Exploring equilibration in Fermionic systems: A connection with minimum distance codes

Jose Alejandro Montaña Cortes

Universidad de los Andes
Faculty of Science, Department of Physics
Bogotá, Colombia
2020

Exploring equilibration in Fermionic systems: A connection with minimum distance codes

Jose Alejandro Montaña Cortes

Supervisor:

Alonso Botero Mejía

A thesis submitted in fulfilment of the requirements for the degree of:

Master's Degree in Physics

In the research group of:

Quantum Mechanics and Information Theory

Universidad De Los Andes

Faculty of Science , Departament of Physics

Bogotá, Colombia

2020

"*To my family, to my friends and specially to Valentina*"

Resumen

La tipicidad canónica ha surgido como una alternativa a los fundamentos de la mecánica estadística, logrando explicar desde una perspectiva de la información cuántica, fenómenos como el de la termalización, el cual emerge como consecuencia del entrelazamiento entre el sistema y su ambiente. Posteriores trabajos han mostrado que esta particular forma de abordar el problema proporciona una idea del mecanismo por el cual se alcanza el equilibrio en sistemas cuánticos. Dentro del marco de estas ideas, promedios temporales juegan un rol principal a la hora de entender como estados reducidos obtenidos a partir de un estado puro, alcanzan el equilibrio. Esto nos conduce a preguntarnos: ¿Es posible encontrar estados reducidos provenientes de un estado puro no estacionario, tales que automáticamente estén equilibrados, es decir, estados reducidos cuyo equilibrio se alcance de forma inmediata?, y en caso de ser así, ¿Es posible determinar el tamaño del conjunto de estados cumpliendo esta propiedad?.

En el presente trabajo, el uso de herramientas de teoría de código, específicamente, códigos aleatorios fermiónicos de distancia mínima, proporcionan una forma diferente de entender los sistemas fermiónicos reducidos. Confirmando así, la existencia de estados reducidos que permanecen en equilibrio para estos sistemas. Se explora así, el espacio de Hilbert generado por el conjunto de estos estados, concretamente, se caracteriza el espacio de Hilbert asociado a estos estados por medio de un exponente de error, el cual proporciona una ley de grandes desviaciones en el sistema. De esta forma, se prueba que el tamaño del espacio de Hilbert asociado a los estados, tales que cumplen la propiedad de equilibrarse de forma inmediata es exponencialmente grande.

Palabras Clave: Tipicidad, Sistemas Fermiónicos, códigos aleatorios fermiónicos de distancia mínima .

Abstract

Canonical typicality has emerged as an alternative to the foundations of statistical mechanics, being able to explain from a viewpoint of quantum information theory, how thermalisation results from entanglement between system and environment. Later results have shown that the ideas used in typicality, provide an explanation form the mechanism of the evolution towards equilibrium for quantum systems. Within this framework of ideas, time averages plays a major rol when it comes to understanding how reduced states obtained from a pure states, reach equilibrium. This led us to ask ourselves, is it possible to find reduced states obtained from a non stationary pure state, such that they equilibrate instantly, this is, reduced states such that immediately reach its equilibrium state?, and if so, is it possible to measure the size of the space fulfilling this property?

In the present work the use of code theory, specifically Fermionic minimum distance codes, provide an alternative to understand fermionic systems. Specifically, we characterised the Hilbert space generated by the states, such that fulfil this property, throughout an error exponent, which provide a characterisation of the large deviation law within the system. Therefore, we prove that the Hilbert space of states fulfilling the property of reaching its equilibrium instantly is indeed exponentially large.

Key Words: Typicality, Fermionic systems, Fermionic minimum distance codes.

Agradecimientos

Declaration of Authorship

I, Jose Alejandro Montaña Cortes, declare that this thesis titled,“ *Exploring equilibration in Fermionic systems: A connection with minimum distance codes*” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

_____

Date:

_____

Content

Introduction

Alternative considerations about the foundations of statistical mechanics proposed by Posescu et. al. [1], have shown that reliability on subjective randomness [2], ensemble averaging [3] or time averaging [4] are not required to understand the emergence of thermalisation; instead, a quantum information perspective [5] provides an alternative answer to the foundations of statistical mechanics, from a uniquely quantum point of view, which does not rely on any ignorance probabilities in the description of the state. This precise way of tackling the problem is intimately related to empiricism. Plan your day, go out for a walk, whenever we plan something, the last thing that goes through our mind is to get smacked by a meteor, we know it may happen, but we know that it is not something normal to happen. This is why it is "typical" to plan our lives without even bother by getting hit by a meteor. Similarly, what Popescu et. al. proved is that if we consider a quantum pure state, subject to a global constraint, the "typical" thing to happen is that the reduced state of the system is very close to the canonical mixed state. That is, not only is the state of the system mixed, but it is precisely in the state we would expect from standard statistical arguments [1], which is achieved by considering subjective randomness [6, 7]. To provide a more precise argument, consider as the universe the system together with a sufficiently large environment, in a quantum pure state. Due to entanglement with the rest of the universe and concentration properties of high dimensional spaces. System thermalisation appears as a local generic property of pure states of the universe subjected to a global constraint. This result is known as"General Canonical Principle", or informally as "Canonical Typicality" and is considered to be an important result when understanding statistical mechanics. Specifically, what this principle tells us is that whenever we look at a sufficiently small system, compared to its environment, the reduced state of the system will approximately correspond to the thermal state [1, 8–10], therefore suggesting that thermalisation occurs as a generic local property of pure states of the universe. It should be emphasised that the results in typicality apart of providing a general viewpoint of thermalisation, those results are kinematic, rather than

dynamical. That is, the particular unitary evolution of the global state is not considered at any moment, and thermalisation is not proven to happen; instead, the key ingredient is Levy's Lemma [11, 12], which plays a similar role to the law of large numbers and governs the properties of typical states in large-dimensional Hilbert spaces [1], and thus provides a powerful tool to evaluate functions of randomly chosen quantum states. We stress here that these ideas were not only proposed by Popescu et. al.; contemporaneously with them, Gemmer. et al [10], as well as Goldstein et. al [9], proposed similar ideas, in which heuristic arguments are used to prove canonical typicality, and exhibit an explicit connection between reduced states an the micro-canonical density matrix at a suitable total energy E. Thus, the result obtained by Popescu et. al. establishes canonical typicality under a great generality by invoking the levy Lemma [1, 11, 12]. For that reason, the viewpoint we discuss here is mostly based on the one proposed by Popescu.

With the purpose of extending typicality beyond the kinematic viewpoint and address the dynamics of thermalisation. In particular, enquire under what conditions the state of the universe will evolve into the large region of its Hilbert Space in which its subsytems are thermalised and remain in that space for most of its evolution. Motivated by previous results heading this direction [13, 14] and the fact that from typicality, is possible to show that the overwhelming majority of states in the universe bring the system to the canonical mixed state, Linden et. al. [15] explore weather thermalisation could happen as a universal property of quantum systems. Thus, by using arguments based on ideas of typicality, reaching equilibrium can be showed to be a typical property of large quantum systems. In this framework, dynamical aspects are addressed to explore the evolution that drives systems to equilibrate, and moreover, to study under what circumstances systems reach equilibrium and how much they fluctuate about the equilibrium state. A series of results in [15–17] suggest that under mild conditions, any subsystem of a sufficiently large system will reach equilibrium and fluctuate around it at almost all times. The only conditions required are that the Hamiltonian has no degenerate energy gaps, and that the state of the universe contains sufficiently many energy eigenstates. These conditions are fulfilled for most physical situations, all but a set of measure zero of Hamiltonians have non-degenerate energy gaps.

Even though thermalisation seems to be a very straightforward process, it is quite difficult to formalise an explanation to it. A closer look, reveals that thermalisation is composed of many different aspects that have to be inspected in detail, and where equilibration, bath

state independence, subsystem state independence and the Boltzmann form of the equilibrium state play a prominent role [15]. First, *equilibration* is as the process in which the system reaches a particular state and remains in that state or close or to it for almost all times. Whenever we refer to equilibration, note that any particular state is not inferred and in general it does not need to be a thermal state. Second, *bath state independence* refers to the fact that the equilibrium state of the system should not depend on the precise initial state of the Bath. That is, only macroscopic parameter are needed to describe the bath [15]. For example, its temperature; in the moment equilibrium is reached, that state of equilibrium should only depend on the temperature of the bath. Third, *subsystem state independence* which tells us that the equilibrium state reached by the system, should be independent of its initial state. Finally the *Boltzmann form of the equilibrium* describes the Boltzmannian form of the equilibrium state ($\rho_{\mathcal{S}} = \frac{1}{Z}\exp(-\frac{H_{\mathcal{S}}}{k_B T})$). Note that equilibration is then a more general process which can depend on different parameters such as initial conditions in an arbitrary way, whereas thermalisation does not.

Realising that thermalisation is compound for the afore mentioned elements, let us address each aspect separately. Moreover, it let us clarify an important distinction between thermalisation and equilibration, indeed, we will consider equilibration as a general quantum phenomenon that may occur in situations other than whose associated with thermalisation. By using this decomposition of thermalisation, Linden et. al. are able to prove the first two elements mentioned above (Equilibration and bath state independence). Namely, they prove not only that reaching equilibrium is a universal property of quantum systems but that this equilibrium state does not depend on the precise details of the bath state, but rather its macroscopic parameters [15].

Up to this point, we have shown that typicality has been proven to be an extremely useful alternative way of studying thermalisation in quantum systems and understand the foundations of statistical mechanics. Nonetheless, a closer look to typicality will derive in a property that we consider could yield to further insights to understand equilibration in quantum systems. To illustrate our ideas, consider two different orthogonal pure states living in the same Hilbert subspace ($|\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}_R$), the one that is obtained by imposing a global constraint over the universe. From typicality we know that the reduced state of $|\psi_1\rangle$ and $|\psi_2\rangle$ approximately leads to the same state, that is,

$$\mathrm{Tr}_{\mathcal{E}} |\psi_1\rangle \langle\psi_1| \approx \mathrm{Tr}_{\mathcal{E}} |\psi_2\rangle \langle\psi_2| \approx \Omega_{\mathcal{S}}, \tag{1}$$

where $\Omega_{\mathcal{S}}$ corresponds to the canonical state of the system. Thus, we could consider a third state $|\psi_3\rangle$ which is a generic linear combination of $|\psi_1\rangle$ and $|\psi_2\rangle$, if the condition of typicality is also imposed over the third state. We have therefore that its reduced state will also lead us to the canonical state, meaning that cross terms obtained in the reduced state associated with $|\psi_3\rangle$ should somehow vanish. Explicitly, when we compute the density matrix associated with the state $|\psi_3\rangle$ we have

$$|\psi_3\rangle\langle\psi_3| = |c_1|^2 |\psi_1\rangle\langle\psi_1| + |c_2|^2 |\psi_2\rangle\langle\psi_2| + c_1^* c_2 |\psi_2\rangle\langle\psi_1| + c_2^* c_1 |\psi_1\rangle\langle\psi_2|, \tag{2}$$

and therefore its reduced state reads

$$\mathrm{Tr}_{\mathcal{E}} |\psi_3\rangle\langle\psi_3| \approx \Omega_{\mathcal{S}} \approx \Omega_{\mathcal{S}} + c_1^* c_2 \,\mathrm{Tr}_{\mathcal{E}} |\psi_2\rangle\langle\psi_1| + c_2^* c_1 \,\mathrm{Tr}_{\mathcal{E}} |\psi_1\rangle\langle\psi_2|, \tag{3}$$

where the condition of normalisation was used ($|c_1|^2 + |c_2|^2 = 1$). Notice that the cross terms in equation (3) should therefore approximately vanish in order to satisfy the relation, namely, the condition which has to be satisfied in order to keep the equality is $\mathrm{Tr}_{\mathcal{E}} |\psi_2\rangle\langle\psi_1|$ $= \mathrm{Tr}_{\mathcal{E}} |\psi_1\rangle\langle\psi_2| = 0$. Since this condition tell us that off-diagonal terms approximately vanish this bring to our mind some previous ideas proposed by Srednicki et. al. [18–20] in ETH, where the off-diagonal terms are expected to be stochastic quantities with mean zero and an amplitude that is exponentially small on the degrees of freedom of the system, that is, the off-diagonal terms are expected to be near zero for a system with a large degree of freedom. For our case, we have something similar but instead of having that the off diagonal terms are near to zero, we have that after taking the partial trace, those terms become approximately zero. Although we are aware that the condition $\mathrm{Tr}_{\mathcal{E}} |\psi_2\rangle\langle\psi_1| = \mathrm{Tr}_{\mathcal{E}} |\psi_1\rangle\langle\psi_2| = 0$ might be related with ETH in some way, due to the impossibility of making an explicit connection between them, we name it *ultra-orthogonality*.

$$\mathrm{Tr}_{\mathcal{E}} |\psi_i\rangle\langle\psi_j| \approx 0, \qquad i \neq j. \tag{4}$$

This particular name was given since when we compute the partial trace over the exterior product of $|\psi_i\rangle$ and $|\psi_j\rangle$ it becomes zero, thus we consider this name provides the idea of having orthogonality over partial traces.

Ultra-orthogonality, can be shown to be related with the equilibration of the system. Consider a time dependent state $|\Psi(t)\rangle$, we can expand this state in its energy eigenstates as

$$|\Psi(t)\rangle = \sum_k c_k e^{-iE_k t} |E_k\rangle, \tag{5}$$

where $\sum_k |c_k|^2 = 1$; hence,

$$\rho(t) = \sum_{k,\ell} c_k c_\ell^* e^{-i(E_k - E_\ell)t} |E_k\rangle \langle E_\ell| = \underbrace{\sum_k |c_k|^2 |E_k\rangle \langle E_k|}_{\omega} + \sum_{k,\ell} c_k c_\ell^* e^{-i(E_k - E_\ell)t} |E_k\rangle \langle E_\ell| , \quad (6)$$

so when we look at the reduced state of $\rho(t)$, defined as $\rho_{\mathcal{S}}(t) = \mathrm{Tr}_{\mathcal{E}} |\psi\rangle \langle\psi|$,

$$\rho_{\mathcal{S}}(t) = \omega_{\mathcal{S}} + \sum_{k,\ell} c_k c_\ell^* e^{-i(E_k - E_\ell)t} \mathrm{Tr}_{\mathcal{E}} |E_k\rangle \langle E_\ell| , \quad (7)$$

where $\omega_{\mathcal{S}} = \mathrm{Tr}_{\mathcal{E}}\, \omega$. Thus all states in equation (7) satisfy the ultra-orthogonality property, then the reduced state would automatically be stationary, meaning that it will be time independent.

Remembering that equilibration is a more general quantum phenomenon and thermalisation can be decomposed in 4 different independent aspects. We notice then that the motivation behind ultra-orthogonality is quite related with equilibration as an immediate process, and since for most cases we expect thermal typicality to be present in the system, the equilibrium state will coincide with the thermal state, meaning that for most cases we will be studying thermal equilibrium. It would therefore be quite interesting to full study this property to provide a better comprehension of what is behind ultra-orthogonality, and provide what may appear to be an alternative mechanism for equilibration. However, in this work we are not going to talk about the general problem, instead we will be discussing a particular case of ultra-orthogonality. Here we are going to describe the reduced state of a pure and full dynamical states of the universe, such that its reduced states, sufficiently small compared the environment, are automatically stationary. Namely, we will be discussing the case when the cross terms in (7) are *exactly* equal to zero ($\mathrm{Tr}_{\mathcal{E}} |E_k\rangle \langle E_\ell| = 0$) in the special case of systems that can be mapped to a Fermionic-like system via some non local transformation such as Jordan Wigner transformation [21].

One of the first questions one might wonder is, "are there few or conversely , many states that fulfil the property that the cross terms in (7) are exactly equal to zero, and if so, what is the size of the Hilbert subspace associated with these states?". We will provide the answers to these questions by using some ideas taken from random minimum distance codes adapted to Fermionic systems, and we will show through a rate exponent that there are exponentially large Hilbert subspaces in which all its states fulfil the condition $\mathrm{Tr}_{\mathcal{E}} |E_k\rangle \langle E_\ell| = 0$.

An insight to clarify how techniques from code theory can be used to study ultra-orthogonality starts by noting that Fermionic states can be treated as binary sequences of excitations, when we interpret Fermionic systems this way, we are able to show that the terms $\mathrm{Tr}_{\mathcal{E}} |E_k\rangle \langle E_\ell|$ is zero whenever the hamming weight (the number of ones in the sequence) of the vector $\vec{E_k} + \vec{E_\ell}$ is larger than the size of the subsystem L. Where $\vec{E_i}$ corresponds to the sequence of excitations corresponding the state $|E_i\rangle$. Observe that higher the energy, the more excitations we have in the sequences, so the vector $\vec{E_k} + \vec{E_\ell}$ will be very likely have more ones when we increase the energy. For our purpose we will look at the the Hilbert space $\mathcal{H}_{\mathcal{C}}$ spanned by mutually ultra-orthogonal vectors $\mathcal{C} = \{x^{(1)}, x^{(2)}, \ldots, x^{(2^k)}\}$. Thus a state $|\psi\rangle \in \mathcal{H}_{\mathcal{C}}$ can be written as

$$|\psi\rangle = \sum_{\vec{n} \in \mathcal{C}} \psi(\vec{n}) \, |\vec{n}\rangle \,, \tag{8}$$

implying then

$$\rho_{\mathcal{S}}(\psi) = \mathrm{Tr}_{\mathcal{E}}(|\psi\rangle \langle\psi|) = \sum |\psi(\vec{n})|^2 \, |\vec{n}\rangle \langle\vec{n}| \,. \tag{9}$$

The fact that these cross terms become zero reminds us an error correction code of minimum distance, because we can see that ultra-orthogonality holds whenever the error vector $\vec{E_k} + \vec{E_\ell}$ has larger hamming weight than L. We could then ask ourselves about the largest Hilbert space that can be constructed, by alternative answering what is the biggest code of minimum distance L that can be built. Previous works [22] compute Minimum distances, distance distribution and error exponents on a binary-symmetric channel for typical code from Shannon's random code ensemble and for typical codes from a random linear code ensemble; calculations that can easily be modified to our purpose and that will let us prove that indeed the size of the Hilbert space in which super-orthogonality holds is exponentially large.

This document is divided in four chapters: In the first chapter, we will introduce canonical typicality and its relation to statistical mechanics. Starting with the postulate of equal a priori probability and the controversy around it. Next we introduce typicality and the foundations of it, and finally a detailed explanation of the problem of ultra-orthogonality. In the second chapter, we will state general definitions concerning Fermionic systems, in particular, we will be discussing the case of the systems that can be mapped to Fermionic systems such as the one dimensional XY. In the third chapter we will provide some definitions with regard to error correction codes and specifically some of the well-known bounds for minimum distance codes; and present our main results which are an explicit connection between the

Fermionic systems and random minimum distance codes, and an estimate of the size of the Hilbert subspace associated with the states that fulfil ultra-orthogonality. Finally, in the fourth chapter we present our conclusions as well as further perspectives of our work.

Chapter 1

Canonical typicality and its connection to thermodynamics

Fundamental questions concerning the foundations of quantum statistical mechanics have been discussed and remain a debatable subject [7]. In these questions the role of probabilities, entropy, the relevance of time averages and ensemble averages to individual physical systems is discussed thoroughly [10] to answer weather or not they are needed to formalise statistical mechanics. One of the most controversial issue is the the validity of the postulate of equal a priori probability, postulate which can not be proven [7] and was used for Boltzmann [23] to introduce the emergence of lack of knowledge to formalise the ideas of classical statistical mechanics. Along this chapter we will discuss some of the ideas based on typicality addressed by several authors [8–10], who have abandoned the unprovable aforementioned postulate and have replaced it with a new viewpoint, which is uniquely quantum, and which does not rely on any ignorance probabilities in the description of the state.

This chapter is divided in three parts: First, we are going to introduce the postulate of equal prior probability, discussing its role in the foundations of statistical mechanics, where the idea of an ensemble as a collection of identical systems will be introduced and the postulate of equal prior probability will be translated to a particular version of ensemble, the microcanonical ensemble. Also a quantum version of the above-mentioned postulate will be exhibited in terms of the random phase postulate [24] and the derivation of the canonical ensemble for a weakly interacting system. The second part of the chapter, will be dedicated to understand the role of entanglement in connection with Statistical mechanics. We introduce the concentration phenomenon of canonical typicality, from which thermalisation will emerge

as a consequence of typicality [1,8]. Finally, we use the previous results an approach studied by *Linden et. al.* [15] in which rather than a kinematic viewpoint of thermalisation, they address thermalisation as a dynamical process in which the unitary evolution is took into account to show that equilibration is a generic property of quantum systems. This leads us to the main subject of this work, the property we named ultra-orthogonality. We will look under what conditions ultra-orthogonality is expected to happen and we will enclose the problem we are going to tackle.

1.1 Postulate of equal a priori probability

Statistical physics has been known to describe with high accuracy macroscopic systems by describing the dynamics of its microscopic constituents. The formalisation of classical statistical mechanics was first proposed by Boltzmann, Gibbs and Maxwell. However, this formalisation has been a subject of strong debates, because in their theory, probability was needed to derive the laws of thermodynamics, causing then some controversy [25] in the scientific community. We will introduce the main theoretical concepts concerning the description of the microscopic world that were proposed at that time (19th century) to understand why the introduction of probabilities was needed.

Consider a mechanical system of N ($N \ggg 1$) particles subject to a time-independent potential V. Supposing every particle in the system has mass m and coordinates in its phase space $(\vec{q}_i, \vec{p}_i)$ we easily see that the state of the system is represented by a point in the phase space Γ $(\vec{q}, \vec{p}) = (\vec{q}_1, \ldots, \vec{q}_N, \vec{p}_1, \ldots, \vec{p}_N)$, and its time evolution follows the dynamics induced by its Hamiltonian. Moreover, imposing the fact that the energy E is conserved in the system, the evolution is then confined to a set of Hamiltonians such that fulfil this restriction ($\Sigma_E = \{(\vec{q}, \vec{p}) \in \Gamma | H(\vec{q}, \vec{p}) = E\}$). However, the equations of motion are deterministic, that is, if we know the state of the universe at a given time we can then know it at any time through the its correspondent equations of motion [25]. In [23] Boltzmann proposed an alternative viewpoint, he proposed to interpret the probability associated to a particular state as the relative time spent by the system in that state. Namely, the probability that a phase point lies in an infinitesimal region of Σ_E is given by

$$\rho(q,p)\mathrm{d}q\mathrm{d}p = f(q,p)\delta(H(q,p) - E)\mathrm{d}q\mathrm{d}p, \tag{1.1}$$

where $dpdq$ is the Lebesgue measure of Γ and f is a suitable function. Certainly, this is an

assumption and it relies on the hypothesis that the total time of a measurement is extremely long compared with the intrinsic time scales of the evolution.

By using the Liouville's theorem, it is possible to conclude that f is a constant function on all the allowed trajectories on Σ_E. In addition, if we assume that the trajectory of a single point in phase space fills densely Σ_E (Ergodic hypothesis), we have that,

$$\rho(\vec{q},\vec{p}) \propto \delta(H(\vec{q},\vec{p}) - E). \tag{1.2}$$

Based on this argument, Boltzmann was able to prove that thermal equilibrium could be described in terms of Maxwell's distribution of velocities. In fact in both of his works [26], and [27], Boltzmann mentioned his dissatisfaction about the Ergodic hypothesis and slowly abandoned it. It seems that he only considered it an useful assumption for his general result

> For any system for which the hypothesis is true, its equilibrium state is characterized by
>
> $$\rho_{\mathrm{mc}}\left(p|\vec{q}_1,\ldots,\vec{q}_N\right)dp = \frac{1}{\sqrt{2m\pi}}\frac{\Gamma\left(\frac{3N}{2}\right)}{\Gamma\left(\frac{3N-1}{2}\right)}\frac{\left(E-U-\frac{p^2}{2m}\right)^{(3N-2)/2}}{(E-U)^{(3N-3)/2}}dp, \tag{1.3}$$
>
> from which an analogy to the Maxwell distribution may be recovered in the limit $N \to \infty$, regardless of any details of the inter-particle interactions, or indeed whether the system represented is a gas, fluid, solid or any other thermal body [23].

At that time, his work was quite illustrative for others since it was the first time where probability considerations were applied to the state of a mechanical system as whole, being that an inspiration for subsequent ideas.

1.1.1 postulate of equal a priori probability and the microcanonical ensemble

The introduction of probability as an ingredient associated to the state of a single systems was an important assumption that Boltzmann had to include in its description in order to fit his theory with thermodynamics. Based on this idea, alternatively Gibbs introduce the probability as a distribution function on a collection of identical systems [28]. In his pioneering treatment of statistical mechanics, Gibbs considered *ensembles*. His ensembles

were infinite sets of macroscopically identical systems, each represented by a point in its phase space, also called *microstates*, being compatible with a single macroscopic configuration, called *macrostate*. In Gibbs's perspective, temporal evolution is not concerned, and instead, the distribution of all the available microscopic configurations will be the focus. Namely, an ensemble is introduced as a probability density function on the phase space Γ, such that the average number of microstates in a region R of Γ is given by $\int_R \rho(\vec{q},\vec{p})d\vec{q}d\vec{p}$. Thus, the expectation value of an observable $f : \Gamma \to \mathbb{R}$ is the average of f over Γ

$$\langle f\rangle = \int_\Gamma f((\vec{q},\vec{p})\rho(\vec{q},\vec{p})d\vec{q}d\vec{p}. \tag{1.4}$$

In this approach, stationarity in the ensembles is assumed to derive the microcanonical ensemble, that is,

$$\frac{\partial \rho_t}{\partial t} = \{H, \rho_t\} = 0. \tag{1.5}$$

So, in order to to make the microcanonical ensemble compatible with the conservation of energy we need,

$$\rho_{\mu c}(\vec{q},\vec{p}) = \frac{1}{|\Sigma_E|}\delta(H(\vec{q},\vec{p}) - E), \tag{1.6}$$

with $|\Sigma_E|$ the measure of the set Σ_E. This precise result is known as the equal probability postulate, which is interpreted by Gibbs as follows [28]:

> *All microstates accessible to an isolated system are equally probable, because there is no evidence that certain microstate should be more probable than others.*

Namely, when a macroscopic system is at equilibrium, every state compatible with the constrain of the system, is equally available (likely). Which formally is translated into the choice of a constant density function, called the microcanonical ensemble.

One of the most important results of this formalism of ensembles i the fact that we can compute temporal averages of an observable through averages in the phase space by equation (1.6). Since this is an important assumption in which ensembles is built on, we will dedicate the next part to explain the details behind the assumption of computing temporal averages in through averages in the phase space.

1.1.2 The Ergodic hypothesis

We dedicate this section to discuss the fact that we can compute temporal averages through averages in the phase space, assumption that is known as the ergodic hypothesis. Particularly, we will discuss some of the arguments provided by Boltzmann [26, 27] to formally

argument the reason why independently of the initial distribution of the system, the systems has to evolve towards a uniform distribution, leading then to the equal a prior probability postulate, and an idea of why ergodicity was underneath assumed.

Ergodicity provides us an equality between ensembles and time averages of observables, and it is accomplish by realising that each measurement of an observable f at time t_0 will take a certain time to accomplish. In this period of time the observable f samples different values, so that effectively, the measurement we are looking at is really a time average, that is

$$\frac{1}{t}\int_0^t f\left(T_s\left(\vec{q}_0,\vec{p}_0\right)\right)\mathrm{d}s, \tag{1.7}$$

with $\vec{q}_0,\vec{p}_0$, the initial microstate at $t=0$, and $\{T_s\}_{s\in\mathbb{R}}$ is the Hamiltonian flow generated by H, a Hamiltonian of N particles which interact through a time independent potential V, meaning that the Hamiltonian will look as

$$H(\vec{q},\vec{p}) = \frac{1}{2m}\sum_{i=1}^{N}\vec{p^2}_i + V\left(\vec{q}_1,\ldots,\vec{q}_N\right). \tag{1.8}$$

Thus we are sampling f trajectories whose initial points are $(\vec{q}_0,\vec{p}_0)$. Due to the difference in scales of time between the measurement and the intrinsic time of the microscopical constituents, it is possible to take the limit

$$f^*\left(\vec{q}_0,\vec{p}_0\right) = \lim_{t\to\infty}\frac{1}{t}\int_0^t f\left(T_s\left(q_0,p_0\right)\right)\mathrm{d}s. \tag{1.9}$$

The main problem with the ergodic hypothesis is to determine when the average ensemble average equals the time averages, since the function $f^*(\vec{q}_0,\vec{p}_0)$ explicitly depends on the initial conditions, whereas $\langle f\rangle$ does not. Perhaps the most well-known, interpretation to this problem is the one provided by the Ehrenfests [29]. In which they suggest that Boltzmann somehow relied on the ergodic hypothesis in his argument, even though he never explicitly mentioned it. It is indeed evident that if the ergodic hypothesis holds, a state will spend time in the various regions of the energy hypersurface in phase space in proportion to their volume [25, 30–33]. In other words, during the evolution of the system along its trajectory, regions with small volume, corresponding to highly non-uniform distributions of state, are visited only sporadically, and regions with larger volume, corresponding to more uniform distributions of state, more often. This makes plausible the idea that if a system starts out from a very small region (a very improbable state), it will display a tendency to evolve

towards the overwhelmingly larger equilibrium state.

This arguments are quite important to understand the reason why classical statistical mechanics rely on an introduced ignorance in the system. However, the previous arguments are completely classical and they have to change in order to fit in a quantum theory of statistical mechanics.

1.1.3 The quantum postulate of equal a priori probability

Since our world is quantum, a quantum formulation of statistical mechanics is needed, thus, instead of probability distributions that live in phase space, one should consider density matrices, which encode the whole physical content of a quantum system, therefore, the way we write the postulate of equal a priori probability has to be adapted to the formalism of quantum mechanics.

In quantum mechanics a system is described in a Hilbert space $\mathcal{H}$ and its evolution is generated by a Hamiltonian operator $\hat{H}$. So in order to formulate the quantum version of the microcanonical ensemble, we first need to fix the energy of the system, to do so, we can constraint the energy to live around the value E ($[E, E+\delta]$), with $\delta \ll E$, but δ large enough in order to have an interval that contains many energy eigenvalues of $\hat{H}$. Rather than considering a phase space region Γ_E, as we did in the classical case, we consider a subspace spanned by all eigenvectors with eigenenergies belonging to the interval($\mathcal{H}_R = \mathcal{H}_{[E,E+\delta]}$). The postulate of equal prior probability affirms that all the states compatible with the energy E are equiprobable, which for the case of the quantum mechanical case, we say will say that the states belongs to $\mathcal{H}_R$. However, only the postulate of equal a priori probability is not enough to describe the microcanonical ensemble. In fact the condition that the states have to be in an incoherent superposition has to be added to build the microcanonical ensemble. This condition is known as the *random phase postulate* and its a added condition which is purely quantum and has no classical equivalent. The idea is can be expressed formally by imposing that the Fourier coefficients of a state $|\psi\rangle \in \mathcal{H}_R$ should have equal probability and completely random phases. We clarify that this condition is guaranteed when the environment is sufficiently complex, the states of the system numerous and its transitions so rapid that phase relationships between different states of the system, can not be maintained over microscopically long time intervals [34].

Using these two postulates we conclude that the state of the universe is described in terms of a projector of the Hamiltonian $\hat{H}$ in the constrained Hilbert space $\mathcal{H}_R$. We define $\mathcal{E}_R$, the equiprobable state of the universe corresponding to the restriction R by

$$\mathcal{E}_R = \frac{\mathbb{I}_R}{d_R}, \tag{1.10}$$

where $\mathbb{I}_R$ is the identity (projection) operator on $\mathcal{H}_R$, and d_R is the dimension of $\mathcal{H}_R$. $\mathcal{E}_R$ is the maximally mixed state in $\mathcal{H}_R$, in which each pure state has equal probability, thus, it will correspond to the standard intuition of assigning equal a priori probabilities to all states of the universe consistent with a constrained. This particular state is of great importance in quantum mechanics, since it is possible to prove, under some general assumptions that if the universe is in the equiprobable state, every small part of it will coincide with the canonical state Ω_S, which is characterised by a Boltzmann distribution among its eigenstates at a given temperate β^{-1},

$$\Omega_S \propto e^{-\beta H_S}. \tag{1.11}$$

1.1.4 The canonical ensemble

This section is dedicated to understand the importance of the quantum equiprobable state defined in (1.10).As we mentioned before, if the universe is in the equiprobable state $\mathcal{E}_R$, it is posible to prove that every small part of it will then correspond to the canonical state. To prove it, we make a division of the universe in two parts. The system S and its environment E. So the Hilbert space will then be described by $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$. So the Hamiltonian of the universe will be written as

$$\hat{H} = \hat{H}_S \otimes \hat{\mathbb{I}}_E + \hat{\mathbb{I}}_S \otimes \hat{H}_E + \hat{H}_{\text{int}}, \tag{1.12}$$

where $\hat{H}_S$, $\hat{H}_E$ act separately on the system and its environment respectively, $\hat{\mathbb{I}}_S$ and $\hat{\mathbb{I}}_E$ are the identity operators on $\mathcal{H}_S$, and the terms $\mathcal{H}_E$, $\hat{H}_{\text{int}}$ correspond to the Hamiltonian of the environment and the interaction between the system and its environment respectively. If we assume that $||\hat{H}_{\text{int}}|| \lll ||\hat{H}_S||, ||H_E||$ (weak interaction) and that the dimension of $\mathcal{H}_E$ is much larger than the dimension of $\mathcal{H}_S$ ($d_E = \dim \mathcal{H}_E \gg d_S = \dim \mathcal{H}_S$). We suppose the macroscopic energy E belongs to a small energy interval ($[E, E+\delta]$), where $\delta \lll E$ is large enough to contain many eigenvalues of H_E. Let $\mathbb{I}_R$ be the identity (projector) operator of

$\hat{H}$ on the energy shell .

$$\mathcal{H}_R = \mathcal{H}_{[E,E+\delta]}. \tag{1.13}$$

Assuming the universe is in the equiprobable state $\mathcal{E}_R$ given by (1.10). The state of the system S can be obtained by partial trace the state of the universe over the environment, that is,

$$\Omega_S = \mathrm{Tr}_E\, \mathcal{R}. \tag{1.14}$$

From now on the main idea will be to show that Ω_S corresponds to the thermal state at a given temperature[1].

Let $\{|E_k\rangle\}_{k=1}^{d_E} \subset \mathcal{H}_E$ and $\{|\varepsilon_\alpha\rangle\}_{k=1}^{d_S} \subset \mathcal{H}_S$ be the energy basis of $\hat{H}_E$ and $\hat{H}_S$ respectively. In other words what this means is that the Hilbert space $\mathcal{H}$ associated with the universe has eigenbasis given by

$$\left\{|\varepsilon_\alpha\rangle \otimes |E_k\rangle : 1 \leq \alpha \leq d_S, 1 \leq k \leq d_E\right\}. \tag{1.15}$$

Since we use the hypothesis that the interaction is weak, we have that the equiprobable state $\mathcal{E}_R$ is written as

$$\mathcal{E}_R = \frac{\mathbb{I}_R}{d_R} \approx \frac{1}{d_R} \sum_{\alpha,k} \chi_{[E,E+\delta]}\left(\varepsilon_\alpha + E_k\right) |\varepsilon_\alpha\rangle \langle\varepsilon_\alpha| \otimes |E_k\rangle \langle E_k|, \tag{1.16}$$

where the sum is over indexes k and α such that $\varepsilon_\alpha + E_k \in [E, E+\delta]$, which is equivalent to say that we sum where the characteristic function does not vanish. By tracing over the environment we get

$$\Omega_S = \mathrm{Tr}_E\, \mathcal{E}_R = \frac{1}{d_R} \sum_{\alpha,k} \chi_{[E,E+\delta]}\left(\varepsilon_\alpha + E_k\right) |\varepsilon_\alpha\rangle \langle\varepsilon_\alpha| = \frac{1}{d_R} \sum_{\alpha} d_\alpha^{(E)} |\varepsilon_\alpha\rangle \langle\varepsilon_\alpha|, \tag{1.17}$$

where

$$d_\alpha^{(E)} = \sum_k \chi_{[E,E+\delta]}\left(\varepsilon_\alpha + E_k\right) = \sum_k \chi_{[E-\varepsilon_\alpha, E-\varepsilon_\alpha+\delta]}\left(E_k\right). \tag{1.18}$$

Replacing $\hat{H}_E = \sum_k E_k |E_k\rangle \langle E_k|$, we get that

$$d_\alpha^{(E)} = \mathrm{Tr}\, \chi_{[E-\varepsilon_\alpha, E-\varepsilon_\alpha+\delta]}\left(H_B\right) = \dim \mathcal{H}^{(E)}_{[E-\varepsilon_\alpha, E-\varepsilon_\alpha+\delta]}, \tag{1.19}$$

where $\mathcal{H}^{(E)}_{[E_1,E_2]} \subset \mathcal{H}_E$ is the subspace generated by all eigenstates with energy in $[E_1, E_2]$. Thus $d_\alpha^{(E)}$ is a non-negative integer.

[1]The following prove was taken from [34]

Defining the bath entropy at energy E as the logarithm of the number of energy levels in the bath,

$$S_E(E) = \ln\left(\dim \mathcal{H}^{(E)}_{[E,E+\delta]}\right), \tag{1.20}$$

we deduce,

$$d^{(E)}_\alpha = \dim \mathcal{H}^{(E)}_{[E-\varepsilon_\alpha, E-\varepsilon_\alpha+\delta]} = \mathrm{e}^{S_E(E-\varepsilon_\alpha)}. \tag{1.21}$$

As the dimension of $\mathcal{H}_E$ is very large ($d_R \ggg 1$), we can therefore assume that the spectrum of $\hat{H}_E$ is quasi-continuous, so that $S_E(E)$ can be considered a continuous differentiable function of E. Thus, the microscopic energy is much smaller than the macroscopic energy E ($\varepsilon \lll E$), and we can write

$$S_E\left(E-\varepsilon_\alpha\right) \approx S_E(E) - \frac{\mathrm{d}S_E(E)}{\mathrm{d}E}\varepsilon_\alpha, \tag{1.22}$$

and,

$$\Omega_S = \frac{1}{d_R}\sum_\alpha d^{(E)}_\alpha \left|\varepsilon_\alpha\right\rangle\left\langle\varepsilon_\alpha\right| \approx \frac{1}{Z}\sum_\alpha \mathrm{e}^{-\beta\varepsilon_\alpha}\left|\varepsilon_\alpha\right\rangle\left\langle\varepsilon_\alpha\right| = \frac{1}{Z}\mathrm{e}^{-\beta\hat{H}_S}, \tag{1.23}$$

with $Z = \mathrm{Tr}\left(e^{-\beta\hat{H}_S}\right)$, and

$$\beta = \frac{\mathrm{d}S_E(E)}{\mathrm{d}E}. \tag{1.24}$$

corresponds to the thermodynamical expression of the inverse temperature of the bath. The result shown in equation (1.23) gives an explicit form of the canonical state in terms of the Boltzmann distribution. In the next section we will dedicate discussion about this result and we will start to dive into the details of canonical typicality.

1.2 Entanglement and canonical typicality

The purpose of this section is to understand why in quantum mechanics, the postulate of equal a priori probability is not needed and will be abandoned by an objective lack of knowledge which emerges as a consequence of the structure of quantum mechanics and is quite related with quantum entanglement.

1.2.1 Entanglement and the foundations of statistical mechanics

Before we start, it is better to show what Schrödinger understood as entanglement,

> When two systems, of which we know the states by their respective representatives, enter into temporary physical interaction due to known forces between them, and when after a time of mutual influence the systems separate again, then they can no longer be described in the same way as before, viz. by endowing each of them with a representative of its own (E. Schrödinger 1995 [35]).

To understand what Schrödinger meant, we can be illustrate a particular situation. Consider a composite system of two spins on the Hilbert space $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$, where $\mathcal{H}_S = \mathcal{H}_E = \mathbb{C}^2$. Denoting $\{|\uparrow\rangle, |\downarrow\rangle\}$, our computational basis of $\mathbb{C}^2$ [36–38]. If the bipartite system is described in terms of a factorized state, for example,

$$|\phi\rangle = |\uparrow\rangle_S \otimes |\uparrow\rangle_E, \tag{1.25}$$

then the system and its environment can be described independently. On the contrary, when the global state is not factorized,

$$\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle_S \otimes |\uparrow\rangle_E + |\downarrow\rangle_S \otimes |\downarrow\rangle_E\right), \tag{1.26}$$

then it is said that the state is entangled. In fact the density matrix associated with this system is given by

$$\rho_S = \mathrm{Tr}_E \left(\left|\Phi^+\right\rangle \left\langle\Phi^+\right|\right) = \frac{1}{2}\hat{\mathbb{I}}_2, \tag{1.27}$$

where $\hat{\mathbb{I}}$ is the identity operator on $\mathcal{H}_S = \mathbb{C}^2$. The latter equation tells us that the system S is in a totally mixed state, meaning that it is uniformly randomly distributed in $\mathcal{H}$. Now consider the general parametrization of the latter example,

$$|\Phi_\alpha\rangle = \sqrt{\alpha}|\uparrow\rangle_S \otimes |\uparrow\rangle_E + \sqrt{1-\alpha}|\downarrow\rangle_S \otimes |\downarrow\rangle_E, \quad \alpha \in [0,1], \tag{1.28}$$

which condensate the separable state $|\phi\rangle$ ($\alpha = 0$), and the Bell state$|\Phi^+\rangle = \left|\Phi_{1/2}\right\rangle$ ($\alpha = 1/2$). The reduced density matrix is then given by

$$\rho_S^\alpha = \alpha|\uparrow\rangle\langle\uparrow| + (1-\alpha)|\downarrow\rangle\langle\downarrow|. \tag{1.29}$$

which is a state whose mixture depends on α. Summarising, if the state of the composite system is factorised, the information on the whole state, and on every subsystem, is completely accessible. On the other hand, if the global state is entangled, even though we have complete knowledge of the state in the universe a prior, only a partial knowledge of the subsystem can be obtained. Particularly, if the global state is maximally entangled ($\alpha = 1/2$) one

has no information at all on the subsystem introducing then an objective lack of knowledge which source comes form the structure of quantum mechanics. Mathematically speaking, the information content of a state ρ is described by the Von Neumann entropy [39].

$$S(\rho) = -\operatorname{tr}(\rho \ln \rho) = -\sum_k p_k \ln p_k, \tag{1.30}$$

where p_k are the eigenvalues of ρ, subject to the conditions

- $0 \leq p_k \leq 1$,
- $\sum_k p_k = 1$.

It is straightforward to realise that every pure state has 0 entropy, meaning that the information encoded in the state is completely available. In our particular case, we see that $S(|\Phi_\alpha\rangle \langle\Phi_\alpha|) = 0$ for every α, and contrarily, for the general parametrised reduced state we get

$$S\left(\rho_S^\alpha\right) = -\alpha \log \alpha - (1-\alpha)\log(1-\alpha), \tag{1.31}$$

that is a positive metric function of $\alpha \in [0, 1]$, which is 0 for separable global states ($\alpha = 0, 1$)

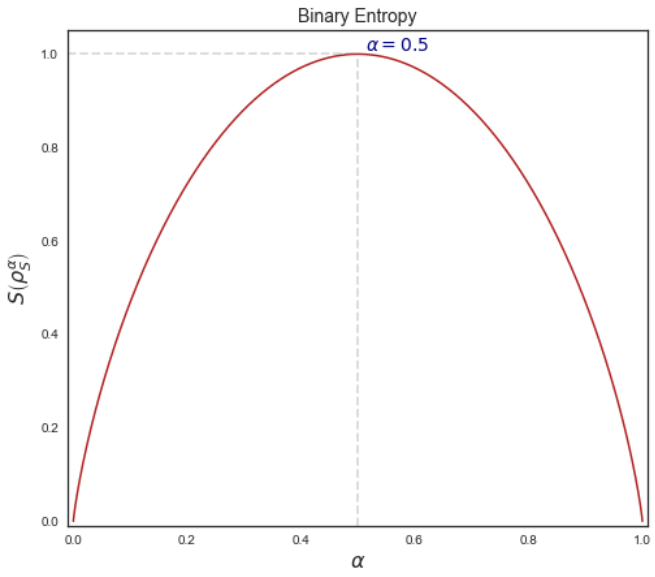

Figure 1.1: Binary Entropy as a function of binary outcome probability α.

and reaches its maximum for the maximally entangled Bell state ($\alpha = 1/2$). Observe that the expression given in (1.31) is nothing but the binary entropy function [40] which is equivalent

to the Shannon entropy [41] of the probability vector $(\alpha, 1-\alpha)$. By using the latter example, it is easy to generalise these ideas to an arbitrary given pure state $|\Psi\rangle$ of a composite system $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$ with generic dimensions $d_S = \dim \mathcal{H}_S \leq d_E = \dim \mathcal{H}_E$, and conclude that

$$0 \leq S(\rho_S) \leq \ln d_S. \tag{1.32}$$

Here $S(\rho_S) = 0$ for separable state $|\Psi\rangle = |u\rangle \otimes |v\rangle$, whereas $S(\rho_S) = \ln d_S$ for maximally entangled states, where

$$|\Psi\rangle = \frac{1}{\sqrt{d_S}} \sum_{k=1}^{d_S} |u_k\rangle \otimes |v_k\rangle \,, \tag{1.33}$$

with $\{u_k\}, \{v_k\}$ the orthogonal basis for each of system and environment. The von Neumann entropy is a measure of entanglement which leads to an objective lack of knowledge. In general entropy provides a tool that can be used to quantify entanglement, even though other entanglement measures exist [42], in the case the overall system is pure, the entropy of one subsystem can be used to measure its degree of entanglement with the other subsystems. Particularly, for bipartite pure states, the von Neumann entropy of reduced states is the unique measure of entanglement in the sense that it is the only function on the family of states that satisfies certain axioms required of an entanglement measure, being this the reason why we only introduce this entropy. This objective "lack of knowledge" is related to the state of the system, because even if we knew everything about the state of the universe, every subsystem (small portion of the universe) could be mixed [43, 44]. In fact, this objective "lack of knowledge" is related to the state of the system, because even if we knew everything about the state of the universe, every subsystem (small portion of the universe) could be mixed [43, 44]. Note that this lack of knowledge came form nothing but the true nature of quantum mechanics, and no randomness is introduced as in the case of classical mechanics[2].

As a conclusion to this section, we can say that the postulate of equal a prior probability, which refers to ensembles, time averages of states of the universe, can be disregard and one can refer only to pure states of the universe. Therefore, supported on the fact that no artificial introduction of probability is needed in quantum mechanics, in the following part

[2]Classically the complete knowledge of the state of the universe implies a complete knowledge of the state of any subsystem. Understanding why in the case of classical mechanics, randomness was artificially add to the description.

we will show that thermalisation appear as a local generic property of pure states of the universe, meaning that for the overwhelming majority of pure states, the reduced state of the system is approximately the canonical mixed state.

1.2.2 Canonical Typicality

Consider a large quantum mechanical system, "The Universe", which we decompose in two parts, the system S and the environment E[3]. Now, suppose the universe has to obey some global constraint R, which translates into the choice of a subspace of the total Hilbert space, say

$$\mathcal{H}_R \subset \mathcal{H}_S \otimes \mathcal{H}_E, \tag{1.34}$$

where we will denote the dimensions of $\mathcal{H}_S$, $\mathcal{H}_E$ and $\mathcal{H}_R$ by d_S, d_E and d_R respectively. In the standard approach to statistical mechanics, as we aforementioned, the restriction is imposed on the total energy. However, as Popescu et. al. emphasise in [1,8], this restriction can be completely arbitrary and not necessarily referring to the energy. Using the definition of the equiprobable state in (1.10), we know that $\mathcal{E}_R$ is the maximally mixed state in $\mathcal{H}_R$.

As we discussed before, the canonical state of the system S is defined as the trace over the environment,

$$\Omega_S = \mathrm{Tr}_E\, \mathcal{E}_R. \tag{1.35}$$

Now, instead of considering the universe in the equiprobable state $\mathcal{E}_R$, which describes subjective ignorance, we consider the universe to be in a random pure state $|\phi\rangle \in \mathcal{H}_R$. In such case, the system will be described by its reduced density matrix

$$\rho_S = \mathrm{tr}_B(|\phi\rangle\langle\phi|). \tag{1.36}$$

Here we ask ourselves, how different is ρ_S form the canonical state Ω_S. The answer to this is provided by Popescu et. al. in [1,8], which states that ρ_S is very close to Ω_S for every pure state compatible with the constraint R. That is, for almost every pure state of the universe, the system behaves as if the universe were actually in the equiprobable mixed state $\mathcal{E}_R$. One might be tempted to think that Ω_S refers to the canonical state given in (1.23), but it in not necessarily true, since the restriction R general and it does not refer to the energy.

[3]Here it is implicit that the dimension of the environment is much larger than the one of the system.

As an illustrative analogy to typicality, consider our universe obeying the global constraint R as a map chart in which we see nothing but a vast ocean representing the pure states such that its reduced state is approximately the canonical state and some islands representing those states who differ drastically from the canonical state. Particularly we represent with a boat a random picked state in the universe, as one can see, for the vast majority of random choices we have that the boat would land in a portion of this ocean instead of an island with very high probability.

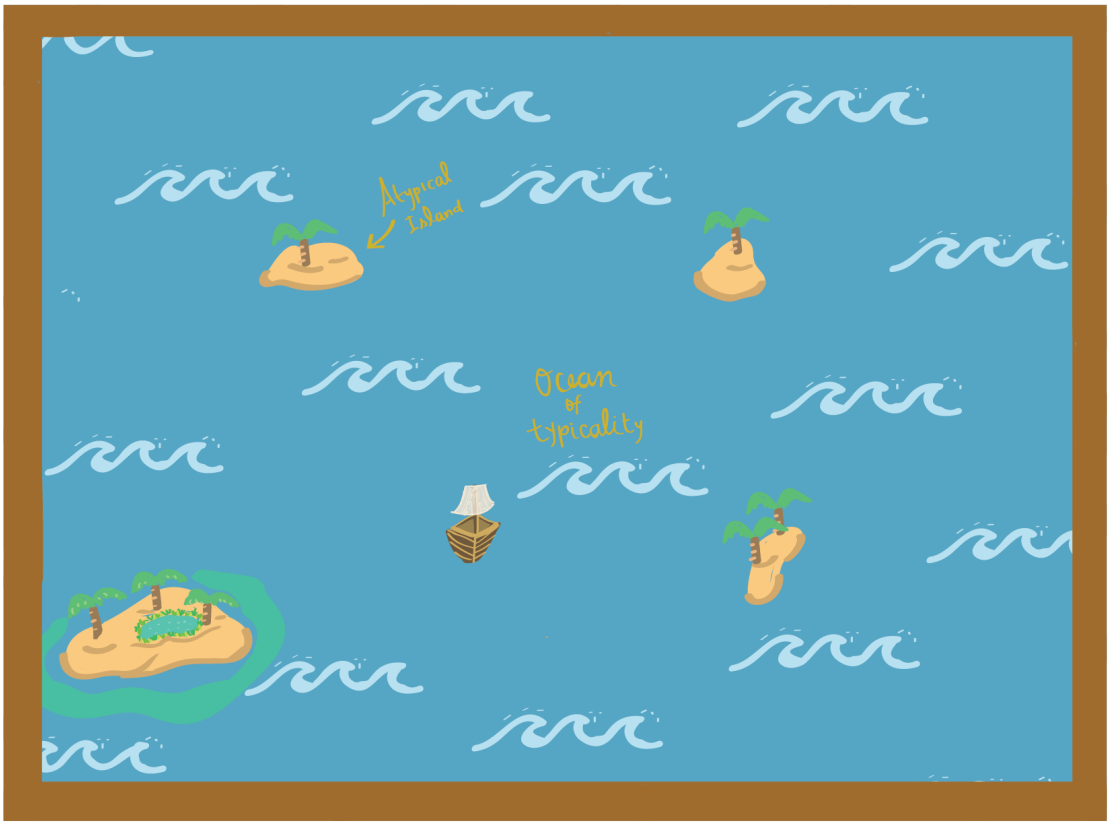

Figure 1.2: Illustration of the canonical typicality in a map chart. The boat here refers to a random picked state in our universe, and the islands refer to those states who radically differ from the Canonical state.

Up to this point we have only provided qualitative arguments about canonical typicality, in the following section we are going to present the respective quantitative arguments to fully understand these ideas.

1.2.3 Quantitative Arguments behind Typicality

In order to formally express canonical typicality, it is important provide more details. To start, it is necessary to first define a notion of distance between states ρ_S and Ω_S, as well as a measure over which pure states $|\phi\rangle$ are defined. Defining the distance of two operators is not a trivial task, as we will see depending on the distance we choose connections to a physical quantity could emerge. We start by defining the trace distance between ρ_S and the canonical state Ω_S, by $||\rho_S - \Omega||_1$, which is defined via the trace norm

$$||\rho||_1 = \mathrm{Tr}\,|\rho| = \mathrm{Tr}\left(\sqrt{\rho^\dagger \rho}\right). \tag{1.37}$$

An interesting fact about this distance is that it represents the maximal difference in probability of obtaining any outcome for any measurement performed on the states ρ_S and Ω_S. This can be clearly seen if we write explicitly the definition of an induced norm

$$\|\rho\| = \sup_{\|M\|=1} |\,\mathrm{Tr}(\rho M)|. \tag{1.38}$$

So, we can say that the trace distance measures how difficult is to know ρ_S and Ω_S via a measurement M, explicitly this is

$$|\mathrm{Tr}\,(\rho_S M) - \mathrm{Tr}\,(\Omega_S M)| \leq \|\rho_S - \Omega_S\|_1 \,\|M\|, \tag{1.39}$$

Having decided to use the trace norm, consider $|\phi\rangle$ to be a pure state in $\mathcal{H}_R$, with respective dimension d_R. As the state is normalized ($\langle\phi|\phi\rangle = 1$) we know that the pure state $|\phi\rangle$ lives in a $(2d_R - 1)$-dimensional real sphere. The states we are interested on, live over the surface of a sphere of d_R dimensions, thus if we are interested in uniformly random sample pure states, we will have to sample with the measure $\sigma(\mathbb{S}^{2d_R-1})$, which is known as the Haar measure.

Once we chose the distance to use as well as the measure over we are sampling our states, we are ready to announce the general result in typicality.

Theorem 1.2.1 (Theorem of Canonical Typicality [1, 8].)**.** *For a random chosen state, sampled with the Haar measure, $|\phi\rangle \in \mathcal{H}_R \subset \mathcal{H}_S \otimes \mathcal{H}_B$ and arbitrary $\varepsilon > 0$ the distance between the reduced density matrix $\rho_S = \mathrm{Tr}_E(|\phi\rangle\langle\phi|)$ and the canonical state $\Omega_S = \mathrm{Tr}_E\,\mathcal{E}_R$ is given probabilistically by:*

$$\mathrm{Prob}\left(\|\rho_S - \Omega_S\|_1 \geq \eta\right) \leq \eta', \tag{1.40}$$

where

$$\eta = \varepsilon + \sqrt{\frac{d_S}{d_E^{\mathrm{eff}}}}, \quad \eta' = 2\exp\left(-C d_R \varepsilon^2\right), \tag{1.41}$$

with

$$C = \frac{1}{18\pi^3}, \quad d_E^{\text{eff}} = \frac{1}{\operatorname{Tr}\Omega_E^2} \geq \frac{d_R}{d_S}, \quad \Omega_E = \operatorname{Tr}_S \mathcal{E}_R \tag{1.42}$$

Notice that η and η' are small quantities, so, it means that the state will be close to the canonical state with high probability, whenever $d_E^{\text{eff}} \ggg d_S$ and $d_R\varepsilon^2 \ggg 1$. What the typicality is telling us is that probabilistically speaking, if the dimension of the accessible space (d_R) is large enough, we will have that for the overwhelming majority of choices of random pure states, will have almost certainly that every, with small enough dimension, will be indistinguishable from the canonical state. Even more, Popescu et. al. get an expression to show how the fluctuations around the average behave, showing explicitly, via the levy lemma, that

$$\langle \|\rho_S - \Omega_S\|_1 \rangle \leq \sqrt{\frac{d_S}{d_E^{\text{eff}}}} \leq \sqrt{\frac{d_S^2}{d_R}}, \tag{1.43}$$

which tells us, if the dimension of our accessible space d_R is large enough, compared to the dimension of the system, the fluctuations on the system are also very small. Thus these two results give us the quantitative arguments of typicality.

Up to this moment we have provided the qualitative as well quantitative arguments to show why typicality gives us a different approach to the one used in equal a priori probability postulate. However, despite this result explains very well the reason why by randomly choosing a state $|\phi\rangle$ over the Haar measure, it coincides with the canonical state in almost all cases, it does not explain the way a state out of equilibrium (atypical state), reaches equilibrium, because as we saw, no particular evolution was considered, and only probabilistic arguments were used. In the next section we are going to see how from a typicality viewpoint, Linden et. al. were able to show that thermalisation can occur in a system reliant on a given unitary dynamic.

1.3 Evolution Towards Equilibrium.

As we pointed out, typicality is a result which is only valid for a given time and a given state, meaning that this can not be used to study how thermalisation occurs as a process. Specifically, we are interested in states that are atypical, in the sense that are states that locally differ drastically from the canonical state. One of the motivations behind treating

the problem of thermalisation through typicality's ideas is that since typicality holds for a huge amount of states, we might think that then, most evolutions will quickly drive us from a state in which the system is not thermalised into one that is, and that system will remain in this state of "thermalisation" for most of its evolution. To prove thermalisation is a much complicated problem and a closer look to it will allow us to picture a roadmap of what is needed to prove that themalisation could occur [15].

- **Equilibration**: It is possible to affirm that a system will equilibrate if its states evolve towards a particular state, which can be in the more general case a mixed state, and remains in that state, or at least quite close to it, for every time[4]

- **Environment state independence**: The equilibrium state that the system reaches has to be independent of the initial state of the environment, this is, when the system reaches its state of equilibrium, this state should depend only on macroscopic parameters of the environment, like temperature or similar macroscopic parameters.

- **System state independence**: If the system is much smaller than the environment, the state of equilibrium should not depend of its initial state.

Having this in mind, it is possible to tackle these problems one by one in order to prove thermalisation in a system. The result provided by Linden et. al. in [15], managed to prove that with relatively full generality, equilibrium is an universal property of Quantum systems and even more that equilibrium state does not depend on the state of the environment, and thus checking two of the listed elements.

To understand the work done by Linden et. al., it is necessary to first understand a couple of definitions they provide in their paper [15].

- **Universe:** Here we will refer always to a large quantum universe living in a Hilbert space $\mathcal{H}$. As previously, we are considering a universe that can be decomposed in two, in this decomposition we refer to the system S as a small part of the Hilbert space and the rest we will call it the environment. Explicitly we decompose the Hilbert space of the universe as a tensor product of the Hilbert space of the system and the

[4]Even though this definition does not specify the sort of equilibrium state of the system, the state of equilibrium will strongly depend on the initial condition of the system as well as the initial conditions over its environment.

environment, $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$, where d_S and d_E the respective dimension of the system and the environment. Notice that here the environment nor the system have been provided with any special property, meaning that for this formulation, the system could be a single particle or even a section of a lattice.

- **Hamiltonian:** The evolution of the universe will be governed by a Hamiltonian given by

$$\hat{H} = \sum_k E_k \left|E_k\right\rangle \left\langle E_k\right|. \tag{1.44}$$

with $|E_k\rangle$ the eigenstate in the energy basis with energy E_k. The main assumption we will require is that the Hamiltonian have non-degenerate energy gaps.

Expressing the last condition in a more explicit way, it is said that a Hamiltonian has non-degenerate energy gaps if any non-zero difference of eigenvalues of energy determine the two energy values involved. That is, for any four eigenstates with energy E_k, E_ℓ, E_m, E_n, satisfy that if $E_k - E_\ell = E_m - E_n$, then $m = n$ and $k = \ell$, or $k = m$ and $\ell = n$. Which turns out to be the same condition of imposing that the energy levels have to be non-degenerate. Conditions that implies that the energy levels are non-degenerate. Notice that the restriction imposed to the Hamiltonian is an extremely natural constraint, this is because all Hamiltonians that lack of symmetries have non-degenerate energies, so we refer to a set of Hamiltonians with measure 1 that fulfil this condition.

Notation: We will work here with pure time dependent states of the universe, states that will be represented by $|\Psi(t)\rangle$ with a time dependent density matrix given by $\rho(r) = |\Psi(t)\rangle \langle\Psi(t)|$. Thus the state of the system at a time t can be found by tracing out the environment, that is, $\rho_S(t) = \mathrm{Tr}_E\, \rho(t)$, and similarly we define the state of the environment as $\rho_E(t) = \mathrm{Tr}_S\, \rho(t)$.

We define a convenient quantity which is known as the transient state, or the time averaged state ω

$$\omega = \langle \rho(t) \rangle_t = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \rho(t) dt, \tag{1.45}$$

and similarly, we define ω_s and ω_E as the time averaged state of the system and the environment respectively.Finally, it is also convenient to re introduce a concept we have already used, the effective dimension of a mixed state ρ:

$$d^{\text{eff}}(\rho) = \frac{1}{\mathrm{Tr}\,(\rho^2)}, \tag{1.46}$$

which is generally a better measurement of the effective dimension than the dimension of the support of ρ. What this measure tells us is how many states contribute to the mixture, carrying the probabilistic weight of different states in the mixture, and different than the support, it is a continuous measure.

With the concepts aforementioned Linden et. al. [15] are able to mathematically prove

> *Every pure state of a quantum universe, composed by a large number of eigenstates of energy*[5] *such that evolves under an arbitrary Hamiltonian, is such that every system small enough will equilibrate.*

Where the reason of requiring the universe to have many changes in its time evolution, is because for equilibration to take place it is needed that part of the information of the initial state of the system leaves the system and enters in the environment. This notion of evolving through many states can be mathematically encapsulated via the effective dimension of the time average state $\omega = \langle \rho(t) \rangle_t$, and the connection between this and the number of eigenstates is with ease seen by expanding $|Psi(t)\rangle$ as

$$|\Psi(t)\rangle = \sum_k c_k e^{-iE_k t} |E_k\rangle \tag{1.47}$$

where $\sum_k |c_k|^2 = 1$ and hence

$$\rho(t) = \sum_{k,l} c_k c_l^* e^{-i(E_k - E_l)t} |E_k\rangle \langle E_l| , \tag{1.48}$$

which can be expanded and written as

$$\begin{aligned} \rho(t) &= \underbrace{\sum_n \|c_n\|^2 |E_n\rangle \langle E_n|}_{\omega} + \underbrace{\sum_{m \neq n} c_n c_m^* |E_n\rangle \langle E_m| e^{-it(E_n - E_m)}}_{\lambda(t)} \\ &= \omega + \lambda(t), \end{aligned} \tag{1.49}$$

Which for the case of non-degeneracy of the energy levels we have that the cross-terms vanish,

$$\omega = \langle \rho(t) \rangle_t = \sum_k |c_k|^2 |E_k\rangle \langle E_k| , \tag{1.50}$$

[5] The reason why we need the global state to have many eigenstates of energy is because by imposing this we can assure that there will be a large quantity of changes throughout the evolution of the system.

which will lead us to

$$d^{\text{eff}}(\omega) = \frac{1}{\text{Tr}\left(\omega^2\right)} = \frac{1}{\sum_k |c_k|^4}. \tag{1.51}$$

The above-mentioned result can be mathematically written in terms of central quantity $D\left[\rho_S(t), \omega_S\right]$, the trace distance between $\rho_S(t)$, the state of the system at a time t, and its time average, $\omega_S = \langle \rho_S(t) \rangle_t$. The main difference between $\rho_S(t)$ and ω_S in terms of the energy eigenstates can be written as

$$\rho_S(t) - \omega_S = \sum_{m \neq n} c_m c_n^* e^{-i(E_m - E_n)t} \, \text{Tr}_E \left| E_m \right\rangle \left\langle E_n \right|. \tag{1.52}$$

Since in general we know that $\rho_S(t)$ fluctuates around the state ω_S, it is evident that the distance between them will change over time. Thus, we can characterise the fluctuations of this quantity, and we will be interested in the time average of distance $\langle D\left[\rho_S(t), \omega_S\right] \rangle_t$, thus the value this average takes will tell us about where the system is spending most of its time. In other words $\langle D\left[\rho_S(t), \omega_S\right] \rangle_t$ will be small when the system equilibrates to ω_S. To be able to prove what is announced as the *Theorem 1* in [15] it is useful to relate the trace distance to the square of the Hilbert-Schmidt distance using a standard bound provided in [45]

$$D\left(\rho_1, \rho_2\right) = \frac{1}{2} \, \text{Tr}_S \sqrt{\left(\rho_1 - \rho_2\right)^2} \leq \frac{1}{2} \sqrt{d_S \, \text{Tr}_S \left(\rho_1 - \rho_2\right)^2}. \tag{1.53}$$

Which combined with the concavity of the square-root function, we have,

$$\langle D\left[\rho_S(t), \omega_S\right] \rangle_t \leq \sqrt{d_S \left\langle \text{Tr}_S \left[\rho_S(t) - \omega_S\right]^2 \right\rangle_t}, \tag{1.54}$$

which will provide us the bound we need to proof the theorem. Now using (1.52) we write

$$\left\langle \text{Tr}_{\mathcal{S}} \left[\rho_{\mathcal{S}}(t) - \omega_{\mathcal{S}}\right]^2 \right\rangle_t = \sum_{m \neq n} \sum_{k \neq l} \mathcal{T}_{klmn} \, \text{Tr}_{\mathcal{S}} \left(\text{Tr}_E \left| E_k \right\rangle \left\langle E_l \right| \text{Tr}_E \right| E_m \rangle \left\langle E_n \right| \right), \tag{1.55}$$

where $\mathcal{T}_{klmn} = c_k c_l^* c_m c_n^* e^{-i(E_k - E_l + E_m - E_n)t}$. Computing this time average taking into account

that the Hamiltonian has non-degenerate energy gaps, we find that

$$\begin{aligned}
\left\langle \mathrm{Tr}_S \left[\rho_S(t) - \omega_S\right]^2 \right\rangle_t &= \sum_{k\neq l} |c_k|^2 |c_l|^2 \,\mathrm{Tr}_S \left(\mathrm{Tr}_E \left|E_k\right\rangle \left\langle E_l\right| \mathrm{Tr}_E \left|E_l\right\rangle \left\langle E_k\right|\right) \\
&= \sum_{k\neq l} |c_k|^2 |c_l|^2 \sum_{ss'bb'} \langle sb|E_k\rangle \langle E_l|s'b\rangle \langle s'b'|E_l\rangle \langle E_k|sb'\rangle \\
&= \sum_{k\neq l} |c_k|^2 |c_l|^2 \sum_{ss'bb'} \langle sb|E_k\rangle \langle E_k|sb'\rangle \langle s'b'|E_l\rangle \langle E_l|s'b\rangle \\
&= \sum_{k\neq l} |c_k|^2 |c_l|^2 \,\mathrm{Tr}_E \left(\mathrm{Tr}_S \left|E_k\right\rangle \left\langle E_k\right| \mathrm{Tr}_S \left|E_l\right\rangle \left\langle E_l\right|\right) \\
&= \sum_{k\neq l} \mathrm{Tr}_E \left[\mathrm{Tr}_S \left(|c_k|^2 \left|E_k\right\rangle \left\langle E_k\right|\right) \mathrm{Tr}_S \left(|c_l|^2 \left|E_l\right\rangle \left\langle E_l\right|\right)\right] \\
&= \mathrm{Tr}_E\, \omega_E^2 - \sum_k |c_k|^4 \,\mathrm{Tr}_S \left[\left(\mathrm{Tr}_E \left|E_k\right\rangle \left\langle E_k\right|\right)^2\right] \\
&\leq \mathrm{Tr}_E\, \omega_E^2,
\end{aligned} \tag{1.56}$$

where $\omega_E = \mathrm{Tr}_S\, \omega$. To obtain a further bound, we invoke weak sub-additivity of the Reényi entropy [46]

$$\mathrm{Tr}\left(\omega^2\right) \geq \frac{\mathrm{Tr}_E\left(\omega_E^2\right)}{\mathrm{rank}\left(\rho_S\right)} \geq \frac{\mathrm{Tr}_E\left(\omega_E^2\right)}{d_S}, \tag{1.57}$$

and therefore combining (1.54), (1.56) and (1.57) we get

$$\langle D\left[\rho_S(t), \omega_S\right]\rangle_t \leq \frac{1}{2}\sqrt{d_S \,\mathrm{Tr}_E\left(\omega_E^2\right)} \leq \frac{1}{2}\sqrt{d_S^2 \,\mathrm{Tr}\left(\omega^2\right)}, \tag{1.58}$$

which by taking the definition of effective dimension, we get the main result shown in [15]

$$\langle D\left[\rho_S(t), \omega_S\right]\rangle_t \leq \frac{1}{2}\sqrt{\frac{d_S}{d^{\mathrm{eff}}\left(\omega_E\right)}} \leq \frac{1}{2}\sqrt{\frac{d_S^2}{d^{\mathrm{eff}}(\omega)}}. \tag{1.59}$$

As we can see, the result obtained by Linden et. al. tell us that the vast majority of quantum systems, in which the dynamic of the universe is governed by a Hamiltonian with no gaps, will spend most of its time close to its equilibrium state independently of its initial state. Note that this result is not necessarily considering that the state of equilibration will coincide with the canonical state, here only the state of equilibrium is taken into account.

Similarly as we pictured in canonical typicality,we can imagine a universe as the one described before, in which the islands refer to systems regions of the universe such that are atypical and the ocean the region which represent the regions where the system equilibrates. As shown in the figure 1.3 we can start from a region in the island (out of the equilibrium)

and going out to the ocean (region of equilibrium), as is shown we jump from the island to the ocean so that in average we spend most of the time out of the island.

Figure 1.3: Illustration of the result from Linden et. al. drawn in a map chart. The ship here refers to the state sailing in our universe, whereas the island refers to the region in which our ship is out of equilibrium.

This results is quite illustrative, since it is telling us that quantum mechanics has the intrinsic property to take systems and bring them to its correspondent equilibrium state.

We will now switch gears to the main goal of our work, once we have discussed the ideas behind typicality we are ready to state our problem properly. Consider two different orthogonal pure states living in the same Hilbert space ($|E_n\rangle, |E_m\rangle \in \mathcal{H}_R$), the Hilbert space associated with the global restriction R, From typicality, we know that the each of the reduced states of $|E_n\rangle, |E_m\rangle$ approximately leads to the canonical state, that is

$$\mathrm{Tr}_{\mathcal{E}} |E_n\rangle \langle E_n| \approx \mathrm{Tr}_{\mathcal{E}} |E_m\rangle \langle E_m| \approx \Omega_{\mathcal{S}}. \tag{1.60}$$

Thus, if we consider a third state $|\Psi\rangle$ to be a generic linear combination of $|E_n\rangle$ and $|E_m\rangle$,

fulfilling the same restriction as the universe. In this case we will have that the density matrix associated with the third state will read

$$\begin{aligned}\rho &= |\Psi\rangle\langle\Psi| \\ &= \|c_n\|^2 |E_n\rangle\langle E_n| + \|c_m\|^2 |E_m\rangle\langle E_m| \\ &+ c_n c_m^* |E_n\rangle\langle E_m| + c_m c_n^* |E_m\rangle\langle E_n| .\end{aligned} \tag{1.61}$$

If we take the partial trace of the equation (1.61), we will end up with the state of the system,

$$\begin{aligned}\rho_S &= \mathrm{Tr}_E\,\rho = \mathrm{Tr}_E\,|\Psi\rangle\langle\Psi| \\ &= \|c_n\|^2 \mathrm{Tr}_E\,|E_n\rangle\langle E_n| + \|c_m\|^2 \mathrm{Tr}_E\,|E_m\rangle\langle E_m| \\ &+ c_n c_m^* \mathrm{Tr}_E\,|E_n\rangle\langle E_m| + c_m c_n^* \mathrm{Tr}_E\,|E_m\rangle\langle E_n| .\end{aligned} \tag{1.62}$$

Notice that the way we constructed the state $|\Psi\rangle$, makes the non-crossed terms in (1.62) to approximately coincide with $\mathrm{Tr}_E\,|E_n\rangle\langle E_n| = \mathrm{Tr}_E\,|E_m\rangle\langle E_m| \approx \Omega(E)$, with $\Omega(E)$ the canonical state. Thus, the equation (1.62) can be written as

$$\rho_S = \Omega(E) + c_n c_m^* \mathrm{Tr}_E\,|E_n\rangle\langle E_m| + c_m c_n^* \mathrm{Tr}_E\,|E_m\rangle\langle E_n| . \tag{1.63}$$

So in order get the equality,the cross terms in (1.63) have to approximately vanish, this property of vanishing partial traces of exterior products of states is what we name *ultra-orthogonality* holds. that is,

$$\mathrm{Tr}_E\,|E_n\rangle\langle E_m| = \mathrm{Tr}_E\,|E_m\rangle\langle E_n| \approx 0. \tag{1.64}$$

Notice that this property appears naturally by just using the results of typicality, and more importantly, the property of getting vanishing partial traces over the crossed terms could explain the path to equilibrium as an instant process. Inspired on this interesting phenomena, we decided to explore ultra-orthogonality for the special case in which the terms in (1.64) of the left hand side are exactly equal to zero. Thus implying that

$$\mathrm{Tr}_{\mathcal{E}}\,\rho(t) \equiv \rho_{\mathcal{S}} = \omega_{\mathcal{S}}, \tag{1.65}$$

meaning that the correspondent reduced state of a fully interacting universe will be immediately constant. Our idea will be to explore if we can find Hilbert spaces in which ultra-orthogonality holds, and more importantly, to find the size of those Hilbert spaces, so that we will be able to answer questions such as "what is the biggest Hilbert Space in

which ultra-orthogonality will holds?". In the next chapters we will reduce our problem to Fermionic systems and we will show that for this specific case, we are able to find exponentially large Hilbert spaces where ultra-orthogonality holds.

Chapter 2

Fermionic Systems. A full Characterisation and Its Connection to the Grassmann Approach.

Up to this this point, we have discussed some interesting properties of quantum mechanics which lead to thermalisation as a consequence of the concentration measure in high dimensional spaces and entanglement, and by using these results we emphasised that ultra orthogonality could emerge as a consequence of typicality. As we mentioned before we are interested in the case when the cross terms of (1.64) are exactly equal to zero for the special case of Fermionic systems. In this Chapter we are Going to provide a background to understand how Fermionic states are usually treated and why is useful to work with them. More specifically, we provide the overview of solvable Fermionic systems, its connection to Majorana fermions and Gaussian states and the formalism of Grassmann for anticommuting variables. In this chapter, we are going to provide a background in how Fermionic systems are usually treated and why is useful to work with them. Particularly we will see an overview of solvable Fermion systems, we will discuss about the Jordan Wigner transformation and its importance to map systems composed by spins to Fermionic systems.

2.1 Overview

In many areas of physics one has to deal with solving quantum many body problems. this is often a computationally difficult if not impossible task. Nonetheless, the cases which can be analytically solved are very well known, and have been a subject of study [21, 47–58].

In spite of the different suppositions over each system, it has been found that a wide class of complicated Hamiltonians with many-body interactions can be often be mapped onto Hamiltonians that are quadratic in annihilation and creation operators and have the generic form [52]

$$\hat{H} = \sum_{ij} C_{ij} \hat{a}_i^\dagger \hat{a}_j + \sum_{ij} \left(A_{ij} \hat{a}_i^\dagger \hat{a}_j^\dagger + \text{h.c.} \right), \tag{2.1}$$

where i, j run from 1 to the number of modes in the system (N) and $\hat{a}_i$, $\hat{a}_i^\dagger$ are Fermionic annihilation and creation operators which satisfy the canonical anti-commutation relations (CAR) [59]

$$\{\hat{a}_k, \hat{a}_l\} = \left\{\hat{a}_k^\dagger, \hat{a}_l^\dagger\right\} = 0, \quad \left\{\hat{a}_k, \hat{a}_l^\dagger\right\} = \delta_{kl}. \tag{2.2}$$

A convenience when working with these kind of Hamiltonians is that can be diagonalized via a Bogoliubov - Valantin transformations transformation (i.e., canonical transformations), which maps Fermionic creation and annihilation operators on the creation and annihilation operators of non-interacting quasi-particles [60, 61]. Explicitly the transformation looks like

$$\begin{aligned} \hat{a}_i &\mapsto \gamma_i \hat{q}_i + \kappa_i \hat{q}_i^\dagger, \\ \hat{a}_i^\dagger &\mapsto \bar{\gamma}_i \hat{q}_i^\dagger + \bar{\kappa}_i \hat{q}_i. \end{aligned} \tag{2.3}$$

where γ_i, κ_i are complex numbers such that preserves the canonical anti-commutation relations given by (2.2) for $\hat{q}$, $\hat{q}^\dagger$[1].

Many relevant physics models are diagonalizable via a Bogoliubov-Valantin transformations, some examples are the Hubbard model, BCS theory of superconductivity in the mean field or Hartree-Fock approximation, and certain solvable spin-chain models (After a Jordan-Wigner transformation) [56–59]. Particularly will explain in detail the case of the 1-dimensional XY model to illustrate how is possible to understand systems of spins as fermions through a map between operators. In general, the XY Hamiltonian model refers

[1]This relation can also be expresses as a condition over γ_i, κ_i,

$$\gamma_i^2 + \kappa_i^2 = 1,$$

and

$$\{\hat{q}_k, \hat{q}_l\} = \left\{\hat{q}_k^\dagger, \hat{q}_l^\dagger\right\} = 0, \quad \left\{\hat{q}_k, \hat{q}_l^\dagger\right\} = \delta_{kl}.$$

to a set of N spin 1/2 particles located on the sites of a d-dimensional lattice. Nevertheless, whenever we refer to the XY model, we will be referring to the $1-$dimensional XY model.

2.1.1 XY model.

A chain of N spins where each spin is able to interact with its nearest neighbours in the X and Y coordinate as well as an external magnetic field, will be described by the Hamiltonian of the form

$$H_{XY} = -\frac{1}{2}\sum_{l=0}^{N-1}\left(\frac{1+\gamma}{2}\sigma_l^x\sigma_{l+1}^x + \frac{1-\gamma}{2}\sigma_l^y\sigma_{l+1}^y + \lambda\sigma_l^z\right), \tag{2.4}$$

where γ is so-called the anisotropy parameter and represents the difference between the strength of the XX interaction and the YY interaction, λ is the intensity of the external magnetic field and σ_l^i is the Pauli matrix ($i = x, y, z$) acting over the i site of the chain. The XY model is a model that has been widely studied for a variety of values of λ and γ and in some limits it has a correspondence to other models of interest in condensed matter [56–58], being the hard boson limit of the Boson Hubbard model one of them as well as the Kitaev chain, under a proper identification of the parameters μ, t and Δ with γ and λ [56, 57].

To explicitly show how is possible to map from an interacting system of spins to a free Fermionic system, we will show the details of the Jordan-Wigner transformation and we will see how it is possible to make the map between spins and free fermions in the case of the XY model.

2.1.2 Jordan-Wigner Transformation

The Jordan-Wigner transformation an important transformation used mainly in Fermionic systems. Jordan-Wigner transformation provides a bridge between spins and fermions through a non-local transformation that maps spin operators onto fermionic creation and annihilation operators, showing us then that there is no distinction between spins$-1/2$ particles and fermions [62]. Particularly, we will show how in the case of the XY model, the Jordan-Wigner transformation can be used to analytically find the spectrum of the model.

We first consider the non-local transformation given by

$$\hat{b}_l = \left(\prod_{m<l} \sigma_m^z\right) \sigma_l^-, \quad \sigma_l^- = \frac{\sigma_l^x - i\sigma_l^y}{2}, \tag{2.5}$$

where these b_l represent spin-less Fermionic operators, and its canonical anticommutation relation (CAR) is given by [21]

$$\left\{\hat{b}_i^\dagger, \hat{b}_j^\dagger\right\} = \left\{\hat{b}_i, \hat{b}_j\right\} = 0, \quad \left\{\hat{b}_i^\dagger, \hat{b}_j\right\} = \delta_{i,j}. \tag{2.6}$$

So inverting the transformation we get

$$\begin{aligned} \sigma_l^z &= 1 - 2\hat{b}_l^\dagger \hat{b}_l \\ \sigma_l^x &= \left(\prod_{m<l} \left(1 - 2\hat{b}_m^\dagger \hat{b}_m\right)\right)\left(\hat{b}_l^\dagger + \hat{b}_l\right) \\ \sigma_l^y &= i\left(\prod_{m<l} \left(1 - 2\hat{b}_m^\dagger \hat{b}_m\right)\right)\left(\hat{b}_l^\dagger - \hat{b}_l\right). \end{aligned} \tag{2.7}$$

The terms of interaction in the Hamiltonian will look as

$$\begin{aligned} \hat{\sigma}_l^x \hat{\sigma}_{l+1}^x &= \left(\hat{b}_l^\dagger - \hat{b}_l\right)\left(\hat{b}_{l+1}^\dagger + \hat{b}_{l+1}\right) \\ \hat{\sigma}_l^y \hat{\sigma}_{l+1}^y &= -\left(\hat{b}_l^\dagger + \hat{b}_l\right)\left(\hat{b}_{l+1}^\dagger - \hat{b}_{l+1}\right), \end{aligned} \tag{2.8}$$

and the Hamiltonian will look like,

$$H_{XY} = -\frac{1}{2}\sum_l \left[\left(\hat{b}_{l+1}^\dagger \hat{b}_l + \hat{b}_l^\dagger \hat{b}_{l+1}\right) + \gamma\left(\hat{b}_l^\dagger \hat{b}_{l+1}^\dagger - \hat{b}_l \hat{b}_{l+1}\right)\right] - \frac{\lambda}{2}\sum_l \left(1 - 2\hat{b}_l^\dagger \hat{b}_l\right), \tag{2.9}$$

after this transformation. The term of $-\lambda N/2$ is usually ignored since it cause only a gauge of the spectrum in the energy [21]. Thus, Note that we ended up with a Hamiltonian which only depends on creation and annihilation operators and that has a similar shape of (2.1). An important feature about the class of Hamiltonians described by (2.1), which we will discuss later on, is that not only the ground state, but all eigenstates describing a excitation in a set of quasi-particles, belong to the class of so-called Fermionic Gaussian states [52]. The importance about this class of states is that are fully characterized by its second order correlations, result that are known as the as Wick theorem [51, 63].

Now that we have discussed the Jordan-Wigner transformations, we will switch back to describe the standard way of describe Fermionic systems, we will talk about the Majorana fermions, which have an interpretation of being particles such that are their own antiparticle [50], and particularly we will see how this formalism helps to diagonalise Hamiltonians of the form of (2.1).

2.2 Majorana Fermions

For the kind of systems we are interested in it has been proven that Majorana operators are an useful tool to diagonalise fermionic systems as the one in (2.1) [53]. In this formalism, instead of $2N$ creation and annihilation operators is convenient to introduce $2N$ Hermitian operators, that is

$$\hat{c}_{2j-1} = \hat{a}_j^\dagger + \hat{a}_j, \quad \hat{c}_{2j} = (-i)\left(a_j^\dagger - a_j\right), \tag{2.10}$$

where these operators are analogous to coordinate and momentum operators for bosonic modes. From the Fermi-Dirac commutation relation, it follows that

$$\{\hat{c}_k, \hat{c}_l\} = 2\delta_{kl}. \tag{2.11}$$

Thus, the algebra generated by the operators $\{\hat{c}_i\}$ is known as the Clifford algebra and is denoted by $\mathcal{C}_{2N}$. By inspection, we can see that any linear transformation of the form $\tilde{\gamma}_\alpha = O_{\alpha\beta}\gamma_\beta$, where $O \in SO(2N)$ belongs to the special orthogonal group in $2N$ dimensions. The transformation between Fermionic and Majorana operators is achieved by a block matrix of form

$$\Omega = \begin{pmatrix} \mathbb{I} & \mathbb{I} \\ i\mathbb{I} & -i\mathbb{I} \end{pmatrix}, \tag{2.12}$$

so the map from Fermionic ($\vec{\hat{a}}^T = (\hat{a}_1, \ldots, \hat{a}_1^\dagger, \ldots)$) and Majorana ($\vec{\hat{c}}^T = (\hat{c}_1, \ldots, \hat{c}_1^\dagger, \ldots)$) operators will be written as $\Omega\vec{\hat{a}} = \vec{\hat{c}}$.When we change from Fermionic operators to Majorana operators, it is convenient to define de Fermionic covariance matrix which as we mentioned before, will fully characterise Gaussians states. In comparison to its boson counterpart the fermion Gaussian states have the property that correlation functions for the creation/annihilation operators are completely determined by the two-point functions as stated in Wick's theorem [63], and since this property is extensible to correlation function pertaining to a reduced subset of the modes, it follows that any partial (reduced) density matrix obtained from ρ remains also Gaussian.

2.3 Fermionic Covariance matrix

A system of N fermion modes, is described by a set of creation and annihilation operators $\hat{a}^\dagger, \hat{a}$ that satisfies the canonical anti-commutations relations given in (2.2). It is said that

the system is Gaussian whenever the correspondent density matrix ρ can be written as [64]

$$\rho = \bigotimes_{k=1}^{N} \tilde{\rho}_k, \quad \tilde{\rho}_k = \frac{1}{2}\left(1 - \lambda_i \left[\tilde{a}_i^\dagger, \tilde{a}_i\right]\right), \tag{2.13}$$

for a certain choice of mode basis $\tilde{a} = u_i^j \hat{a}_j + v_i^j \hat{a}_j^\dagger$, and with $|\lambda_i| \leq 1$ (where the equality holds for pure states). Equivalently, Gaussian states are fully characterized by their second moments, so an equivalent form of writing ρ is

$$\rho = \frac{1}{Z} \cdot \exp\left[-\frac{i}{4}\hat{c}^T G \hat{c}\right], \tag{2.14}$$

with $\hat{c} = (\hat{c}_1, \hat{c}_2, \ldots, \hat{c}_{2N})$, the vector of Majorana operators (2.10), Z a normalization constant and G real anti-symmetric $2N \times 2N$ matrix. Since G is a skew-symmetric matrix, it can always be brought to the block diagonal form

$$OGO^T = \bigoplus_{i=1}^{N} \begin{pmatrix} 0 & -\beta_j \\ \beta_j & 0 \end{pmatrix} \quad \text{with} \quad O \in \text{SO(2N)}, \tag{2.15}$$

by a special orthogonal matrix $O \in SO(2N)$ where the β_j are called the Williamson eigenvalues of the matrix G. From equation (2.14) note that Gaussian states have an interpretation as thermal (Gibbs) states corresponding to a Hamiltonian of the form

$$\hat{H} = \frac{i}{4}\hat{c}^T G \hat{c} = \frac{i}{4}\sum_{k>l} G_{kl}\left[\hat{c}_k, \hat{c}_l\right]. \tag{2.16}$$

Equation (2.15) shows that every Gaussian state has a normal mode decomposition in terms of N single modes of the form (2.13) ($\sim \exp(-\beta \hat{a}^\dagger \hat{a})$) [65]. Realising that a Gaussian state can be decomposed in this way, provides an intuition of how a state can be fully determined by the expectation values of quadratic operators ($\hat{a}_i^{(\dagger)}\hat{a}_j^{(\dagger)}$ and $\hat{a}_i^\dagger \hat{a}_j$). When we collect these expectation values in a real and skew-symmetric, we will construct the covariance matrix Γ, defined via

$$\Gamma_{kl} = \frac{i}{2}\,\text{tr}\left(\rho\left[c_k, c_l\right]\right). \tag{2.17}$$

Thus, we can bring this anti-symmetric matrix to its block diagonal form, via a canonical transformation, as

$$\tilde{\Gamma} = O\Gamma O^T = \bigoplus_{i=1}^{M} \begin{pmatrix} 0 & \lambda_i \\ -\lambda_i & 0 \end{pmatrix}, \tag{2.18}$$

with $\lambda_i \geq 0$ the Williamson eigenvalues. In terms of the creation/annihilation operators obtained from the transformed $\tilde{c} = O\hat{c}$, the Gaussian state ρ takes the form (2.14) with $\tilde{\Gamma}$ as

its Fermionic covariance matrix. It is easy to check that the relation between G and Γ will be given by $\lambda_i = \tanh(\beta_i/2)$, for $i = 1, 2, \ldots, N$ [65].

The equivalence between the special orthogonal group in $2N$ dimensions ($SO(2N)$) and the Fermionic Gaussian states, drives to an interesting property about states describing multi-particles excitations. If $|vacuum\rangle$ is the ground state of some Hamiltonian, with annihilation operators $\hat{a}_i$ in a given quasi-particle basis, then $\hat{a}_i^\dagger |vacuum\rangle = \hat{c}_{2i} |vacuum\rangle$. Meaning that if any multi-particle state of this kind is obtained from the ground state $|vacuum\rangle$ through some transformation, such that preserves the canonical anti-commutation, the state will remain Gaussian. In other words, Gaussian states are preserved under any unitary transformation that preserves anti-commutation relations. The fact that all eigenstates of the Hamiltonian in (2.1) are Gaussian is an important property, because it means that excited states can also be treated with the Covariance matrix formalism.

We have presented details about some generalities in quadratic Hamiltonians, the equivalence between $1/2-$spin systems and fermions, and how fermionic systems described by (2.1) can be brought to its diagonal form, to be analytically solved. In the next chapter we will expound some historical background concerning error correcting code theory, as well as some of the most relevant results in the bounds in minimum distance codes. Once we have explained all the tools of code theory that will be needed, we will show how is possible to connect the formalism we discussed in this chapter with code theory, and interpret Fermionic states as binary sequences of excitations, to prove that is possible to find exponentially large Hilbert spaces in which ultra-orthogonality will holds.

Chapter 3

Exploiting The Tools From Theory Code in Fermionic Systems.

In the last chapter we talked about Fermionic systems and the formalism used when dealing with it. In this chapter we first are going to mention some concepts and definitions in order to understand how the coding theory can be linked to Fermionic states, particularly we will discuss the case of random minimum codes and well known results about this subject. We will talk about the exponent errors for random minimum codes and we will show how all this theory can be linked to the case Fermionic random minimum codes and compute its correspondent exponent error.

3.1 Error correcting Code Theory

In 1948, Claude Shannon presents his extraordinary work named "A Mathematical Theory of Communication" [41] in which he provided a precise measure of the information content of a random variable in terms of its entropy. His work is divided in two parts the noiseless coding theorem and the noisy channel theorem. For the purpose of our needs we are going to focus only in the second part of his work, which states that a reliable communication is possible if we use schemes such that its rate is less than the capacity of the channel. Even though he never provided an idea of how this schemes could be found, his work is considered one of the most relevant discovery of the century. Here we will consider codes in communication scenario, as the one showed in figure 3.1, meaning that there will be a sender who wants to send k message symbols over a noisy channel and there will be a receiver who has to correct possible errors over the sent code to fully interpret it. The sender will first encodes the k

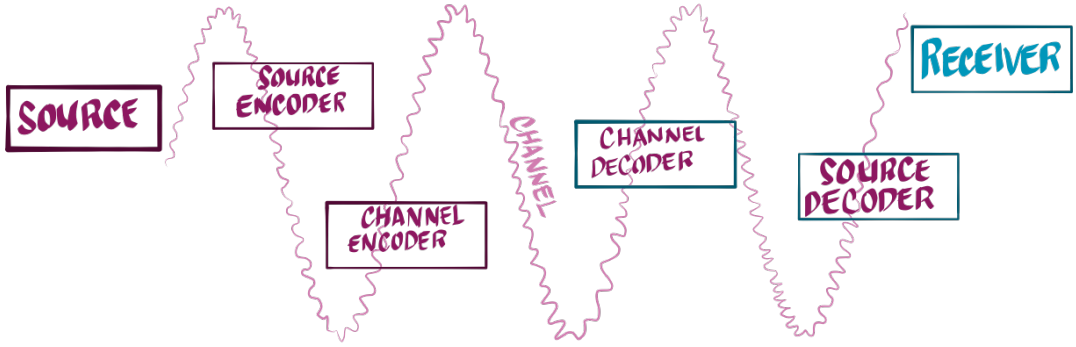

Figure 3.1: Representation of the scheme of communication. In the image the noise in the channel is represented by the noise wavy connection between the parts in the communication.

message symbols into n symbols. The receiver then tries to recover the original k message symbols. thus, encoding is the process of adding redundancy and decoding is the process of removing errors and the communication can only be done over the channel [40]. The most fundamental question one can ask is what will be the relation between the amount of redundancy and the errors that can be corrected, and in order to answer this question we will provide some useful definitions.

3.2 Some basic definitions

Definition 3.2.1 (Code)**.** *A code of block C length n over an alphabet Σ is a subset of Σ^n. If $|\Sigma| = q$, we say that C is a q-ary code.*

It is worth mention that, associated with a code there is also an encoding map E which maps the message set $\mathcal{M}$, identified in some canonical way with $\{1, 2, \ldots, |C|\}$ say, to code words belonging to Σ^N, and thus we have to understand the code as the image of the encoding map [66].

Definition 3.2.2 (Dimension of a code)**.** *Given a code $C \subset \Sigma^n$, its dimension is given by*

$$k \stackrel{\text{def}}{=} \log_q |C|, \tag{3.1}$$

An interesting fact about defining the dimension of the code in this way is that implicitly it is telling us that when working with codes exponential growth will be always taken into account.

We have to provide here a way to measure the amount of redundancy in a given message.

Definition 3.2.3 (Rate of a code)**.** *The rate of a code with dimension k and block length n is given by*

$$R \stackrel{def}{=} \frac{k}{n}, \tag{3.2}$$

This definition is nothing but the average amount of non redundant information each of the n symbols transmitted over the channel.
However, an alternative, and more general way of defining this is via the size of the code and the alphabet as

$$R(C) = \frac{\log |C|}{n \log |\Sigma|}. \tag{3.3}$$

Definition 3.2.4 (Hamming distance)**.** *The Hamming distance between two strings x and y of the same length over a finite alphabet Σ, denoted $\Delta(x, y)$, is defined as the number of positions at which the two strings differ, i.e, $\Delta(x, y) = |\{i|x_i \neq y_i\}|$. The fractional Hamming distance or relative distance between $x, y \in \Sigma^n$ is given by $\delta(x, y) = \frac{\Delta(x,y)}{n}$.*

It is trivial to check that the Hamming distance defines a metric on Σ^n.

Definition 3.2.5. ***Hamming weight:*** *The Hamming weight of a string x over alphabet Σ is defined as the number of non-zero symbols in the string. More formally, the Hamming weight of a string $\mathcal{W}(x) = |\{i|x_i \neq 0\}|$. Note that $\mathcal{W}(x - y) = \Delta(x, y)$.*

Given a string $x \in \Sigma^N$, the Hamming ball of radius r around x is the set $\{y \in \Delta^n | \Delta(x, y) \leq r\}$.

The minimum distance, or simply distance, of a code C, denoted $\Delta(C)$, is defined to be the minimum Hamming distance between two distinct code words of C. That is

Definition 3.2.6 (Minimum distance)**.** *The minimum distance, or simply distance, of a code C, denoted $\Delta(C)$, is defined to be the minimum Hamming distance between two distinct code words of C. That is*

$$\Delta(C) = \min_{\substack{c_1, c_2 \in C \\ c_1 \neq c_2}} \Delta(c_1, c_2). \tag{3.4}$$

In particular, for every pair of distinct code words in C the Hamming Distance between them is at least $\Delta(C)$

The relative distance of C, denoted $\delta(C)$, is the quantity $\frac{\Delta(C)}{N}$, where N is the block length of C. Thus any two code words of C differ in at least a fraction $\Delta(C)$.

Definition 3.2.7 (Notation)**.** *A q-ary code of block length N and dimension k will be referred to as an $[N,k]_q$ code. Further, if the code has minimum distance d, it will be referred to as an $[N,k,d]_q$ code. When the alphabet size q is clear from the context, or not very relevant to the discussion, we omit the subscript.*

Up to this point we have only described specific codes, codes with fixed block length and dimension. However, since we are interested in the asymptotic behaviour, it turns out to be more useful the study of families of codes instead of an specific code.

Definition 3.2.8 (Family of Codes)**.** *Let $q \geq 2$. let $\{n_i\}_{i\geq 1}$ be and increasing sequence of block lengths and suppose there exists sequences $\{k_i\}_{i\geq 1}$ and $\{d_i\}_{i\geq 1}$ such that for all $i \geq 1$ there exist an $[n_i,k_i,d_i]_q$ code C_i. then the sequence $C = \{C_i\}_{i\geq 1}$ is a family of codes. The rate of C is defined as*

$$R(C) = \lim_{i\to\infty} \left\{ \frac{k_i}{n_i} \right\}, \tag{3.5}$$

and the relative distance of C is defined as

$$\delta(C) = \lim_{i\to\infty} \left\{ \frac{d_i}{n_i} \right\}, \tag{3.6}$$

from now on whenever we talk about a code we are implicitly referring to a family of codes.

For the purpose of this work, we are going to focus on a particular class of codes, name minimum distance code. This special kind of codes came to our interest since they seem to appear naturally when study typical Fermionic states. We will then show some of the main results on this particular kind of codes to after show how this results could be extended to our particular case.

When we talk about codes of minimum distance d we refer to codes which have the property that for every pair x, y of codewords we have that

$$\Delta(x,y) = d, \tag{3.7}$$

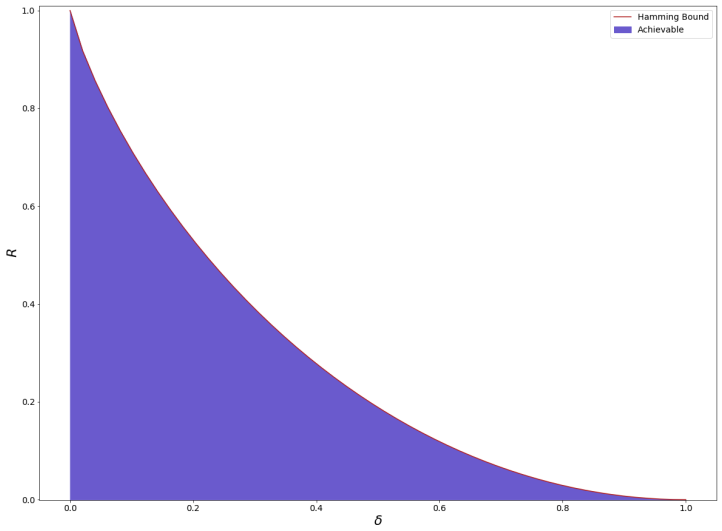

Figure 3.2: An illustration of the Hamming bound for the case of $q = 2$. Note any code above this bound could exist. The shaded region shows the codes that could exist.

and therefore, the its relative distance δ is given by $N\delta = d$.

When working with this kind of codes one may wonder what is the best rate we can achieve. Particularly we are going to show 3 results, 1 positive and 2 negative results[1]. Even though there are other known bounds, we will not talk about others but Gilbert-Varshamov, Hamming and Plotkin bounds, the reason for this is due to the fact that the other bound apply for large enough alphabets, so for binary codes we are not interested at all in these kind of bounds [40].

3.2.1 Hamming Bound (Sphere packing bound)

Definition 3.2.9 (Volume of Hamming ball)**.** *Let $q \geq 2$ and $n \geq r \geq 1$ be integers. Then the volume of a Hamming ball of radius r is given by*

$$\mathrm{Vol}_q(r, n) = |B_q(\mathbf{0}, r)| = \sum_{i=0}^{r} \binom{n}{i} (q-1)^i, \tag{3.8}$$

[1]Note that a negative result refer to an upper bound on the rate, meaning that the maximum achievable rate we could get can not exceed some value. Whereas a positive result refer to lower bound on the rates we can achieve.

where the choice of $\mathbf{0}$ *as the center of the Hamming ball is chosen arbitrary, since the volume of the Hamming ball is independent of its center.*

It is simple to show that

$$\frac{k}{n} \leq 1 - \frac{\log_q \mathrm{Vol}_q\left(\left\lfloor \frac{d-1}{2} \right\rfloor, n\right)}{n}, \tag{3.9}$$

where the volume in (3.9) correspond to the definition in(3.8). With some algebra and using the Stirling asymptotic approximation one can show that

$$\mathrm{Vol}_q\left(\left\lfloor \frac{d-1}{2} \right\rfloor, n\right) \geq q^{H_q\left(\frac{\delta}{2}\right)n - o(n)}, \tag{3.10}$$

where the latter inequality immediately provide us an upper bound on the rate

$$R \leq 1 - H_q\left(\frac{\delta}{2}\right) + o(1). \tag{3.11}$$

The inequality in (3.11) is known as the Hamming Bound.

3.2.2 Plotkin Bound

Definition 3.2.10 (Plotkin Bound)**.** *The following holds for any code* $\mathcal{C} \subset [q]^n$

- *If* $d = \left(1 - \frac{1}{q}\right) n, |\mathcal{C}| \leq 2qn.$
- *If* $d > \left(1 - \frac{1}{q}\right) n, |\mathcal{C}| \leq \frac{qd}{qd-(q-1)n}.$

Note that the Plotkin Bound implies that a code with relative distance $\delta \geq 1 - 1/q$*, must necessarily have* $R = 0$*.*

Definition 3.2.11. *For any q-ary code with relative distance* $0 \leq \delta \leq 1 - \frac{1}{q}$*,*

$$R \leq 1 - \left(\frac{q}{q-1}\right)\delta + o(1). \tag{3.12}$$

To illustrate the proof of this bound we can consider the distance $d = n\delta$. So we can shorten the codewords and group them in a way such that they agree on the first $n - n'$ symbols, with $n' = \left\lfloor \frac{qd}{q-1} \right\rfloor - 1$. Then in particular for any $x \in [q]^{n-n'}$, define the prefix code

$$\mathcal{C}_{\mathbf{x}} = \{(c_{n-n'+1}, \ldots c_n) \mid (c_1 \ldots c_N) \in \mathcal{C}, (c_1 \ldots c_{n-n'}) = \mathbf{x}\}. \tag{3.13}$$

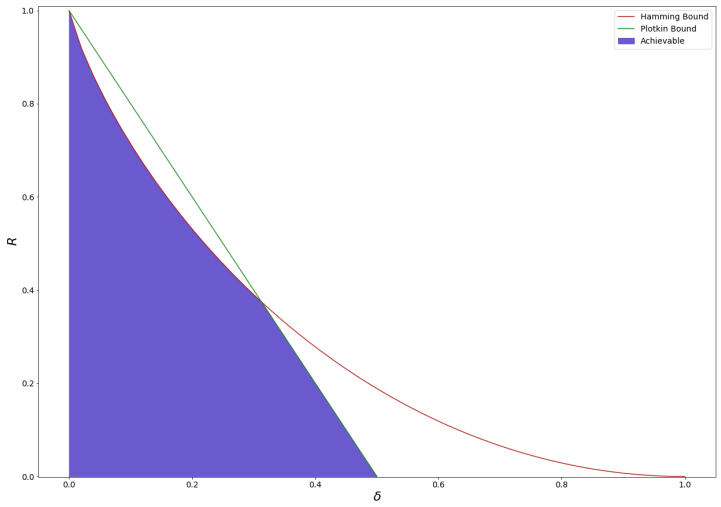

Figure 3.3: An illustration of Plotkin and Hamming bounds for the case of $q = 2$. For this case the shaded region changes indicating us that there are not codes with $R > 0$ when $\delta = 1 - 1/q$.

for all x, $\mathcal{C}_x$, has distance d as $\mathcal{C}$ has distance d. Additionally, it has block length $n' < \left(\frac{q}{q-1}\right) d$, and thus $d > \left(1 - \frac{1}{q}\right) n'$. From the Plotkin bound, this implies that

$$|\mathcal{C}_{\mathbf{x}}| \leq \frac{qd}{qd - (q-1)n'} \leq qd, \tag{3.14}$$

where the second inequality follows from the fact that $qd - (q-1)n'$ is an integer.

Note that from the definition of $\mathcal{C}_x$

$$|\mathcal{C}| = \sum_{\mathbf{x} \in [q]^{n-n'}} |\mathcal{C}_{\mathbf{x}}|\,, \tag{3.15}$$

which tell us that

$$|\mathcal{C}| \leq \sum_{\mathbf{x} \in [q]^{n-n'}} qd = q^{n-n'} \cdot qd \leq q^{n - \frac{q}{q-1}d + o(n)} = q^{n\left(1 - \delta \cdot \frac{q}{q-1} + o(1)\right)}\,, \tag{3.16}$$

in other words this provides another upper bound to the rate given by

$$R \leq 1 - \left(\frac{q}{q-1}\right)\delta + o(1) \tag{3.17}$$

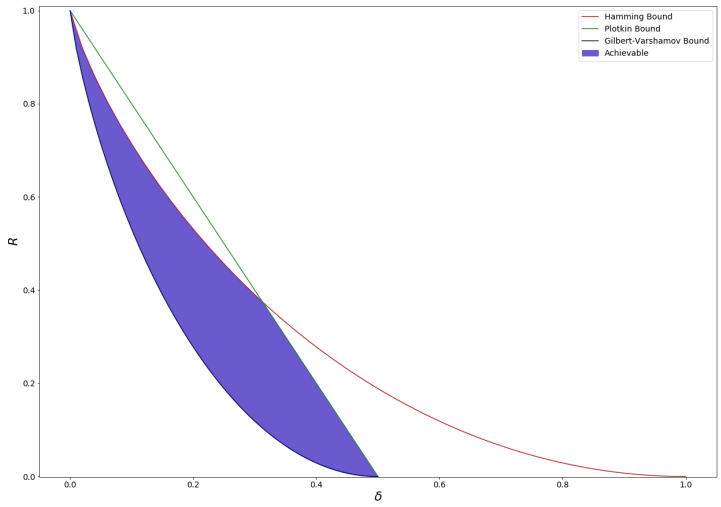

Figure 3.4: An illustration of 3 bounds Hamming, Plotkin and Gilbert-Varshamov for the case of $q = 2$. The lower bound correspond to the Gilvert-Varshamov Bound whereas the other 2 are the upper bound for the rate.

To close this part we will show the latter but positive result which provide us lower bound on the code rates, Gilbert Varshamov bound.

3.2.3 Gilbert Varshamov Bound

We now switch gears to provide a positive result. We will only provide the main ideas for the proof of this result an we will discuss why this result turn out to be one of the most important results.

Definition 3.2.12 (Gilbert-Varshamov Bound)**.** *Let $q \geq 2$. For every $0 \leq \delta < 1 - \frac{1}{q}$ and $0 < \varepsilon \leq 1 - H_q(\delta)$. There exist a code with rate $R \geq 1 - H_q(\delta) - \varepsilon$ and relative distance δ.*

To provide a main idea about the proof we can consider a greedy approach. First we start with an empty code $\mathcal{C}$ and we keep adding vectors that are not in $\mathcal{C}$ and that have Hamming distance at least d from all the existing codewords in $\mathcal{C}$. Notice that by doing so we can assure that we will never add a vector c that will make that will make the distance of $\mathcal{C}$ fall below d. Indeed it is easy to see that after doing so, we have

$$\bigcup_{\mathbf{c} \in C} B(\mathbf{c}, d-1) = [q]^n, \tag{3.18}$$

this is easily checked, because if it was not true, then there would exist a vector $\mathbf{v} \in [q]^n \backslash C$, such that $\Delta(\mathbf{v}, \mathbf{c}) \geq d$ and therefore $\mathbf{v}$ can be added. How ever, this would contradict the fact that we have finished the procedure. So

$$\bigcup_{c \in C} B(\mathbf{c}, d-1) \mid = q^n. \tag{3.19}$$

It isn't hard to see that

$$\sum_{\mathbf{c} \in C} |B(\mathbf{c}, d-1)| \geq \left| \bigcup_{\mathbf{c} \in C} B(\mathbf{c}, d-1) \right|, \tag{3.20}$$

which implies that

$$\sum_{\mathbf{c} \in C} |B(\mathbf{c}, d-1)| \geq q^n, \tag{3.21}$$

but as mentioned before, the volume of the Hamming ball is translation invariant,

$$\sum_{\mathbf{c} \in C} \mathrm{Vol}_q(d-1, n) \geq q^n. \tag{3.22}$$

Since $\sum_{\mathbf{c} \in C} Vol_q(d-1, n) = Vol_q(d-1, n) \cdot |C|$

$$\begin{aligned} |C| &\geq \frac{q^n}{\mathrm{Vol}_q(d-1, n)} \\ &\geq \frac{q^n}{q^{nH_q(\delta)}} \\ &= q^{n(1-H_q(\delta))}. \end{aligned} \tag{3.23}$$

Therefore concluding the proof. It is worth mention that this way of proceeding the code have not any special structure but as one might think this algorithm will take exponentially long time to finish. How ever, one may wonder if there is a special kind of code that also achieve this rate. Indeed is possible to show that random linear codes lies, with high probability, on the Gilbert-Varshamov Bound. To pick a linear random code we only need pick a random $k \times n$ matrix, in which each entry is chosen uniformly and independently at random according to its alphabet [40,66]. Aside from these result providing some bounds on the rate of the minimum distance codes. We are interested in the performance of a special kind of random codes, more specifically binary codes over a binary-symmetric channel (BSC). In here we derive the minimum distance, distance distribution and error exponent of a typical random code (TRC) from a random code ensemble (RCE), as well as the one correspondent to a typical linear code (TLC) from a linear code ensemble (LCE) [22]. As mentioned by *A*

Barg, and G. D. Forney, Jr most of the important of these results are expressed in terms of the Gilbert-Varshamov distance $\delta_{GV}(R)$.

3.2.4 Error Exponents for Random Minimum Distance Codes

It is very well known that on a BSC with crossover probability p, the channel capacity is $C = 1 - H_2(p)$ [2]. The error coding exponent $E_r(R)$ is positive for $0 \leq R < C$ and given by [67, 68].

$$E_r(R) = \begin{cases} R_0 - R, & 0 \leq R \leq R_{\text{crit}} \\ E_{\text{sp}}(R), & R_{\text{crit}} \leq R \leq C \end{cases}, \tag{3.24}$$

where R_0, R_{crit} and $E_{\text{sp}}(R)$, are known as cutoff rate, critical rate and the sphere-packing exponent respectively. In [69] Gallager has shown that the random coding exponent is the true error exponent for the RCE on any discrete memoryless channel. Here we will show the main results provided in [22] and we will provide the ideas to derive these results, as we will show later this ideas of error exponents will be quite helpful to understand how it is possible to make a connection between minimum distance codes and typical Fermionic states.

Random Binary Codes

Consider a binary code C of length N and rate R bits per symbol is a set of $M = 2^{NR}$. For the case of RCE one compute the probability that by taking a random codeword $\mathbf{x}_i$ of length N it would have Hamming distance $d = N\delta$ from an arbitrary binary $N-$tuple $\mathbf{b}$ and see that it will be independent of $\mathbf{b}$ and equals to

$$\Pr\{d_H(\mathbf{x}_i, \mathbf{b}) = d\} = \binom{N}{d} \tilde{p}^d (1-\tilde{p})^{N-d}, \tag{3.25}$$

where $\tilde{p}$ corresponds to the probability of having a one. Under this RCE, two distances $d_H(\mathbf{x}_i, \mathbf{x}_j)$ and $d_H(\mathbf{x}_{i'}, \mathbf{x}_{j'})$ are independent random variables. So if we consider the number of unordered pairs of codewords $(\mathbf{x}_i, \mathbf{x}_j)$ with $i \neq j$ in C at a distance d apart

$$S_{\mathcal{C}}(d) = \sum_{i=0}^{M-1} \sum_{j=0}^{i-1} \Phi\left\{d_H\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = d\right\}, \tag{3.26}$$

where $\Phi\left\{d_H\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = d\right\}$ is equal to 1 if the condition $d_H(\mathbf{x}_i, \mathbf{x}_j)$ is satisfied and 0 otherwise. For the case of RCE on a BSC $S_{\mathcal{C}}(d)$ is a sum of $\binom{M}{2}$ pairwise independent, identically

[2] We will be using the notation $\mathcal{H}$ to refer to the binary entropy $\mathcal{H} \equiv H_2$

distributed random variables, so we have

$$\mathrm{E}S_{\mathcal{C}}(d) = \begin{pmatrix} M \\ 2 \end{pmatrix} \mathrm{E}\Phi \doteq 2^{N(2R-1+\mathcal{H}(\delta))}. \tag{3.27}$$

Therefore we are ready to state the following theorem

Theorem 3.2.1 (Minimum distance in RCE)**.** *For $0 \leq R < 1/2$ and any $\varepsilon > 0$, the probability that a code length N and rate R from the RCE has relative minimum distance less than $\delta_{GV}(2R) - \varepsilon$ goes to zero exponentially as $N \to \infty$. For $0 \leq R < 1$, if $d = N\delta$ is such that*

$$\delta_{GV}(2R) + \varepsilon \leq \delta \leq 1 - \delta GV(2R) - \varepsilon, \tag{3.28}$$

then the probability that the number of codeword pairs at a distance d satisfies $S_{\mathcal{C}}(d) \doteq 2^{N(2R-1+\mathcal{H}(\delta))}$ goes to one as $N \to \infty$.

Proof. For a given value of the code value of the code rate R we can choose d such that $d/N \to \delta \leq \delta_{\mathrm{GV}}(2R) - \varepsilon$. Then

$$\Pr\{S_{\mathcal{C}}(d) \geq 1\} \leq \mathrm{E}S_{\mathcal{C}}(d) \doteq 2^{-N(1-\mathcal{H}(\delta)-2R)} \to 0, \tag{3.29}$$

which in other words it tell us that with probability differing from 1 by an exponentially falling quantity, there will be no pairs at distance d. Notwithstanding this result, if $\delta_{\mathrm{GV}}(2R) + \varepsilon < \delta < 1 - \delta_{\mathrm{GV}}(2R) - \varepsilon$, then $1 - \mathcal{H}(\delta) < 2R$ and the average of number of pairs $\mathrm{E}S_{\mathcal{C}}(d)$ at a distance d is exponentially large. To see this, we can use the Chebyshev inequality, so for any $\delta > 0$, we have

$$\Pr\left\{|S_{\mathcal{C}}(d) - \mathrm{E}S_{\mathcal{C}}(d)| \geq \begin{pmatrix} M \\ 2 \end{pmatrix} \alpha\right\} \leq \frac{\mathrm{E}\Phi}{\begin{pmatrix} M \\ 2 \end{pmatrix} \alpha^2}, \tag{3.30}$$

by choosing $\alpha \doteq 2^{-N(1-\mathcal{H}(\delta)+\Delta)} < \mathrm{E}\Phi$ for any $\Delta > 0$, we have

$$\Pr\left\{|S_{\mathcal{C}}(d) - \mathrm{E}S_{\mathcal{C}}(d)| > \begin{pmatrix} M \\ 2 \end{pmatrix} \alpha\right\} \leq \frac{2\mathrm{E}\Phi}{M(M-1)\alpha^2} \doteq 2^{-N(2R-1+\mathcal{H}(\delta)-2\Delta)}. \tag{3.31}$$

The exponent on the right-hand side can be made positive by choosing Δ small enough. This establishes the fact that $S_{\mathcal{C}}(d) \doteq 2^{N(2R-1+\mathcal{H}(\delta))}$ for the chosen value of d with probability tending to one as $N \to \infty$. □

Random Linear Codes

A binary linear code C of length N and rate K/N is a set of $M = 2^K$ binary N-tuples that is generated by K $N-$tuples $\mathbf{g}_j$, $1 \leq k \leq K$. This is

$$\mathbf{x}(u) = \Sigma_k \mathbf{u}_k \mathbf{g}_j, \tag{3.32}$$

where $\mathbf{u}$ is an arbitrary binary $K-$tuple. Here we will be considering the case in which each of the 2^{NK} matrices are chosen with equal probability. For the case of linear codes, the distribution of distance $\{d_H(\mathbf{x}_i, \mathbf{x}_j), i \neq j\}$ from any given codeword $\mathbf{x}_i$ is independent of i. The average distance distribution of a linear code C therefore reduces to

$$\mathcal{N}_{\mathcal{C}}(d) = \sum_{j \neq i} \Phi \left\{ d_H \left(\boldsymbol{x}_i, \boldsymbol{x}_j \right) = d \right\} \quad (d = 1, 2, \ldots, N), \tag{3.33}$$

where $\mathbf{x}_i$ is an arbitrary codeword. Typically, $\mathbf{x}_i$ is taken as the all zero codeword $\mathbf{0} = \mathbf{x}(\mathbf{0})$. If $(\mathbf{u}_j, \mathbf{u}_k)$ is any pair distinct non-zero $K-$tuples, then the corresponding codewords $(\mathbf{x}(\mathbf{u}_j), \mathbf{x}(\mathbf{u}_k))$ are a pair of independent random binary $N-$tuples. It follows that two distinct distances $d_H(\mathbf{x}_i, \mathbf{x}_j)$ and $d_H(\mathbf{x}_i, \mathbf{x}_k)$ from a given codeword $\mathbf{x}_i$ are pairwise-independent and distributed as in the RCE. In particular

$$\Pr \left\{ d_H \left(\boldsymbol{x}_i, \boldsymbol{x}_j \right) = N\delta \right\} \doteq 2^{-N(1-\mathcal{H}(\delta))}, \tag{3.34}$$

for any $d = N\delta$, the quantity $\mathcal{N}_{\mathcal{C}}(d)$ in (3.33) is a sum of $M - 1 \doteq 2^{NR}$ pairwise-independent, identically distributed random variables with mean $E\Phi \doteq 2^{-N(1-mathcalH(\delta))}$. Its mean value is therefore equal to

$$\mathcal{N}_{\mathrm{LCE}}(d) \doteq 2^{N(R-1+\mathcal{H}(\delta))}. \tag{3.35}$$

Therefore we say that the relative minimum distance of a code chosen at random from the LCE will be, with probability $1 - 2^{-\Omega(N)}$, approximately equal to the GV relative distance $\delta_{GV}(R)$.

In summary, the typical minimum distance in the LCE is better than that in the RCE because the minimum of only $M - 1$ pairwise-independent distances, whereas in the RCE it is the minimum of $\binom{M}{2}$ pairwise independent distances.

Up to this point one may be wonder how all this theory of correcting errors and minimum distance codes are connected to our problem. In following section we are going to show how

this connection emerge as a natural consequence of the structure in the Clifford algebra and therefore provide an explanation of what we named as "Super-Orthogonality".

3.3 Mechanism Behind Typicality as a Random Minimum Code.

As we mentioned in the first chapter, we are interested in study the behaviour of states which are close in energy. First we will show why it is possible to study "ultra-orthogonality" for Fermionic states in two possibilities. The first one is in terms of minimum distance codes and we will see that "ultra-orthogonality" turns out to be an exact result. For the second part, we deduce its correspondent exponent error. For the last part, we will tackle the problem when we can not work over a minimum distance and we will show how "ultra-orthogonality" can be acting over this specific case. Even though for this case we were not able to generalised this to every Fermionic system, we will discuss why this result should holds in general.

3.3.1 "ultra-orthogonality" on Fermions

To start this part we want to recapitulate a couple of things. First, as we showed in equation (1.52), when we take the partial trace the remaining cross terms will be of our interest. When we discuss the property of ultra-orthogonality, we stated that the cross terms of $\mathrm{Tr}_E |E_n\rangle \langle E_m|$ should be the ones that had to be near to zero, in order for thermalisation to occur on the specific case we discussed in chapter 1. Here we are going to formalise all these ideas.

We can start by choosing an arbitrary state such that it can be decomposed in its vectors of the Fock basis as[3]

$$|\psi\rangle = \sum_{\vec{n}} \psi(\vec{n}) |\vec{n}\rangle . \tag{3.36}$$

Therefore we are interested in studying the terms of the form[4]

$$\hat{X}_{ij} \doteq \mathrm{Tr}_E \left(|\vec{n}_i\rangle \langle \vec{n}_j| \right) \equiv \mathrm{Tr}_{N/L} \left(|\vec{n}_i\rangle \langle \vec{n}_j| \right) . \tag{3.37}$$

[3]Note that we can use the Fock basis since it is naturally the basis in which we diagonalise our Hamiltonian, therefore it is an energy basis as well.

[4]Here we denote the partial trace by $\mathrm{Tr}_{N/L}$ meaning that for the corresponding sequence of length N in the Fock Space we take the respective L elements, meaning that $N - L$ elements will be regarded as our environment.

What ultra-orthogonality tell us is that this term has to be zero or in the worst scenario something of the order of fluctuations. Since we are working on the very special case of Quadratic Hamiltonian describing Fermionic systems, we recall the fact that the operators are generated by the Majorana operators, and form the so called Clifford algebra, described in section 2.2. We also saw that these operators can be mapped to the Grassman variables, which allow us to compute things like observables. Taking into account that we started with an space in which we had to deal with N Pauli operator and we changed to a new space in which we work with $2N$ spinless operators and the fact that the most general function we can built over the Grassman algebra is polynomial of the Grassmann variables. Thus we construct a function of the Grassman variables which takes two binary sequences $(\vec{x}, \vec{y})$, $x_i, y_i \in \{0, 1\}$, lets call it $\gamma(\vec{x}, \vec{y})$. The reason for defining this function, is that we are going to describe the system in the space of size L with its correspondent operators, meaning that these sequences can not be arbitrary, they must have to be sequences such that after its first L elements, they must have only zeros. To illustrate this, consider the following example, let $\vec{x} = (0100\ldots0)$, $\vec{y} = (1100\ldots0)$, here $L = 3$, so our function will be described by[5]

$$\gamma(\vec{x}, \vec{y}) = \gamma_1^{x_1=0}\gamma_4^{y_1=1}\gamma_2^{x_2=1}\gamma_5^{y_2=1}\gamma_3^{x_3=0}\gamma_6^{y_3=1} = \gamma_4\gamma_2\gamma_5\gamma_6, \tag{3.39}$$

Note that in this way we are able to write any product of Grassmann operators. in general, for two sequences, and a fixed size L, this function will be given by

$$\gamma(\vec{x}, \vec{y}) = \gamma_1^{x_1}\gamma_{L+1}^{y_1}\gamma_2^{x_2}\gamma_{L+2}^{y_2}\ldots\gamma_L^{x_L}\gamma_{2L}^{y_L}, \tag{3.40}$$

where this function has the property that

$$\gamma(\vec{x}, \vec{y})\gamma(\vec{x}', \vec{y}') = e^{i\phi(\vec{x},\vec{y},\vec{x}',\vec{y}')}\gamma(\vec{x} + \vec{x}', \vec{y} + \vec{y}'). \tag{3.41}$$

The phase appear as a consequence of the anti-commutation relation and $\phi(\vec{x}, \vec{y}, \vec{x}', \vec{y}')$ is a function that would depend on the weight of the sequences $\vec{x}, \vec{y}, \vec{x}'$and $\vec{y}'$.

This provide a set of operators that live in the space of size L, and that will allow us to

[5]Here we have changed the original notation given in (2.10) and change it by

$$\hat{c}_{2j-1} \to \gamma_j, \qquad \hat{c}_{2j} \to \gamma_{N+j}, \tag{3.38}$$

for the case of N operators.

expand our operator $\hat{X}_{ij}$,[6]

$$\hat{X}_{ij} = \sum_{\vec{x},\vec{y}} f(\vec{x},\vec{y})\gamma(\vec{x},\vec{y}). \tag{3.42}$$

Our task will then be to find the coefficient $f(\vec{x},\vec{y})$, this can be achieve by multiplying on both sides by $\gamma(\vec{x'},\vec{y'})$ and taking the trace over L

$$\mathrm{Tr}_L\left(\hat{X}_{ij}\gamma^\dagger(\vec{x}',\vec{y}')\right) = \sum_{\vec{x},\vec{y}} f(\vec{x},\vec{y})\,\mathrm{Tr}_L\left(\gamma(\vec{x},\vec{y})\gamma^\dagger(\vec{x}',\vec{y}')\right). \tag{3.43}$$

The right hand side of this equation can be combined with equation (3.41) and deduce that it will give us a delta ($\delta_{\vec{x}+\vec{x}'}\delta_{\vec{y}+\vec{y}'}$). Thus the coefficients $f(\vec{x},\vec{y})$ are given by

$$f(\vec{x}',\vec{y}') = \frac{1}{2^L}\,\mathrm{Tr_L}\left(\hat{X}_{ij}\gamma^\dagger(\vec{x}',\vec{y}')\right) = \frac{1}{2^L}\left\langle\vec{n}_j\right|\gamma^\dagger(\vec{x}',\vec{y}')\left|\vec{n}_i\right\rangle. \tag{3.44}$$

Thus we first have to know how the operators γ acts over the states $|\vec{n}_i\rangle$, for this we recall the fact that the Hamiltonians are diagonalised via an orthogonal transformation which links what we call the spacial modes and the normal modes

$$\overbrace{\gamma_{i_1}\gamma_{i_2}\ldots\gamma_{i_k}}^{\text{Spacial modes}} = O_{i_1\alpha_1}O_{i_2\alpha_2}\ldots O_{i_k\alpha_k}\underbrace{\gamma_{\alpha_1}\gamma_{\alpha_2}\ldots\gamma_{\alpha_k}}_{\text{Normal modes}}. \tag{3.45}$$

Note that this operators are diagonal over our $\vec{x}$'s and$\vec{y}$'s, which will allow us to operate over the states $|\vec{n}_i\rangle$.[7]. The equation (3.44) will be simplified to

$$\left\langle\vec{n}_j\right|\gamma(\vec{x},\vec{y})\left|\vec{n}_i\right\rangle = \delta_{\vec{n}_i+\vec{x}+\vec{y},\vec{n}_j}e^{i\phi(\vec{n}_i,\vec{n}_j,\vec{x},\vec{y})}, \tag{3.47}$$

whereas the coefficients $f(\vec{x},\vec{y})$ will be given by

$$f(\vec{x},\vec{y}) = \frac{1}{2^L}\sum_{\vec{x}',\vec{y}'}\mathcal{U}_{\vec{x}\vec{x}'}\mathcal{V}_{\vec{y}\vec{y}'}\underbrace{\left\langle\vec{n}_j\right|\gamma(\vec{x},\vec{y})\left|\vec{n}_i\right\rangle}_{\propto\delta_{\vec{n}_i+\vec{n}_j,\vec{x}+\vec{y}}}. \tag{3.48}$$

In equation (3.47) the term with the delta can be change by $\delta_{\vec{n}_i+\vec{n_j},\vec{x}+\vec{y}}$, an the reason to do so is because we are working with arithmetic mod 2 and we can see that the term $\vec{n}_i+\vec{n}_j$

[6]For illustrate these ideas we first suppose we have the operators in order. However, we will generalise it later.

[7]By using the function $\gamma(\vec{x},\vec{y})$ there are two ways of getting an specific state $|\vec{n}_i\rangle$

$$|n_i\rangle = \gamma(\vec{n}_i,0)\,|0\rangle, \qquad |n_i\rangle = \gamma(0,\vec{n}_i)\,|0\rangle\,e^{i\phi(\vec{n_i})}, \tag{3.46}$$

therefore we can say that the $\vec{x}$'s takes the 0 and turn them into a 1, and the $\vec{y}$'s take 0 and transform it into a one multiplied by a phase.

is nothing but the vector of differences, which means that the only values different than zero in this vector are when $n_{i_k} \neq n_{j_k}$. This result is extremely important because it tell us that when ever we work with states like $\hat{X}_{ij}$, if the vector of differences have more ones than the vector of differences $\vec{x}+\vec{y}$. Note that the vector of difference given by $\vec{x}+\vec{y}$ can have at most L errors, which means that whenever the number of ones in the difference vector defined by the states $\vec{n}_i + \vec{n}_j$ exceeds L the state $\hat{X}_{ij}$ will be immediately zero. This turns out to be an astonishing result and bring even more questions, like how likely is to have more than L errors when $N \gg L$?, what happen when we have less errors is this quantity still small enough as we expected?. These questions will be tackled in a moment but something to stress is the fact that it is the branch point of our study. We will be first addressing the first question and afterwards we will talk about the second one.

3.4 Fermionic Random Minimum Codes.

Lets recapitulate a little more what we have done in the latter chapter. We define two binary sequences of excitations $\vec{n}_i$, $\vec{n}_j$, and the vector of differences $\vec{e}_{ij} = \vec{n}_i + \vec{n}_j$, we will denote the distance between these two binary sequences by

$$\begin{aligned} d &= W(\vec{e}_{ij}) \rightarrow \text{Weight of } \vec{e}_{ij} \\ &= d_H(\vec{n}_i, \vec{n}_j) \rightarrow \text{Hamming distance.} \end{aligned} \tag{3.49}$$

we show that there is a specific distance $d > L$ at which the state $\hat{X}_{ij} = \mathrm{Tr}_{N/L}\left(\langle \vec{n}_i| \, |\vec{n}_j\rangle\right)$ is equal to zero $\hat{X}_{ij}$. This might sound quite familiar since it is connected to minimum distance codes and to see this more clearly consider a set of binary codewords

$$\mathcal{C} = \{\vec{x}^{(1)}, \vec{x}^{(2)}, \ldots, \vec{x}^{(2^k)}\} \quad \vec{x} \in \{0,1\}^N, \tag{3.50}$$

and size $|\mathcal{C}| = 2^{NR}$, with $R = k/N$ as defined in (3.2). Analogously, consider the Hilbert space

$$\mathcal{H}_{\mathcal{C}} = \mathrm{Span}\left(|\vec{x}^{(1)}\rangle, |\vec{x}^{(2)}\rangle, \ldots, |\vec{x}^{(2^k)}\rangle\right). \tag{3.51}$$

It is easy to check that $|\mathcal{H}_{\mathcal{C}}| = |\mathcal{C}| = 2^{NR}$. From the previous results, we can assure that there exist a Hilbert space with dimension $\dim|\mathcal{H}_{\mathcal{C}}| = 2^{N(1-\mathcal{H}(\ell))}$, where $\ell = L/N$, the relative distance. More specifically, $\forall\, |psi\rangle \in \mathcal{H}_{\mathcal{C}}$,

$$|\psi\rangle = \sum_{\vec{n}\in\mathcal{C}} \psi(\vec{n}) \, |\vec{n}\rangle, \tag{3.52}$$

where the code $\mathcal{C}$ refers to a code (N, k, ℓ), $d > \ell$. So

$$\rho_L(\psi) = \mathrm{Tr}_{N/L}\left(|\psi\rangle\langle\psi|\right) = \sum |\psi(\vec{n})|^2 |\vec{n}\rangle\langle\vec{n}| . \tag{3.53}$$

Thus, the states belonging to $\mathcal{H}_\mathcal{C}$ are like frozen states, meaning that

$$\frac{d\rho_L}{dt} = 0. \tag{3.54}$$

One of the most important question one could ask is, what is the biggest code with relative distance ℓ constrained with certain value of energy?. Namely, denoting by

$$S(\ell) = \bigcup_{\mathcal{C},\delta>\ell} \mathcal{H}_\mathcal{C}, \tag{3.55}$$

the set of codes of minimum distance ℓ. So we would like to know how big this set is and estimate its size, and for doing that, its correspondent error exponent will be needed. To do so, we first define our problem in terms of random variables. Let $X_i(\theta_k)$ be a random probability that takes the value 1 with probability $p(\theta_k)$ and the value 0 with probability $1 - p(\theta_k)$, and let $X = \sum_i X_i(\theta_k)$ be the sum of these random variables, or equivalently the number of errors. Then we ask ourselves about the probability of having certain number of errors in our Fermionic code. Particularly we ask the probability of having a quantity of error greater of equal than a certain quantity. To address this question we can make use of the Chernoff inequality[8]

$$\begin{aligned} P(e^{SX} \geq e^{Sd}) &\leq \min_S \left\langle e^{SX} \right\rangle e^{-Sd} \\ P(e^{-SX} \geq e^{-Sd}) &\leq \min_S \left\langle e^{-SX} \right\rangle e^{Sd}. \end{aligned} \tag{3.56}$$

We then compute the expected value $\left\langle e^{SX} \right\rangle$ taking into account that we are working with independent variables

$$\left\langle e^{SX} \right\rangle = \prod_k E(e^{SX(\theta_k)}) = \prod_k \left(1 + p(\theta_k)(e^S - 1)\right) = e^{\sum_k \log(1+p(\theta_k)(e^S-1))} \equiv e^{-Nr(\delta)}, \tag{3.57}$$

where $r(\delta)$ corresponds to the correspondent error exponent

$$r(\delta) = \min_S \frac{1}{N}\left(\log\langle e^{SX}\rangle - S\delta\right). \tag{3.58}$$

[8]In this part we do not specify what distribution is the one we are choosing, one might guess that it is related to the Fermi-Dirac distribution, but our result will be in terms of the this general distribution $p(\theta_k)$. In the next chapter we will provide an expression for the case of the XY model.

Since we are interested in the case when $N \to \infty$, the error exponent can be written as

$$r(\delta) \overset{N\to\infty}{\equiv} \min_{S} \oint \frac{d\theta}{2\pi} \log\left(1 + p(\theta)(e^{S} - 1)\right) - S\delta, \tag{3.59}$$

which even though can not be analytically solved, it can be numerically solved by deriving and obtaining

$$\delta = \oint \frac{d\theta}{2\pi} \frac{p(\theta)e^{S}}{1 - p(\theta) + p(\theta)e^{S}}. \tag{3.60}$$

Thus we conclude that the mean number of codes at distance d is given by

$$\langle S_{\mathcal{C}}(d) \rangle = \binom{M}{2} 2^{-Nr(\delta)}, \tag{3.61}$$

with $M \equiv 2^{NR}$, we conclude that when ever we work with rates lower than $r(\delta)/2$ the average number of pairs at a distance d goes to zero exponentially. Other wise the number of pairs that have minimum distance d are exponentially large

$$\langle S_{\mathcal{C}}(d) \rangle \doteq 2^{N(2R - r(\delta))}. \tag{3.62}$$

With the latter equation we answer one of our main questions, showing that indeed there exist exponentially large Hilbert subspaces such that its reduced states are automatically constant. However, we have not seen what happens in the case when we work with distances less than L. In the next section we will discuss a little bit more about it, and we will provide the intuition behind this case in order to have the full picture of how ultra-orthogonality may work in Fermionic systems.

3.4.1 Case when the numbers of errors is less than L

For this case we know that the term $\hat{X}_{ij}$ was different than zero, however, we are going to show that nevertheless it is not necessarily zero, this quantity will be intuitively small. To study this we will consider the norm of $||\hat{X}_{ij}||$, if we take into account the relation found in (3.48) we will find that

$$\begin{aligned} ||\hat{X}_{ij}||^2 &= \mathrm{Tr}\left(\hat{X}_{ij}\hat{X}_{ij}^{\dagger}\right) = \frac{1}{2^L}\sum_{\vec{x},\vec{y}} |f(\vec{x},\vec{y})|^2 \\ &= \frac{1}{2^L}\sum_{\substack{\vec{x},\vec{y} \\ \vec{x}'',\vec{y}''}} \left(\mathcal{U}_{\vec{x}\vec{x}'}\mathcal{U}_{\vec{x}\vec{x}''}\right)\left(\mathcal{V}_{\vec{y}\vec{y}'}\mathcal{V}_{\vec{y}\vec{y}''}\right) e^{\phi(\vec{x},\vec{y}')\phi(\vec{x}'',\vec{y}'')} \delta_{\vec{n}_i+\vec{n}_j,\vec{x}'+\vec{y}'}\delta_{\vec{n}_i+\vec{n}_j,\vec{x}''+\vec{y}''}. \end{aligned} \tag{3.63}$$

Our purpose will then be to bound the terms $\sum_{\vec{x}} \left(\mathcal{U}_{\vec{x}\vec{x}'}\mathcal{U}_{\vec{x}\vec{x}''}\right)$, the reason to this is because if we can bound these terms by some quantity, this bound will also holds for the part containing

the $\vec{y}$'s, thus for the rest of this work we will working with this quantity instead of working with (3.63). First, notice that the expression in (3.48) will not be too useful to us, since we are more interested in changing basis, for example we will be constantly changing from the spacial modes to the normal modes, so, by writing in this way the products of the operators we are not taking into account the geometric meaning of these terms. In the expansion of an operator, the resulting terms from $\gamma(\vec{x},\vec{y})$ have to be interpreted as $p-$forms, meaning that this will correspond to a volume generated by some vectors. So if we want to take into account the fact that volumes change over basis, we should have to understand that the products of the operators in $\gamma(\vec{x},\vec{y})$ should transform in a particular way to take this into account. It is not trivial, but is simple to check that the way this products should change in order to transform as Grassmann variables and take into account the change on the volumes over transformation is via the antisymmetrizing operator. So when we see the products such as $\gamma_1\gamma_2\gamma_3$ we have to understand it as $\frac{1}{3!}\gamma_{[1}\gamma_2\gamma_{3]}$. Generalising these ideas we consider a general $p-$form, lets say $\gamma(\vec{\alpha})$

$$\gamma(\vec{\alpha}) =_{[} \gamma_{\alpha_1}\gamma_{\alpha_2}\gamma_{\alpha_2}\cdots\gamma_{\alpha_p]}, \tag{3.64}$$

if each of these elements transform as $\gamma_{\alpha_i} = O_{\alpha_i j}\tilde{\gamma}_j$. The $p-$ form will transform as

$$\gamma(\vec{\alpha}) = \det\left(O\big|_{\vec{\alpha},\vec{\beta}}\right)\gamma_{\vec{\beta}}, \tag{3.65}$$

where the term $\det\left(O\big|_{\vec{\alpha},\vec{\beta}}\right)$ refers to the minor of the matrix O. Turning back to our main problem look that the term of $\sum_{\vec{x}}(\mathcal{U}_{\vec{x}\vec{x}'}\mathcal{U}_{\vec{x}\vec{x}''})$ can be written in terms of this determinants as

$$\sum_{\vec{x}}(\mathcal{U}_{\vec{x}\vec{x}'}\mathcal{U}_{\vec{x}\vec{x}''}) = \det\left[\left(O\Pi O^T\right)\big|_{\vec{x}',\vec{x}''}\right]. \tag{3.66}$$

Therefore if we are able to find a bound for this determinant we could show that the quantity $||\hat{X}_{ij}||$ is indeed small as we have been saying. Nonetheless, to bound this quantity in general is not an easy task and we can not show a general bound for this quantity, from the previous arguments based on typicality we expect $||\hat{X}_{ij}||$ to be bounded by a small quantity as we expected from the arguments aforementioned. Of course, one could explore this quantity for specific systems and find bounds for those particular cases and find the physical meaning associated with it[9], however, it is not on the scope of this work to explore this for particular cases and we leave it a a future work.

[9]Specifically we have computed numerically the values of (3.66) for the case of the one dimensional XY model and we have found that these values are for most cases extremely small.

Chapter 4

Conclusions and Perspectives

Throughout all this document, we have exposed a series of arguments to show, the way Ultra Orthogonality could be an alternative mechanism for understanding equilibration. As we mentioned, this is connected to thermalisation problem, since for most systems thermal typicality is expected, thus the state of equilibrium we end with has to be nothing but the canonical state. Additionally, we tackle the problem when Ultra Orthogonality holds exactly, that means, the equilibrium state is automatically reached by computing the correspondent reduced state of the Universe. By this, we mean that we have led all our efforts to show that for the case of Fermionic systems, there exists a particular sort of pure dynamical states, such that when taking its correspondent partial trace over the environment, these resultant reduced states are automatically in its correspondent equilibrium, and the way to understand this phenomena, can be through an information theory viewpoint, whereby, the fact that we work over a finite algebra, as the one in fermions, let us formulate this particular problem in terms of minimum distance codes.

Not satisfied with finding this result in a virtual general way for Fermionic systems, we decided to estimate the size of the correspondent Hilbert subspace associated with the states that fulfil Ultra Orthogonality, in the sense that all reduced states, will be automatically stationary states (Constant states over time). Thanks the formalism of exponent errors, used in Code Theory , it was possible to provide an expression to the large deviation present in the Fermionic systems, namely, this exponent tell us the an estimate about the expected number of random Fermionic minimum codes that fulfil Ultra Orthogonality, Which as we show, whenever this exponent is larger than zero, we therefore are able to show that the correspondent Hilbert subspace to these states is exponentially large. In other words, aside

of our result being quite general, we found that whenever the correspondent exponent, which depends on the parameters of the system, is larger than zero, we will have a huge Hilbert subspace (exponentially large), which contains states such that when taking the partial trace over its correspondent environment are automatically equilibrated.

This particular result bring more questions than answers. One of the first questions we may ask is, what happens if instead of thinking this problem only in terms of random codes, we impose a condition over the energy?. From our point of view, we consider that this could be the main key to connect all this results with the thermalisation problem, indeed we think that even though, the the exponent error must change, we could still find situations in which it is grater than zero, and therefore, find that the expected number of codes at certain energy will also follow a large deviation law. The problem we stumble upon was that when imposing a restriction over the energy for the codewords, liner superposition may end up not conserving the restriction of energy, so another study will be needed to confirm if this can be done and provide the connection to thermalisation.

Finally, we want to point out that the physical meaning of these quantities was not discussed, and maybe the question of what exactly is the physical meaning of these exponents? could have emerge at some point without any answer. What we wanted to stress here is that indeed, there was an attempt to find some physical interpretation about these quantities. Specifically, we tried to tackled this matter by studying the one dimensional XY model of spins 1/2. The choice for this particular system, was due to the fact that this particular system can be analytically solved with efficient algorithms. In spite of our efforts, hitherto, a complete interpretation or an intuition of our result have not been found. Thus we emphasise this point as a future work to be done in this subject.

Bibliography

[1] S. Popescu, A. J. Short, and A. Winter. 1, 2, 9, 20, 22

[2] S.-K. Ma, *Statistical Mechanics*, Wspc, illustrated edition edition. 1

[3] I. E. Farquhar, *Ergodic Theory in Statistical Mechanics*, John Wiley & Sons Inc, 1st ed. edition edition. 1

[4] R. Jancel, *Foundations of Classical and Quantum Statistical Mechanics: International Series of Monographs in Natural Philosophy*, Pergamon. 1

[5] M. Horodecki, J. Oppenheim, and A. Winter, **436**, 673. 1

[6] J. M. Deutsch, **12**, 075021. 1

[7] N. Singh, **27**, 1330003, Publisher: World Scientific Publishing Co. 1, 8

[8] S. Popescu, A. J. Short, and A. Winter, **2**, 754. 1, 8, 9, 20, 22

[9] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, **96**, 050403. 1, 2, 8

[10] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics: Emergence of Thermodynamic Behavior Within Composite Quantum Systems*, volume 657 of *Lecture Notes in Physics*, Springer Berlin Heidelberg. 1, 2, 8

[11] V. D. Milman and G. Schechtman, *Asymptotic Theory of Finite Dimensional Normed Spaces: Isoperimetric Inequalities in Riemannian Manifolds*, Springer, Google-Books-ID: q_RqCQAAQBAJ. 2

[12] M. Ledoux, *The Concentration of Measure Phenomenon*, American Mathematical Society. 2

[13] H. Tasaki, **80**, 1373. 2

[14] J. Gemmer and M. Michel, **73**, 1. 2

[15] N. Linden, S. Popescu, A. J. Short, and A. Winter, **79**, 061103. 2, 3, 9, 24, 26, 27, 28

[16] N. Linden, S. Popescu, A. J. Short, and A. Winter, **12**, 055021. 2

[17] A. S. L. Malabarba, L. P. García-Pintos, N. Linden, T. C. Farrelly, and A. J. Short, **90**, 012121. 2

[18] M. Srednicki, **50**, 888. 4

[19] J. M. Deutsch, **43**, 2046. 4

[20] M. Rigol and M. Srednicki, **108**. 4

[21] A. F. Reyes-Lega, page 1. 5, 32, 35

[22] A. Barg and G. Forney, **48**, 2568, Conference Name: IEEE Transactions on Information Theory. 6, 47, 48

[23] L. Boltzmann, *Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie: (vorgelegt in der Sitzung am 8. Februar 1866)*, Staatsdruckerei, Google-Books-ID: 8Hc5AAAAcAAJ. 8, 9, 10

[24] L. D. Landau and E. M. Lifshitz, *Statistical Physics: Volume 5*, Butterworth-Heinemann, 3 edition edition. 8

[25] J. Uffink, *Handbook for Philsophy of Physics 924-1074*, Elsevier, Amsterdam. 9, 12

[26] L. Boltzmann, **219**, 211, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.187121906 10, 11

[27] L. E. Boltzmann, *Einige allgemeine sätze über Wärmegleichgewicht*, K. Akad. der Wissensch. 10, 11

[28] J. W. Gibbs, *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*, C. Scribner's Sons. 10, 11

[29] P. Ehrenfest, *The conceptual foundations of the statistical approach in mechanics, by Paul and Tatiana Ehrenfest. Translated by Michael J. Moravcsik*, Ithaca, N.Y., Cornell University Press [1959]. 12

[30] C. R. d. Oliveira and T. Werlang, **29**, 189, Publisher: Sociedade Brasileira de Física. 12

[31] L. Engelhardt, M. Lopez del Puerto, and N. Chonacky, **83**, 787, Publisher: American Association of Physics Teachers. 12

[32] G. Gallavotti, *Statistical Mechanics: A Short Treatise*, Theoretical and Mathematical Physics, Springer-Verlag. 12

[33] S. Goldstein, Boltzmann's approach to statistical mechanics, in *Chance in Physics: Foundations and Perspectives*, edited by J. Bricmont et al., Lecture Notes in Physics, pages 39–54, Springer. 12

[34] K. Huang, *Statistical Mechanics, 2nd Edition*, John Wiley & Sons, 2 edition edition. 13, 15

[35] E. Schrödinger, **31**, 555, Publisher: Cambridge University Press. 17

[36] A. Ekert, P. Hayden, and H. Inamori, version: 1. 17

[37] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Google-Books-ID: 65FqEKQOfP8C. 17

[38] K. W. Edwin and R. J. Taud, Optimal operating system composition using integer programming and statistic reserve determination, in *Electrical Load-Curve Coverage*, pages 13–30, Pergamon. 17

[39] J. Von Neumann, *Mathematical foundations of quantum mechanics*, Princeton, N.J. : Princeton University Press. 18

[40] D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, 1 edition edition. 18, 40, 43, 47

[41] C. E. Shannon, **27**, 379, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1538-7305.1948.tb01338.x. 19, 39

[42] M. B. Plenio and S. Virmani. 19

[43] J. Gemmer and G. Mahler, **31**, 249. 19

[44] D. N. Page, **71**, 1291. 19

[45] C. Fuchs and J. van de Graaf, **45**, 1216, Conference Name: IEEE Transactions on Information Theory. 27

[46] G. Camilo, G. T. Landi, and S. Eliëns, **99**, 045155. 28

[47] Density operators for fermions, in *Quantum Theory of Optical Coherence*, pages 595–631, John Wiley & Sons, Ltd, Section: 16 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9783527610075.ch16. 32

[48] Y. Niu et al., **85**, 035110. 32

[49] M.-C. Chung and I. Peschel, **64**, 064412. 32

[50] M. Leijnse and K. Flensberg, **27**, 124003. 32, 35

[51] L. G. Molinari. 32, 35

[52] A. Botero and B. Reznik, **331**, 39 . 32, 33, 35

[53] S. Bravyi. 32, 36

[54] E. Lieb, T. Schultz, and D. Mattis, **16**, 407. 32

[55] J. I. Latorre, E. Rico, and G. Vidal, **4**, 48. 32

[56] S. Katsura, **127**, 1508. 32, 33, 34

[57] E. Barouch and B. M. McCoy, **3**, 786. 32, 33, 34

[58] E. Barouch, B. M. McCoy, and M. Dresden, **2**, 1075. 32, 33, 34

[59] E. Fradkin, *Field Theories Of Condensed Matter Systems Vol#82 On-demand Printing Of #52253*, CRC Press, 1 edition edition. 33

[60] F. A. Berezin, *The method of second quantization*, Number v. 24 in Pure and applied physics, Academic Press. 33

[61] N. N. Bogoljubov, **7**, 794. 33

[62] M. A. Nielsen. 34

[63] B. Westwanski and A. Pawlikowski, **43**, 201. 35, 36

[64] S.-A. Cheong and C. L. Henley. 37

[65] C. V. Kraus, M. M. Wolf, J. I. Cirac, and G. Giedke, **79**, 012306. 37, 38

[66] E. T. Jaynes, *Probability Theory: The Logic of Science*. 40, 47

[67] R. Gallager. 48

[68] R. G. Gallager, *Information Theory and Reliable Communication*. 48

[69] R. Gallager, **19**, 244. 48

[70] E. Knill.

[71] B. M. Terhal and D. P. DiVincenzo, **65**, 032325.

[72] E. Knill, R. Laflamme, and G. Milburn.

[73] E. Knill.

[74] E. Knill, R. Laflamme, and G. J. Milburn, **409**, 46, Number: 6816 Publisher: Nature Publishing Group.

[75] J. M. Deutsch, **81**, 082001.

[76] J. M. Deutsch, H. Li, and A. Sharma, **87**, 042135.

[77] S. Lloyd, Quantum information ProcessingQuantum information processing, in *Computational Complexity: Theory, Techniques, and Applications*, edited by R. A. Meyers, pages 2496–2533, Springer.

[78] P. Facchi and G. Garnero, **14**, 1740001.

[79] S. Bravyi, version: 1.

[80] I. Peschel, **36**, L205.

[81] *Field States*, chapter 11, pages 291–319, John Wiley & Sons, Ltd, 2005.

[82] P. Francesco, P. Mathieu, and D. Sénéchal, *Conformal Field Theory*, Graduate Texts in Contemporary Physics, Springer-Verlag.

[83] D. E. Soper, **18:12**, Institution: Institute of Theoretical Science, University of Oregon, Eugene, Oregon 97403.

[84] N. Merhav.

[85] Y. Domb, R. Zamir, and M. Feder, **62**, 121, Conference Name: IEEE Transactions on Information Theory.

[86] C. P. Robert and G. Casella, The multi-stage gibbs sampler, in *Monte Carlo Statistical Methods*, edited by C. P. Robert and G. Casella, Springer Texts in Statistics, pages 371–424, Springer.

[87] Markov chain monte carlo in practice.

[88] Gibbs sampler - an overview | ScienceDirect topics.

[89] J. Eisert, M. Cramer, and M. B. Plenio, **82**, 277, version: 4.

[90] The fisher-hartwig conjecture and toeplitz eigenvalues - ScienceDirect.

[91] A. Böttcher and B. Silbermann, *Introduction to Large Truncated Toeplitz Matrices*, Universitext, Springer-Verlag.

[92] D. Bump, *Lie Groups*.

[93] D. Bump and P. Diaconis, **97**, 252.

[94] G. Szegö, **76**, 490.