nu}\partial_\nu\phi\right), \tag{5.95}$$

and where the metric g has positive Euclidean signature.

If the manifold is compact, then we can diagonalise $-\Delta$ to obtain eigenvalues λ_n and eigenfunctions f_n. If we write ϕ in the eigenbasis of $-\Delta$ as $\phi(x) = \sum_n c_n f_n(x)$, then one can evaluate the following path integral:

$$\begin{aligned} Z &:= \int \mathcal{D}\phi\, e^{-S_E[\phi]} \\ &= \int \left(\prod_n \frac{dc_n}{\sqrt{2\pi}}\right) e^{-\frac{1}{2}\sum_n \lambda_n |c_n|^2} \\ &= \prod_n \lambda_n^{-1/2}. \end{aligned} \tag{5.96}$$

Note that this partition function is only determined up to an ambiguous measure constant, which we chose for each dimension to be $dc_n/\sqrt{2\pi}$. We will formally define the **functional determinant** of $-\Delta$ as

$$\det(-\Delta) := \prod_n \lambda_n = Z^{-1/2}. \tag{5.97}$$

One then proceeds to perform the following sequence of formal manipulations. First, we write

$$\begin{aligned} \log\det(-\Delta) &= \log\prod_n \lambda_n \\ &= \sum_n \log\lambda_n. \end{aligned} \tag{5.98}$$

Then, we define the **spectral zeta function**:

$$\zeta_{-\Delta}(s) := \sum_n \frac{1}{\lambda_n^s}, \tag{5.99}$$

where $\{\lambda_n\}_n$ are the eigenvalues of $-\Delta$. Note that

$$\begin{aligned} -\left.\frac{d}{ds}\right|_{s=0} \zeta_{-\Delta}(s) &= \sum_n \log \lambda_n \\ &= \log\det(-\Delta) \end{aligned} \tag{5.100}$$

Now, recall the definition of the gamma function:

$$\Gamma(s) = \int_0^\infty dt\, t^{s-1} e^{-t}. \tag{5.101}$$

We then rescale the integration variable by λ_n to get

$$\Gamma(s) = \lambda_n^s \int_0^\infty dt\, t^{s-1} e^{-\lambda_n t}. \tag{5.102}$$

Then rearranging this formula, we get

$$\zeta_{-\Delta}(s) := \sum_n \frac{1}{\lambda_n^s} = \frac{1}{\Gamma(s)} \int_0^\infty dt\, t^{s-1} h(t), \tag{5.103}$$

where $h(t)$ is the **heat trace**,

$$h(t) := \sum_n e^{-\lambda_n t} = \mathrm{tr}[e^{-(-\Delta)t}]. \tag{5.104}$$

The final step is observing that

$$\left.\frac{d}{ds}\right|_{s=0} \frac{t^s}{\Gamma(s)} = 1, \tag{5.105}$$

then if we formally permit the interchange of $(d/ds)|_{s=0}$ and integration by t, we obtain

$$\log\det(-\Delta) = -\int_0^\infty \frac{dt}{t}\, h(t). \tag{5.106}$$

Although the above sequence of manipulations was only formal, one typically uses Eq. (5.106) as a definition of the functional determinant of the operator $-\Delta$. One then determines the corresponding partition function as

$$\log Z = -\frac{1}{2} \log\det(-\Delta). \tag{5.107}$$

One is now faced with the problem of calculating the heat trace $h(t)$ of the operator $-\Delta$. Notice that if the spectrum of $-\Delta$ is known, then it is straightforward to calculate the heat trace. Suppose we represent $-\Delta$ by an operator with the same spectrum on a Hilbert space with a position-like basis (cf., Ref.[62]) such that

$$\langle x|y\rangle = \delta^{(n)}(x-y) \tag{5.108}$$

and

$$\langle x|\,(-\Delta)\,|y\rangle = -\partial_x^2\delta^{(n)}(x-y). \tag{5.109}$$

Let us then define the **heat kernel**, K, as the operator solution to the equation

$$\frac{\partial}{\partial t}K(t) = -\Delta K(t), \tag{5.110}$$

with initial condition $K(x,y;0) := \langle x|\,K(0)\,|y\rangle = \delta^{(n)}(x-y)$. The formal solution to the above equation is

$$K(t) = e^{-(-\Delta)t}, \tag{5.111}$$

so that we see

$$h(t) = \mathrm{tr}\,K(t). \tag{5.112}$$

The operator K is called a heat kernel because for flat space, it is a Green's function for the heat equation:

$$K(x,y;t) = \frac{1}{(4\pi t)^{n/2}}e^{-\frac{(x-y)^2}{4t}}. \tag{5.113}$$

For general spaces, DeWitt [28] proposed the following ansatz, in Riemann normal coordinates centred at y [28]:

$$K(x,y;t) = \frac{1}{(4\pi t)^{n/2}}g^{-1/4}(x)e^{-\frac{x^2}{4t}}\sum_{p=0}^{\infty}b_p(x)t^p, \tag{5.114}$$

which we can see is simply the flat space heat kernel with curvature corrections parametrised by the sum

$$g^{-1/4}(x)\sum_{p=0}^{\infty}b_p(x)t^p. \tag{5.115}$$

Inserting this ansatz into the heat equation (Eq. (5.110)), one eventually obtains the following recursion relations for the coefficients b_p [28]:

$$\begin{aligned}(p+1)b_{p+1}(x) + x^\mu\partial_\mu b_{p+1}(x) &= g^{1/4}(x)\Delta(g^{-1/4}(x)b_p(x)) \\ x^\mu\partial_\mu b_0(x) &= 0.\end{aligned} \tag{5.116}$$

Using the initial condition $K(x,y;0) = \delta^{(n)}(x-y)$, one can conclude $b_0 = 1$.

Since we are only interested in the diagonal elements of $K(x,y;t)$ (for the heat trace), one can solve for the coefficients perturbatively in the limit $x \to y$. Since we are using Riemann normal coordinates, in this limit one finds:

$$g_{\mu\nu}(x) = \delta_{\mu\nu} - \frac{1}{3} R_{\mu\alpha\nu\beta} x^\alpha x^\beta + \mathcal{O}(x^4). \tag{5.117}$$

In this limit, the recursion relation for the coefficient b_1 gives

$$b_1 = \Delta g^{-1/4} = \frac{1}{6} R. \tag{5.118}$$

If we were to proceed with the recursion, we would eventually obtain the following asymptotic expansion for the heat trace at small t:

$$\begin{aligned} h(t) &= \int_{\mathcal{M}} d^n x \sqrt{g}\, K(x,x;t) \\ &\sim \frac{1}{t^{n/2}} \left(a_0 + a_2 t + \cdots \right), \end{aligned} \tag{5.119}$$

where

$$a_{2p} := \frac{1}{(4\pi)^{n/2}} \int_{\mathcal{M}} d^n x \sqrt{g}\, b_p(x). \tag{5.120}$$

The first two coefficients are:

$$a_0 = \frac{1}{(4\pi)^{n/2}} \int_{\mathcal{M}} d^n x \sqrt{g} = \frac{\mathrm{Vol}(\mathcal{M})}{(4\pi)^{n/2}} \tag{5.121}$$

$$a_2 = \frac{1}{(4\pi)^{n/2}} \int_{\mathcal{M}} d^n x \sqrt{g}\, \frac{R}{6}. \tag{5.122}$$

An important property of the heat trace [73] that will be used below is that for a product manifold, $\mathcal{M} \times \mathcal{N}$, the heat trace factors as

$$h_{\mathcal{M}\times\mathcal{N}}(t) = h_{\mathcal{M}}(t) \times h_{\mathcal{N}}(t). \tag{5.123}$$

5.4.3 Scaling behaviour

Let $\partial\Sigma$ represent the boundary of the region of space for which we wish to calculate the entanglement entropy. Then the conical singularity we encountered in the replica trick

resides on this surface, and near this surface, the manifold looks like the direct product $\partial\Sigma \times C_\beta$ (where C_β is a cone with an angular coordinate which ranges from 0 to β) [73]. Therefore, we can evaluate the heat trace for the entire manifold by determining the heat trace for $\partial\Sigma$ and C_β separately.

For the cone C_β, the heat trace coefficients are [80]:

$$a_0(C_\beta) = \frac{\beta V}{4\pi} \tag{5.124}$$

$$a_2(C_\beta) = \frac{1}{4\pi}\frac{4\pi^2 - \beta^2}{6\beta}. \tag{5.125}$$

Consider the 4-dimensional case, so that $\partial\Sigma$ is a 2-dimensional surface. Recall that the entanglement entropy for Σ can be written as:

$$\begin{aligned} S(\Sigma) &= \left(1 - \beta\frac{\partial}{\partial\beta}\right)\Big|_{\beta=2\pi} \log Z(\beta) \\ &= -\frac{1}{2}\left(1 - \beta\frac{\partial}{\partial\beta}\right)\Big|_{\beta=2\pi} \log\det(-\Delta) \\ &= \frac{1}{2}\left(1 - \beta\frac{\partial}{\partial\beta}\right)\Big|_{\beta=2\pi} \int_0^\infty \frac{dt}{t}\, h(t) \\ &= \frac{1}{2}\left(1 - \beta\frac{\partial}{\partial\beta}\right)\Big|_{\beta=2\pi} \int_0^\infty \frac{dt}{t^3} \left(\frac{\beta V}{4\pi} + \frac{1}{4\pi}\frac{4\pi^2 - \beta^2}{6\beta}t\right) \\ &\qquad \times \left[a_0(\partial\Sigma) + a_2(\partial\Sigma)t + \cdots\right]. \end{aligned} \tag{5.126}$$

Terms proportional to β vanish after taking the β derivative, and

$$\left(1 - \beta\frac{\partial}{\partial\beta}\right)\Big|_{\beta=2\pi} \frac{1}{\beta} = \frac{1}{\pi}, \tag{5.127}$$

therefore

$$S(\Sigma) = \frac{1}{12}\int_0^\infty \frac{dt}{t^2}\, \left[a_0(\partial\Sigma) + a_2(\partial\Sigma)t + \cdots\right]. \tag{5.128}$$

If we integrate termwise, we see that the integrals are divergent. To remedy this, one typically introduces a cutoff ϵ^2 on the lower limit of the integral. Note that t has units of (length)^2 (see Eq. (5.110)), so the divergence of the lower limit of the integral is an ultraviolet divergence. The upper limit of the integral gives an infrared divergence. However,

here we are only interested in the ultraviolet behaviour of the entropy, so we shall simply ignore contributions from this integration limit. Thus, we have

$$\begin{aligned} S(\Sigma) &= \frac{1}{12}\int_{\epsilon^2}\frac{dt}{t^2}\ [a_0(\partial\Sigma) + a_2(\partial\Sigma)t + \cdots] \\ &= \frac{1}{12}\left[\frac{a_0(\partial\Sigma)}{\epsilon^2} + a_2(\partial\Sigma)\log\epsilon^2\right] + (\text{UV finite terms}\ldots). \end{aligned} \tag{5.129}$$

Since $\partial\Sigma$ is a 2-dimensional surface and $a_0(\mathcal{M}) \sim \text{Vol}(\mathcal{M})$, let us write $A(\Sigma) := \text{Vol}(\partial\Sigma)$ (i.e. the area of the region Σ). Then we find that the leading order behaviour of the entropy is

$$S(\Sigma) \sim \frac{A(\Sigma)}{48\pi\epsilon^2}. \tag{5.130}$$

Thus, we see both that the entanglement entropy of a region Σ is both proportional to its surface area (the area law) and is ultraviolet divergent. We also see that in four dimensions there is a subleading logarithmically divergence.

5.4.4 Example: Two-dimensional Flat space

Consider a two-dimensional flat space with coordinates (x^0, x^1) (this would correspond to the Wick-rotated 1+1-dimensional Minkowski spacetime). We can easily calculate the entanglement entropy of the half-space $x^1 \geq 0$ using the above method.

After using the replica trick, the corresponding β-sheeted space we must consider is simply the cone C_β. Therefore we can readily calculate the entanglement entropy:

$$\begin{aligned} S &= \frac{1}{2}\left(1 - \beta\frac{\partial}{\partial\beta}\right)\Big|_{\beta=2\pi}\int_{\epsilon^2}\frac{dt}{t}\ h(t) \\ &= \frac{1}{2}\left(1 - \beta\frac{\partial}{\partial\beta}\right)\Big|_{\beta=2\pi}\int_{\epsilon^2}\frac{dt}{t^2}\ \left(\frac{\beta V}{4\pi} + \frac{1}{4\pi}\frac{4\pi^2 - \beta^2}{6\beta}t\right) \\ &= \frac{1}{12}\int_{\epsilon^2}\frac{dt}{t} \\ &= \frac{1}{6}\log\left(\frac{1}{\epsilon}\right). \end{aligned} \tag{5.131}$$

This is the well-known logarithmic scaling behaviour in 1+1-dimensions.

For a finite interval of length L, it turns out that the entanglement entropy of this region has a similar scaling behaviour [42]:

$$S(L) = \frac{1}{3} \log \left(\frac{L}{\epsilon} \right) . \tag{5.132}$$

The doubling of the coefficient of the logarithm is due to the fact that the major contributions to the entanglement entropy come from correlations near the boundary of the region, and for a finite interval L there are two boundaries as opposed to a single boundary in the half-space case.

In 1+1-dimensions, the area law for an interval of length L would imply that the entropy should be constant, since the volume of the boundary (two points) does not change as the size of L changes. The interpretation of the logarithmic scaling behaviour is due to the fact that for a 1+1-dimensional massless field (as considered here), there is an accumulation of long-range correlations which causes a deviation from the area law.

5.4.5 Example: Four-dimensional Schwarzschild

When the spacetime is curved, determining the remaining heat trace coefficients is tedious. Here we will only state an interesting result shown in Ref.[73] that for a four dimensional Schwarzschild spacetime, one can determine the entropy for the horizon Σ:

$$S(\Sigma) = \frac{A(\Sigma)}{48\pi\epsilon^2} + \frac{1}{45} \log \left(\frac{2GM}{\epsilon} \right) . \tag{5.133}$$

The leading order behaviour scales with the area similar to the Bekenstein-Hawking entropy, however we see that there is also a subleading logarithmic correction.

Chapter 6

Entanglement Entropy in Bandlimited Quantum Field Theory

6.1 Bandlimited Quantum Field Theory

As we mentioned in the introduction, there are many indications that there should exist an ultraviolet cutoff or minimum length at the Planck scale. In this chapter, we apply Shannon's sampling theory discussed in Chapter 3 to quantum fields as a model for such a cutoff. The sampling theory model first arose from examining the implications of corrections to the Heisenberg uncertainty relation of the form,

$$\Delta X \Delta P \geq \frac{1}{2}\left(1 + \beta(\Delta P)^2 + \dots\right), \tag{6.1}$$

(for some constant $\beta > 0$) which was suggested in several studies of quantum gravity and string theory (e.g., [29, 54, 87, 51, 43]). Such a generalised uncertainty principle implies a constant lower bound on the uncertainty in X, $\Delta X_{\text{min}} = \sqrt{\beta}$. In the case where the position operator X has such a lower bound on its uncertainty, it is not self-adjoint but merely symmetric, and one can construct self-adjoint extensions of the position operator with discrete spectra [60, 51].

Therefore, fields on a space whose coordinates exhibit such a finite minimum uncertainty will possess an ultraviolet cutoff since they possess only a finite density of degrees of freedom [51]. This cutoff can be naturally modelled by a bandlimit on the fields, thus they become subject to sampling theory. The sampling theory model is superior to regular lattice theories because one can represent a function on a lattice while preserving translational

and rotational symmetry of the theory (since the function has an equivalent continuous representation). The purpose of this chapter is to study the behaviour of the entanglement entropy and the nature of the nonlocality of the degrees of freedom in the bandlimited quantum field theory.

The work presented in this chapter is based on work which will be published in Ref.[69].

6.1.1 Sampling Theorem for Quantum Fields

In this work we will restrict our attention to 1+1 dimensional, free, real, scalar fields $\hat{\phi}$. Consider such a field along with a corresponding Hilbert space $\mathcal{H}$. Let $|\phi\rangle \in \mathcal{H}$ be an eigenstate of the field operator, such that for all $x \in \mathbb{R}$ we have

$$\hat{\phi}(x)\,|\phi\rangle = \phi(x)\,|\phi\rangle \tag{6.2}$$

Now consider a subspace of the Hilbert space, $\mathcal{H}_{(-\Omega,\Omega)}$, defined to be the span of the eigenvectors $|\phi\rangle$ with eigenvalues $\phi(x)$ which represent functions (field configurations) bandlimited by Ω. Then since $\phi(x)$ is a bandlimited function, it can be reconstructed from its values on a sampling lattice with Beurling density π/Ω. Since these values are determined by the action of the operator $\hat{\phi}$ on this sampling lattice, the action of $\hat{\phi}$ at any $x \in \mathbb{R}$ is determined by its action on the sampling lattice. For a Nyquist lattice $\{x_n\}_n$, this works out explicitly as:

$$\begin{aligned}
\hat{\phi}(x)\,|\phi\rangle &= \phi(x)\,|\phi\rangle \\
&= \sum_n \mathrm{sinc}[(x-x_n)\Omega]\phi(x_n)\,|\phi\rangle \\
&= \sum_n \mathrm{sinc}[(x-x_n)\Omega]\hat{\phi}(x_n)\,|\phi\rangle\,.
\end{aligned} \tag{6.3}$$

Since this holds for any eigenstate $|\phi\rangle$ in $\mathcal{H}_{(-\Omega,\Omega)}$, and since $\mathcal{H}_{(-\Omega,\Omega)}$ is the span of these eigenstates, then the action of the operator $\hat{\phi}(x)$ at any $x \in \mathbb{R}$ is determined entirely in terms of the action of the operators $\{\hat{\phi}(x_n)\}_n$ on the Nyquist lattice. Therefore, when acting on the subspace $\mathcal{H}_{(-\Omega,\Omega)}$, we can represent the action of $\hat{\phi}(x)$ as

$$\hat{\phi}(x) = \sum_n \mathrm{sinc}[(x-x_n)\Omega]\hat{\phi}(x_n). \tag{6.4}$$

Therefore, we see that the operators on the Nyquist lattice form a complete set of commuting observables for the system. Furthermore, using the generalisation of the sampling

theorem to non-Nyquist sampling lattices, one can equivalently determine the action of $\hat{\phi}(x)$ from the action of the $\hat{\phi}(x_n)$'s on the sampling lattice using an appropriate reconstruction kernel. Recall also that any lattice can be used as a sampling lattices, provided that the average density of the sample points (the Beurling density) is at or above the Nyquist rate Ω/π.

The assumption that the field is free is important since this will ensure that the Hilbert space $\mathcal{H}_{(-|Omega,\Omega)}$ is preserved under time evolution. In interacting theories, different spatial modes of the field mix due to scattering, which implies that the Hilbert space $\mathcal{H}_{(-\Omega,\Omega)}$ would not be preserved in time. One could always project the state at a later time onto this subspace in order to preserve the cutoff, but then the evolution would not be unitary. Currently it is not clear how to implement such a cutoff in an interacting theory.

Note that we can regard Eq. (6.4) as a second-quantised version of the sampling theorem, in the sense that the sampling theorem here applies to quantum fields whereas that of Chapter 3 was derived for wavefunctions in the language of first-quantisation. However, in this work we will only be using the sampling theorem to provide a discrete mathematical representation of a field, which will be convenient for computations, as opposed to considering a scheme for reconstructing a field from a set of measurements (though this would be of interest for future work).

The sampling theorem for quantum fields (Eq. (6.4)) implies that the degrees of freedom of the field are nonlocal. We see this immediately from the reconstruction formula since the observable at x represented by the operator $\hat{\phi}(x)$ can be written in terms of the field operators on the sampling lattice, $\{\hat{\phi}(x_n)\}_n$. This will also become obvious in Subsection 6.1.2 when we calculate the commutation relations for a 1+1 dimensional Klein-Gordon field and in Subsection 6.3.1 when we consider the implications this has for the Hilbert space factorisation of the system.

6.1.2 Two-point Functions

Here we will begin to examine the effects of bandlimitation by calculating its impact on the two-point functions of a 1+1 dimensional Klein-Gordon field. The free, massless Klein-Gordon Hamiltonian is given by

$$H[\phi,\pi] := \frac{1}{2}\int_{\mathbb{R}} dx \ \left(\pi^2(x) + \phi(x)(-\Delta)\phi(x)\right), \tag{6.5}$$

where $\Delta = \partial_x^2$ is the scalar Laplacian operator. Hereafter we will drop the $\hat{}$ from the operators. The fields ϕ and π can be expressed in the usual mode expansion (see Chapter 4)

in terms of the creation and annihilation operators a_k and $a_k^\dagger$ (with commutation relations $[a_k, a_{k'}^\dagger] = (2\pi)\delta(k-k')$ and $[a_k, a_{k'}] = [a_k^\dagger, a_{k'}^\dagger] = 0$), as

$$\phi(x) = \int_{\omega<|k|<\Omega} \frac{dk}{2\pi} \frac{1}{\sqrt{2|k|}} \left(a_k e^{ikx} + a_k^\dagger e^{-ikx} \right), \text{ and} \tag{6.6}$$

$$\pi(x) = \int_{\omega<|k|<\Omega} \frac{dk}{2\pi} \frac{1}{i} \sqrt{\frac{|k|}{2}} \left(a_k e^{ikx} - a_k^\dagger e^{-ikx} \right), \tag{6.7}$$

where the bandlimit Ω is implemented simply by restricting the possible values of k. We also introduce an infrared cutoff ω as a temporary technical tool to tame the infrared divergence exhibited by 1+1 massless scalar fields. In this section, the impact of the infrared divergence in the two-point functions will be removed by an appropriate renormalisation procedure. For the entropy calculations below, the infrared divergence will remain, but will will consider its impact on the behaviour of the entanglement entropy in Section 6.5. A hard infrared cutoff is used instead of a mass to tame the infrared divergence in order to calculate certain integrals analytically. We will find in Section 6.5 that such an infrared cutoff will affect the entanglement entropy in a similar manner as would a mass.

These mode expansions and the commutation relations for a_k and $a_k^\dagger$ allow us to calculate the two-point functions for the field. The equal-time commutation relations between two arbitrary points of the field are

$$[\phi(x), \pi(x')] = i \left(\frac{\Omega}{\pi} \operatorname{sinc}(\Omega \Delta x) - \frac{\omega}{\pi} \operatorname{sinc}(\omega \Delta x) \right), \tag{6.8}$$

$$[\phi(x), \phi(x')] = 0, \text{ and} \tag{6.9}$$

$$[\pi(x), \pi(x')] = 0, \tag{6.10}$$

where we have written $\Delta x := x - x'$. We see that the ϕ's still commute among themselves (as well as the π's) but the ϕ-π commutator becomes nonlocal.

We recall that typically in quantum field theory, locality is enforced by ensuring operators in spacelike-separated regions commute. This results in the equal-time ϕ-π commutator,

$$[\phi(x), \pi(x')] = i\delta(x - x'), \tag{6.11}$$

along with the ϕ's and π's commuting among themselves. Therefore, the bandlimited theory violates locality in the sense that operators in spacelike-separated regions do not commute.

Recall also that this notion of locality in the usual quantum field theory ensures that causality is preserved. It is known that, without violating causality, one can introduce

nonlocality to quantum theory in addition to the nonlocality due to entanglement (see, e.g., [6]). However, the nonlocality in the bandlimited theory violates causality. Thus, the bandlimited theory may exhibit nonlocalities beyond those allowed for in the causal theories.

The violation of causality in the bandlimited theory may not be so surprising since the imposition of a cutoff only on the spatial wavenumbers is not a Lorentz-covariant operation. Although this could be seen as a limitation of this model, we anticipate that many of the features of the model, in particular in regards to the nonlocal behaviour, which may be instructive for future studies of nonlocality in more complicated models.

Although our method of imposing the bandlimit was simply to restrict the spatial momenta to the range $\omega < |k| < \Omega$ in the mode expansion, one also obtains the same commutators as in Eq. (6.8) by constraining the phase space to the surface $\tilde{\phi}_k = \tilde{\pi}_k = 0$ for $\omega < |k| < \Omega$ and employing the constrained quantisation procedure of Dirac [22] to obtain the Dirac brackets of the constrained system.

One can also calculate the equal-time correlation functions for various states of the quantum field. For example, for the ground state of the field, we obtain for two distinct points,

$$\langle 0|\, \phi(x)\phi(x')\, |0\rangle = \frac{1}{2\pi}[\mathrm{Ci}(\Omega\Delta x) - \mathrm{Ci}(\omega\Delta x)], \tag{6.12}$$

which in the coincidence limit becomes

$$\langle 0|\, \phi^2(x)\, |0\rangle = \frac{1}{2\pi} \log\left(\frac{\Omega}{\omega}\right). \tag{6.13}$$

Here we see explicitly the infrared divergence of the field. Similarly, for the conjugate momentum, for two distinct points,

$$\langle 0|\, \pi(x)\pi(x')\, |0\rangle = \frac{\cos(\Omega\Delta x) - \cos(\omega\Delta x)}{2\pi\Delta x^2} + \frac{\Omega\sin(\Omega\Delta x) - \omega\sin(\omega\Delta x)}{2\pi\Delta x}, \tag{6.14}$$

which in the coincidence limit becomes

$$\langle 0|\, \pi^2(x)\, |0\rangle = \frac{\Omega^2 - \omega^2}{4\pi}. \tag{6.15}$$

The equal-time correlation function between ϕ and π always vanishes: $\frac{1}{2}\langle 0|\, \{\phi(x), \pi(x')\}\, |0\rangle = 0$.

For the purposes of plotting these functions, we wish to remove the infrared divergence in order to distill the effect of the ultraviolet cutoff Ω. The π-π correlator (6.14) is not

infrared divergent, so for this function we can directly take the limit $\omega/\Omega \to 0$. The second term of Eq. (6.12) contains the infrared divergence since as $\omega/\Omega \to 0$ (for fixed Ω), we have

$$\mathrm{Ci}\left(\frac{\omega}{\Omega}\Omega\Delta x\right) \sim \gamma + \log(\Omega\Delta x) + \log(\omega/\Omega) + \mathcal{O}(\omega/\Omega). \tag{6.16}$$

Therefore, we see that if we add a counterterm $\log(\omega/\Omega)/(2\pi)$ to this correlator, we get an infrared-finite function

$$\langle 0|\, \phi(x)\phi(x')\, |0\rangle \mapsto \frac{1}{2\pi}[\mathrm{Ci}(\Omega\Delta x) - \gamma - \log(\Omega\Delta x)]. \tag{6.17}$$

However, if we want this function to agree with the non-bandlimited correlation function,

$$\langle 0|\, \phi(x)\phi(x')\, |0\rangle = \int_{\mathbb{R}} \frac{dk}{2\pi}\, \frac{1}{2|k|} e^{ik\Delta x} = \frac{-1}{2\pi}(\gamma + \log|\Delta x|), \tag{6.18}$$

as $\Omega \to \infty$, then we must add another $\log(\Omega)/(2\pi)$ to this function. Finally, we obtain:

$$\langle 0|\, \phi(x)\phi(x')\, |0\rangle = \frac{1}{2\pi}[\mathrm{Ci}(\Omega\Delta x) - \gamma - \log|\Delta x|], \tag{6.19}$$

and at a single point we have

$$\langle 0|\, \phi^2(x)\, |0\rangle = \frac{1}{2\pi}\log(\Omega). \tag{6.20}$$

In Figures 6.1 and 6.2 we present plots of the ϕ-ϕ and π-π correlators (respectively). These plots show that the correlations in both the ϕ and π variables decay with distance, but acquire oscillations on the order of the ultraviolet scale.

We will also calculate the correlators for a thermal state of the field. That is, we put each mode of the field in the thermal state (with inverse temperature $\beta := T^{-1}$),

$$\rho_k = \frac{1}{Z_k} \sum_{n_k=0}^{\infty} e^{-\beta\omega_k\left(n_k+\frac{1}{2}\right)}\, |n_k\rangle\, \langle n_k|\,. \tag{6.21}$$

Formally, the total density matrix for the field is then

$$\rho = \bigotimes_{\omega<|k|<\Omega} \rho_k. \tag{6.22}$$

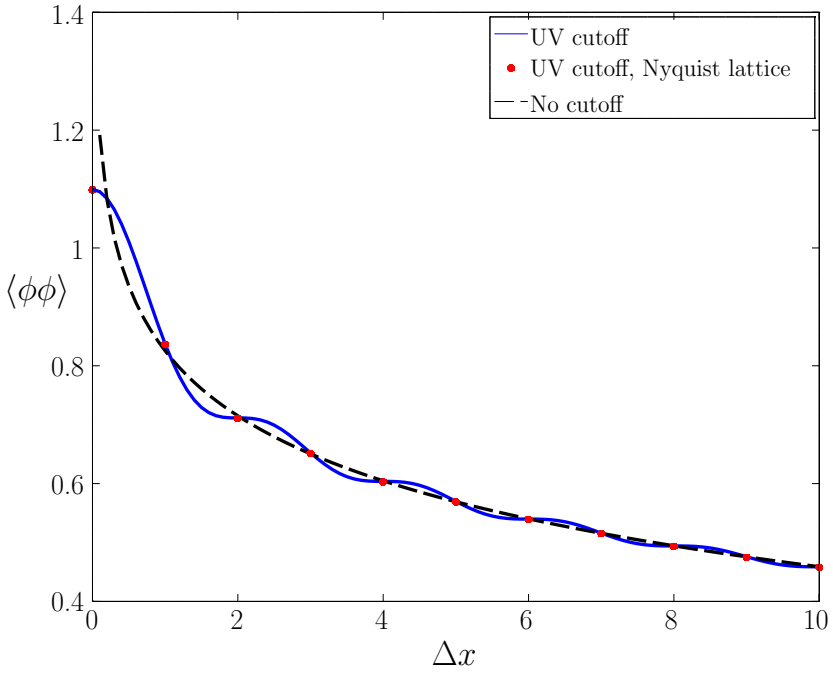

Figure 6.1: ϕ-ϕ correlations as a function of their separation. The horizontal axis is scaled by Ω/π so that integer values correspond to Nyquist spacings. The bandlimited correlations are blue with the Nyquist spacings indicated by red dots. The black dashed line shows the ultraviolet-divergent correlations without the ultraviolet bandlimit. We see that for points on the Nyquist lattice, the bandlimited correlators are in closer agreement to the correlation functions without the bandlimit. A counterterm of $\log(\omega)/(2\pi)$ is added to $\langle\phi\phi\rangle$ to cancel the infrared divergence.

From this density matrix, one can readily calculate the power spectra of ϕ and π as

$$\langle|\phi_k|^2\rangle = \mathrm{tr}(|\phi_k|^2\rho) = \frac{1}{\omega_k}\left(\frac{1}{e^{\beta\omega_k}-1}+\frac{1}{2}\right), \tag{6.23}$$

$$\langle|\pi_k|^2\rangle = \mathrm{tr}(|\pi_k|^2\rho) = \omega_k\left(\frac{1}{e^{\beta\omega_k}-1}+\frac{1}{2}\right). \tag{6.24}$$

Note that in the high-temperature limit $T/\Omega \gg 1$, we have

$$\frac{1}{e^{\beta\omega_k}-1} \approx \frac{1}{\beta\omega_k} - \frac{1}{2}, \tag{6.25}$$

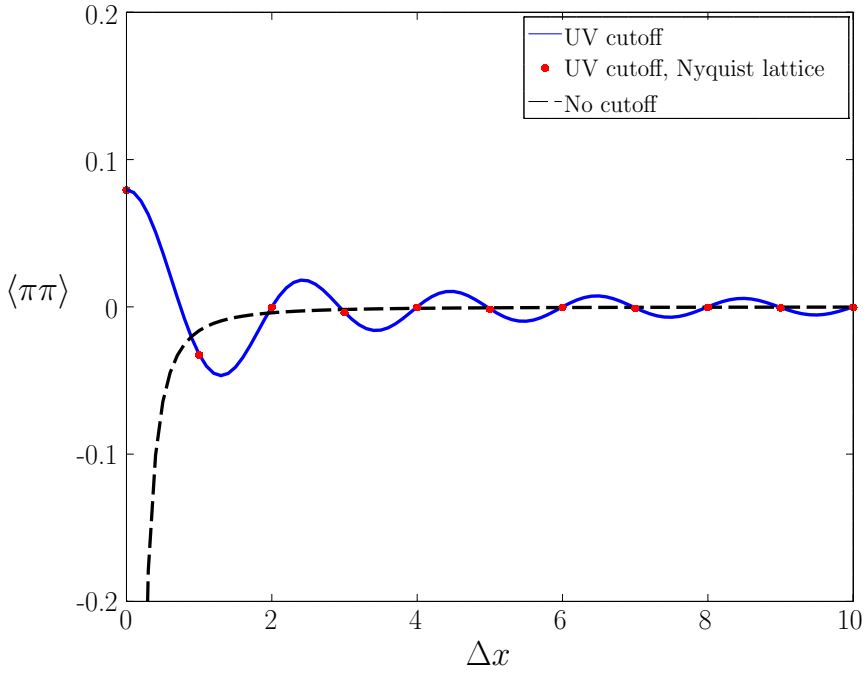

Figure 6.2: π-π correlations as a function of their separation. The horizontal axis is scaled by Ω/π so that integer values correspond to Nyquist spacings. The vertical axis is scaled by $1/\Omega^2$. The bandlimited correlations are blue with the Nyquist spacings indicated by red dots. The black dashed line shows the ultraviolet-divergent correlations without the ultraviolet bandlimit.

so that the power spectra become

$$\langle|\phi_k|^2\rangle_{T/\Omega\gg 1} = \frac{1}{\beta\omega_k^2}, \tag{6.26}$$

$$\langle|\pi_k|^2\rangle_{T/\Omega\gg 1} = \frac{1}{\beta}. \tag{6.27}$$

Here we see a recovery of the equipartition theorem, since

$$\begin{aligned}\langle H_k\rangle_{T/\Omega\gg 1} &= \frac{1}{2}\langle|\pi_k|^2\rangle_{T/\Omega\gg 1} + \frac{1}{2}\langle\omega_k^2|\phi_k|^2\rangle \\ &= \frac{1}{2}\frac{1}{\beta} + \frac{1}{2}\omega_k^2\frac{1}{\beta\omega_k^2} \\ &= T.\end{aligned} \tag{6.28}$$

We then use the fact that the correlation functions are obtained by the inverse Fourier transform of the power spectra. Thus, we obtain

$$\langle \phi(x)\phi(x')\rangle = \int_{\omega}^{\Omega} \frac{dk}{\pi} \cos(k\Delta x)\frac{1}{k}\left(\frac{1}{e^{\beta k}-1}+\frac{1}{2}\right), \tag{6.29}$$

$$\langle \pi(x)\pi(x')\rangle = \int_{\omega}^{\Omega} \frac{dk}{\pi} \cos(k\Delta x)k\left(\frac{1}{e^{\beta k}-1}+\frac{1}{2}\right). \tag{6.30}$$

Again, for the purposes of plotting these functions we shall add counterterms to remove the infrared divergence. As in the vacuum case, the π correlator is not divergent in the infrared. In the limit $\omega/\Omega \to 0$, the ϕ correlator goes as

$$\langle \phi(x)\phi(x')\rangle \sim \int_{\omega} \frac{dk}{\pi}\frac{1}{\beta k^2} = \frac{1}{\pi\beta\omega}. \tag{6.31}$$

Thus, we subtract the term $1/(\pi\beta\omega)$ from the ϕ correlator in order to remove the infrared divergence. Figures 6.3 and 6.4 plot these correlators as a function of Δx, along with the classical correlators calculated below. The correlators are plotted for various temperatures approaching the high temperature classical limit (the low temperature behaviour is similar to the vacuum correlations plotted in Figures 6.1 and 6.2 but with an overall shift due to the temperature). We see the expected convergence with the classical correlators as the temperature approaches $T/\Omega \to 1$.

6.1.3 A first look at Locality

We noted above that the field degrees of freedom are nonlocal due to the fact that the field at a generic point $x \in \mathbb{R}$ can be written in terms of the field variables on a sampling lattice as well as the observation that the commutation relations are nonzero at spacelike separation. One goal we would like to achieve is to localise a subset of N degrees of freedom, namely the field at N distinct points $\{\phi(x_n)\}_{n=1}^{N}$, to a volume of space. The localisation of field degrees of freedom will be considered in detail Section 6.4 using the results of the entropy calculations performed in Section 6.3. Here we will attempt to gain some insight by examining the set of first-quantised position vectors located at the points $\{x_n\}_n$, which we shall denote $\{|x_n)\}_{n=1}^{N}$. These position vectors are simply those considered in Section 3.2 as eigenvectors from various self-adjoint extensions of the bandlimited position operator.

For a general subsystem comprised of these N points, the vectors $\{|x_n)\}_{n=1}^{N}$ are not orthogonal since they are eigenvectors belonging to different self-adjoint extensions of the

position operator. However, any set of N distinct vectors spans an N-dimensional function space. It is possible to write down a projector onto this subspace, and as we shall show, it is possible to use this projector to construct a function which characterises the volume occupied by these N sample points. We begin by mapping the N position vectors to an orthogonal basis (though not uniquely):

$$|e_i) = \sum_j B_{ij}|x_j). \tag{6.32}$$

The projector onto this subspace can be written as

$$\begin{aligned} \mathbb{1}_N &:= \sum_i |e_i)(e_i| \\ &= \sum_{j,k} \left(\sum_i B_{ij}(B^\dagger)_{ik} \right) |x_j)(x_k|. \end{aligned} \tag{6.33}$$

Recall that the space of bandlimited functions has a reproducing kernel $K(x,x') = \frac{\Omega}{\pi}(x|x') = \frac{\Omega}{\pi}\,\mathrm{sinc}[(x-x')\Omega]$. It turns out we can use this reproducing kernel to express the matrix elements of the projector $\mathbb{1}_N$ in the nonorthogonal basis $\{|x_n)\}_{n=1}^N$ in the following way:

$$\begin{aligned} K(x_j,x_k) &:= \frac{\Omega}{\pi}(x_j|x_k) \\ &= \frac{\Omega}{\pi}\left(\sum_i (e_i|(B^{-1\dagger})_{ji} \right) \left(\sum_l (B^{-1})_{kl}|e_l) \right) \\ &= \frac{\Omega}{\pi} \sum_i (B^{-1\dagger})_{ji}(B^{-1})_{ki}. \end{aligned} \tag{6.34}$$

Viewing $K(x_j,x_k)$ as the $(j,k)^{th}$ element of an $N \times N$ matrix K_N, we can write

$$\mathbb{1}_N = \frac{\Omega}{\pi} \sum_{j,k} (K_N^{-1})_{kj}|x_j)(x_k|. \tag{6.35}$$

Also recall that this space admits an overcomplete resolution of identity:

$$\frac{\Omega}{\pi} \int_{\mathbb{R}} dx\, |x)(x| = \mathbb{1}. \tag{6.36}$$

Inserting this resolution of identity on either side of the projector, one can obtain the continuum representation

$$\mathbb{1}_N = \int_{\mathbb{R}^2} dx dx' \left(\frac{\Omega}{\pi} \sum_{jk} K(x, x_j)(K_N^{-1})_{kj} K(x_k, x') \right) |x)(x'|. \tag{6.37}$$

Now, if we think of $\mathbb{1}_N$ as an indicator function for the subspace spanned by the vectors $\{|x_n)\}_{n=1}^N$, then the trace of the operator will correspond to the amount of volume occupied by this subspace in the continuum $x \in \mathbb{R}$. Since the trace is performed by integrating over the diagonal elements of $\mathbb{1}_N$, we propose that these diagonal elements, as a function of x, can be used to characterise the spatial profile of this subspace. Therefore, the proposal is that the function

$$f(x) := \frac{\Omega}{\pi} \sum_{jk} K(x, x_j)(K_N^{-1})_{kj} K(x_k, x) \tag{6.38}$$

is the spatial profile corresponding to the degrees of freedom located at the points $\{x_n\}_{n=1}^N$.

It is now possible to visualise this spatial profile for various configurations of N points. Figure 6.5 illustrates this for $N = 5$ equally spaced points at various values of the spacing. We notice that when the spacing between the points is larger than the Nyquist spacing (1 in the plot), each degree of freedom occupies a volume centred at the point and on the order of a Nyquist spacing in size. When the spacing between the points approaches the Nyquist spacing, the points converge to form a volume of length 5 (in Nyquist units, i.e., units such that the Nyquist spacing is 1). Surprisingly, when the spacing between the points is decreased below the Nyquist spacing, the volume remains the same size as it was at the Nyquist spacing. This result can be anticipated, since the trace of the operator $\mathbb{1}_N$ is always N because it is a projector onto an N-dimensional subspace. However, it would mean that N closely spaced points would describe the same volume of space independently of their precise positions within this volume (only somewhat independently, since this volume will always be centred at the centre of the points). We shall return to examine this behaviour in Section 6.4, where we will find that this spatial profile is indeed that which characterises these degrees of freedom.

6.2 Entropy for Classical Analog

Before proceeding to calculate the entanglement entropy of subsystems of the bandlimited quantum fields, we first consider a classical analog in order to see how the bandlimit affects

the entropy already at the classical level. For the classical case, we will examine bandlimited fields drawn from some statistical distribution and calculate the entropy of a subsytem comprised of a field located at N sample points after marginalising over the remaining degrees of freedom of the field. For example, for bandlimited Gaussian white noise, N samples on a Nyquist lattice represent N uncorrelated degrees of freedom. Therefore, the entropy will simply be additive and thus the entropy will be proportional to N. In order to compare with the quantum calculations for the entropy, the statistical distribution which will be of interest here is a thermal distribution of a 1+1-dimensional Klein-Gordon Hamiltonian.

It is convenient to begin with the Hamiltonian in momentum space, with the ultraviolet and infrared cutoffs imposed,

$$H[\tilde{\phi},\tilde{\pi}] = \frac{1}{2}\int_{\omega<|k|<\Omega}\frac{dk}{2\pi}\left(|\tilde{\pi}(k)|^2 + k^2|\tilde{\phi}(k)|^2\right). \tag{6.39}$$

The Boltzmann probability distribution for inverse temperature β is given by

$$p[\tilde{\phi},\tilde{\pi}] = \frac{1}{Z}e^{-\beta H[\tilde{\phi},\tilde{\pi}]}. \tag{6.40}$$

One can easily obtain the power spectra of the fields from this probability distribution,

$$\langle|\tilde{\phi}(k)|^2\rangle = \frac{1}{\beta\omega_k^2}, \text{ and} \tag{6.41}$$

$$\langle|\tilde{\pi}(k)|^2\rangle = \frac{1}{\beta}, \tag{6.42}$$

which correspond to the high temperature limit of the quantum power spectra (Eq.(6.26)). We can also obtain the correlation functions by taking the inverse Fourier transform of the power spectra:

$$\langle\phi(x)\phi(x')\rangle = \frac{1}{\beta}\left[\frac{\cos(\omega\Delta x)}{\pi\omega} - \frac{\cos(\Omega\Delta x)}{\pi\Omega}\right] + \frac{\Delta x}{\beta}\left[\mathrm{Si}(\omega\Delta x) - \mathrm{Si}(\Omega\Delta x)\right], \tag{6.43}$$

$$\langle\pi(x)\pi(x')\rangle = \frac{1}{\beta}\left[\frac{\Omega}{\pi}\,\mathrm{sinc}(\Omega\Delta x) - \frac{\omega}{\pi}\,\mathrm{sinc}(\omega\Delta x)\right]. \tag{6.44}$$

These are the classical correlators plotted in Figures 6.3 and 6.4 alongside the quantum correlators. An infrared counterterm of $1/(\pi\beta\omega)$ is subtracted from the ϕ-ϕ correlator in these plots to remove the infrared divergence, as it was for the quantum ϕ-ϕ correlator. Thus the plotted function is

$$\langle\phi(x)\phi(x')\rangle = \frac{-1}{\beta}\left[\frac{\cos(\Omega\Delta x)}{\pi\Omega} + \Delta x\,\mathrm{Si}(\Omega\Delta x)\right]. \tag{6.45}$$

The limit $\omega/\Omega \to 0$ (fixed Ω) was also taken in the π-π classical correlator for the plot.

We can change the degrees of freedom of the system from momentum space to sample space (where the degrees of freedom are the field amplitudes on a sampling lattice) via the transformation:

$$\begin{aligned} \phi(x_n) &= \int_{\omega<|k|<\Omega} \frac{dk}{2\pi} e^{ikx_n} \tilde{\phi}(k), \\ \pi(x_n) &= \int_{\omega<|k|<\Omega} \frac{dk}{2\pi} e^{ikx_n} \tilde{\pi}(k). \end{aligned} \tag{6.46}$$

Since this is a linear change of variables and the momentum space probability distribution (Eq. (6.40)) is Gaussian, then the sample space probability distribution will also be Gaussian. Thus, we can write the distribution as

$$p(\{\phi_n, \pi_n\}_n) = \frac{1}{Z} e^{-\frac{1}{2}\sum_{m,n} \pi_m (\sigma_\pi^{-1})_{mn} \pi_n - \frac{1}{2}\sum_{m,n} \phi_m (\sigma_\phi^{-1})_{mn} \phi_n}, \tag{6.47}$$

where we have written $\phi_n := \phi(x_n)$ and $\pi_n := \pi(x_n)$. The elements of the matrices σ_ϕ and σ_π are determined from Eqs. (6.43) and (6.44):

$$(\sigma_\phi)_{mn} := \langle \phi_m \phi_n \rangle, \tag{6.48}$$

$$(\sigma_\pi)_{mn} := \langle \pi_m \pi_n \rangle. \tag{6.49}$$

For the entropy calculation of N samples of the field, we seek the reduced probability distribution after marginalising over all but these degrees of freedom. In order to perform such a marginalisation for any state, we must first split the phase space into a subspace describing these N degrees of freedom and a complementary subspace describing the rest. As we shall see, this is not automatic in the bandlimited theory because the degrees of freedom are nonlocal. The procedure of splitting the phase space will be to perform a change of variables which splits the Poisson bracket and symplectic form into a direct sum of the brackets and forms on the subspace of the N samples and a complementary set of degrees of freedom. First, it will be convenient to write the phase space variables in a vector $\vec{X} = (\phi_1, \pi_1, \phi_2, \pi_2, \dots)$, where the first N indices correspond to the samples in the subsystem we wish to keep. Then the Poisson brackets can be encoded in the anti-symmetric matrix $\Lambda_{ij} := \{X_i, X_j\}_{PB}$, where

$$\begin{aligned} \{\phi_i, \pi_j\}_{PB} &= \frac{\Omega}{\pi} \operatorname{sinc}(\Omega \Delta x_{ij}) - \frac{\omega}{\pi} \operatorname{sinc}(\omega \Delta x_{ij}), \\ \{\phi_i, \phi_j\}_{PB} &= \{\pi_i, \pi_j\}_{PB} = 0. \end{aligned} \tag{6.50}$$

Note that these are consistent with the matrix elements as obtained for the quantum commutation relations (Eq. (6.8)). Next, we write the Poisson bracket in block matrix form,

$$\Lambda = \begin{bmatrix} \alpha & \eta \\ -\eta^T & \gamma \end{bmatrix}, \tag{6.51}$$

where $\alpha^T = -\alpha$ and $\gamma^T = -\gamma$, ensuring that $\Lambda^T = -\Lambda$. The blocks are split so that the upper and left portions run through the indices $1, \ldots, 2N$, corresponding to the ϕ and π samples for the N sample points of the subsystem. The remaining indices correspond to samples taken at the remaining points of any sampling lattice from which the field can be perfectly reconstructed.

Now consider a change of phase space variables $\vec{X}' = Q\vec{X}$, where

$$Q := \begin{bmatrix} I & 0 \\ \eta^T\alpha^{-1} & I \end{bmatrix}. \tag{6.52}$$

The transformed Poisson bracket is

$$\Lambda' := Q\Lambda Q^T = \begin{bmatrix} \alpha & 0 \\ 0 & \eta^T\alpha^{-1}\eta + \gamma \end{bmatrix}. \tag{6.53}$$

We see that the Poisson bracket is now split into a subspace consisting of the first N degrees of freedom and a complementary subspace. Note that the Poisson bracket elements between degrees of freedom in the first subspace has not changed, i.e., $\alpha \overset{Q}{\mapsto} \alpha$. Furthermore, we also note that the complementary subsystem does not generally correspond to a complementary set of lattice points, but is simply a complementary set of degrees of freedom which characterise the remainder of the entire system. After the transformation Q, the symplectic form of the system is

$$\Lambda^{-1} = \sum_{n,m=1}^{N} (\alpha^{-1})_{nm} d\phi'_n \wedge d\pi'_n + \sum_{n,m\notin\{1,\ldots,N\}} [(\eta^T\alpha^{-1}\eta + \gamma)^{-1}]_{nm} d\phi'_n \wedge d\pi'_n. \tag{6.54}$$

Note that this procedure requires the matrix α to be invertible. For a generic set of lattice points in the bandlimited theory this is always the case. Note that we can write $\alpha_{ij} = K(x_i, x_j) = (\Omega/\pi)(x_i|x_j)$. If α were not invertible, then we could find a vector $\vec{v}$ such that

$$0 = \vec{V}^T\alpha\vec{V} = (\Omega/\pi)\sum_{ij}(x_i|V_iV_j|x_j) = (\Omega/\pi)\|\sum_i V_i|x_i)\|^2. \tag{6.55}$$

Thus, $\sum_i V_i|x_i) = 0$. However, this would imply that for any bandlimited function, $|f)$, we would have

$$0 = (f|\sum_i V_i|x_i) = \sum_i V_i f^*(x_i). \tag{6.56}$$

As shown in Ref.[50], given a finite set of points x_i and target values y_i, it is possible to find a bandlimited function $|f)$ such that $f(x_i) = y_i$. Therefore, we have a contradiction and so α must be invertible.

Now that we have split the phase space and have a symplectic form, we must choose a measure for the phase space before we can perform the marginalisation over the degrees of freedom complementary to the N samples. As we noted in Chapter 5, a natural measure to choose is that obtained from the symplectic volume form. Recall that the ambiguity in the definition of the entropy is resolved with the choice of this measure. We shall also include a factor of $1/(2\pi)$ for each dimension of the phase space so that the entropy matches the high temperature limit of the von Neumann entropy calculated in the quantum setting. Therefore, we find that the reduced probability distribution for the N samples is

$$\begin{aligned} p_{\text{red}}(\{\phi_n, \pi_n\}_{n=1}^N) &= \int \det(\eta^T \alpha^{-1} \eta + \gamma)^{-1} \left(\prod_{i \notin \{1,\ldots,N\}} \frac{d\phi_i' d\pi_i'}{2\pi} \right) p(\{\phi_n, \pi_n\}_n) \\ &= \frac{1}{Z_{\text{red}}} e^{-\frac{1}{2}\sum_{m,n=1}^N \pi_m (\sigma_\pi|_N^{-1})_{mn} \pi_n - \frac{1}{2}\sum_{m,n=1}^N \phi_m (\sigma_\phi|_N^{-1})_{mn} \phi_n}. \end{aligned} \tag{6.57}$$

Note that by construction of the matrix Q to perform the phase space splitting, the samples $1, \ldots, N$ are not changed, i.e., $\phi_n \overset{Q}{\mapsto} \phi_n' = \phi_n$. Therefore, the statistics of these variables do not change after marginalising:

$$\begin{aligned} \langle \phi_m \phi_n \rangle &= \int \det(\Lambda)^{-1} \left(\prod_i \frac{d\phi_i d\pi_i}{2\pi} \right) \phi_m \phi_n p(\{\phi_n, \pi_n\}_n) \\ &= \int \det(\alpha)^{-1} \left(\prod_{i \in \{1,\ldots,N\}} \frac{d\phi_i' d\pi_i'}{2\pi} \right) \phi_m \phi_n \\ &\qquad \times \int \det(\eta^T \alpha^{-1} \eta + \gamma)^{-1} \left(\prod_{i \notin \{1,\ldots,N\}} \frac{d\phi_i' d\pi_i'}{2\pi} \right) p(\{\phi_n, \pi_n\}_n) \\ &= \int \det(\alpha)^{-1} \left(\prod_{i \in \{1,\ldots,N\}} \frac{d\phi_i' d\pi_i'}{2\pi} \right) \phi_m \phi_n p_{\text{red}}(\{\phi_n, \pi_n\}_{n=1}^N), \end{aligned} \tag{6.58}$$

and similar for $\langle \pi_m \pi_n \rangle$. Therefore, we see explicitly that the matrices $\sigma_\phi|_N$ and $\sigma_\pi|_N$ are simply the matrices σ_ϕ and σ_π with indices restricted to the subspace $1, \ldots, N$.

We can now calculate the entropy of the reduced probability distribution. The partition function is

$$\begin{aligned} Z_{\text{red}} &= \int \det(\alpha)^{-1} \left(\prod_{i=1}^{N} \frac{d\phi_i d\pi_i}{2\pi} \right) e^{-\frac{1}{2}\sum_{m,n=1}^{N} \pi_m (\sigma_\pi|_N^{-1})_{mn} \pi_n - \frac{1}{2}\sum_{m,n=1}^{N} \phi_m (\sigma_\phi|_N^{-1})_{mn} \phi_n} \\ &= \det(\alpha)^{-1} \sqrt{\det\left(\sigma_\pi|_N^{-1}\right) \det\left(\sigma_\phi|_N^{-1}\right)}. \end{aligned} \tag{6.59}$$

Therefore the entropy is

$$\begin{aligned} S &= \int \det(\alpha)^{-1} \left(\prod_{i=1}^{N} \frac{d\phi_i d\pi_i}{2\pi} \right) p_{\text{red}}(\{\phi_n, \pi_n\}_{n=1}^{N}) \, \log\left[p_{\text{red}}(\{\phi_n, \pi_n\}_{n=1}^{N})\right] \\ &= N + \log\left[\det(\alpha)^{-1} \sqrt{\det\left(\sigma_\pi|_N^{-1}\right) \det\left(\sigma_\phi|_N^{-1}\right)} \right]. \end{aligned} \tag{6.60}$$

We see that the coupling of the degrees of freedom appearing in the second term causes the entropy to deviate from simply scaling exactly proportional to N as in the bandlimited white noise case. The entropy is plotted in Figure 6.6 for the massless, bandlimited Klein-Gordon field in a thermal state. The plot shows the entropy how the entropy scales with the number N of Nyquist samples at various temperatures. We see that the dominant behaviour is the linear scaling with the number of points N.

It is possible make the entropy formula (6.60) look more like the entropy formula for quantum Gaussian states shown in Section 5.3. First, we permute the indices in the subspace $1, \ldots, N$ so that all of the ϕ_n's appear before the π_n's. Then the Poisson bracket matrix can be written as

$$\Lambda|_N := \begin{bmatrix} 0 & \lambda \\ -\lambda & 0 \end{bmatrix}, \tag{6.61}$$

$$\Sigma|_N := \begin{bmatrix} \sigma_\phi|_N^{-1} & 0 \\ 0 & \sigma_\pi|_N^{-1} \end{bmatrix}, \tag{6.62}$$

where λ is simply the matrix α after the index permutation. It is then straightforward to verify that the entropy can be written as

$$\begin{aligned} S &= N + \log\left[\sqrt{\det(\Lambda|_N^{-1} \Sigma|_N)} \right] \\ &= \sum_{i=1}^{N} (1 + \log(d_i)), \end{aligned} \tag{6.63}$$

where $\{d_i\}_{i=1}^N$ are the positive imaginary parts of the eigenvalues of $\Lambda|_N^{-1}\Sigma|_N$ which come in pairs $\pm id_i$ (i.e., the symplectic eigenvalues of $\Sigma|_N$).

We can compare this formula to the formula obtained in Section 5.3. Recall that the corresponding quantum formula for the entropy is:

$$S(\rho) = \sum_j f(d_j), \tag{6.64}$$

where

$$f(d_j) := (d_j + 1/2)\log(d_j + 1/2) - (d_j - 1/2)\log(d_j - 1/2). \tag{6.65}$$

When the symplectic eigenvalues are large (for example, this happens for thermal states at high temperature), we find

$$f(d_j) \approx 1 + \log(d_j). \tag{6.66}$$

Thus we see agreement with the classical formula. This is also illustrated in Figure 6.7, which compares the classical and quantum formulas for one symplectic eigenvalue (plotted in the x-axis).

6.3 Entropy in Bandlimited Quantum Field Theory

We shall now proceed to perform the analogous entropy calculation for N samples of the bandlimited quantum field theory. However, first we must generalise the Gaussian state method presented in Section 5.3 to deal with situations where the commutation relations between the degrees of freedom are not canonical (e.g., in the bandlimited theory). This is accomplished in Subsection 6.3.1.

Then we will calculate the entropy of N samples of the field in various situations. Since in sampling theory, the samples one takes are not confined to any particular lattice, we are free to choose samples from any configuration of points as long as the Beurling density of the lattice exceeds the Nyquist rate. If we are only interested in the positions of a finite number N of sample points, then we can place them anywhere in $\mathbb{R}$ by choosing the complementary set of lattice points appropriately. We will calculate the entropy for N samples taken at the Nyquist density (Nyquist sampling), below the Nyquist density (undersampling), and above the Nyquist density (oversampling). In particular, we are interested in the scaling of the entropy with N in different sampling density regimes. In Section 6.4, we will use the scaling behaviour to make conclusions about the locality of the degrees of freedom in the bandlimited theory.

6.3.1 Hilbert space factorisation

Suppose we choose a sampling lattice $\{x_n\}_n$ on which we represent the fields by their samples $\{\phi_n := \phi(x_n), \pi_n := \pi(x_n)\}_n$. Let us identify a subset of these points $n \in \{1, \ldots, N\}$ and denote the corresponding subsystem as W_N. Let us also write the subsystem corresponding to the lattice points $n \notin \{1, \ldots, N\}$ by $W_{\overline{N}}$.

Recall that in the case where we have a bipartite system $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, the entanglement entropy of the subsytem A for a state ρ on $\mathcal{H}$ can be calculated by performing the partial trace on subsystem B,

$$\rho_A := \mathrm{tr}_B\, \rho, \tag{6.67}$$

and calculating the von Neumann entropy of ρ_A,

$$S(\rho_A) := -\,\mathrm{tr}(\rho_A \log \rho_A). \tag{6.68}$$

However, the Hilbert space of the bandlimited theory, $\mathcal{H}$, does not factor into the Hilbert spaces for the two subsystems W_N and $W_{\overline{N}}$. That is,

$$\mathcal{H} \neq \mathcal{H}_{W_N} \otimes \mathcal{H}_{W_{\overline{N}}}. \tag{6.69}$$

This is because, as we found in Subsection 6.1.2, the commutation relations for the bandlimited Klein-Gordon field were nonlocal at the cutoff scale:

$$[\phi(x), \pi(x')] = i\left(\frac{\Omega}{\pi}\,\mathrm{sinc}(\Omega\Delta x) - \frac{\omega}{\pi}\,\mathrm{sinc}(\omega\Delta x)\right), \tag{6.70}$$

$$[\phi(x), \phi(x')] = 0, \text{ and} \tag{6.71}$$

$$[\pi(x), \pi(x')] = 0, \tag{6.72}$$

One way of seeing this is by taking the point of view of algebraic quantum field theory (see, e.g., Ref.[34]). From this perspective, we define our fundamental observables $\{\phi_n, \pi_n\}_n$ which, along with the commutation relations (6.70), generates an algebra. The Hilbert space $\mathcal{H}$ is then simply a Hilbert space upon which we represent this algebra.

(Note that when we have infinitely many degrees of freedom the Stone-von Neumann theorem breaks down so there exist unitarily inequivalent representations of the same algebra [34]. The full range of practical consequences of this fact in quantum field theory is still unknown. However, since here we will only be dealing with finite dimensional subsystems this fact does not concern us.)

We can now see that the Hilbert space of the field $\mathcal{H}$ does not factor as $\mathcal{H}_{W_N} \otimes \mathcal{H}_{W_{\overline{N}}}$, where $\mathcal{H}_{W_N}$ is the Hilbert space of the algebra generated by $\{\phi_n, \pi_n\}_{n=1}^N$ and$\mathcal{H}_{W_{\overline{N}}}$ is the

Hilbert space of the algebra generated by $\{\phi_n, \pi_n\}_{n\notin\{1,\dots,N\}}$. This is because if it could be factored in this way, then for ϕ_m in subsystem W_N and π_n in subsystem $W_{\overline{N}}$ we could write the action on $\mathcal{H}_{W_N} \otimes \mathcal{H}_{W_{\overline{N}}}$ as

$$\phi_m \leftrightarrow \phi_m \otimes \mathbb{1}, \tag{6.73}$$

$$\pi_n \leftrightarrow \mathbb{1} \otimes \pi_n. \tag{6.74}$$

However, by the definition of the tensor product, this would imply that ϕ_m and π_n commute, but for points x_n and x_m which do not lie on a Nyquist lattice they clearly should not because of Eq. (6.70). In the special case where we remove the infrared cutoff ω and take all of the points on a Nyquist lattice $\{x_n^{(\alpha)} = (2\pi n - \alpha)/(2\Omega)\}_n$ (for fixed $\alpha \in [0, 2\pi)$) we note that we obtain canonical commutation relations. Therefore, in this case we do get a Hilbert space factorisation $\mathcal{H} = \otimes_n \mathcal{H}_n$. However, for a generic sampling lattice this will not be the case.

This Hilbert space factorisation issue is similar to a situation occurring in lattice gauge theory. If we choose a subset of lattice points N, with corresponding algebra of observables $\mathcal{A}_N$, and the complementary set of lattice points with algebra of observables $\mathcal{A}_{\overline{N}}$, then the total algebra of the gauge field does not factor:

$$\mathcal{A} \neq \mathcal{A}_N \otimes \mathcal{A}_{\overline{N}}. \tag{6.75}$$

However, in gauge theory the factorisation is prevented due to the fact that the algebra $\mathcal{A}_N$ has a non-trivial centre due to nonlocal constraints [17, 34], and the resolution of this problem is more complicated than the situation we have here.

The fact that the Hilbert space for the bandlimited theory does not factor can be resolved simply by finding a subsystem $\overline{W_N}$ complementary to W_N in such a way that we do get a factorisation

$$\mathcal{H} = \mathcal{H}_{W_N} \otimes \mathcal{H}_{\overline{W_N}}. \tag{6.76}$$

Generally the subsystem $\overline{W_N}$ will not correspond to a set of sample points, but this will allow us to define the entanglement between the subsystem W_N and the rest of the system. This factorisation can be accomplished in the same manner that we used to split the phase space in the classical scenario in Section 6.2. Let us write the fields in a vector $\vec{X} = (\phi_1, \pi_1, \phi_2, \pi_2, \dots)$, so that we can define the matrix of commutators $i\Lambda_{mn}\mathbb{1} := [\phi_m, \pi_n]$. Then we will use the matrix

$$Q := \begin{bmatrix} I & 0 \\ \eta^T \alpha^{-1} & I \end{bmatrix} \tag{6.77}$$

to perform a change of variables, $\vec{X}' = Q\vec{X}$, so that the new commutator matrix takes the form:

$$\Lambda' := Q\Lambda Q^T = \begin{bmatrix} \alpha & 0 \\ 0 & \eta^T\alpha^{-1}\eta + \gamma \end{bmatrix}. \tag{6.78}$$

We see that we have split the observables into two subsystems whose observables commute, so that we have the desired tensor factorisation. Note also that the variables in the subsystem W_N have not changed under Q.

Now that we have the factorisation, by Darboux's theorem, we can perform another change of variables in each subspace so that in these subspaces the observables obey the canonical commutation relations. Let this transformation be represented as the matrix

$$T = \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix}, \tag{6.79}$$

so that the commutator and correlator matrices become

$$\Lambda'' := T\Lambda' T^T = \bigoplus_i \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{6.80}$$

$$\Sigma'' := T\Sigma' T^T. \tag{6.81}$$

If the state we are interested in is Gaussian, we can now employ the procedure presented in Section 5.3 for calculating the entanglement entropy of W_N. First, we restrict the commutator and correlator matrices to the subspace corresponding to W_N,

$$\Lambda''|_N = T_1\Lambda|_N T_1^T \tag{6.82}$$

$$\Sigma''|_N = T_1\Sigma|_N T_1^T. \tag{6.83}$$

Then we must find the symplectic eigenvalues of $\Sigma''|_N$. Note that

$$\begin{aligned} \{\pm id_j\}_{j=1}^N &= \mathrm{spec}(\Lambda''|_N^{-1}\Sigma''|_N) \\ &= \mathrm{spec}(T_1^{-T}\Lambda|_N^{-1}T_1^{-1}T_1\Sigma|_N T_1^T) \\ &= \mathrm{spec}(\Lambda|_N^{-1}\Sigma|_N) \end{aligned} \tag{6.84}$$

where $\mathrm{spec}(A)$ denotes the spectrum of the operator A. Therefore, we see that we do not need to explicitly perform the change of variables using the matrices Q and T, since the symplectic eigenvalues can simply be determined from the original matrices Λ and Σ, restricted to the subspace of samples which defined W_N. The phase space method for calculating the entanglement entropy of Gaussian states thus generalises easily to the case where the commutation relations between the degrees of freedom are not canonical.

6.3.2 Nyquist sampling

First we calculate the entanglement entropy for N contiguous Nyquist spaced samples in a thermal state. For this purpose, we simply need the matrix of commutators and correlators between the N sample points. These matrix elements were calculated above in Eqs. (6.8), (6.29), and (6.30). The entropy as a function of the number of sample points is illustrated in Figure 6.8 for a range of temperatures. The plot shows a transition between linear scaling of the entropy with N at high temperatures, similar to the classical case, and logarithmic scaling of the entropy at low temperatures.

The logarithmic scaling behaviour for the ground state is a well-known result from conformal field theories (such as the massless 1+1-dimensional Klein-Gordon field). The expected scaling behaviour of the entanglement entropy for a region of length L in a conformal field theory is $S \sim \frac{1}{3}\log(\Omega L)$ [42]. For the ground state (i.e., zero temperature thermal state) data plotted in Figure 6.8, the fitted curve we obtain is

$$S(N) = 0.334 \log N + 3.25, \tag{6.85}$$

which is in agreement with the expected result.

6.3.3 Undersampling

As we mentioned above, we are free to place a finite number N of sample points anywhere in $\mathbb{R}$. We will first consider the case when adjacent sample points are equidistantly spaced by a distance Δx. When the samples are pushed farther than a Nyquist spacing apart, i.e., $\Delta x > \pi/\Omega$, it will be referred to as undersampling the field in this region. If the samples are pushed together so that $\Delta x < \pi/\Omega$, it will be referred to as oversampling the field in this region. Here we will consider the ground state of the field in the case of undersampling, and in the next subsection we will examine the case of oversampling.

Since the entanglement entropy is dominated by local correlations of the field, when the samples are taken far from one another, the correlations of each sample contributing to the entanglement entropy will be roughly independent. Therefore, we expect that we should recover a volume law scaling behaviour for the entropy (i.e., the entropy should scale proportionally to the number of sample points) when the samples are placed far apart, similar to the volume law obtained at high temperatures.

Since the ground state commutator and correlator matrices (Eqs. (6.8), (6.12), and (6.14)) were calculated for arbitrary spacings Δx, the entropy calculation is formally the

same as in the Nyquist spacing case. Figure 6.9 shows the dependence of the entanglement entropy of N points as a function of N for various choices of spacing Δx.

In the figure, we see that for spacings at or close to the Nyquist spacing, the entropy scales logarithmically with the number of points. We also see a transition to a linear scaling behaviour (volume law) as the spacing between adjacent samples approaches 2 Nyquist spacings. This supports the intuition that the entanglement entropy is generated by local correlations.

6.3.4 Oversampling

In the case of oversampling ($\Delta x < \pi/\Omega$), a similar entropy calculation as performed in the previous subsections becomes numerically unstable. One can only reliably calculate the entropy for sample spacings between 97%-100% of the Nyquist spacing. Throughout this range, we continue to see the logarithmic scaling behaviour of the entanglement entropy, with fitted curves,

$$S(N) = c_0 \log N + c_1 \tag{6.86}$$

where in each case $c_0 \in [0.333, 0.335]$ and $c_1 = 3.25$. (This result is for an infrared scale of $\omega/\Omega = 10^{-300}$.)

Despite this numerical instability, we are able to extend arbitrarily far into the oversampling regime if we only consider a small number of points. In Figure 6.10, we show the entanglement entropy for 5 sample points as a function of their separation. In this plot, we see that in the oversampling regime the entropy of the sample points does not depend significantly on the particular value of the spacing. This suggests that the entropy retains the logarithmic scaling behaviour when oversampling.

Furthermore, the plateau behaviour of the entropy is neither simply a feature of the ground state nor only a feature of the quantum nature of the field. In Figure 6.11, we show the results of a similar calculation of the entropy for 5 samples of a thermal state of the field as a function of their separation for various temperatures. The result of the analogous classical calculation is also shown. Again, one sees that in the oversampling regime the behaviour of the entropy does not depend significantly on the spacing between adjacent points.

6.3.5 Derivative sampling

In this subsection we will examine an alternative method for studying the behaviour in the oversampling regime based on sampling the field and its spatial derivatives at a single

point.In classical sampling theory, it is known that instead of only sampling the amplitudes of a function, if the derivatives of the function are also sampled, then the samples need only be taken at a fraction of the Nyquist rate [41]. For instance, if one samples the function and its first $N-1$ derivatives, then one only needs to take samples at the rate $\Omega/(N\pi)$.

If we consider a subsystem of two points, then we can perform the following symplectic transformation on the phase space variables:

$$\begin{pmatrix} \phi(x) \\ \phi(x+\Delta x) \\ \pi(x) \\ \pi(x+\Delta x) \end{pmatrix} \rightarrow \begin{pmatrix} \frac{1}{\sqrt{2}}(\phi(x+\Delta x)+\phi(x)) \\ \frac{1}{\sqrt{2}}(\phi(x+\Delta x)-\phi(x)) \\ \frac{1}{\sqrt{2}}(\pi(x+\Delta x)+\pi(x)) \\ \frac{1}{\sqrt{2}}(\pi(x+\Delta x)-\pi(x)) \end{pmatrix}. \tag{6.87}$$

Thus, we see that sampling two points in the limit where their separation vanishes, $\Delta x \rightarrow 0$, can equivalently be viewed as sampling the field and its derivative at a single point (as well as the conjugate momentum and its derivative at that point). This suggests that sampling the field and its first $N-1$ derivatives at a single point, $\{\phi, \phi', \phi'', \ldots, \phi^{(N-1)}\}$, (as well as the conjugate momentum and its derivatives) is equivalent to sampling N amplitudes of the field which are very closely spaced. This could be considered an extreme case of oversampling the field.

The entropy calculation can be performed in the same manner as before, with the following commutation relations and correlation functions between these new variables:

$$[(\partial_x)^n\phi(x),(\partial_x)^m\pi(x)] = \begin{cases} \frac{i}{\pi}(-1)^{\frac{n+3m}{2}}\frac{1}{n+m+1}\left(\Omega^{n+m+1}-\omega^{n+m+1}\right) & \text{if } (n+m)=0 \mod 2 \\ 0 & \text{if } (n+m)=1 \mod 2 \end{cases}, \tag{6.88}$$

$$\langle(\partial_x)^n\phi(x)\cdot(\partial_x)^m\phi(x)\rangle = \begin{cases} (-1)^{\frac{n+3m}{2}}\frac{1}{2\pi}\log\left(\frac{\Omega}{\omega}\right) & \text{if } n+m=0 \\ (-1)^{\frac{n+3m}{2}}\frac{1}{2\pi}\frac{1}{n+m}\left(\Omega^{n+m}-\omega^{n+m}\right) & \text{if } n+m\neq 0, (n+m)=0 \mod 2 \\ 0 & \text{if } (n+m)=1 \mod 2 \end{cases}, \tag{6.89}$$

and

$$\langle(\partial_x)^n\pi(x)\cdot(\partial_x)^m\pi(x)\rangle = \begin{cases} (-1)^{\frac{n+3m}{2}}\frac{1}{2\pi}\frac{1}{n+m+2}\left(\Omega^{n+m+2}-\omega^{n+m+2}\right) & \text{if } (n+m)=0 \mod 2 \\ 0 & \text{if } (n+m)=1 \mod 2 \end{cases}. \tag{6.90}$$

The results of this calculation are illustrated in Figure 6.12. The figure presents the results of the derivative sampling calculation alongside the curve obtained from the Nyquist spaced sample calculation for the ground state. We see that both curves exhibit the same leading order behaviour $S \sim \frac{1}{3} \log N$, though differ by a constant offset. However, we are only interested in the leading order scaling behaviour of the entropy since the constant term is not universal.

The observation that the entanglement entropy also scales as $\frac{1}{3} \log N$ in this extreme case of oversampling supports the conclusion that the scaling behaviour of the entanglement entropy is the same throughout the oversampling regime.

6.4 Locality of Information

In the previous section, we saw that the scaling behaviour of the entanglement entropy depends on the spacing between adjacent samples. When the samples are placed far apart, the entropy exhibits a volume law. When the samples are placed at a Nyquist spacing or closer, the entropy exhibits a logarithmic scaling behaviour. Although the degrees of freedom corresponding to sample points are nonlocal, as we have discussed, we would like to determine to what extent they can be localised. In this section, we will use the scaling behaviour of the entropy in different spacing regimes to provide evidence that one can think of the degrees of freedom as occupying a volume of space on the order of a Nyquist spacing in size, which we shall call a Planck volume.

Suppose each degree of freedom does in fact occupy a Planck volume centred at the sample point. Then what behaviour would we expect for the entanglement entropy of a set of sample points? Since entanglement entropy can be thought to be generated by correlations at the boundary of a region traced out, when the samples are placed far apart, these correlations will mainly be with degrees of freedom in the complementary subsystem. Therefore in this regime we should indeed expect a volume law for the scaling of the entanglement entropy. When the samples are placed at the Nyquist spacing, since the Planck volumes are approximately a Nyquist spacing in size, we should expect that N contiguous samples would cover a single interval of N Nyquist spacings in length. In the non-bandlimited 1+1-dimensional massless scalar field theory, the entanglement entropy of such an interval grows logarithmically with the length of the interval, thus we should expect our observed logarithmic scaling of the entanglement entropy with the number of samples N.

The case of oversampling, however, is peculiar. If we were to place N sample points with a spacing smaller than the Nyquist spacing, then we would expect that the N Planck

volumes centred at each sample point should overlap. However, even when the points are very close together (e.g., in the case of derivative sampling), we find that the entropy grows logarithmically in the number of samples. This seems to contradict the notion that bandlimited fields have finite information density, since if we take N samples within a single Planck volume, we are still probing N degrees of freedom.

We will now show that there is no contradiction: the N samples spaced below the Nyquist spacing are actually probing N independent Planck volumes of space nonlocally. The setup is as follows. First, we take N samples which are equidistantly spaced, which we will call subsystem W_N. Then we take an additional sample point, denoted by subsystem P, which we shall use to probe the locality of the subsystem W_N. How will this be accomplished? Consider the mutual information between subsystems W_N and P:

$$I(W_N : P) := S(W_N) + S(P) - S(W_N, P), \tag{6.91}$$

where $S(W_N)$ and $S(P)$ are the entanglement entropies of subsystems W_N and P (respectively) and $S(W_N, P)$ is the entanglement entropy of the combined system. When the probe point P is placed far from the subsystem W_N, it will not share significant correlations with W_N, thus we will have:

$$S(W_N, P) \approx S(W_N) + S(P), \tag{6.92}$$

so the mutual information will be low. If we begin to move the probe point closer to the subsystem W_N, then they will begin to share correlations, so the entanglement entropy $S(W_N, P)$ should decrease and the mutual information $I(W_N : P)$ should increase. Therefore, we should be able to use the mutual information $I(W_N : P)$ as an indicator to determine whether the probe point P is close to W_N. We can then plot the mutual information as a function of the position of the probe point P for various spacings between the samples in W_N to determine the effective volume of space occupied by subsystem W_N.

A numerical demonstration of this procedure is shown in Figure 6.13. We see that indeed for spacings larger than a Nyquist spacing, the sample points occupy a Planck volume centred at the sample point. At the Nyquist spacing, we see that the sample points occupy a single interval of approximately length N. Surprisingly, we see that as the spacing between the samples in W_N is decreased below the Nyquist spacing, the occupied volume does not change. This would indicate that the N samples continue to occupy a volume of length N, regardless of their precise positions within this volume (although the volume will always be centred at the centre of the sample points). In this sense, the Planck volumes described by the N degrees of freedom are incompressible. Therefore, we see that there is consistency with the notion that the field contains finite information density, since

if one attempts to probe the field at two points closer than a Nyquist spacing apart, then one will simply access the same degrees of freedom as in the case where they are spaced at the Nyquist spacing.

We also note the similarity of the mutual information plot in Figure 6.13 with the spatial profile of Figure 6.5.

Also recall that the plateau behaviour below the Nyquist spacing that we observed in the entropy was also a feature of the classical theory. It may not be surprising that the classical bandlimited theory also exhibits the same localisation properties as the quantum theory because the peculiar properties of the degrees of freedom we have observed are due to the new kinematical structure introduced by the bandlimitation. As we mentioned in Section 4.3, much of the kinematical structure of quantum mechanics is inherited directly from the phase space structure of the corresponding classical theory. This nonlocality of the degrees of freedom is evidently one of those features.

6.5 Infrared behaviour

Whereas above we were concerned with the ultraviolet behaviour of the entanglement entropy, here we briefly discuss the impact of the infrared cutoff. In Refs.[16, 59] it was determined that for an infrared cutoff ω in a 1+1-dimensional free, scalar field, the entanglement entropy acquires a subleading double-logarithm in ω:

$$S = \frac{1}{3}\log(N) + \frac{1}{2}\log\left(\log\left(\frac{1}{\omega}\right)\right) - \log\pi. \tag{6.93}$$

We see explicitly that the entanglement entropy is infrared divergent, since the double-log term diverges as $\omega \to 0$.

Using the above methods, we can calculate the entanglement entropy of the bandlimited Klein-Gordon field with a varying infrared cutoff ω. This calculation was done by performing a numerical fit to the curve

$$S(N) = c_1 \log N + c_2. \tag{6.94}$$

For the infrared cutoff varying throughout the range $\omega/\Omega \in [10^{-50}, 10^{-300}]$, we find that the leading order behaviour of c_2 is

$$c_2 = \frac{1}{2}\log\left(\log\left(\frac{\Omega}{\omega}\right)\right), \tag{6.95}$$

which agrees with the result of [16, 59]. As demonstrated in Ref.[16], we see that the cutoff ω has a similar effect on the entanglement entropy as a mass.

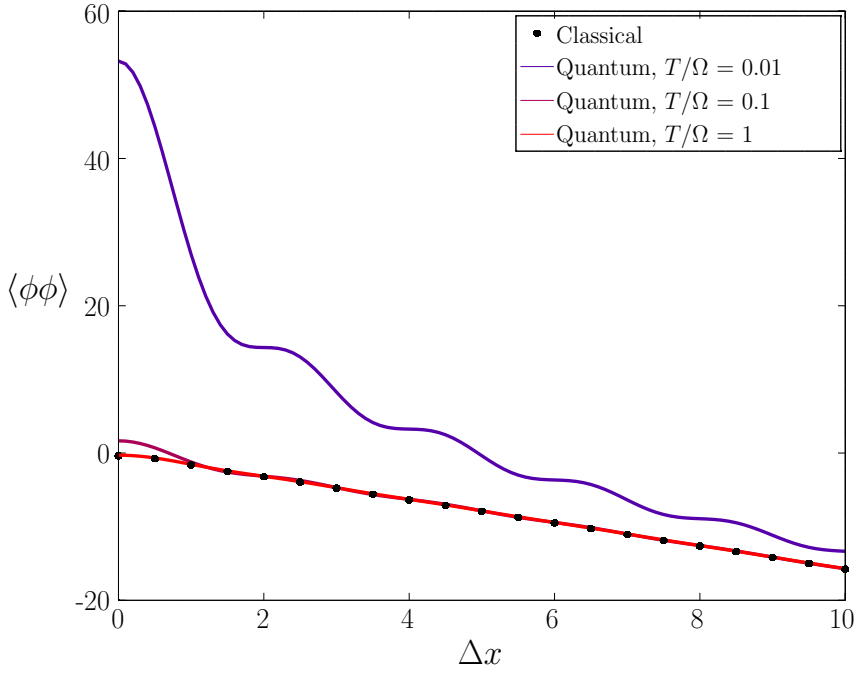

Figure 6.3: Classical and quantum ϕ-ϕ correlations as a function of their separation at various temperatures. The horizontal axis is scaled by Ω/π so that integer values correspond to Nyquist spacings. The vertical axis is scaled by Ω/T. The classical correlator is plotted as black dots. The scaling of the vertical axis absorbs all of the temperature dependence of the classical correlator (since it grows proportionally to T/Ω), thus the single graph of the classical correlator completely characterises its behaviour. The quantum correlators are shown as lines for temperatures up to $T/\Omega = 1$, at which point the quantum correlator converges to the classical correlator.

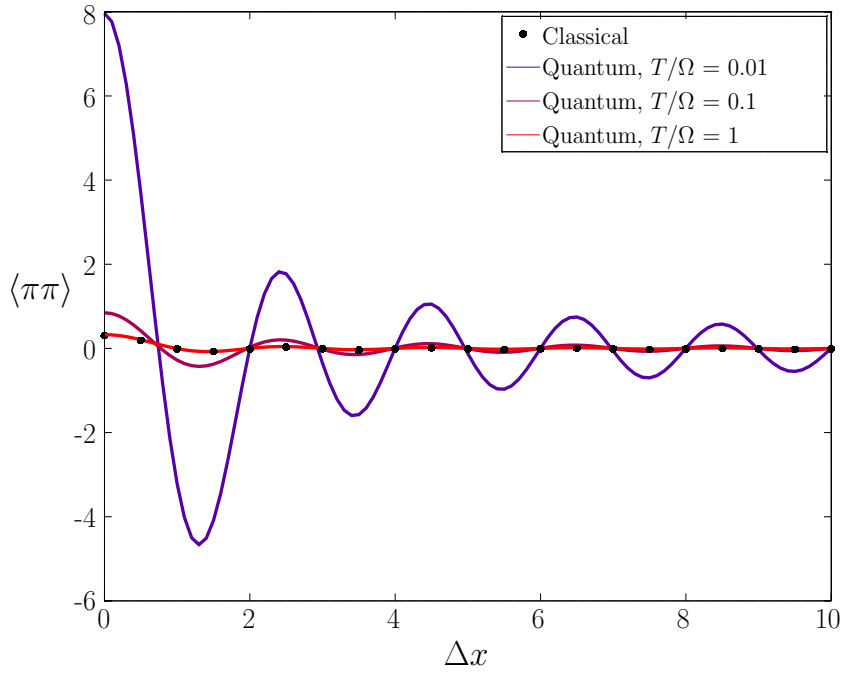

Figure 6.4: Classical and quantum π-π correlations as a function of their separation at various temperatures. The horizontal axis is scaled by Ω/π so that integer values correspond to Nyquist spacings. The vertical axis is scaled by Ω/T. The classical correlator is plotted as black dots. The scaling of the vertical axis absorbs all of the temperature dependence of the classical correlator (since it grows proportionally to T/Ω), thus the single graph of the classical correlator completely characterises its behaviour. The quantum correlators are shown as lines for temperatures approaching $T/\Omega = 1$, at which point the quantum correlator converges to the classical correlator.

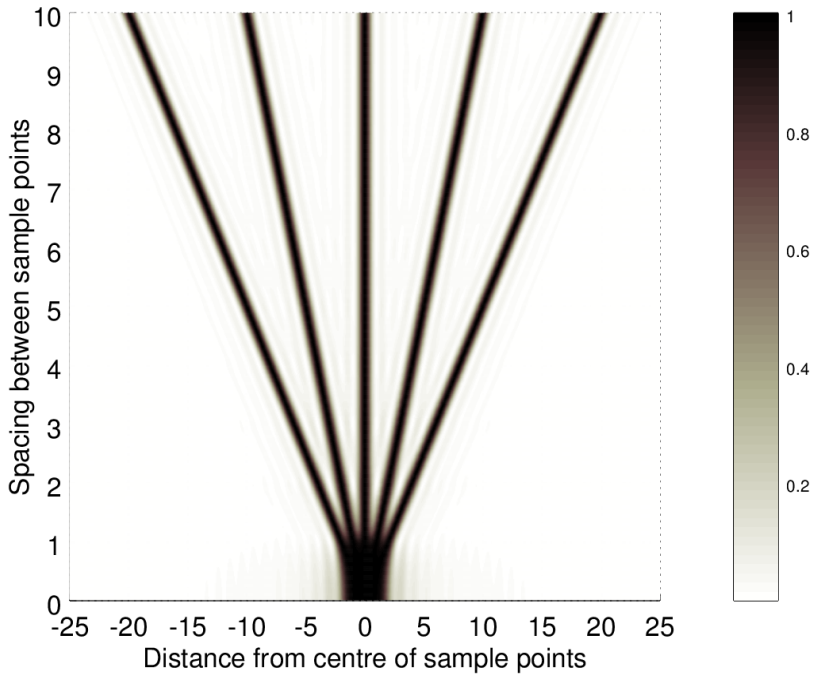

Figure 6.5: Spatial profile (in the horizontal axis) for 5 degrees of freedom as a function of their spacing (vertical axis). Both axes are scaled so that the cutoff length, or Nyquist spacing, π/Ω is equal to 1. We see that for spacings above the Nyquist spacing, the degrees of freedom occupy a region consisting of 5 disjoint intervals of length ~ 1 surrounding each sample point. Below the Nyquist spacing, the degrees of freedom merge to occupy a single interval of length ~ 5. This interval does not decrease in size even as the sampling points are taken on top of one another.

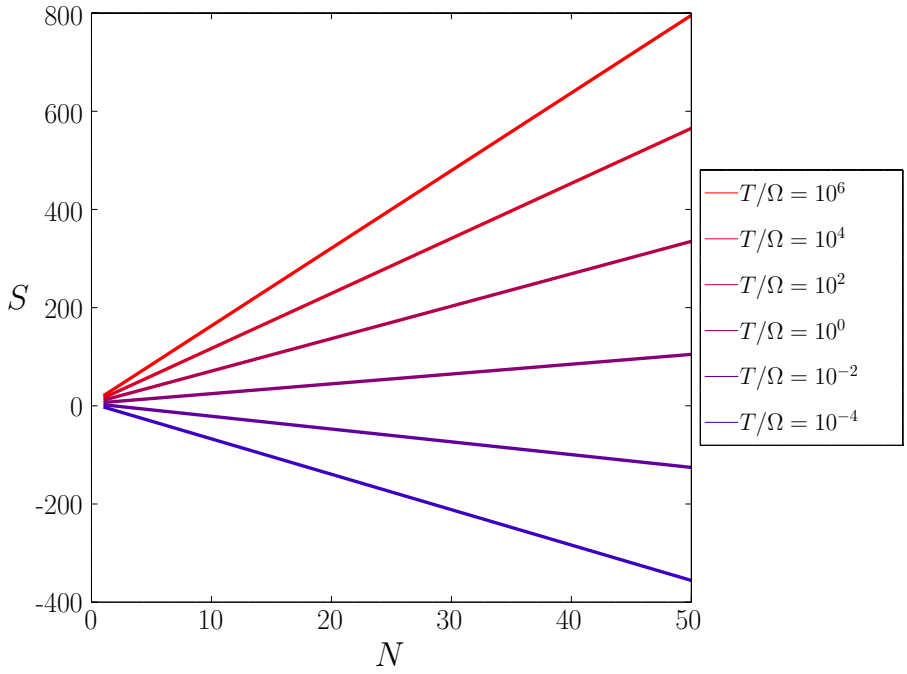

Figure 6.6: Entropy of a set of sample points for a thermally-distributed classical bandlimited Klein-Gordon field for several temperatures. We see that the entropy grows linearly with the number of points, illustrating a volume law. Notice also that this figure shows negative entropy at low temperatures, which simply reflects the fact that continuous probability distributions can have negative entropy. Physically, this represents a breakdown of the classical approximation in statistical mechanics at low temperatures.

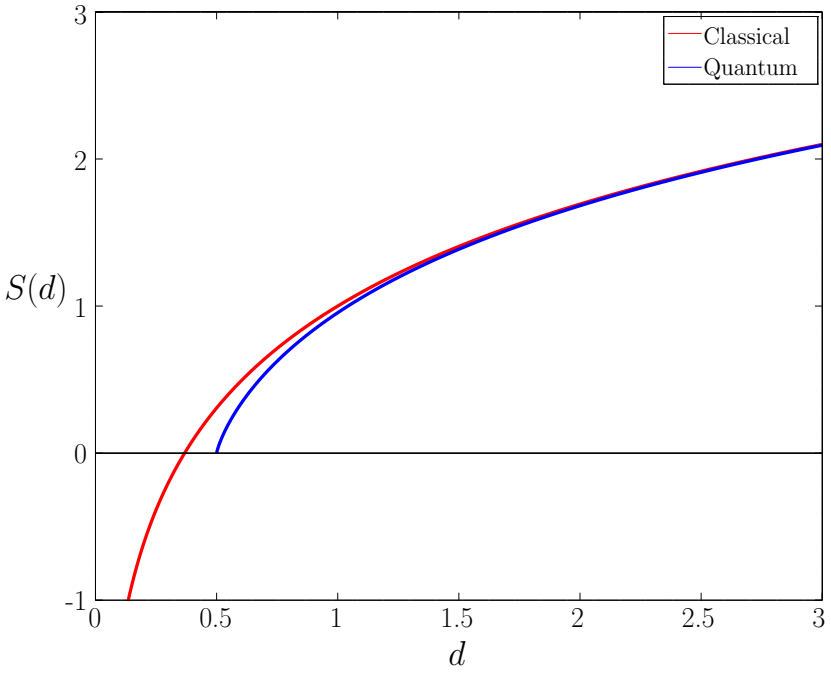

Figure 6.7: The entropy of a harmonic oscillator as a function of the symplectic eigenvalue, both classically and quantum-mechanically. In quantum mechanics, the uncertainty principle requires $d \geq 1/2$, which is saturated by the vacuum state for which the entropy is zero. Classically, the symplectic eigenvalue can be any positive number, but the entropy becomes negative for small d. This is a consequence of the fact that the entropy of a continuous probability distribution is not bounded from below.

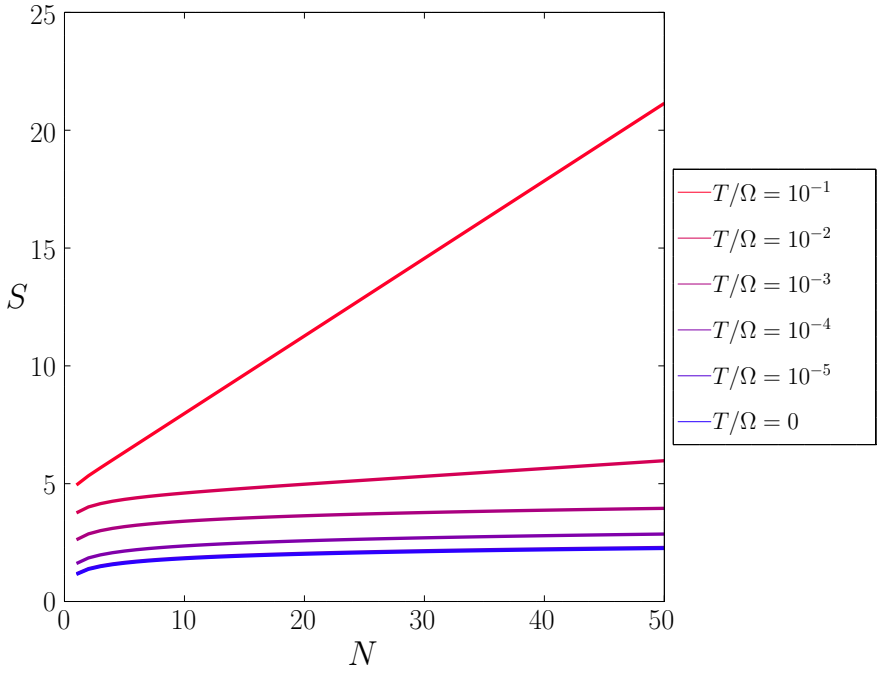

Figure 6.8: Entropy of a set of sample points for a thermal state of a quantum Klein-Gordon field. The plot illustrates the transition between the logarithmic behaviour at low temperature to linear behaviour at higher temperatures where the entropy becomes an extensive variable. The infrared scale in this plot is $\omega/\Omega = 10^{-300}$.

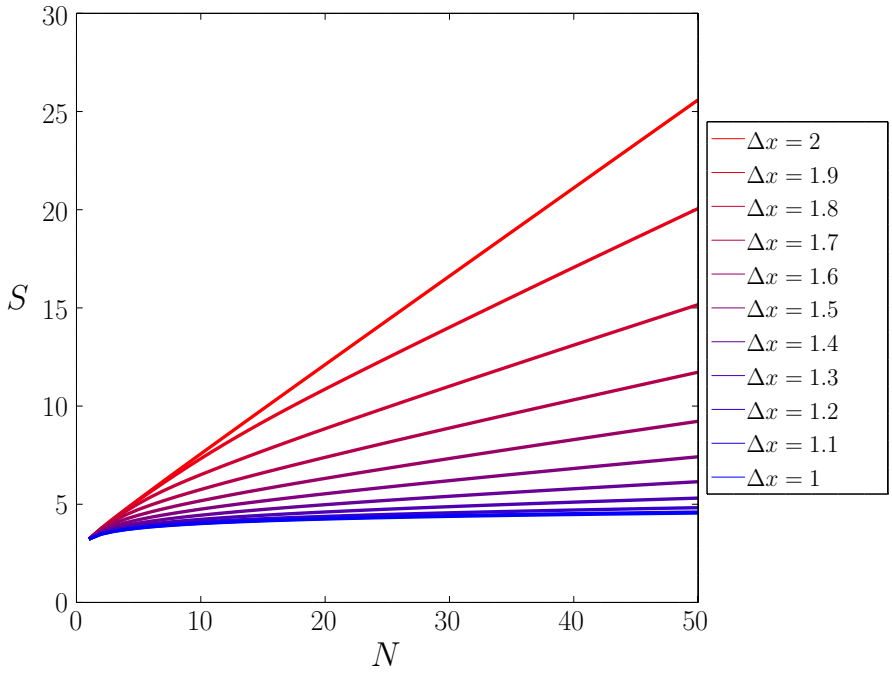

Figure 6.9: Entanglement entropy dependence on number of points traced out for sample spacings between 1 and 2 Nyquist spacings. The Δx labels in the legend are scaled by Ω/π so that the Nyquist spacing is 1. We see that once the adjacent point spacing has reached twice the Nyquist spacing, the entropy has transitioned from the logarithmic scaling law to a volume law. The infrared scale is $\omega/\Omega = 10^{-300}$.

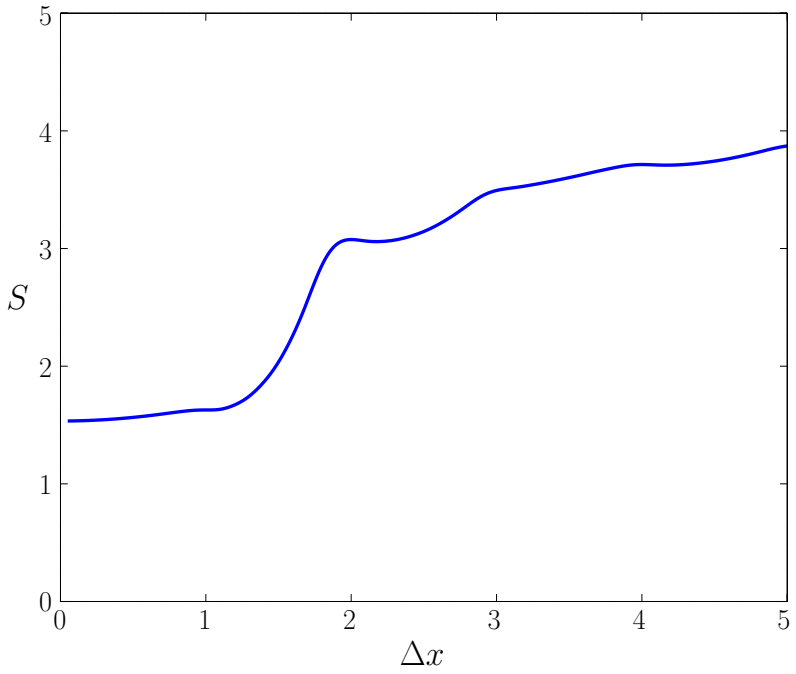

Figure 6.10: Entanglement entropy of five points as a function of the spacing between them. The horizontal axis is scaled by Ω/π so that the Nyquist spacing is 1. For spacings below the Nyquist spacing, we see a plateauing effect indicating that the entanglement entropy is not sensitive to the spacing between the points for spacings below the Nyquist spacing. Above the Nyquist spacing the entropy tends to increase as the spacing increases, due to the fact that as the points are placed farther apart, the local correlations are increasingly with the complementary subsystem. In this plot the infrared to ultraviolet ratio is $\omega/\Omega = 10^{-5}$.

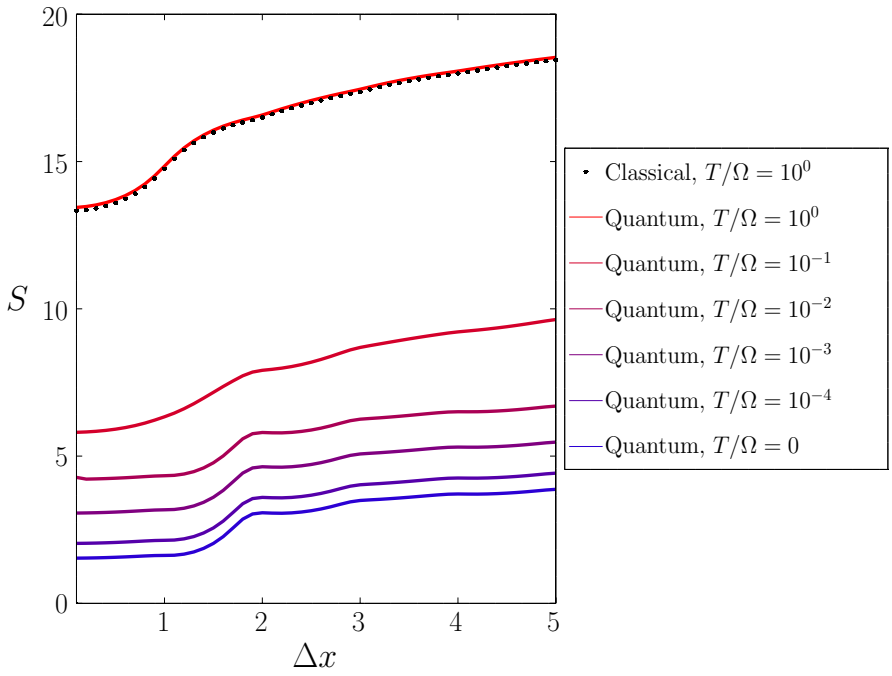

Figure 6.11: Entanglement entropy of five points as a function of the spacing between them at various temperatures. The horizontal axis is scaled by Ω/π so that the Nyquist spacing is 1. We see here a plateau in the entropy at small spacings, similar to the plateau in the ground state entropy. The analogous classical calculation is also shown as black dots for temperature $T/\Omega = 1$. At temperatures $T/\Omega > 1$, the entropy behaves the same as for $T/\Omega = 1$ but shifted vertically by a constant $\sim \log(T/\Omega)$. In this plot the infrared to ultraviolet ratio is $\omega/\Omega = 10^{-5}$.

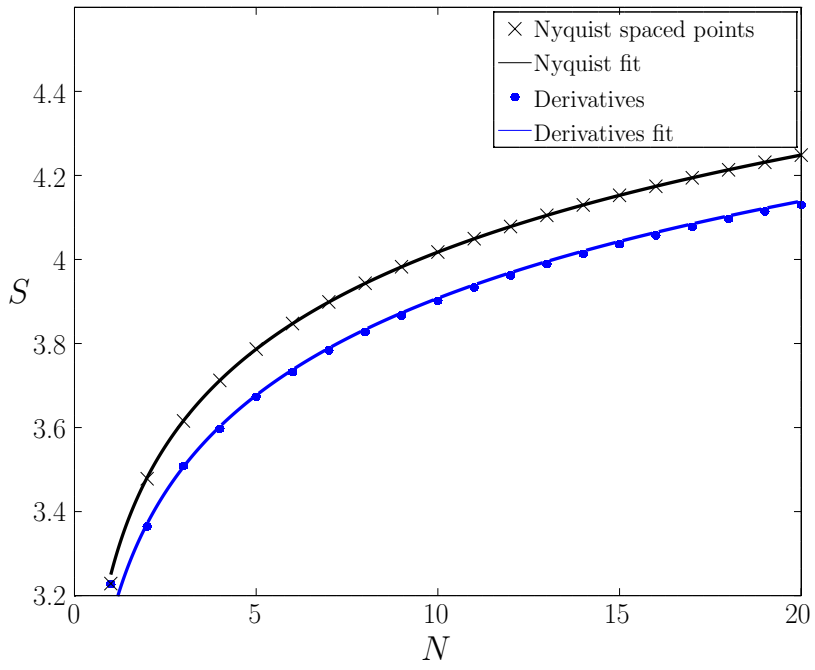

Figure 6.12: Entanglement entropy dependence of number of derivatives traced out at a single point. Fitted curve to Nyquist spaced points is $S(N) = 0.334\log(N)+3.25$, where N is the number of sampling points. Fitted curve to derivative sampling points is $S(N) = \frac{1}{3}\log(N) + 3.14$, where $N-1$ is the number of derivatives sampled (which corresponds to N sampled points). We see that both curves differ by a constant ≈ 0.11. The infrared to ultraviolet ratio is $\omega/\Omega = 10^{-300}$.

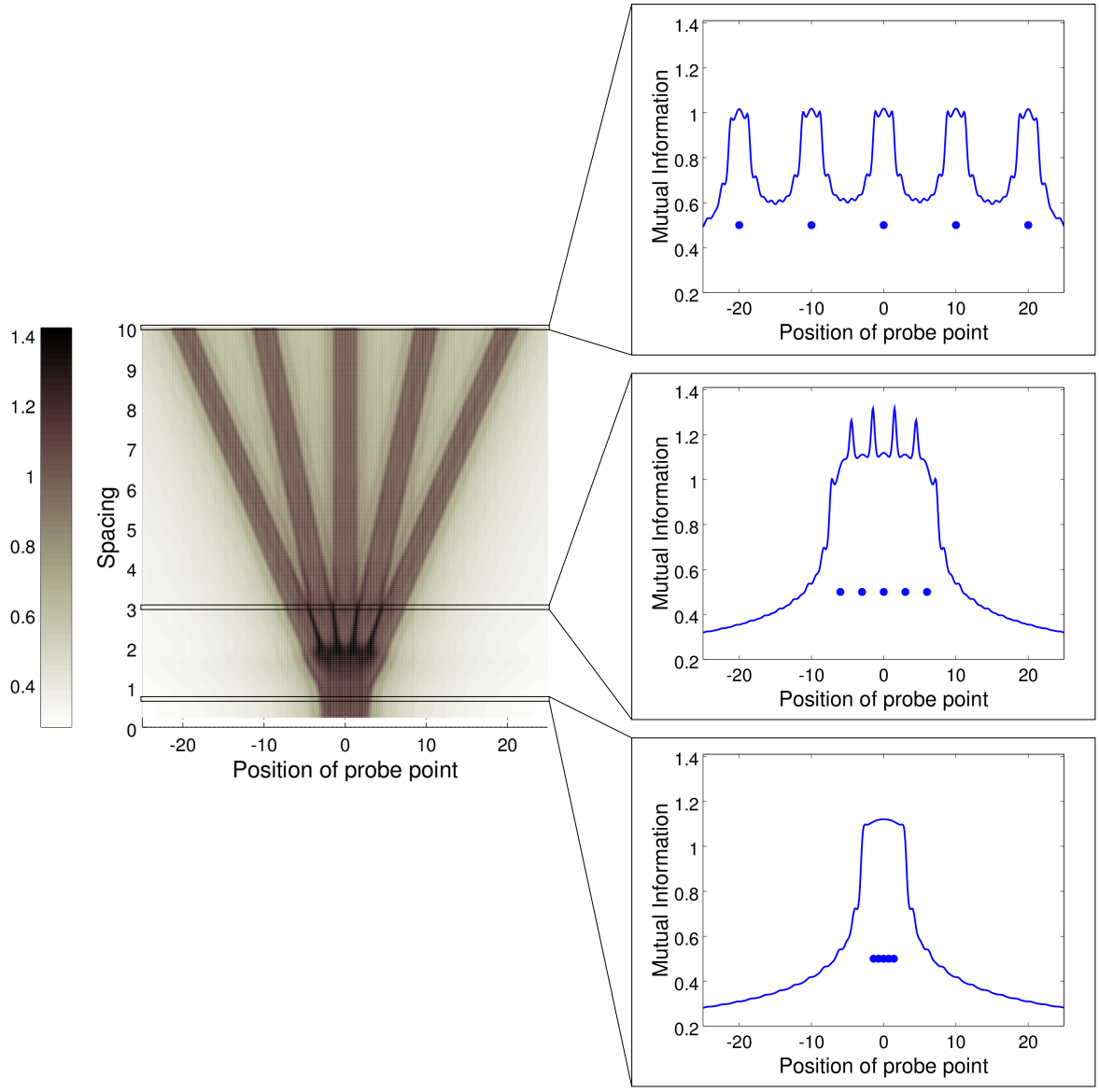

Figure 6.13: Mutual information between subsystem A_N and probe point P. The spatial axes are scaled so that the Nyquist spacing is 1. For spacings much larger than the Nyquist spacing, the degrees of freedom occupy independent intervals of the order of the Nyquist spacing in size. For spacings at or below the Nyquist spacing, the points in the interval describe a fixed volume of size N.

Chapter 7

Conclusions and Future Work

In this work we applied Shannon's sampling theory to free, scalar quantum fields in order to obtain a quantum field theory which does not suffer from ultraviolet divergences, yet retains Euclidean symmetries. However, this came with the price of introducing nonlocality to the degrees of freedom of the field, which manifested itself through nonlocal commutation relations. By studying the scaling behaviour of the entanglement entropy, we were able to determine that each degree of freedom of the bandlimited field occupies and independent, incompressible Planck-scale volume in space.

There are many directions in which one could take this work and much research to be done regarding the role of locality in quantum gravity at large. The tools discussed herein are amenable to further studies along these lines.

One direction would be to examine the robustness of the qualitative behaviour of the nonlocality of the degrees of freedom to changes in the cutoff model. For example, one could implement a softer cutoff by only exponentially suppressing large wavelengths. In this case, one does not obtain a sampling theorem, but it would still be possible to apply the Gaussian formalism for calculating the entanglement entropy from the two-point functions.

We mentioned above that a bandlimited interacting theory would introduce non-unitarity to the temporal evolution. However, the motivation for imposing a bandlimit on the fields was due to the strong interaction of the fields with gravity at very high energies. Therefore, the non-unitarity is a consequence of the fact that we are not explicitly modelling these interactions (which would in principle require a working theory of quantum gravity). One could then use this non-unitarity of the interacting theories as a feature of the bandlimited theory since the qualitative features of this non-unitarity may provide some insight into the nature of these gravitational interactions and thus the quantum nature of gravity. For

instance, the Gaussian methods used in this work could be applied to quadratic couplings between different fields to investigate the nature of this non-unitarity.

There is an important question as to the nature of the locality of the degrees of freedom of a bandlimited field on curved spacetimes as well as spacetimes of higher dimension. Since the bandlimited quantum field theory model is based on the conjecture that gravity will somehow cause the field to become bandlimited, a natural next step would be to understand bandlimitation when one begins to turn on gravity.

A direction of great interest is to perform the cutoff in such a way that it preserves Lorentz symmetry as well as the spatial Euclidean symmetries. The idea proposed in [49] is to apply a bandlimit to the spectrum of a covariant operator, such as the d'Alembertian. The bandlimited functions would then be limited to those in this subspace.

Another potential application is in regards to studies of other nonlocal or redundant encoding of degrees of freedom such as that proposed in Ref. [4] for quantum error correction in the context of AdS/CFT. The nature of bulk locality is a big question in this field, and the kind of study done in this work may be applicable to a connection to holography based on wavelets [70, 83].

References

[1] Gerardo Adesso and Fabrizio Illuminati. Entanglement in continuous variable systems: Recent advances and current perspectives. *J.Phys.*, A40:7821–7880, 2007.

[2] Yakir Aharonov, Jeeva Anandan, Sandu Popescu, and Lev Vaidman. Superpositions of time evolutions of a quantum system and a quantum time-translation machine. *Phys. Rev. Lett.*, 64:2965–2968, Jun 1990.

[3] N. I. Akhiezer and I. M. Glazman. *Theory of Linear Operators in Hilbert Space*. Dover Publications, Inc., New York, 1993.

[4] Ahmed Almheiri, Xi Dong, and Daniel Harlow. Bulk Locality and Quantum Error Correction in AdS/CFT. *JHEP*, 1504:163, 2015.

[5] K. Audenaert, J. Eisert, M. B. Plenio, and R. F. Werner. Entanglement properties of the harmonic chain. *Phys. Rev. A*, 66:042327, Oct 2002.

[6] J. Barrett and S. Pironio. Popescu-rohrlich correlations as a unit of nonlocality. *Phys. Rev. Lett.*, 95:140401, Sep 2005.

[7] J. J. Benedetto and P. J. S. G. Ferreira, editors. *Modern Sampling Theory*. Birkhauser, Boston, 2001.

[8] C. Beny. *Information Flow at the Quantum-Classical Boundary*. PhD thesis, University of Waterloo, Waterloo, Ontario, Canada, 2008.

[9] M V Berry. Evanescent and real waves in quantum billiards and gaussian beams. *J. Phys. A: Math. Gen.*, 27(11):L391.

[10] Eugenio Bianchi and Robert C. Myers. On the Architecture of Spacetime Geometry. 2012.

[11] Luca Bombelli, Joe Henson, and Rafael D. Sorkin. Discreteness without symmetry breaking: A Theorem. *Mod.Phys.Lett.*, A24:2579–2587, 2009.

[12] Luca Bombelli, Rabinder K. Koul, Joohan Lee, and Rafael D. Sorkin. Quantum source of entropy for black holes. *Phys. Rev. D*, 34(2):373–383, 1986.

[13] G. Bonneau, J. Faraut, and G. Valent. Self-adjoint extensions of operators and the teaching of quantum mechanics. *American Journal of Physics*, 69:322–331, 2001.

[14] Jr. Callan, Curtis G. and Frank Wilczek. On geometric entropy. *Phys.Lett.*, B333:55–61, 1994.

[15] L. Carleson, P. Malliavin, J. Neuberger, and J. Wermer, editors. *The collected works of Arne Beurling*, volume 2. Birkhauser, Boston, MA, 1989.

[16] H. Casini and M. Huerta. Entanglement entropy in free quantum field theory. *J.Phys.*, A42:504007, 2009.

[17] Horacio Casini, Marina Huerta, and José Alejandro Rosabal. Remarks on entanglement entropy for gauge fields. *Phys. Rev. D*, 89:085012, Apr 2014.

[18] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2nd edition, 2006.

[19] L. Crispino, A. Higuchi, and G. Matsas. The unruh effect and its applications. *Reviews of Modern Physics*, 80:787–838.

[20] T. Demarie. Pedagogical introduction to the entropy of entanglement for Gaussian states. 2012.

[21] P. A. M. Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, New York, 4 edition, 1987.

[22] P. A. M. Dirac. *Lectures on Quantum Mechanics*. Dover Publications, New York, 2001. (Reprint of the 1964 Belfer Graduate School of Science edition).

[23] William Donnelly and Steven B. Giddings. Diffeomorphism-invariant observables and their nonlocal algebra. 2015.

[24] J. Eisert, M. Cramer, and M.B. Plenio. Area laws for the entanglement entropy - a review. *Rev.Mod.Phys.*, 82:277–306, 2010.

[25] P.J.S.G. Ferreira and A. Kempf. The energy expense of superoscillations. In *Signal Processing Conference, 2002 11th European*, pages 1–4, Sept 2002.

[26] Brendan Z. Foster and Ted Jacobson. Quantum field theory on a growing lattice. *JHEP*, 0408:024, 2004.

[27] S. A. Fulling. Aspects of Quantum Field Theory in Curved Space-time. *London Math. Soc. Student Texts*, 17:1–315, 1989.

[28] D. Fursaev and D. Vassilevich. *Operators, Geometry and Quanta*. Springer, 2011.

[29] Luis J. Garay. Quantum gravity and minimum length. *Int.J.Mod.Phys.*, A10:145–166, 1995.

[30] A. Gezerlis. Nuclear theory (course notes), 2014.

[31] Steven B. Giddings. Hilbert space structure in quantum gravity: an algebraic perspective. 2015.

[32] P. B. Gilkey. *Asymptotic Formulae in Spectral Geometry*. CRC Press, Boca Raton, 2004.

[33] S. J. Gustafson and I. M. Sigal. *Mathematical Concepts of Quantum Mechanics*. Springer-Verlag, Berlin, Heidelberg, New York, 2003.

[34] R. Haag. *Local Quantum Physics*. Springer, Berlin Heidelberg, 1996.

[35] B. C. Hall. *Quantum Theory for Mathematicians*. Springer-Verlag, New York, 2013.

[36] Y. Hao and A. Kempf. On a non-fourier generalization of shannon sampling theory. *Proc. of the 10th Canadian Workshop on Information Theory*, pages 193–196, 2007.

[37] Y. Hao and A. Kempf. On the stability of a generalized shannon sampling theorem. *Proc. of the 2008 International Symposium on Information Theory and its Applications*, pages 784–789, 2008.

[38] Y. Hao and A. Kempf. Filtering, sampling and reconstruction with time-varying bandwidths. *IEEE Signal Processing Letters*, 17:241–244, 2010.

[39] Joe Henson. The Causal set approach to quantum gravity. 2006.

[40] J. R. Higgins. *Sampling Theory in Fourier and Signal Analysis: Foundations*. Oxford University Press Inc., New York, 1996.

[41] J. R. Higgins and Stens R. L. *Sampling Theory in Fourier and Signal Analysis: Advanced Topics*. Oxford University Press Inc., New York, 1999.

[42] Christoph Holzhey, Finn Larsen, and Frank Wilczek. Geometric and renormalized entropy in conformal field theory. *Nucl.Phys.*, B424:443–467, 1994.

[43] Sabine Hossenfelder. Minimal Length Scale Scenarios for Quantum Gravity. *Living Rev. Rel.*, 16:2, 2013.

[44] Ted Jacobson. Thermodynamics of space-time: The Einstein equation of state. *Phys.Rev.Lett.*, 75:1260–1263, 1995.

[45] Ted Jacobson. Entanglement equilibrium and the Einstein equation. 2015.

[46] Ted Jacobson and Renaud Parentani. Horizon entropy. *Found.Phys.*, 33:323–348, 2003.

[47] A. J. Jerri. The shannon sampling theorem – its various extensions and applications: A tutorial review. *Proceedings of the IEEE*, 65:1565–1596, 1977.

[48] A. Kempf. Quantum field theory for cosmology (course notes), 2014.

[49] A. Kempf, A. Chatwin-Davies, and R.T.W. Martin. A fully covariant information-theoretic ultraviolet cutoff for scalar fields in expanding FRW spacetimes. *J.Math.Phys.*, 54:022301, 2013.

[50] Achim Kempf. Black holes, bandwidths and Beethoven. *J.Math.Phys.*, 41:2360–2374, 2000.

[51] Achim Kempf. Fields over unsharp coordinates. *Phys. Rev. Lett.*, 85:2873–2876, Oct 2000.

[52] Achim Kempf. Mode generating mechanism in inflation with cutoff. *Phys. Rev.*, D63:083514, 2001.

[53] Achim Kempf and Larissa Lorenz. Exact solution of inflationary model with minimum length. *Phys. Rev.*, D74:103517, 2006.

[54] Achim Kempf, Gianpiero Mangano, and Robert B. Mann. Hilbert space representation of the minimal length uncertainty relation. *Phys.Rev.*, D52:1108–1118, 1995.

[55] E. Kreyszig. *Introductory Functional Analysis with Applications*. John Wiley & Sons. Inc., New York, 1989.

[56] H. J. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Math.*, 117:37–52, 1967.

[57] David A. Lowe, Joseph Polchinski, Leonard Susskind, Larus Thorlacius, and John Uglum. Black hole complementarity versus locality. *Phys.Rev.*, D52:6997–7010, 1995.

[58] G. Ludwig. *Foundations of Quantum Mechanics I.* Springer-Verlag, New York, Heidelberg, Berlin, 1983.

[59] Krishnanand Mallayya, Rakesh Tibrewala, S. Shankaranarayanan, and T. Padmanabhan. Zero modes and divergence of entanglement entropy. 2014.

[60] R. T. W. Martin and A. Kempf. Quantum uncertainty and the spectra of symmetric operators. *Acta Applicandae Mathematicae*, 106(3):349–358, 2009.

[61] C. Alden Mead. Possible connection between gravitation and fundamental length. *Phys. Rev.*, 135:B849–B862, Aug 1964.

[62] V. Mukhanov and S. Winitzki. *Introduction to Quantum Effects in Gravity.* Cambridge University Press, New York, 2007.

[63] A. W. Naylor and G. R. Sell. *Linear Operator Theory in Engineering and Science.* Springer-Verlag, New York, 2 edition, 2000.

[64] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, New York, 10 anv. edition, 2010.

[65] H. Nyquist. Certain topics in telegraph transmission theory. *AIEE Trans.*, 47:617–644, 1928.

[66] M. Ohya and I. Volovich. *Mathematical Foundations of Quantum Information and Computation and Its Applications to Nano- and Bio-systems.* Springer, Dordrecht, Heidelberg, London, New York, 2011.

[67] A. Peres and D. R. Terno. Quantum information and relativity theory. *Rev. Mod. Phys.*, 76:93–123, Jan 2004.

[68] M.E. Peskin and D.V. Schroeder. *An Introduction to Quantum Field Theory.* Advanced book classics. Addison-Wesley Publishing Company, 1995.

[69] J. Pye, W. Donnelly, and A. Kempf. Locality and entanglement in bandlimited quantum field theory. 2015. (To be published.).

[70] Xiao-Liang Qi. Exact holographic mapping and emergent space-time geometry. 2013.

[71] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 1948.

[72] C. E. Shannon. Commnication in the presence of noise. *Proceedings of the IRE*, 37:10–21, 1949.

[73] Sergey N. Solodukhin. Entanglement entropy of black holes. *Living Reviews in Relativity*, 14(8), 2011.

[74] Rafael D Sorkin. On the entropy of the vacuum outside a horizon. In *Tenth International Conference on General Relativity and Gravitation (held Padova, 4-9 July, 1983), Contributed Papers*, volume 2, pages 734–736, 1983.

[75] Rafael D. Sorkin. Expressing entropy globally in terms of (4D) field-correlations. 2012.

[76] M. Srednicki. Entropy and area. *Phys. Rev. Lett.*, 71:666–669, 1993.

[77] L. Susskind and J. Lindesay. *An Introduction to Black Holes, Information, and the String Theory Revolution*. World Scientific Publishing Co. Pte. Ltd., Singapore, 2005.

[78] Leonard Susskind and John Uglum. Black hole entropy in canonical quantum gravity and superstring theory. *Phys.Rev.*, D50:2700–2711, 1994.

[79] M. Unser. Sampling - 50 years after shannon. *Proceedings of the IEEE*, 88:569–587, 2000.

[80] D.V. Vassilevich. Heat kernel expansion: user's manual. *Physics Reports*, 388(56):279 – 360, 2003.

[81] R. M. Wald. *General Relativity*. The University of Chicago Press, Chicago and London, 1984.

[82] John A. Wheeler. On the nature of quantum geometrodynamics. *Annals of Physics*, 2.

[83] Steven White. Wavelets and mera. Perimeter Institute condensed matter seminar, 2013.

[84] F. Wilczek. Quantum field theory. *Rev. Mod. Phys.*, 71:S85–S95, Mar 1999.

[85] John Williamson. On the algebraic problem concerning the normal forms of linear dynamical systems. *American Journal of Mathematics*, pages 141–163, 1936.

[86] John Williamson. On the normal forms of linear canonical transformations in dynamics. *American Journal of Mathematics*, pages 599–617, 1937.

[87] Edward Witten. Reflections on the fate of space-time. *Phys. Today*, 49N4:24–30, 1996.

[88] J. L. Yen. On nonuniform sampling of bandwidth-limited signals. *IRE Trans. on Circuit Theory*, 3:251–257, 1956.

[89] A. I. Zayed. *Advances in Shannon's Sampling Theory*. CRC Press, 1993.

[90] Nikolay I. Zheludev. What diffraction limit? *Nature Materials*, 7(6):420–422, Jun 2008.