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

.....

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, que 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 cómo 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, 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, proporciona una forma diferente de entender los sistemas fermiónicos reducidos. Confirmando así, la existencia de estados reducidos que permanecen en equilibrio para estos sistemas. En este sentido, se explora 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, 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. It has been able to explain from a viewpoint of quantum information theory, how thermalisation emerge from entanglement between the system and its environment. Later results have shown that the ideas used in typicality, provide an explanation to the mechanism of the evolution towards equilibrium for large quantum systems. Within this framework of ideas, time averages plays a major rol when it comes to understand how reduced states, obtained from a pure states, can 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, that is, reduced states such that immediately reach its equilibrium state? If so, is it possible to measure the size of the space fulfilling this property?

In the present work, tools of code theory, specifically fermionic minimum distance codes, provide an alternative understanding of fermionic systems. Specifically, we characterised the size of the Hilbert space generated by the states, such that fulfil this property, throughout an error exponent, which provides a characterisation of the large deviation law within the system. Hence, 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

In the last two decades, a wave of works in quantum information theory has proven to be a new and objective form of understand the foundations of statistical mechanics [1–7]. Particularly, alternative considerations about the foundations of statistical mechanics proposed by Posescu et. al. [1], have shown that reliability on subjective randomness [8], ensemble averaging [9] or time averaging [10] are not required to understand the emergence of thermalisation. Instead, a quantum information perspective [11] 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. For example: 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, almost every reduced state obtained from a quantum pure state will approximately coincide with the canonical thermodynamic state [12, 13].

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 the *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, 2, 4, 14], 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 [15, 16], 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 [14], as well as Goldstein et. al [2], 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. However, the result obtained by Popescu et. al. establishes canonical typicality under a great generality by invoking the levy Lemma [1, 15, 16]. 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, we enquire under what conditions the state of the universe will evolve into the large region of its Hilbert Space in which its subsystems are thermalised and remain in that space for most of its evolution. Motivated by previous results heading this direction [3, 17] 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. [18] explore whether 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 [18–20] 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 [18]. First, *equilibration* is the process in which the system reaches a particular state and remains in that state or close 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 parameters are needed to describe the bath [18]. For example, its temperature: In the moment equilibrium is reached, that state should only depend on the temperature of the bath. Third, *subsystem state independence* refers to the fact 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 by the afore-mentioned elements, 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 those 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 [18].

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 then 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 hence 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. [21–23] 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. 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 [24].

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 presents in 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 to 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 [25] 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 a discussion of the thermodynamic entropy and the main differences between Boltzmann's and Gibb's approach to thermodynamics. 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 provide all the theoretical background concerning fermionic systems and code theory, in particular, we will be discussing the case of the the one dimensional XY model, and we will finish with the introduction of some important results in error correction codes. In the third chapter, we present our main results which are an explicit connection between the Fermionic

systems and random minimum distance codes. We will show an estimate size of the Hilbert subspace associated with the states that fulfil ultra-orthogonality. Finally, chapter 4 contains the conclusions and 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 [13]. In these questions the role of probabilities, entropy, the relevance of time averages and ensemble averages to individual physical systems is discussed thoroughly [14] 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 [13] and has been used since 1902 by Gibbs [26] 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 [2, 4, 14], 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 discuss the thermodynamic entropy. To do this, we are going to compare the Gibbs entropy as well as the Boltzmann entropy, and we are going to argue that the correct thermodynamic entropy correspond to the Boltzmann entropy. Also the quantum extension of these ideas are going to be discussed for the case of the quantum Gibbs entropy. 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, 4]. Finally, we will study a dynamical point of view of thermalisation proposed by Linden et. al. [18], to show that equilibration emerge as a generic property of local quantum systems. This will lead us to the main subject

our 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 The idea of thermodynamic entropy.

Science is often presented as a collection of universally accepted knowledge and discoveries, in which disagreement among scientists is often downplayed. Specifically, in physics many questions have been a matter of disagreement, where not only questions related with new discoveries have played a main role in the discussions, but also questions concerning concepts that have been taken for granted in the books, such as the interpretation of the quantum mechanics. Particularly, when we look at the foundations of statistical mechanics, two different formulations of entropy are often presented in the literature. The first one, proposed by Boltzmann [27], which provides a definition of thermodynamic entropy for an individual system, and the other one, proposed by Gibbs [26], which gives an entropy definition of a probability distribution over the phase space, i.e., of an ensemble.

It is stated that the Gibbs entropy gives the correct thermodynamic entropy [28], since it yeilds to the correct thermodynamic predictions, while Boltzmann H expression is correct only in the case of ideal gases. However, there is a school of thought which holds that Boltzmann expression is directly related to the entropy, and the Gibbs' one is simply erroneous and misleading [29].

Is not new to that statistical physics, based on the Gibbs interpretation, has provided high accurate results that has yielded to the correct thermodynamic predictions. Hence, the problem behind the discussions relating the foundations of statistical mechanics are not referring to its usefulness, but instead, with subtle differences in the the interpretations behind these theories.

1.1.1 Boltzmann entropy

Let $X = (\vec{q}_1, \ldots, \vec{q}_N; \vec{p}_1, \ldots, \vec{p}_N)$ be the *microstate* of a classical system at a time t, consisting of a large number N of identical particles forming a gas in a box Λ.The evolution of the system is then determined via Hamilton's equations of motion

$$\frac{\mathrm{d}\vec{q}_i}{\mathrm{d}t} = \frac{\partial H}{\partial \vec{p}_i}, \qquad \frac{\mathrm{d}\vec{p}_i}{\mathrm{d}t} = -\frac{\partial H}{\partial \vec{q}_i}, \tag{1.1}$$

with $H(\vec{q}_1, \ldots, \vec{q}_N; \vec{p}_1, \ldots, \vec{p}_N)$ the Hamiltonian function. Since the energy E is a constant, the evolution of the system is then confined to a set of Hamiltonians Ω_E such that fulfil the restriction $\Omega_E = \{(\vec{q}, \vec{p}) \in \Omega(\Lambda) | H(\vec{q}, \vec{p}) = E\}$, where $\Omega(\Lambda)$ is the set of all possible states in the phase space. Now consider the subset Γ of Ω_E, as the set of all phase points that "look macroscopically similar" to X. In other words, every phase point X has an associated *macrostate* $\Gamma(X)$ consisting of phase points that are macroscopically similar to X. Thus, if we reticulate the one-particle phase space ($\vec{q}, \vec{p}$-space) into macroscopically small but microscopically large cells Δ_α. Over each cell we can specify the number n_α of particles on each cell. By doing so, we will end up looking at a histogram which is a deterministic function of the microstate of the system and that specifies a macrostate over time. Notice that the histogram we are building is not a probability distribution, in fact, it corresponds to an instantaneous occupation number which tells us how many particles are in certain state, hence probabilities are not needed to interpret a macrostate. However, the value of the histogram remains unknown for us, because we have no clue what the initial conditions were at the beginning, hence, this is the moment probability plays a crucial role in Boltzmann's description.

One of the Boltzmann's great achievement in [27], was to arrive at an understanding of the meaning of the *Boltzmann entropy* as a measure of the set of all phase points that look macroscopically like X ($\Gamma(X)$). Explicitly he found that the quantity[1]

$$S(X) = k \log |\Gamma(X)|, \tag{1.2}$$

gives an expression for the thermodynamic entropy, to which the Second law refers. Where k is the Boltzmann's constant, and $|.|$ denotes the volume given by the Lesbege measure onto Ω_E.

In his paper [27] Boltzmann ask himself about the most common histogram that appears for a given macrostate. What he proved is that the vast majority of the points in the phase space have the property to end up looking as the histogram found when maximizing S. Thus Ω_E will consist almost entirely of phase points in the equilibrium macrostate Γ_{Eq}, with few exceptions whose totality has volume of the order 10^N relative to that of Ω_E. Moreover, for non-equilibrium phase points X of energy E, the Hamiltonian dynamics governing the motion of X_t arising from X would have to be exceptionally special to avoid reasonably quick carrying X_t to Γ_{Eq} and keep it there for a long time.

[1] Boltzmann used the notation H to refer to the entropy, we do not use it to not confuse it with the Hamiltonian.

Even though Boltzmann's ideas provide the correct definition of thermodynamic entropy, Boltzmann entropy fails when we consider interacting systems [30, 31]. The work of Boltzmann in the next years [32, 33] was dedicated to include the interactions in his description. However, he never could include interactions to his theory, and that was reflected in the fact that his predictions differed from the experiments, thus, it opened the path for alternative formulations of statistical mechanics.

1.1.2 Gibbs entropy

For the sake of providing an alternative foundation of statistical mechanics, Gibbs proposed in his work [26] a form of study statistical mechanics, where probability densities and *ensembles* play the main role. In his development, ensembles were infinite sets of macroscopically identical systems, each represented by a correspondent microstate, being compatible with a macrostate, and the probabilities appeared when he decided to explain the state of the system through probability density ρ on its phase space.

$$S_{\mathcal{G}}(\rho) = -k \int \rho(X) \log \rho(X). \tag{1.3}$$

This particular identification of the state of the system and the success of this theory, lead to link the *Gibbs entropy* with the thermodynamic entropy. However, this particular identification, as pointed by Goldstein et. al. [34, 35], is not correct. To understand the point of Goldstein et. al., note that first, the Gibbs entropy is computed by a function of N particles instead of 1 particle as in Boltzmann's idea, and second, Gibbs entropy is a constant of motion. That is, if we write ρ_t for the evolution on densities induced by the motion on phase space, we have that $S_{\mathcal{G}}(\rho_t)$ is independent of t,

$$\frac{\mathrm{d}S_{\mathcal{G}}}{\mathrm{d}t} = 0. \tag{1.4}$$

It is frequently asked how this can be compatible with the second law if the entropy does not change. Ideas to answer these kind of questions have been proposed by Jaynes in [36] in which he argues that even though Gibbs entropy does not change over time, the distributions ρ_t for posterior measurements will lead to an entropy $S'_{\mathcal{G}}$ such that $S_{\mathcal{G}} \leq S'_{\mathcal{G}}$ and thus, the second law will be recovered. This is position has been strongly criticised by S. Goldstein et. al. [34, 35], where they stated that the real thermodynamic entropy is the one provided by Boltzmann and that the Gibbs entropy is not even an entity of the right sort to describe what should be understood as the thermodynamic entropy, because in Gibbs entropy, ρ

corresponds to a probability distribution, that it a probability of an ensemble of systems, and it is not a function on phase space, a function of the actual state X of an individual system.

One could wonder what is so attractive about the Gibbs entropy since it does not provide a correct interpretation of thermodynamics. the answer to this is simple, Gibbs approach is simple and elegant and produce the correct answers when predicting thermodynamic quantities. In Gibbs's perspective, the idea of assign a probability density function on the phase space $\Omega(\Lambda)$, not only allow us to compute expected values of observables, but it also let us understand the state of a subsystem by simply computing the corresponding marginal probability of probability density describing the universe. Moreover, Gibbs's achievement [26] was to realize that there is a canonical measure over the phase space that allow us to conveniently define our probabilities. The measure that allow us to do that is given by the Darboux theorem [37, 38], and is known as the measure of the *symplectic form*. Specifically, when Gibbs looked at the problem of the gas with constant energy and a fixed volume, he assign equal probability to all possible systems that look macroscopically like X, in his own words [26]:

> *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, Gibbs considered that whenever a macroscopic system is at equilibrium, every state compatible with the constraints of the system has to be equally available (likely) compared to the others. Mathematically this translates into the choice of a constant density function, called the *micro-canonical* ensemble.

The two notions of Gibbs and Boltzmann entropy are parallel to two notions of thermal equilibrium, notions that are described by Goldstein et. al. [35] as the *ensemblist* and the *Individualist* point of view. In the view of the ensemblist, a system is in thermal equilibrium if and only if its phase point X is random with the appropiate distribution, such as the micro-canonical distribution. In the Individualist view, in contrast, a system is in thermal equilibrium if and only if its phase point X lies in a certain subset Γ_{Eq} of phase space. These two positions allow us to compare in a clearer way the main differences between Boltzmann and Gibbs entropy for classical systems. However, our world is quantum and the arguments we have discussed so far are classic. In the next section we are going to discuss how Gibbs can be implemented in the quantum case.

1.2 The Quantum case

In quantum mechanics, a system is described by a vector in Hilbert space $\mathcal{H}$ and its evolution is generated by a Hamiltonian operator $\hat{H}$. Hence, when we consider a macroscopic quantum system with Hilbert space $\mathcal{H}$, we can think on translate Gibbs' ideas in a quantum mechanical context. First, any description of state of knowledge in a quantum mechanical systems has to be in terms of the maximum available information [39]. That is, the quantum systems has to be in a quantum pure state $|\psi_1\rangle$ with probability p_1 or it may be in the state $|\psi_2\rangle$ with probability p_2, etc.All the alternatives $|\psi_i\rangle$ are not necessarily mutually orthogonal, but each may be expanded in terms of a complete orthonormal set of functions $|\phi_k\rangle$

$$|\psi_i\rangle = \sum_k c_{ki} \, |\phi_k\rangle \, . \tag{1.5}$$

This state of knowledge is interpreted by a point P_i with coordinates c_{ki}. At each point P_i, we place a weight p_i, such that at the end we have a collection of weights p_i assigned to a state. Since each of the possible wave functions is normalised to unity,

$$\langle \psi_i, \psi_i \rangle = \int |\psi_i|^2 d\tau = 1, \tag{1.6}$$

we have that

$$\sum_k |c_{ik}|^2 = 1, \tag{1.7}$$

and all points P_i lie over the unit hypersphere. If each of the possible states $|\psi_i\rangle$ satisfies the same Schrödinger equation, then as time goes on, the dynamics of the points P_i can be seen as a rigid rotation hypersphere. Hence, the measure over which we assign the probabilities have to be an invariant measure over unitary transformations and a uniform measure over the sphere. This measure is known as the *Haar measure*. Moreover, p_i are not in general the probabilities of mutually exclusive events. In quantum mechanics, if a state is known to be in $|\psi_i\rangle$, then the probability of finding it upon measurement $|\psi_j\rangle$ is given by $|\langle \psi_i, \psi_i \rangle|^2$. Therefore, the probabilities p_i refer to independent mutually exclusive events only when the states $|\psi_i\rangle$ are orthogonal states to each other. Since nothing assures that we only work with orthogonal states, it is convenient to define an entropy that takes into account in the quantum mechanical case as

$$S_{vN}(\hat{\rho}) = -k \, \mathrm{Tr}(\hat{\rho} \log \hat{\rho}), \tag{1.8}$$

where $\hat{\rho}$ corresponds to the density matrix. This expression is known as the *quantum Gibbs entropy* or simply as the *von Neumann entropy* [40]. Note that this definition of entropy

assigns zero entropy to any quantum pure state, and that similar as in the classical case, this entropy is a constant in time, meaning that it can not account for the second law of thermodynamics.

This ensemblist viewpoint has been criticised [34, 35, 41, 42] to rely on subjective randomness and ensemble averages that in certain occasions do not have a clear physical meaning. However, in the last decades, there has been a wave of works dedicated to quantum thermalisation [1–7], which often are connected with the key words "eigenstate thermalization hypothesis" (ETH), and *canonical typicality*, in which an individualist viewpoint is addressed. A common factor in all these works is that an individual, closed, macroscopic quantum system in a pure state $|\psi(t)\rangle$ that evolves unitarily will, under conditions usually satisfied, behave very much as one would expect a system in thermal equilibrium to behave. In the next section, we will introduce the ideas behind typicality and we will dive in the details of it.

1.3 Canonical typicality

The purpose of this section is to explain why in quantum systems, the ensemblist ideas are not necessary to explain thermalisation, and an individualist view point can be used instead.

Consider a large quantum mechanical system, we will call "The Universe", which we decompose in two parts, the system S and its environment E, where the dimension of the environment $d_{\mathcal{E}}$ is considered to be much larger than the dimension of the system $d_{\mathcal{S}}$. 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

$$\mathcal{H}_R \subset \mathcal{H}_{\mathcal{S}} \otimes \mathcal{H}_{\mathcal{E}}, \tag{1.9}$$

where the dimension of $\mathcal{H}_R$ is denoted by d_R. When we deal with the standard approach in statistical mechanics, the restriction is imposed over the total energy. However, as Popescu et. al. emphasise in [1, 4], this restriction can be completely arbitrary and not necessarily referring to the energy.

Let $\mathcal{E}_R$ be the equiprobable state in $\mathcal{H}_R$,

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

where $\mathbb{I}_R$ corresponds to the projector operator on $\mathcal{H}_R$, and $\mathcal{E}_R$ correspond with the maximally mixed state in $\mathcal{H}_R$, because it corresponds to the state of maximum ignorance.

We define Ω_S, the canonical state of the system corresponding to the restriction R, as the quantum state of the system when the universe is in the equiprobable state $\mathcal{E}_R$. The canonical state of the system Ω_S is therefore obtained by tracing out the environment in the equiprobable state of the universe:

$$\Omega_S = \mathrm{Tr}_E\, \mathcal{E}_R. \tag{1.11}$$

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

$$\rho_S = \mathrm{tr}_B(|\phi\rangle\langle\phi|). \tag{1.12}$$

In the spirit of Boltzmann's ideas one could think that at an individual level, thermalisation may appear as consequence that the vast majority of states have the property that equilibrate. Here we address this ideas from a quantum point of view, thus we ask ourselves, "how different is ρ_S form the canonical state Ω_S". An answer to this question is addressed by Popescu et. al. in [1, 4], 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 $\mathcal{E}_R$.

To formally express canonical typicality, 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.

We define the trace distance between ρ_S and the canonical state Ω_S, by $||\rho_S - \Omega||_1$, this distance is explicitly calculated by

$$||\rho||_1 = \mathrm{Tr}\,|\rho| = \mathrm{Tr}\left(\sqrt{\rho^\dagger \rho}\right). \tag{1.13}$$

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. Thus, the states the state $|\phi\rangle$ lives over the sphere surface of d_R dimensions. Hence, if we are randomly sample pure states, we will have to sample them with the previously discussed Haar measure, which is the measure that is invariant under unitary transformations.
Therefore the theorem of canonical typicality reads,

Theorem 1.3.1 (Theorem of Canonical Typicality [1, 4].)**.** *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:*

$$\mathrm{Prob}\left(\|\rho_S - \Omega_S\|_1 \geq \eta\right) \leq \eta', \tag{1.14}$$

where

$$\eta = \varepsilon + \sqrt{\frac{d_S}{d_E^{\text{eff}}}}, \quad \eta' = 2\exp\left(-Cd_R\varepsilon^2\right), \tag{1.15}$$

with

$$C = \frac{1}{18\pi^3}, \quad d_E^{\text{eff}} = \frac{1}{\operatorname{Tr}\Omega_E^2} \geq \frac{d_R}{d_S}, \quad \Omega_E = \operatorname{Tr}_S \mathcal{E}_R \tag{1.16}$$

Note that η and η' are small quantities, thus, we can assert that whenever $d_E^{\text{eff}} \ggg d_S$ and $d_R\varepsilon^2 \ggg 1$, every reduced state will be close to its correspondent canonical state. What this results tells 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 subsystem, with small enough dimension, will be indistinguishable from the canonical state. Moreover, Popescu et. al [4] find an bound to the average differences between the state of the system (the reduced state) and the canonical state. Explicitly, using the levy lemma, they are able to show

$$\langle \|\rho_S - \Omega_S\|_1 \rangle \leq \sqrt{\frac{d_S}{d_E^{\text{eff}}}} \leq \sqrt{\frac{d_S^2}{d_R}}. \tag{1.17}$$

This bound tells us that most of pure states constraint to a global restriction have the property that when we look at the local state of the system it seem to behave as the thermal state. This result is what motivated the illustration shown in **??**.

In 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. An the reason for that is that no particular evolution was considered here, and only probabilistic arguments were used. this means that typicality is a kinematic description of thermalisation. However, because almost all states of the universe have the property that the system is in a approximately at the canonical state, we anticipate that most evolutions will quickly carry a state in which the system is not thermalised to one in which it is. Namely, in the next section we will see how from a typicality viewpoint, Linden et. al. [18] were able to show that thermalisation can occur in a system reliant on a unitary dynamic.

1.4 Evolution Towards Equilibrium.

Typicality is kinematic result, meaning that is only valid for a given time and a given state, and were unitary evolution does not play a role. Specifically, we are interested in states that

are atypical, in the sense that are states that locally drastically from the canonical state. As pointed out by Linden et. al. [18], to prove thermalisation is a much complicated problem since a closer look of the elements present in thermalisation, shows that equilibration, environment state independence, system state independence and the Boltzmannian form of the state of equilibrium are needed to assure that thermalisation has taken place. Specifically, the result obtained by Linden et. al. in [18] addressed only the first two elements, showing that, equilibrium can be understood as a local universal property of Quantum systems. It is important to stress that when we refer to equilibrium we do not necessarily refer to thermal equilibrium; indeed the equilibrium state can be an arbitrary state with only the property that it does not change over time.

To understand how is possible to prove that equilibrium appear as a "typical behaviour" in quantum mechanics. It is important to first define some concepts that we have used before but that we consider are important to make explicit.

Definition 1.4.1 (Universe, system and environment:)**.** *For us the universe will refer always to a large quantum system living in a Hilbert space $\mathcal{H}$. As before, our universe is always decomposed in two, and in this decomposition we refer to the system S as a small part of the total Hilbert space. The remaining, we will call it the environment. Explicitly we will always decompose the Hilbert space of the universe as a tensor product of the Hilbert space of the system and the environment, $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$, where d_S and d_E are the respective dimension of the system and the environment.*

Note that 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.

For the sake of proving the result in [18], we also define the Hamiltonian of the universe as

Definition 1.4.2 (Hamiltonian:)**.** *The evolution of the universe will be governed by a Hamiltonian given by*

$$\hat{H} = \sum_k E_k \left|E_k\right\rangle \left\langle E_k\right| . \tag{1.18}$$

with $|E_k\rangle$ the eigenstate in the energy basis with energy E_k. Where the main required assumption is that the Hamiltonian has non-degenerate energy gaps.

Expressing this condition in a more explicit way, it is said that a Hamiltonian has no-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 $\ell = n$.

Notice that the restriction imposed to the Hamiltonian is an extremely natural constraint, because all Hamiltonians that lack of symmetries have non-degenerate energies, so we talk about a set of Hamiltonians with measure 1 that fulfils this condition.

Definition 1.4.3 (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 identically, we define the state of the environment as $\rho_E(t) = \mathrm{Tr}_S\, \rho(t)$.

It is convenient to define the transient states of the universe, or the time averaged state ω as

$$\omega = \langle \rho(t) \rangle_t = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \rho(t) dt. \tag{1.19}$$

This definition allows us to also define ω_s and ω_E as the time averaged state of the system and the environment respectively. Finally, we re introduce the concept of the effective dimension of a mixed state ρ:

$$d^{\text{eff}}(\rho) = \frac{1}{\mathrm{Tr}\,(\rho^2)}. \tag{1.20}$$

The meaning of this effective is how many states contribute to the mixture, carrying the probabilistic weight of different states in the mixture, and different than the support of an operator in the Hilbert space, it is a continuous measure.

With the concepts aforementioned, Linden et. al. [18] are able to show that every pure state of a quantum universe, composed by a large number of eigenstates of energy, and that evolves under an arbitrary Hamiltonian, is such that every small system will equilibrate. where the reason of consider the global state to have many eigenstates of energy is because fi there are many eigenstates, we can assure that there will be a large quantity of changes throughout the evolution of the system. The 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

$$|\Psi(t)\rangle = \sum_k c_k e^{-iE_k t} |E_k\rangle \tag{1.21}$$

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

$$\rho(t) = \sum_{k,l} c_k c_l^* e^{-i(E_k - E_l)t} |E_k\rangle \langle E_l| , \tag{1.22}$$

that can be expanded and written as

$$\begin{aligned} \rho(t) &= \underbrace{\sum_n \|c_n\|^2 |E_n\rangle \langle E_n|}_{\omega} + \underbrace{\sum_{m \neq n} c_n c_m^* |E_n\rangle \langle E_m| e^{-it(E_n - E_m)}}_{\lambda(t)} \\ &= \omega + \lambda(t). \end{aligned} \tag{1.23}$$

For the case of non-degeneracy of the energy levels, we have that the cross-terms vanish, that is

$$\omega = \langle \rho(t) \rangle_t = \sum_k |c_k|^2 |E_k\rangle \langle E_k| , \tag{1.24}$$

which will lead us to

$$d^{\text{eff}}(\omega) = \frac{1}{\text{Tr}\,(\omega^2)} = \frac{1}{\sum_k |c_k|^4}. \tag{1.25}$$

In the same way as in typicality, we are going to ask ourselves about the distance between $\rho_S(t)$ and $\omega_S = \langle \rho_S(t) \rangle_t$. To do this, we first compute the difference between $\rho_S(t)$ and ω_S in terms of the energy eigenstates as

$$\rho_S(t) - \omega_S = \sum_{m \neq n} c_m c_n^* e^{-i(E_m - E_n)t} \,\text{Tr}_E\, |E_m\rangle \langle E_n| . \tag{1.26}$$

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 will be interested in the time average of the trace distance $\langle ||\rho_S(t), \omega_S||_1 \rangle_t$. The value this average takes will tell us about where the system is spending most of its time. In other words $\langle ||\rho_S(t), \omega_S||_1 \rangle_t$ will be small when the system equilibrates to ω_S. To be able to prove what is announced as the *Theorem 1* in [18] it is useful to relate the trace distance to the square of the Hilbert-Schmidt distance using a standard bound provided in [43]

$$||\rho_1 - \rho_2||_1 = \frac{1}{2} \,\text{Tr}_S \sqrt{(\rho_1 - \rho_2)^2} \leq \frac{1}{2} \sqrt{d_S \,\text{Tr}_S\, (\rho_1 - \rho_2)^2}. \tag{1.27}$$

Which combined with the concavity of the square-root function, yields

$$\langle ||\rho_S(t), \omega_S||_1 \rangle_t \leq \sqrt{d_S \left\langle \mathrm{Tr}_S \left[\rho_S(t) - \omega_S\right]^2 \right\rangle_t}, \tag{1.28}$$

that provides us the bound we need to proof the theorem. Now using (1.26) we write

$$\left\langle \mathrm{Tr}_{\mathcal{S}} \left[\rho_{\mathcal{S}}(t) - \omega_{\mathcal{S}}\right]^2 \right\rangle_t = \sum_{m\neq n} \sum_{k\neq l} \mathcal{T}_{klmn} \, \mathrm{Tr}_{\mathcal{S}} \left(\mathrm{Tr}_E \left| E_k \right\rangle \left\langle E_l \right| \mathrm{Tr}_E \right| E_m \rangle \left\langle E_n \right| \right), \tag{1.29}$$

where $\mathcal{T}_{klmn} = c_k c_l^* c_m c_n^* e^{-i(E_k - E_l + E_m - E_n)t}$. We compute the time average taking into account that the Hamiltonian has non-degenerate energy gaps, thus, we find that

$$\begin{aligned}
\left\langle \mathrm{Tr}_S \left[\rho_S(t) - \omega_S\right]^2 \right\rangle_t &= \sum_{k\neq l} |c_k|^2 |c_l|^2 \, \mathrm{Tr}_S \left(\mathrm{Tr}_E \left| E_k \right\rangle \left\langle E_l \right| \mathrm{Tr}_E \right| E_l \rangle \left\langle E_k \right| \right) \\
&= \sum_{k\neq l} |c_k|^2 |c_l|^2 \sum_{ss'bb'} \langle sb|E_k\rangle \langle E_l|s'b\rangle \langle s'b'|E_l\rangle \langle E_k|sb'\rangle \\
&= \sum_{k\neq l} |c_k|^2 |c_l|^2 \sum_{ss'bb'} \langle sb|E_k\rangle \langle E_k|sb'\rangle \langle s'b'|E_l\rangle \langle E_l|s'b\rangle \\
&= \sum_{k\neq l} |c_k|^2 |c_l|^2 \, \mathrm{Tr}_E \left(\mathrm{Tr}_S \left| E_k \right\rangle \left\langle E_k \right| \mathrm{Tr}_S \right| E_l \rangle \left\langle E_l \right| \right) \\
&= \sum_{k\neq l} \mathrm{Tr}_E \left[\mathrm{Tr}_S \left(|c_k|^2 \left| E_k \right\rangle \left\langle E_k \right| \right) \mathrm{Tr}_S \left(|c_l|^2 \left| E_l \right\rangle \left\langle E_l \right| \right) \right] \\
&= \mathrm{Tr}_E \, \omega_E^2 - \sum_k |c_k|^4 \, \mathrm{Tr}_S \left[\left(\mathrm{Tr}_E \left| E_k \right\rangle \left\langle E_k \right| \right)^2 \right] \\
&\leq \mathrm{Tr}_E \, \omega_E^2,
\end{aligned} \tag{1.30}$$

where $\omega_E = \mathrm{Tr}_S \, \omega$. To obtain a further bound, we invoke weak sub-additivity of the Reényi entropy [44]

$$\mathrm{Tr}\left(\omega^2\right) \geq \frac{\mathrm{Tr}_E \left(\omega_E^2\right)}{\mathrm{rank}\left(\rho_S\right)} \geq \frac{\mathrm{Tr}_E \left(\omega_E^2\right)}{d_S}. \tag{1.31}$$

Hence, combining (1.28), (1.30) and (1.31) we get

$$\langle ||\rho_S(t), \omega_S||_1 \rangle_t \leq \frac{1}{2} \sqrt{d_S \, \mathrm{Tr}_E \left(\omega_E^2\right)} \leq \frac{1}{2} \sqrt{d_S^2 \, \mathrm{Tr}\left(\omega^2\right)}. \tag{1.32}$$

By taking the definition of effective dimension, we get the main result shown in [18]

$$\langle ||\rho_S(t), \omega_S||_1 \rangle_t \leq \frac{1}{2} \sqrt{\frac{d_S}{d^{\mathrm{eff}}\left(\omega_E\right)}} \leq \frac{1}{2} \sqrt{\frac{d_S^2}{d^{\mathrm{eff}}(\omega)}}. \tag{1.33}$$

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, but since we expect to have thermal typicality in the system, that state of equilibrium will coincide with the thermal equilibrium.

1.5 Consequences of typicality

We presented typicality as an alternative mechanism to understand thermalisation in large quantum systems, and we saw how these ideas let us explain equilibration in large quantum systems. The purpose of this section will be to point some of the consequences within typicality.

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

$$\mathrm{Tr}_{\mathcal{E}} |E_n\rangle \langle E_n| \approx \mathrm{Tr}_{\mathcal{E}} |E_m\rangle \langle E_m| \approx \Omega_{\mathcal{S}}. \tag{1.34}$$

Consider a third state $|\Psi\rangle$ to be a generic linear combination of $|E_n\rangle$ and $|E_m\rangle$, fulfilling the same restriction R of the universe. In this case we will have that the density matrix associated with the third state is

$$\begin{aligned} \rho &= |\Psi\rangle \langle \Psi| \\ &= \|c_n\|^2 |E_n\rangle \langle E_n| + \|c_m\|^2 |E_m\rangle \langle E_m| \\ &+ c_n c_m^* |E_n\rangle \langle E_m| + c_m c_n^* |E_m\rangle \langle E_n|. \end{aligned} \tag{1.35}$$

If we take the partial trace of the equation (1.35), we will end up with the state of the system,

$$\begin{aligned} \rho_S &= \mathrm{Tr}_E\, \rho = \mathrm{Tr}_E\, |\Psi\rangle \langle \Psi| \\ &= \|c_n\|^2 \mathrm{Tr}_E\, |E_n\rangle \langle E_n| + \|c_m\|^2 \mathrm{Tr}_E\, |E_m\rangle \langle E_m| \\ &+ c_n c_m^* \mathrm{Tr}_E\, |E_n\rangle \langle E_m| + c_m c_n^* \mathrm{Tr}_E\, |E_m\rangle \langle E_n|. \end{aligned} \tag{1.36}$$

Notice that the way we constructed the state $|\Psi\rangle$, makes the non-crossed terms in (1.36) 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 (1.36) can be written as

$$\rho_S = \Omega(E) + c_n c_m^* \mathrm{Tr}_E\, |E_n\rangle \langle E_m| + c_m c_n^* \mathrm{Tr}_E\, |E_m\rangle \langle E_n|. \tag{1.37}$$

In order to get the equality, the cross terms in (1.37) have to approximately vanish. This property of vanishing partial traces of exterior products of states is what we name *ultra-orthogonality*. Explicitly, that is,

$$\mathrm{Tr}_E \left|E_n\right\rangle \left\langle E_m\right| = \mathrm{Tr}_E \left|E_m\right\rangle \left\langle E_n\right| \approx 0. \tag{1.38}$$

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. To see this, we can replace the condition (1.38) in equation (1.26) and see that

$$\mathrm{Tr}_{\mathcal{E}}\, \rho(t) \equiv \rho_{\mathcal{S}} \approx \omega_{\mathcal{S}}. \tag{1.39}$$

Notice that whenever ultra-orthogonality holds, temporal averages are not needed and equilibration will appear as an instantaneous process. Inspired on this interesting phenomena, we decided to explore ultra-orthogonality for the special case in which the terms in (1.38) of the left hand side are exactly equal to zero, and thus implying the equality in equation (1.39). This means that the correspondent reduced state of a fully interacting universe will be immediately constant. The idea of this work is to explore if it is possible to find Hilbert spaces in which ultra-orthogonality holds, and more importantly, to find out how big these Hilbert spaces can be.

Specifically we will be interested in the question "what is the biggest Hilbert space in which ultra-orthogonality will holds?". In the next chapters we will answer this question for the specific case of systems that can be mapped to fermionc systems through non-local transformations.

Chapter 2

Theoretical background

At the end of the last chapter we announced the problem we want to study, the ultra-orthogonality property. Particularly we showed how this property is connected to the problem of equilibration as an instantaneous process rather. We also saw that ultra-orthogonality emerged as a direct consequence of typicality, hence, it is a general property of quantum systems. The purpose of this chapter is to provide the necessary theoretical background to study ultra-orthogonality for the specific case of fermionic systems.

In this chapter we introduce the concepts of quasi-free fermionic models on a lattice and present Majorana fermions to define the fermionic covariance matrix and how this formalism can be used to develop some standard calculations on the diagonalising of the Hamiltonian of the one-dimensional XY model, and we will discuss how is possible to treat excited states in as well local reduced states in this model. Since part of this work is dedicated to show an explicit connection between a special case of fermionic systems and code theory, the second part of this chapter will be dedicated to introduce some concepts of code theory that will allows us to make the explicit connection between code theory and fermionic system in the third chapter.

2.1 Fermionic quadratic Hamiltonian

In many areas of physics one has to deal with solving quantum many body problems. This is often a computationally difficult if not an impossible task. However, the cases that can be analytically solved are well-known, and have been a subject of study [24, 45–56]. It has been found that a wide class of complicated Hamiltonians with many-body interactions can often be mapped onto Hamiltonians that are quadratic in annihilation and creation operators and

have the generic form [50]

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

where i, j run from 1 to N, the number of modes in the system and $\hat{a}_i$, $\hat{a}_i^\dagger$ are fermionic annihilation and creation operators that satisfy the canonical anti-commutation relations (CAR) [57]

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

A convenience when working with these kind of Hamiltonians is that they can be diagonalised via a so-called Bogoliubov - Valantin transformation [58], also known as a canonical transformation, which maps fermionic creation and annihilation operators to creation and annihilation operators of non-interacting quasi-particles [58, 59]. Explicitly, the transformation looks like

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

where α_i, κ_i are complex numbers such that preserves the canonical anti-commutation relations given by (2.2) for $\hat{q}$, $\hat{q}^\dagger$. This relation can also be expresses as a condition over γ_i, κ_i,

$$\gamma_i^2 + \kappa_i^2 = 1, \tag{2.4}$$

and

$$\{\hat{q}_k, \hat{q}_l\} = \left\{\hat{q}_k^\dagger, \hat{q}_l^\dagger\right\} = 0, \quad \left\{\hat{q}_k, \hat{q}_l^\dagger\right\} = \delta_{kl}. \tag{2.5}$$

The Bogoliubov-Valantin is relevant in many physics models because this transformation diagonalise many Hamiltonians; some examples of this are the Hubbard model, the BCS theory of superconductivity in the mean field or Hartree-Fock approximation, and certain solvable spin-chain models (After a Jordan-Wigner transformation) [54–57].

Hamiltonians with the generic form of (2.1) have the interesting property that not only the ground state but every eigenstate representing a certain number of excitations of quasi-particles, described by $\hat{a}$ and $\hat{a}^\dagger$, belong to the so-called class of fermionic Gaussian states, which is an interesting property, since it allows us to characterise them in terms of second order correlations, and the reason is because all the higher moments factorize as stated in Wick's theorem [49, 60]. An equivalent but convenient characterization of second order correlations are defined in terms of Majorana fermions as we will see bellow .

2.2 Majorana Fermions

Majorana fermions are represented in terms of $2N$ hermitian operators

$$\hat{\gamma}_j = \hat{a}_j^\dagger + \hat{a}_{j+N}, \quad \hat{\gamma}_{j+N} = (-i)\left(a_j^\dagger - a_j\right), \tag{2.6}$$

where these operators are analogous to coordinate and momentum operators for bosonic modes, and for each fermion labelled by j of the original system we define two operators above. The canonical Fermi-Dirac commutation relation takes the form

$$\{\hat{c}_k, \hat{c}_l\} = 2\delta_{kl}. \tag{2.7}$$

The algebra generated by the operators $\{\hat{\gamma}_i\}$ is known as the Clifford algebra and is denoted by $\mathcal{C}_{2N}$[1]. When we change from the Fermionic operators $a^T := (\hat{a}, \hat{a}_2, \ldots, \hat{a}_N, \hat{a}_1^\dagger, \hat{a}_2^\dagger, \ldots, \hat{a}_N^\dagger)$ to Majorana operators $\gamma^T := (\hat{\gamma}_1, \hat{\gamma}_2, \ldots, \hat{\gamma}_N, \hat{\gamma}_{N+1}, \ldots, \hat{\gamma}_{2N}$, it is convenient to define the Fermionic covariance matrix which will fully characterise Gaussians states.

2.3 Fermionic Covariance matrix

As we mentioned before, Gaussian states are completely characterised by its second moments [49, 60], that is, Gaussian states have a density matrix ρ [61],

$$\rho = \frac{1}{Z} \cdot \exp\left[-\frac{i}{4}\hat{\gamma}^T G \hat{\gamma}\right], \tag{2.8}$$

with $\hat{\gamma} = (\hat{\gamma}_1, \hat{\gamma}_2, \ldots, \hat{\gamma}_{2N})$, the vector of Majorana operators (2.6), 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

$$OGO^T = \begin{pmatrix} 0 & -\tilde{B} \\ \tilde{B} & 0 \end{pmatrix} \quad \text{with} \quad O \in \mathrm{SO(2N)}, \tag{2.9}$$

where $\tilde{B}$ is diagonal, with eigenvalues that we denote by $\tilde{\beta}_k$. The right hand side of (2.9) is known as the Williamson form of the skew-symmetric matrix G, and $\tilde{\beta}_k$ are the Williamson eigenvalues of G [62].

[1] The orthogonal group in $2N$ dimensions $O(2N)$ preserves the Clifford algebra hence the canonical canonical Fermi-Dirac commutation relations of fermionic operators

It is convenient to characterise second order correlations in terms of the so-called *fermionic covariance matrix* (FMC), whose entries are

$$\Gamma_{kl} = \frac{i}{2} \operatorname{Tr}\left(\rho\left[\gamma_k, \gamma_l\right]\right), \tag{2.10}$$

where $[\gamma_k, \gamma_l] := \gamma_k\gamma_l - \gamma_l\gamma_k$. Thus, we can bring this anti-symmetric matrix to its block diagonal form, via a canonical transformation, as

$$\tilde{\Gamma} = O\Gamma O^T = \begin{pmatrix} 0 & -\operatorname{diag}(\lambda_{i)} \\ \operatorname{diag}(\lambda_i) & 0 \end{pmatrix}, \tag{2.11}$$

where $\lambda_k = \tanh(\tilde{\beta}_k/2)$, for $k = 1, 2 \ldots, N$ [62] which determines the connection between the matrix G in (2.9) and the FMC Γ. The Williamson eigenvalues are $\lambda_k = n_k - 1/2$, with n_k the fermion occupation number of the normal mode labelled by k.

The equivalence between the special orthogonal group in $2N$ dimensions ($SO(2N)$) and the Fermionic Gaussian states, leads to an interesting property about states describing multi-particles excitations. If $|0\rangle$ is the ground state of some Hamiltonian, with annihilation operators $\hat{a}_i$ in a given quasi-particle basis, then $\hat{a}_i^\dagger |0\rangle = \hat{c}_{2i} |0\rangle$. Meaning that if any multi-particle state of this kind is obtained from the ground state $|0\rangle$ through some transformation, such that preserves the canonical anti-commutation relation, 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, and since we will be interested in the case of the excited states, it will be a property that we will exploit.

2.4 XY model.

The XY Hamiltonian model is a set of N spin 1/2 particles located on the sites of d-dimensional lattice. For the purpose of this document, whenever we refer to the XY model, we will have in mind the $1D$ XY model.

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

$$H_{XY} = -\frac{1}{2} \sum_{l=0}^{N-1} \left(\frac{1+\gamma}{2} \sigma_l^x \sigma_{l+1}^x + \frac{1-\gamma}{2} \sigma_l^y \sigma_{l+1}^y + \lambda \sigma_l^z \right), \tag{2.12}$$

where γ is so-called the anisotropy parameter and represents the difference between the strength of the XX interaction and the YY interaction in the spin space, λ is the intensity of the external magnetic field and

$$\sigma_l^i = \mathbb{I} \otimes \cdots \otimes \mathbb{I} \otimes \underbrace{\sigma^i}_{\text{site } l} \otimes \mathbb{I} \otimes \cdots \otimes \mathbb{I}, \tag{2.13}$$

where σ^i are Pauli matrices for $i = x, y, z$.

2.4.1 The spectrum

In order to find the spectrum of the of the XY model, it is necessary to perform three different transformations. These results are very standard and we present them to make our discussion self-consistent.

2.4.2 Jordan-Wigner transformation

The Jordan-Wigner transformation is an important transformation used mainly in Fermionic systems [63]. The 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. Consider the next non-local transformation

$$\hat{a}_l = \left(\prod_{m<l} \sigma_m^z \right) \sigma_l^-, \quad \sigma_l^- = \frac{\sigma_l^x - i\sigma_l^y}{2}, \tag{2.14}$$

where $\hat{a}_l$ represent spinless fermionic operators, and its canonical anticommutation relation (CAR) is given by [24]

$$\left\{ \hat{a}_i^\dagger, \hat{a}_j^\dagger \right\} = \{ \hat{a}_i, \hat{a}_j \} = 0, \quad \left\{ \hat{a}_i^\dagger, \hat{a}_j \right\} = \delta_{i,j}. \tag{2.15}$$

Inverting the transformation we get

$$\begin{aligned} \sigma_l^z &= 1 - 2\hat{a}_l^\dagger \hat{a}_l, \\ \sigma_l^x &= \left(\textstyle\prod_{m<l} \left(1 - 2\hat{a}_m^\dagger \hat{a}_m \right) \right) \left(\hat{a}_l^\dagger + \hat{a}_l \right), \\ \sigma_l^y &= i \left(\textstyle\prod_{m<l} \left(1 - 2\hat{a}_m^\dagger \hat{a}_m \right) \right) \left(\hat{a}_l^\dagger - \hat{a}_l \right). \end{aligned} \tag{2.16}$$

The terms of interaction in the Hamiltonian become

$$\begin{aligned} \hat{\sigma}_l^x \hat{\sigma}_{l+1}^x &= \left(\hat{a}_l^\dagger - \hat{a}_l \right) \left(\hat{a}_{l+1}^\dagger + \hat{a}_{l+1} \right), \\ \hat{\sigma}_l^y \hat{\sigma}_{l+1}^y &= - \left(\hat{a}_l^\dagger + \hat{a}_l \right) \left(\hat{a}_{l+1}^\dagger - \hat{a}_{l+1} \right), \end{aligned} \tag{2.17}$$

and the Hamiltonian of the XY model becomes,

$$H_{XY} = -\frac{1}{2}\sum_{l}\left[\left(\hat{a}_{l+1}^{\dagger}\hat{a}_{l} + \hat{a}_{l}^{\dagger}\hat{a}_{l+1}\right) + \gamma\left(\hat{a}_{l}^{\dagger}\hat{a}_{l+1}^{\dagger} - \hat{a}_{l}\hat{a}_{l+1}\right)\right] - \frac{\lambda}{2}\sum_{l}\left(1 - 2\hat{a}_{l}^{\dagger}\hat{a}_{l}\right), \quad (2.18)$$

with the boundary condition of $\hat{a}_N \equiv \hat{a}_1$, and where the term of $-\lambda N/2$ was ignored since it does not affect of the spectrum in the energy [24].

Note that we ended up with a Hamiltonian that only depends only on creation and annihilation operators and that has a similar shape of the Hamiltonian (2.1) presented at the beginning of the chapter.

2.4.3 Fourier Ttansformation

If we consider periodic boundary conditions, that is, we identify the spin in site N with the spin in site 1, then, a Fourier transform can be applied to the operators $\hat{a}_l$ in the following way [24]

$$\hat{d}_k = \frac{1}{\sqrt{N}}\sum_{l=1}^{N}\hat{a}_l e^{-i\phi_k l}, \quad \theta_k = \frac{2\pi}{N}k. \quad (2.19)$$

Note that the Fourier transformation is unitary, so the operators $\hat{d}_k$ are fermionic operators and will preserve the CAR.

In terms of $\hat{d}_k$ operators, the Hamiltonian (2.18) takes the form

$$H_{XY} = \sum_{k=-(N-1)/2}^{(N-1)/2}\left(-\lambda + \cos\phi_k\right)\hat{d}_k^{\dagger}\hat{d}_k + \frac{i\gamma}{2}\sum_{k=-(N-1)/2}^{(N-1)/2}\sin\phi_k\left(\hat{d}_k\hat{d}_{-k} + h.c\right), \quad (2.20)$$

where we have suppressed an additional term that is proportional to $1/N$ [55, 56], and the reason is because we are interested in the thermodynamic limit $N \to \infty$.

2.4.4 Bogoliubov -Valantin transformation

As mentioned in section 2.1, fermionic quadratic Hamiltonians can be easily diagonalised via a Bogoliubov-Valantin transformation over the operators $\hat{d}_k$

$$\tilde{d}_k = u_k\hat{d}_k^{\dagger} + iv_k\hat{d}_{-k}. \quad (2.21)$$

Since we want this transformation to preserve CAR, it is needed that $u_k^2 + v_k^2 = 1$, which implies that an appropriate parametrization will be $u_k = \cos(\psi_k/2)$ and $v_k = \sin(\psi_k/2)$, with

$$\cos\frac{\psi_k}{2} = \frac{-\lambda + \cos\phi_k}{\sqrt{\left(\lambda - \cos\phi_k\right)^2 + \left(\gamma\sin\phi_k\right)^2}}, \quad (2.22)$$

So finally our Hamiltonian will look as

$$H_{XY} = \sum_{-(N-1)/2}^{(N-1)/2} \tilde{\Lambda}(\theta_k) \tilde{d}_k^\dagger \tilde{d}_k, \tag{2.23}$$

with

$$\tilde{\Lambda}(\theta_k) := \sqrt{(\lambda - \cos\phi_k)^2 + (\gamma \sin\phi_k)^2}, \tag{2.24}$$

where the latter expression allow us to identify the critical regions of the model.

2.4.5 Fermionic covariance matrix for the XY model

Since we devote our work to study ultra-orthogonality, we have to be able to study properties of reduced density matrices of eigenstates of Hamiltonians quadratic in fermionic operators, and to be able to do that, it is important to characterise the covariance matrix of the XY model. In order to do this, we need to express the Hamiltonian (2.12) in terms of Majorana fermions using an analogous Jordan-Wigner transformation to the one used to diagonalise the XY Hamiltonian but into $2N$ Majorana fermions,

$$\hat{\gamma}_l = \left(\prod_{m<l} \hat{\sigma}_m^z\right) \hat{\sigma}_l^x, \quad \hat{\gamma}_{l+N} = \left(\prod_{m<l} \hat{\sigma}_m^z\right) \hat{\sigma}_l^y, \tag{2.25}$$

where again $l = 1, 2 \ldots N-1$.

Note that the three following products:

$$\hat{\gamma}_l \hat{\gamma}_{l+N} = \left(\prod_{m<l} \hat{\sigma}_m^z\right) \left(\prod_{m<l} \hat{\sigma}_m^z\right) \hat{\sigma}_l^x \hat{\sigma}_l^y = i\hat{\sigma}_l^z, \tag{2.26}$$

$$\hat{\gamma}_{l+N} \hat{\gamma}_{l+1} = \left(\prod_{m<l} \hat{\sigma}_m^z\right) \hat{\sigma}_l^y \left(\prod_{m<l+1} \hat{\sigma}_m^z\right) \hat{\sigma}_{l+1}^x = \hat{\sigma}_l^y \hat{\sigma}_l^z \hat{\sigma}_{l+1}^x = i\hat{\sigma}_l^x \hat{\sigma}_{l+1}^x, \tag{2.27}$$

and

$$\hat{\gamma}_l \hat{\gamma}_{l+N+1} = \left(\prod_{m<l} \hat{\sigma}_m^z\right) \hat{\sigma}_l^x \left(\prod_{m<l+1} \hat{\sigma}_m^z\right) \hat{\sigma}_{l+1}^y = \hat{\sigma}_l^x \hat{\sigma}_l^z \hat{\sigma}_{l+1}^y = -i\hat{\sigma}_l^y \hat{\sigma}_{l+1}^y. \tag{2.28}$$

Coincide, up to constant factors, with the three terms in (2.12); then we can write the XY Hamiltonian as [53]

$$H_{XY} = \frac{i}{4} \sum_{\alpha,\beta=0}^{2N} \Omega_{\alpha\beta} \left[\hat{\gamma}_\alpha, \hat{\gamma}_\beta\right], \tag{2.29}$$

where Ω is the antisymmetric matrix of the form

$$\Omega = \left[\begin{array}{c|c} 0 & \tilde{\Omega} \\ \hline -\tilde{\Omega}^T & 0 \end{array}\right], \tag{2.30}$$

with

$$\tilde{\Omega} = \begin{pmatrix} \lambda & \frac{1-\gamma}{2} & 0 & 0 & \dots & 0 & \frac{1+\gamma}{2} \\ \frac{1+\gamma}{2} & \lambda & \frac{1-\gamma}{2} & 0 & \dots & 0 & 0 \\ 0 & \frac{1+\gamma}{2} & \lambda & \frac{1-\gamma}{2} & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \dots & \vdots & \vdots \\ \frac{1-\gamma}{2} & 0 & 0 & 0 & \dots & \frac{1+\gamma}{2} & \lambda. \end{pmatrix}. \tag{2.31}$$

Given that $\tilde{\Omega}$ is a circulant matrix, it can be diagonalised by means of a Fourier transformation. Therefore it can be written as

$$\tilde{\Omega}_{mn} = \frac{1}{N} \sum_{\theta_k \in (-\pi, \pi)} \omega\left(\theta_k\right) e^{\phi(\theta_k)} e^{i(m-n)\theta_k}. \tag{2.32}$$

where $\omega\left(\theta_k\right) = \omega\left(\theta_k\right)^* = \omega\left(-\theta_k\right), \phi\left(\theta_k\right) = -\phi\left(\theta_k\right)$ and are given by

$$\omega^2\left(\theta_k\right) := \left(\lambda - \cos\theta_k\right)^2 + \gamma^2 \sin^2\theta_k, \tag{2.33}$$

and

$$\phi\left(\theta_k\right) := \arctan\left(\frac{\lambda - \cos\theta_k}{-\gamma \sin\theta_k}\right). \tag{2.34}$$

The summation in (2.32) is understood over k with $-(N-1)/2 \leq k \leq (N-1)/2$, which is equivalent to $-\pi \leq \theta_k \leq \pi$. So defining the following functions

$$u_m^c\left(\theta_k\right) = \sqrt{\frac{2}{N}} \cos\left(m\theta_k + \phi\left(\theta_k\right)\right), \quad u_m^s\left(\theta_k\right) = \sqrt{\frac{2}{N}} \sin\left(m\theta_k + \phi\left(\theta_k\right)\right), \tag{2.35}$$

$$v_n^c\left(\theta_k\right) = \sqrt{\frac{2}{N}} \cos\left(n\theta_k\right), \quad u_n^s\left(\theta_k\right) = \sqrt{\frac{2}{N}} \sin\left(n\theta_k\right). \tag{2.36}$$

We expand the equation (2.32)

$$\begin{aligned} \tilde{\Omega}_{mn} &= \frac{1}{N}\left[\omega(0) + (-1)^{m-n}\omega(\pi) + 2 \sum_{0<\theta_k<\pi} \omega\left(\theta_k\right) \cos\left(\theta_k(m-n) + \phi\left(\theta_k\right)\right)\right] \\ &= \frac{\omega(0)}{N} + (-1)^{m-n}\frac{\omega(\pi)}{N} + \sum_{0<\theta_k\leq\pi} \omega\left(\theta_k\right)\left(u_m^c\left(\theta_k\right) v_n^c\left(\theta_k\right) + u_m^s\left(\theta_k\right) v_n^s\left(\theta_k\right)\right), \end{aligned} \tag{2.37}$$

Now imposing $u^s(0) = v^s(\pi) = 0 \text{y} u^c(0) = v^c(\pi) = \frac{1}{\sqrt{N}}$, to rewrite $\tilde{\Omega}_{m,n}$

$$\tilde{\Omega}_{mn} = \sum \omega(\theta_k) \left(u_m^c(\theta_k) v_n^c(\theta_k) + u_m^s(\theta_k) v_n^s(\theta_k)\right). \tag{2.38}$$

Therefore, the upper-right block of (2.29) of the Hamiltonian reads

$$H = \sum_{m,n=0}^{N-1} \frac{i}{4} \sum_{\theta_k=0}^{\pi} \omega(\theta_k) \left(u_m^c(\theta_k) v_n^c(\theta_k) + u_m^s(\theta_k) v_n^s(\theta_k)\right) \left[\hat{\gamma}_n, \hat{\gamma}_{m+N}\right], \tag{2.39}$$

and rearranging things, we get

$$H = \sum_{\theta_k=0}^{\pi} \omega(\theta_k) \left(\underbrace{\left[\hat{\gamma}_k^c, \hat{\gamma}_{k+N}^c\right]}_{1-2\sigma_k^z} + \underbrace{\left[\hat{\gamma}_k^s, \hat{\gamma}_{k+N}^s\right]}_{1-2\sigma_k^z}\right), \tag{2.40}$$

where we have used

$$\hat{\gamma}_k^{c,s} := \sum_n u_n^{c,s}(\theta_k) \hat{\gamma}_n, \quad \hat{\gamma}_{k+N}^{c,s} := \sum_n v_n^{c,s}(\theta_k) \hat{\gamma}_{n+N}. \tag{2.41}$$

Now recall that the Fermionic covariance matrix is defined by (2.10), then, the transformation that brings Ω into its Williamson form, does the same on the FMC. Thus the upper-right block of the FMC is position space is

$$\begin{aligned}
\tilde{\Gamma}_{mn} &= \sum_{\theta_k}^{\pi} \left[m^c(\theta_k) u_m^c(\theta_k) v_n^c(\theta_k) + m^s(\theta_k) u_m^s(\theta_k) v_n^s(\theta_k)\right] \\
&= \sum_{\theta_k}^{\pi} \left(\frac{m^c(\theta_k) + m^s(\theta_k)}{2}\right) \left(u_m^c(\theta_k) v_n^c(\theta_k) + u_m^s(\theta_k) v_n^s(\theta_k)\right) \\
&+ \sum_{\theta_k}^{\pi} \left(\frac{m^c(\theta_k) - m^s(\theta_k)}{2}\right) \left(u_m^c(\theta_k) v_n^c(\theta_k) - u_m^s(\theta_k) v_n^s(\theta_k)\right),
\end{aligned} \tag{2.42}$$

where $m^{c,s}(\theta_k) = n^{c,s}(\theta_k) - \frac{1}{2}$, being $n^{c,s}(\theta_k$ the "cosine" ("sine") fermion occupation number of the mode labeled by k.

Now let $m^{\pm}(\theta_k) = \frac{m^c(\theta_k) \pm m^s(\theta_k)}{2}$. We can undo the transformation from (2.32) to (2.38) to have

$$\tilde{\Gamma}_{mn} = \overbrace{\sum_{\theta_k}^{\pi} m^+(\theta_k) e^{i\phi(\theta_k)} e^{i(n-m)\theta_k}}^{\tilde{\Gamma}_{mn}^+} + \underbrace{\sum_{\theta_k}^{\pi} m^-(\theta_k) e^{i\phi(\theta_k)} e^{i(n+m)\theta_k}}_{\tilde{\Gamma}_{mn}^-}. \tag{2.43}$$

We notice that $\tilde{\Gamma}_{mn}^+$ is circulant, whereas $\tilde{\Gamma}_{mn}^-$ is not, However, observe that $\tilde{\Gamma}_{mn}^+ = \tilde{\Gamma}_{mn'}^-$, with n' a change on the index $n \to -n'$. As the figure 2.1 shows, this transformation can be interpreted as a rotation over the circle.

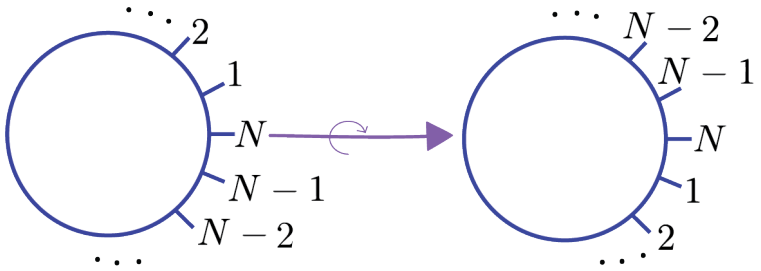

Figure 2.1: Meaning of the relabel done in the circulant matrix, which can be seen as a reflection over the circle.

Explicitly we can write that if $\tilde{\Gamma}^{+}_{mn}$ has the shape

$$\begin{pmatrix} a_0 & a_{-1} & \cdots & a_2 & a_1 \\ a_1 & a_0 & \cdots & a_3 & a_2 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{-1} & a_{-2} & a_{-3} & \cdots & a_0 \end{pmatrix}, \tag{2.44}$$

then $\tilde{\Gamma}^{-}_{mn}$ will be given by

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_{-2} & a_{-1} \\ a_1 & a_2 & \cdots & a_{-1} & a_0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{-1} & a_0 & a_1 & \cdots & a_{-2} \end{pmatrix}, \tag{2.45}$$

which we name anticirculant.

In this way, both the circulant and the anticirculant parts of Γ are computed as Fourier transforms of the vectors $m^{+}(\theta_k)e^{i\phi(\theta_k)}$ and $m^{-}(\theta_k)e^{i\phi(\theta_k)}$ respectively.

Here we spot 3 things. First, the FCM always can be written as a circulant matrix Γ^{+} plus an anticirculant matrix Γ^{-}. Second, in the ground state, the FCM is circulant because the fermion occupation numbers $n^{c}(\theta_k) = n^{s}(\theta_k) = 0, \forall k$. Third, for a generic excited state, we have that in average the FCM matrix is always circulant, because $\langle n^{c}(\theta_k)\rangle = \langle n^{s}(\theta_k)\rangle$.

2.4.6 Local modes in the XY model

For our purpose, it is important to study the behaviour of reduced states in the chain, that is, we want to look into small portions of size L in a translationally invariant chain of size

N ($N > L$). To do this it is convenient to write the modes of the small chain of size L in terms of the modes of whole chain of size N.

Since we work with a one-dimensional, translationally invariant closed chains of N free fermions, with local interactions, it is often useful to expand the annihilation/creation operators $\hat{a}_x$ ($\hat{a}_l^\dagger$) at the site x ($x = 0, 1, \ldots, N-1$) in terms of their counterpart in the plane-wave basis, obtained through

$$\hat{b}_q = \frac{e^{i\eta_q}}{\sqrt{N}} \sum_{x=0}^{N-1} e^{-i\phi_q x} \hat{a}_x, \tag{2.46}$$

and its hermitian conjugate, with $\phi_q = \frac{2\pi}{N} l$, and η_q is a phase to be adjusted. By doing so, we showed that in the case of the XY model, the Hamiltonian (2.18) became diagonal on the operators $\hat{b}_q$, explicitly we showed that the Hamiltonian had the form

$$\hat{H} = \sum_{q=0}^{N} \Lambda(\phi_q) \hat{b}_q^\dagger \hat{b}_q, \tag{2.47}$$

with $\Lambda(\phi_q)$ given by (2.24) in the XY model.

We consider now a set of local plane wave modes for the portion of the chain of length L comprising the sites $x = 0, 1, \ldots, L-1$,

$$\tilde{b}_k = (-1)^k \frac{e^{-i\tilde{\phi}_k/2}}{\sqrt{L}} \sum_{x=0}^{L-1} e^{-i\tilde{\phi}_k x} \hat{a}_x, \tag{2.48}$$

where similarly as in the case of the large chain of size N, we take $\tilde{\phi}_k = \frac{2\pi}{L} k$. By introducing the modes of the local chain in this way, the plain wave modes will diagonalize any Hamiltonian of the form (2.47), with N replaced by L.

We now expand the local operators $\tilde{b}_k$ in terms of the global operators $\hat{b}_q$ that are defined over the chain of size N. The relation between the two becomes

$$\tilde{b}_k = \frac{1}{\sqrt{NL}} \sum_{q=0}^{N-1} D_L(\phi_q - \tilde{\phi}_k) \hat{b}_q, \tag{2.49}$$

where we chose appropriately $\eta_q = \phi_q (L-1)/2$, and

$$D_L(\phi) = \frac{\sin\left(\frac{\phi}{2} L\right)}{\sin\left(\frac{\phi}{2}\right)}, \tag{2.50}$$

is the Dirichlet kernel [64].

Now consider an excited state $|\vec{n}\rangle$ of the chain, described by a set of excitation numbers $\vec{n} \equiv (n_0, n_1, \ldots, n_{N-1})$, where $n_q \in \{0, 1\}$. In particular the state satisfies

$$\langle \vec{n}| \hat{b}_q^\dagger \hat{b}_{q'} |\vec{n}\rangle = \delta_{q,q'} n_q, \quad \langle \vec{n}| \hat{b}_q \hat{b}_{q'} |\vec{n}\rangle = \langle \vec{n}| \hat{b}_q^\dagger \hat{b}_{q'}^\dagger |\vec{n}\rangle = 0, \tag{2.51}$$

for all q, q'. To any excited state, we can associate the correspondent $L \times L$ FCM, that is

$$A_{kk'}(\vec{n}) = \langle \vec{n} | \tilde{b}_k^\dagger \tilde{b}_{k'} | \vec{n} \rangle = \frac{1}{NL} \sum_{q=0}^{N-1} D_L(\phi_q - \tilde{\phi}_k) D_L(\phi_q - \tilde{\phi}_{k'}) n_q. \tag{2.52}$$

Therefore, we have a full characterization of any subchain in the state $|\vec{n}\rangle$, or equivalently in the state

$$\rho_L(\vec{n})) \operatorname{Tr}_{N-L} |\vec{n}\rangle \langle \vec{n}|, \tag{2.53}$$

the corresponding partial density matrix on the subchain.

2.5 Error correcting code theory

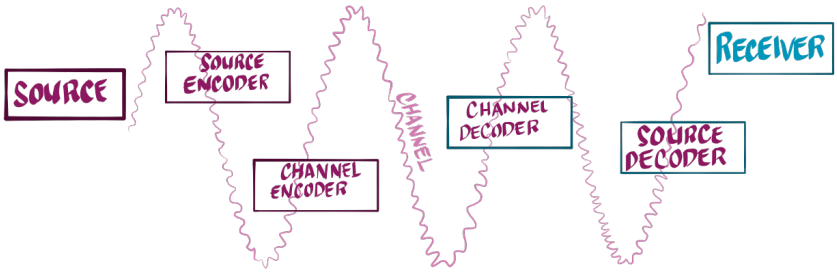

Figure 2.2: Representation of the scheme of communication.The encoding system introduces some redundancy into the transmitted vector $\mathbf{x}$. The decoding system uses this known redundancy to deduce from the received vector $\mathbf{y}$ both the original source vector and the noise introduced by the channel.

We will now turn to study error correcting codes. We start by motivating the problem of error correction codes as a mechanism to understand communication, then we move to introduce basic definitions to formalise the problem of error correcting codes. Afterwards, we will move to study the case of random minimum distance codes and some interesting results about them.

Every day we communicate over noisy channels such as in cellphone lines, over which two devices communicate digital information. When we design these channels the main purpose

is to be able to transmit information in a reliable way while dealing with errors induced by the noise in the channel. Information theory and coding theory offer a way to study communications as C. Shannon pointed out in 1948 [65]. As we illustrate in figure 2.2, we add encoders before the channel and decoders after it. The encoders encode the source message $\mathbf{x}$ into a transmitted message $\mathbf{y}$, adding redundancy to the original message in some way. The channel adds noise to the transmitted message, yielding a received message $\mathbf{y}$. The decoders use the known redundancy introduced by the encoding system to infer both the original signal $\mathbf{s}$ and the added noise.

2.5.1 Channel coding

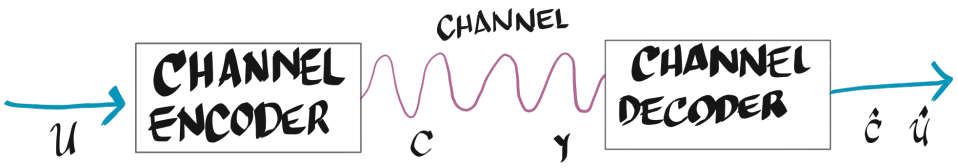

Figure 2.3: Channel coding

For the purpose of our work, the model of channel we will use, is the *discrete probability channel*: a probabilistic channel S is defined as a triple (F, Φ, Prob), where F is a finite *input alphabet*, Φ is a finite *output alphabet*, and Prob is a conditional probability distribution

$$\text{Prob}\{\mathbf{x}|\mathbf{y}\}, \tag{2.54}$$

defined for every pair $(\mathbf{x}, \mathbf{y}) \in F^m \times \Phi^m$, where m ranges over all positive integers and F^m/Φ^m denotes the set of all words of length m over F/Φ. It is important to clarify that we assume that the channel neither deletes nor inserts symbols; that is, the length of the output word $\mathbf{y}$ always equals the length of the input word $\mathbf{x}$.

The input channel encoder is an *information word* $\mathbf{u}$ out of M possible information words (see figure 2.3). The channel encoder generates a *codeword* $\mathbf{c} \in F^n$ that is input to the channel. The resulting output of the channel is a received word $\mathbf{y} \in \Phi^n$, which is fed into the channel decoder. the decoder, in turn, produces a *decoded codeword* $\hat{\mathbf{c}}$ and a *decoded information word* $\hat{\mathbf{u}}$, with the aim of having $\mathbf{c} = \hat{\mathbf{c}}$ and $\mathbf{u} = \hat{\mathbf{u}}$. this implies that the channel encoder needs to be such that the mapping $\mathbf{u} \mapsto \mathbf{c}$ is one to one.

Definition 2.5.1 (Rate)**.** *The rate of the channel encoder is defined as*

$$R = \frac{\log_{|F|} M}{n}. \tag{2.55}$$

If all information words have the same length over F, then this length is given by the numerator, $\log_{|F|} M$, in the expression for R. Since the mapping of the encoder is one-to-one, we have that $R \leq 1$.
In the case where the input alphabet F has the same size as the output alphabet Φ, it will be convenient to assume that $F = \Phi$ and that the elements of F form a finite Abelian group. We then say that the channel is an *additive channel*.

Given an additive channel, let $\mathbf{x}$ and $\mathbf{y}$ be input and output words, respectively, both in F^m. the *error word* is defined as the difference $\mathbf{x} - \mathbf{y}$, where the subtraction is taken component by component. The action of the channel can be described as adding an error word $\mathbf{e} \in F^m$ to the input word $\mathbf{x}$ to produce the output word $\mathbf{y} = \mathbf{x} + \mathbf{e}$, as shown in figure **??**.

When F is an Abelian group, it contains the zero (or unit) element. The *error locations* are indexes of the nonzero entries in the error word $\mathbf{e}$. Those entries are referred to as the *error values*.

2.5.2 Block codes

An (n, M) (*block*) code over a finite alphabet F is a nonepty subset $\mathfrak{C}$ of size M of F^n. The parameter n is called the *code length* and M is the *code size*. The *information length* or *codeword length* of $\mathfrak{C}$ is defined by $k = \log_{|F|} M$, and the *rate* is $R = k/n$. the range of the mapping defined by the channel encoder in figure 2.3 forms an (n,M) code, and this in which the term (n, M) code will be used. The elements of a code are called *codewords*.

In addition to the length and size of a code, we will be interested in quantifying how much the codewords in the code differ from each other. To do this, we will introduce the following definitions.

Definition 2.5.2 (Hamming distance)**.** *Let F be an alphabet. The Hamming distance between two codewords $\mathbf{x}, \mathbf{y} \in F^n$ is the number of coordinates on which $\mathbf{x}$ and $\mathbf{y}$ differ. the Hamming distance will be denoted by $d(\mathbf{x}, \mathbf{y})$.*

It is easy to verify that the Hamming distance satisfies the following properties of a metric for every three words $\mathbf{x}, \mathbf{y}, \mathbf{z} \in F^n$.

(i) $d(\mathbf{x}, \mathbf{y}) =\geq 0$, with equality if and only if $\mathbf{x} = \mathbf{y}$).

(ii) Symmetry: $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$.

(iii) The triangle inequality: $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$.

Definition 2.5.3 (Hamming weight)**.** *Let F be an Abelian group. The Hamming weight of $\mathbf{e} \in F^n$ is the number of nonzero entries in $\mathbf{e}$. We denote the Hamming weight by $w(\mathbf{e})$.*

Notice that for every two words $\mathbf{x}, \mathbf{y} \in F^n$,

$$d(\mathbf{x}, \mathbf{y}) = w(\curvearrowright - \mathbf{x}). \tag{2.56}$$

Turning back now to block codes, let $\mathfrak{C}$ be an (n, M) code over F with $M > 1$. The *minimum distance* of $\mathfrak{C}$ is the minimum Hamming distance between any two distinct codewords of $\mathfrak{C}$; that is, the minimum distance d is given by

$$d = \min_{\mathbf{c}_1, \mathbf{c}_2 \in \mathfrak{C} : \mathbf{c}_1 \neq \mathbf{c}_2} d(\mathbf{c}_1, \mathbf{c}_2). \tag{2.57}$$

An (n, M) with minimum distance d is often called (n, M, d) code. We will sometimes use the notation $d(\mathfrak{C}))$ for the minimum distance of a given code C.

2.5.3 Decoding

Let $\mathfrak{C}$ be an (n, M, d) code over an alphabet F and let S be the channel defined by the triple (F, Φ, Prob). A decoder for the code $\mathfrak{C}$with respect to the channel S is a function

$$\mathcal{D} : \Phi^n \to \mathfrak{C}. \tag{2.58}$$

The *decoding error probability* P_{err} of $\mathcal{D}$ is defined by

$$P_{\text{err}} = \max_{\mathbf{c} \in \mathfrak{C}} P_{\text{err}}(\mathbf{c}), \tag{2.59}$$

where

$$P_{\text{err}}(\mathbf{c}) = \sum_{\mathbf{y} : \mathcal{D}(\mathbf{y}) \neq \mathbf{c}} \text{Prob}\{\mathbf{y} \text{ received } | \mathbf{c} \text{ transmited } \} \tag{2.60}$$

Information theory is concerned with the theoretical limitations of reliable communication, meaning that is always asking. "What is the best error-correcting performance we could

achieve?"

Coding theory is concerned with the creation of practical encoding and decoding systems.

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 [66]. 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.

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 2.5.4 (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 definitions for codes, codes with fixed block length and dimension. However, since we are interested in the asymptotic behaviour, it turns out to be more useful to study families of codes instead of an specific code.

Definition 2.5.5 (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 asymptotic rate of C is defined as*

$$R(C) = \lim_{i\to\infty} \left\{ \frac{k_i}{n_i} \right\}, \tag{2.61}$$

and the asymptotic relative distance of C is defined as

$$\delta(C) = \lim_{i\to\infty} \left\{ \frac{d_i}{n_i} \right\}, \tag{2.62}$$

from now on whenever we talk about a code we will implicitly refer to an associated family of codes.

Our purposes of this work, we are going to focus on a particular class of codes, namely minimum distance codes. This special kind of codes came to our interest because 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

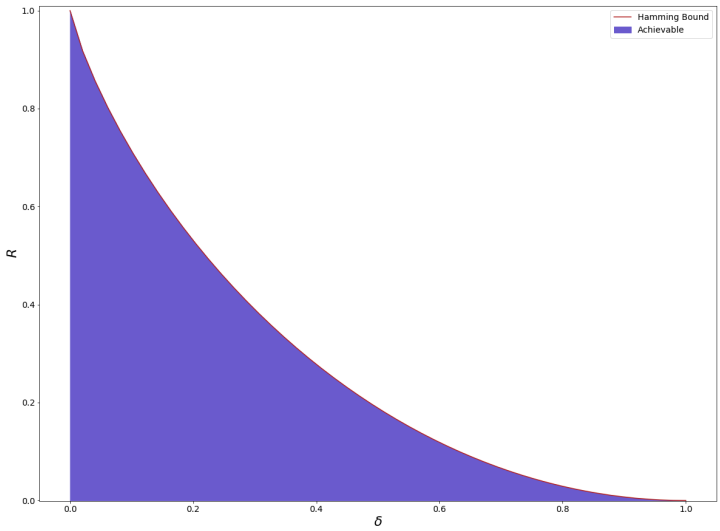

Figure 2.4: An illustration of the Hamming bound for the case of $q = 2$. Note any code above this bound could exist. The shaded region shows the codes that could exist.

to our particular case.

Codes of minimum distance d are codes that have the property that for every pair x, y of codewords we have that

$$\Delta(x, y) \geq d, \tag{2.63}$$

and therefore, all relative distances δ satisfy 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 three results, one positive and two negative results[2]. 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 [66].

[2]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.

2.5.4 Hamming Bound (Sphere packing bound)

Definition 2.5.6 (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*

$$\mathrm{Vol}_q(r,n) = |B_q(\mathbf{0},r)| = \sum_{i=0}^{r} \binom{n}{i} (q-1)^i, \tag{2.64}$$

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

$$\frac{k}{n} \leq 1 - \frac{\log_q \mathrm{Vol}_q\left(\left\lfloor \frac{d-1}{2} \right\rfloor, n\right)}{n}, \tag{2.65}$$

where the volume in (2.65) correspond to the definition in(2.64). With some algebra and using the Stirling asymptotic approximation one can show that

$$\mathrm{Vol}_q\left(\left\lfloor \frac{d-1}{2} \right\rfloor, n\right) \geq q^{H_q\left(\frac{\delta}{2}\right)n - o(n)}, \tag{2.66}$$

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

$$R \leq 1 - H_q\left(\frac{\delta}{2}\right) + o(1). \tag{2.67}$$

The inequality in (2.67) is known as the Hamming Bound.

2.5.5 Plotkin Bound

Definition 2.5.7 (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 2.5.8. *For any q-ary code with relative distance $0 \leq \delta \leq 1 - \frac{1}{q}$,*

$$R \leq 1 - \left(\frac{q}{q-1}\right)\delta + o(1). \tag{2.68}$$

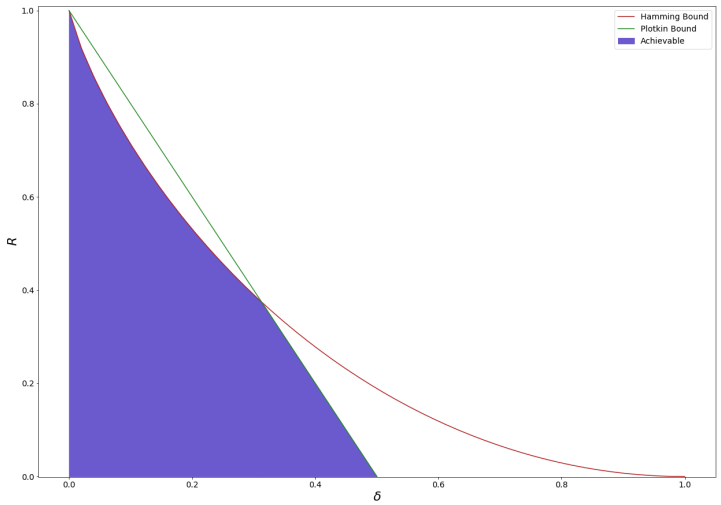

Figure 2.5: An illustration of Plotkin and Hamming bounds for the case of $q = 2$. For this case the shaded region changes indicating us that there are not codes with $R > 0$ when $\delta = 1 - 1/q$.

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

$$\mathcal{C}_{\mathbf{x}} = \{(c_{n-n'+1}, \dots c_n) \mid (c_1 \dots c_N) \in \mathcal{C}, (c_1 \dots c_{n-n'}) = \mathbf{x}\} . \tag{2.69}$$

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

$$|\mathcal{C}_{\mathbf{x}}| \leq \frac{qd}{qd - (q-1)n'} \leq qd, \tag{2.70}$$

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

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

$$|\mathcal{C}| = \sum_{\mathbf{x} \in [q]^{n-n'}} |\mathcal{C}_{\mathbf{x}}| , \tag{2.71}$$

tells us that

$$|\mathcal{C}| \leq \sum_{\mathbf{x} \in [q]^{n-n'}} qd = q^{n-n'} \cdot qd \leq q^{n - \frac{q}{q-1} d + o(n)} = q^{n\left(1 - \delta \cdot \frac{q}{q-1} + o(1)\right)} , \tag{2.72}$$

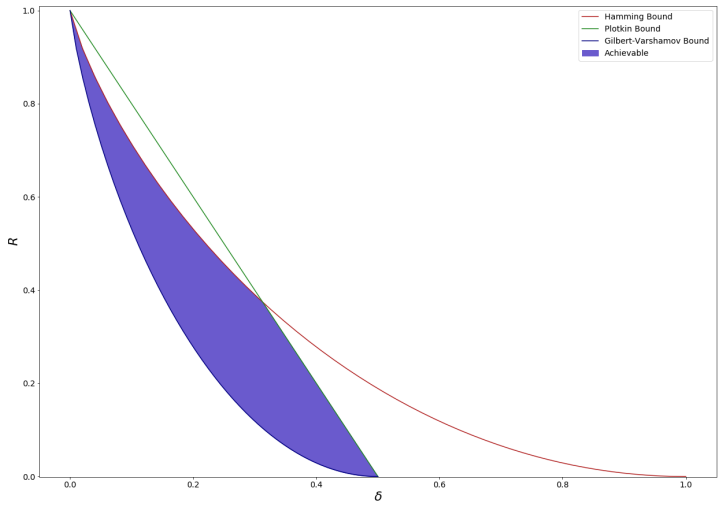

Figure 2.6: An illustration of 3 bounds Hamming, Plotkin and Gilbert-Varshamov for the case of $q = 2$. The lower bound correspond to the Gilvert-Varshamov Bound whereas the other 2 are the upper bound for the rate.

so, this provides another upper bound to the rate given by

$$R \leq 1 - \left(\frac{q}{q-1}\right)\delta + o(1) \tag{2.73}$$

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

2.5.6 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 2.5.9 (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 exists a code with rate $R \geq 1 - H_q(\delta) - \varepsilon$ and relative distance δ.*

To provide a sketch of 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

$$\bigcup_{\mathbf{c}\in C} B(\mathbf{c}, d-1) = [q]^n; \tag{2.74}$$

This is easily checked, since if it were 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 $\mathbf{v}$ could be added. However, this would contradict the fact that we have finished the procedure. So

$$\bigcup_{c\in C} B(\mathbf{c}, d-1) \mid = q^n. \tag{2.75}$$

Also, It isn't hard to see that

$$\sum_{\mathbf{c}\in C} |B(\mathbf{c}, d-1)| \geq \left| \bigcup_{\mathbf{c}\in C} B(\mathbf{c}, d-1) \right|, \tag{2.76}$$

which implies that

$$\sum_{\mathbf{c}\in C} |B(\mathbf{c}, d-1)| \geq q^n, \tag{2.77}$$

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

$$\sum_{\mathbf{c}\in C} \mathrm{Vol}_q(d-1, n) \geq q^n. \tag{2.78}$$

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

$$\begin{aligned} |C| &\geq \frac{q^n}{\mathrm{Vol}_q(d-1, n)} \\ &\geq \frac{q^n}{q^{nH_q(\delta)}} \\ &= q^{n(1-H_q(\delta))}. \end{aligned} \tag{2.79}$$

Therefore concluding the proof. It is worth mentioning that this way of proceeding the code does not any special structure, but as one might think, this algorithm will take exponentially long time to finish. However, one may wonder if there is a special kind of code that also achieves this rate. Indeed is possible to show that *random linear codes* lies, with high probability, on the Gilbert-Varshamov Bound. To pick a *random linear 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 [66, 67]. 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). 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) [25]. As mentioned by A Barg, and G. D. Forney, Jr [25] most of the important results are expressed in terms of the Gilbert-Varshamov distance $\delta_{GV}(R)$.

2.5.7 Error Exponents for Random Minimum Distance Codes

It is very well known that on a BSC with crossover probability p, the channel capacity is $C = 1 - H_2(p)$ [3]. The coding error exponent $E_r(R)$ is positive for $0 \leq R < C$ and given by [68, 69].

$$E_r(R) = \begin{cases} R_0 - R, & 0 \leq R \leq R_{\text{crit}} \\ E_{\text{sp}}(R), & R_{\text{crit}} \leq R \leq C \end{cases}, \tag{2.80}$$

where R_0, R_{crit} and $E_{\text{sp}}(R)$, are known as the cutoff rate, the critical rate and the sphere-packing exponent, respectively. Gallager has shown [70],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 [25] 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 computes 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

$$\Pr\{d_H(\mathbf{x}_i, \mathbf{b}) = d\} = \binom{N}{d} \tilde{p}^d (1 - \tilde{p})^{N-d}, \tag{2.81}$$

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

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

of unordered pairs of codewords $(\mathbf{x}_i, \mathbf{x}_j)$ with $i \neq j$ in C at a distance d apart

$$S_{\mathcal{C}}(d) = \sum_{i=0}^{M-1} \sum_{j=0}^{i-1} \Phi \left\{ d_H \left(\boldsymbol{x}_i, \boldsymbol{x}_j \right) = d \right\}, \tag{2.82}$$

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 distributed random variables, so we have

$$\mathrm{E}S_{\mathcal{C}}(d) = \left(\begin{array}{c} M \\ 2 \end{array} \right) \mathrm{E}\Phi \doteq 2^{N(2R-1+\mathcal{H}(\delta))}. \tag{2.83}$$

Therefore we are ready to state the following theorem

Theorem 2.5.1 (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*

$$\delta_{GV}(2R) + \varepsilon \leq \delta \leq 1 - \delta GV(2R) - \varepsilon, \tag{2.84}$$

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

$$\Pr \left\{ S_{\mathcal{C}}(d) \geq 1 \right\} \leq \mathrm{E}S_{\mathcal{C}}(d) \doteq 2^{-N(1-\mathcal{H}(\delta)-2R)} \to 0, \tag{2.85}$$

which in other words tells 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

$$\Pr \left\{ |S_{\mathcal{C}}(d) - \mathrm{E}S_{\mathcal{C}}(d)| \geq \left(\begin{array}{c} M \\ 2 \end{array} \right) \alpha \right\} \leq \frac{\mathrm{E}\Phi}{\left(\begin{array}{c} M \\ 2 \end{array} \right) \alpha^2}; \tag{2.86}$$

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

$$\Pr \left\{ |S_{\mathcal{C}}(d) - \mathrm{E}S_{\mathcal{C}}(d)| > \left(\begin{array}{c} M \\ 2 \end{array} \right) \alpha \right\} \leq \frac{2\mathrm{E}\Phi}{M(M-1)\alpha^2} \doteq 2^{-N(2R-1+\mathcal{H}(\delta)-2\Delta)}. \tag{2.87}$$

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 $n-$linerarly independent tuples $\mathbf{g}_j$, $1 \leq k \leq K$. This is

$$\mathbf{x}(u) = \Sigma_k \mathbf{u}_k \mathbf{g}_j, \tag{2.88}$$

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 distance distribution $\{d_H(\mathbf{x}_i, \mathbf{x}_j), i \neq j\}$ from any given codeword $\mathbf{x}_i$ is independent of i. The average distance distribution of a linear code C therefore reduces to

$$\mathcal{N}_{\mathcal{C}}(d) = \sum_{j \neq i} \Phi \left\{ d_H \left(\boldsymbol{x}_i, \boldsymbol{x}_j \right) = d \right\} \quad (d = 1, 2, \ldots, N), \tag{2.89}$$

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

$$\Pr \left\{ d_H \left(\boldsymbol{x}_i, \boldsymbol{x}_j \right) = N\delta \right\} \doteq 2^{-N(1-\mathcal{H}(\delta))}, \tag{2.90}$$

for any $d = N\delta$, the quantity $\mathcal{N}_{\mathcal{C}}(d)$ in (2.89) 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

$$\mathcal{N}_{\mathrm{LCE}}(d) \doteq 2^{N(R-1+\mathcal{H}(\delta))}. \tag{2.91}$$

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 emerges as a natural consequence of the structure in the Clifford algebra and therefore provide an explanation of what we named as ultra orthogonality.

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.

3.1 Mechanism Behind Typicality as a Random Minimum Code.

As we mentioned in the first chapter, we are interested in studing the behaviour of states that 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.

3.1.1 ultra orthogonality on fermions

To start this part we want to recapitulate a couple of things: First, as we showed in equation (1.26), 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[1]

$$|\psi\rangle = \sum_{\vec{n}} \psi(\vec{n}) \, |\vec{n}\rangle \, . \tag{3.1}$$

Therefore we are interested in studying the terms of the form[2]

$$\hat{X}_{ij} \doteq \mathrm{Tr}_E \left(|\vec{n}_i\rangle \langle \vec{n}_j| \right) \equiv \mathrm{Tr}_{N/L} \left(|\vec{n}_i\rangle \langle \vec{n}_j| \right) . \tag{3.2}$$

What ultra-orthogonality tells 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 operator is 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 a 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 $\vec{x} = (0100 \ldots 0)$, $\vec{y} = (1100 \ldots 0)$, here $L = 3$, so our function will be described by[3]

$$\gamma(\vec{x}, \vec{y}) = \gamma_1^{x_1=0} \gamma_4^{y_1=1} \gamma_2^{x_2=1} \gamma_5^{y_2=1} \gamma_3^{x_3=0} \gamma_6^{y_3=1} = \gamma_4 \gamma_2 \gamma_5 \gamma_6, \tag{3.4}$$

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

$$\gamma(\vec{x}, \vec{y}) = \gamma_1^{x_1} \gamma_{L+1}^{y_1} \gamma_2^{x_2} \gamma_{L+2}^{y_2} \cdots \gamma_L^{x_L} \gamma_{2L}^{y_L}, \tag{3.5}$$

[1]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.

[2]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.

[3]Here we have changed the original notation given in (2.6) and change it by

$$\hat{c}_{2j-1} \to \gamma_j, \qquad \hat{c}_{2j} \to \gamma_{N+j}, \tag{3.3}$$

for the case of N operators.

where this function has the property that

$$\gamma(\vec{x},\vec{y})\gamma(\vec{x}',\vec{y}') = e^{i\phi(\vec{x},\vec{y},\vec{x}',\vec{y}')}\gamma(\vec{x}+\vec{x}',\vec{y}+\vec{y}'). \tag{3.6}$$

The phase appears 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 provides a set of operators that live in the space of size L, and that will allow us to expand our operator $\hat{X}_{ij}$,[4]

$$\hat{X}_{ij} = \sum_{\vec{x},\vec{y}} f(\vec{x},\vec{y})\gamma(\vec{x},\vec{y}). \tag{3.7}$$

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

$$\mathrm{Tr}_L\left(\hat{X}_{ij}\gamma^\dagger(\vec{x}',\vec{y}')\right) = \sum_{\vec{x},\vec{y}} f(\vec{x},\vec{y})\,\mathrm{Tr}_L\left(\gamma(\vec{x},\vec{y})\gamma^\dagger(\vec{x}',\vec{y}')\right). \tag{3.8}$$

The right hand side of this equation can be combined with equation (3.6) 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

$$f(\vec{x}',\vec{y}') = \frac{1}{2^L}\,\mathrm{Tr_L}\left(\hat{X}_{ij}\gamma^\dagger(\vec{x}',\vec{y}')\right) = \frac{1}{2^L}\left\langle\vec{n}_j\right|\gamma^\dagger(\vec{x}',\vec{y}')\left|\vec{n}_i\right\rangle. \tag{3.9}$$

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 that links what we call the spacial modes and the normal modes

$$\overbrace{\gamma_{i_1}\gamma_{i_2}\ldots\gamma_{i_k}}^{\text{Spacial modes}} = O_{i_1\alpha_1}O_{i_2\alpha_2}\ldots O_{i_k\alpha_k}\underbrace{\gamma_{\alpha_1}\gamma_{\alpha_2}\ldots\gamma_{\alpha_k}}_{\text{Normal modes}}. \tag{3.10}$$

Note that this operators are diagonal over our $\vec{x}$'s and $\vec{y}$'s, which will allow us to operate over the states $|\vec{n}_i\rangle$.[5]. The equation (3.9) simplifies to

$$\left\langle\vec{n}_j\right|\gamma(\vec{x},\vec{y})\left|\vec{n}_i\right\rangle = \delta_{\vec{n}_i+\vec{x}+\vec{y},\vec{n}_j}e^{i\phi(\vec{n}_i,\vec{n}_j,\vec{x},\vec{y})}, \tag{3.12}$$

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

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

$$|n_i\rangle = \gamma(\vec{n}_i,0)\,|0\rangle, \qquad |n_i\rangle = \gamma(0,\vec{n}_i)\,|0\rangle\,e^{i\phi(\vec{n_i})}, \tag{3.11}$$

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.

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

$$f(\vec{x},\vec{y}) = \frac{1}{2^L}\sum_{\vec{x}',\vec{y}'}\mathcal{U}_{\vec{x}\vec{x}'}\mathcal{V}_{\vec{y}\vec{y}'}\underbrace{\langle\vec{n}_j|\,\gamma(\vec{x},\vec{y})\,|\vec{n}_i\rangle}_{\propto\delta_{\vec{n}_i+\vec{n}_j,\vec{x}+\vec{y}}}. \tag{3.13}$$

In equation (3.12) the delta can be changed for $\delta_{\vec{n}_i+\vec{n_j},\vec{x}+\vec{y}}$, and the reason to do so is because we are working with arithmetic modulo 2 and we can see that the term $\vec{n}_i+\vec{n}_j$ 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 whenever 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 differences 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 it 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.

3.2 Fermionic Random Minimum Codes.

Let us recapitulate a little more what we have done in the latter section. 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

$$\begin{aligned} d &= W(\vec{e}_{ij}) \rightarrow \text{Weight of } \vec{e}_{ij} \\ &= d_H(\vec{n}_i,\vec{n}_j) \rightarrow \text{Hamming distance.} \end{aligned} \tag{3.14}$$

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

$$\mathcal{C}=\{\vec{x}^{(1)},\vec{x}^{(2)},\ldots,\vec{x}^{(2^k)}\}\quad \vec{x}\in\{0,1\}^N, \tag{3.15}$$

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

$$\mathcal{H}_{\mathcal{C}}=\mathrm{Span}\left(|\vec{x}^{(1)}\rangle,|\vec{x}^{(2)}\rangle,\ldots,|\vec{x}^{(2^k)}\rangle\right). \tag{3.16}$$

It is easy to check that $|\mathcal{H}_{\mathcal{C}}| = |\mathcal{C}| = 2^{NR}$. From the previous results, we can ensure that there exists 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, $\forall\,|psi\rangle \in \mathcal{H}_{\mathcal{C}}$,

$$|\psi\rangle = \sum_{\vec{n}\in\mathcal{C}} \psi(\vec{n})\,|\vec{n}\rangle\,, \tag{3.17}$$

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

$$\rho_L(\psi) = \mathrm{Tr}_{N/L}\left(|\psi\rangle\,\langle\psi|\right) = \sum |\psi(\vec{n})|^2\,|\vec{n}\rangle\,\langle\vec{n}|\,. \tag{3.18}$$

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

$$\frac{d\rho_L}{dt} = 0. \tag{3.19}$$

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

$$S(\ell) = \bigcup_{\mathcal{C},\delta>\ell} \mathcal{H}_{\mathcal{C}}, \tag{3.20}$$

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 for the probability of having a number of errors greater or equal to a certain quantity. To address this question we can make use of the Chernoff inequality[6]

$$\begin{aligned} P(e^{SX} \geq e^{Sd}) &\leq \min_S \left\langle e^{SX}\right\rangle e^{-Sd} \\ P(e^{-SX} \geq e^{-Sd}) &\leq \min_S \left\langle e^{-SX}\right\rangle e^{Sd}. \end{aligned} \tag{3.21}$$

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

$$\left\langle e^{SX}\right\rangle = \prod_k E(e^{SX(\theta_k)}) = \prod_k \left(1 + p(\theta_k)(e^S - 1)\right) = e^{\sum_k \log(1+p(\theta_k)(e^S-1))} \equiv e^{-Nr(\delta)}, \tag{3.22}$$

[6]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.

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

$$r(\delta) = \min_{S} \frac{1}{N} \left(\log\langle e^{SX} \rangle - S\delta \right). \tag{3.23}$$

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

$$r(\delta) \overset{N\to\infty}{\equiv} \min_{S} \oint \frac{d\theta}{2\pi} \log\left(1 + p(\theta)(e^S - 1)\right) - S\delta, \tag{3.24}$$

which, even though can not be analytically solved, it can be numerically solved by differentiating with respect to S, to obtain

$$\delta = \oint \frac{d\theta}{2\pi} \frac{p(\theta)e^S}{1 - p(\theta) + p(\theta)e^S}. \tag{3.25}$$

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

$$\langle S_{\mathcal{C}}(d) \rangle = \binom{M}{2} 2^{-Nr(\delta)}, \tag{3.26}$$

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. Otherwise the number of pairs that have minimum distance d isexponentially large

$$\langle S_{\mathcal{C}}(d) \rangle \doteq 2^{N(2R - r(\delta))}. \tag{3.27}$$

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.

3.2.1 Case when the numbers of errors is less than L

For this case we know that the term $\hat{X}_{ij}$ is generally different from zero, however, we are going to show that while 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 (3.13), we will find that

$$\begin{aligned} ||\hat{X}_{ij}||^2 &= \mathrm{Tr}\left(\hat{X}_{ij}\hat{X}_{ij}^{\dagger}\right) = \frac{1}{2^L} \sum_{\vec{x},\vec{y}} |f(\vec{x},\vec{y})|^2 \\ &= \frac{1}{2^L} \sum_{\substack{\vec{x},\vec{y} \\ \vec{x}'',\vec{y}''}} \left(\mathcal{U}_{\vec{x}\vec{x}'}\mathcal{U}_{\vec{x}\vec{x}''}\right)\left(\mathcal{V}_{\vec{y}\vec{y}'}\mathcal{V}_{\vec{y}\vec{y}''}\right) e^{\phi(\vec{x},\vec{y}')\phi(\vec{x}'',\vec{y}'')} \delta_{\vec{n}_i+\vec{n}_j,\vec{x}'+\vec{y}'} \delta_{\vec{n}_i+\vec{n}_j,\vec{x}''+\vec{y}''}. \end{aligned} \tag{3.28}$$

Our purpose will then be to bound the terms $\sum_{\vec{x}} (\mathcal{U}_{\vec{x}\vec{x}'}\mathcal{U}_{\vec{x}\vec{x}''})$. The reason for 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 (3.28). First, notice that the expression in (3.13) 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 $\gamma(\vec{\alpha})$

$$\gamma(\vec{\alpha}) =_{[} \gamma_{\alpha_1}\gamma_{\alpha_2}\gamma_{\alpha_2}\cdots\gamma_{\alpha_p]}, \tag{3.29}$$

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

$$\gamma(\vec{\alpha}) = \det\left(O\big|_{\vec{\alpha},\vec{\beta}}\right)\gamma_{\vec{\beta}}, \tag{3.30}$$

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

$$\sum_{\vec{x}} (\mathcal{U}_{\vec{x}\vec{x}'}\mathcal{U}_{\vec{x}\vec{x}''}) = \det\left[\left(O\Pi O^T\right)\big|_{\vec{x}',\vec{x}''}\right]. \tag{3.31}$$

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[7], however, it is not on the scope of this work to explore

[7]Specifically we have computed numerically the values of (3.31) for the case of the one dimensional XY model and we have found that these values are for most cases extremely small.

this for particular cases and we leave it a a future work.

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, *The foundations of statistical mechanics from entanglement: Individual states vs. averages*. 1, 2, 8, 14, 15

[2] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghì, **96**, 050403. 1, 2, 8, 14

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

[4] S. Popescu, A. J. Short, and A. Winter, **2**, 754. 1, 8, 14, 15, 16

[5] **81**, 011109. 1, 14

[6] A. M. Kaufman et al. 1, 14

[7] C. Gogolin and J. Eisert, **79**, 056001.

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

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

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

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

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

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

[14] 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

[15] 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

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

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

[18] N. Linden, S. Popescu, A. J. Short, and A. Winter, **79**, 061103. 2, 3, 8, 16, 17, 18, 19, 20

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

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

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

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

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

[24] A. F. Reyes-Lega, page 1. 5, 23, 27, 28

[25] A. Barg and G. Forney, **48**, 2568, Conference Name: IEEE Transactions on Information Theory. 6, 44

[26] J. W. Gibbs, *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*, C. Scribner's Sons. 8, 9, 11, 12

[27] L. Boltzmann, *Zur Priorität der Auffindung der Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und dem Principe der kleinsten Wirkung*, volume 219, Wiley Online Library, 1871. 9, 10

[28] L.r., **48**, 697. 9

[29] P. Ehrenfest and T. 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]. 9

[30] P. L. Garrido, S. Goldstein, and J. L. Lebowitz, **92**, 050602. 11

[31] S. Goldstein and J. L. Lebowitz, **193**, 53.

[32] L. Boltzmann, *Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie:(vorgelegt in der Sitzung am 8. Februar 1866)*, Staatsdruckerei, 1866. 11

[33] L. Boltzmann, *Über die Beziehung zwischen dem zweiten Hauptsatze des mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung, respective den Sätzen über das Wärmegleichgewicht*, Kk Hof-und Staatsdruckerei, 1877. 11

[34] 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. 11, 14

[35] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghi. 11, 12, 14

[36] E. T. Jaynes, American Journal of Physics **33**, 391 (1965). 11

[37] J. Butterfield, ON SYMPLECTIC REDUCTION IN CLASSICAL MECHANICS, in *Philosophy of Physics*, edited by J. Butterfield and J. Earman, Handbook of the Philosophy of Science, pages 1 – 131, North-Holland, ISSN: 18789846. 12

[38] G. Belot, THE REPRESENTATION OF TIME AND CHANGE IN MECHANICS, in *Philosophy of Physics*, edited by J. Butterfield and J. Earman, Handbook of the Philosophy of Science, pages 133 – 227, North-Holland, ISSN: 18789846. 12

[39] E. T. Jaynes, Phys. Rev. **108**, 171 (1957). 13

[40] J. Von Neumann, *Mathematical foundations of quantum mechanics*, Princeton, N.J. : Princeton University Press. 13

[41] W. H. McCrea, The Mathematical Gazette **23**, 415–417 (1939). 14

[42] D. ter Haar, Chapter 6 - the ensembles in quantum statistics, in *Elements of Statistical Mechanics (Third edition)*, edited by D. ter Haar, pages 216 – 267, Butterworth-Heinemann, Oxford, third edition edition, 1995. 14

[43] C. Fuchs and J. van de Graaf, **45**, 1216, Conference Name: IEEE Transactions on Information Theory. 19

[44] G. Camilo, G. T. Landi, and S. Eliëns, **99**, 045155. 20

[45] 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. 23

[46] Y. Niu et al., **85**, 035110. 23

[47] M.-C. Chung and I. Peschel, **64**, 064412. 23

[48] M. Leijnse and K. Flensberg, **27**, 124003. 23

[49] L. G. Molinari. 23, 24, 25

[50] A. Botero and B. Reznik, **331**, 39 . 23, 24

[51] S. Bravyi. 23

[52] E. Lieb, T. Schultz, and D. Mattis, **16**, 407. 23

[53] J. I. Latorre, E. Rico, and G. Vidal, **4**, 48. 23, 29

[54] S. Katsura, **127**, 1508. 23, 24

[55] E. Barouch and B. M. McCoy, **3**, 786. 23, 24, 28

[56] E. Barouch, B. M. McCoy, and M. Dresden, **2**, 1075. 23, 24, 28

[57] E. Fradkin, *Field Theories Of Condensed Matter Systems Vol#82 On-demand Printing Of #52253*, CRC Press, 1 edition edition. 24

[58] N. N. Bogoljubov, **7**, 794. 24

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

[60] B. Westwanski and A. Pawlikowski, **43**, 201. 24, 25

[61] S.-A. Cheong and C. L. Henley. 25

[62] C. V. Kraus, M. M. Wolf, J. I. Cirac, and G. Giedke, **79**, 012306. 25, 26

[63] M. A. Nielsen. 27

[64] A. Bashirov, Chapter 12 - fourier series and integrals, in *Mathematical Analysis Fundamentals*, edited by A. Bashirov, pages 307–345, Elsevier. 33

[65] C. E. Shannon, **27**, 379, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1538-7305.1948.tb01338.x. 35

[66] D. J. C. MacKay, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, 1 edition edition. 38, 39, 43

[67] E. T. Jaynes, *Probability Theory: The Logic of Science*. 43

[68] R. Gallager. 44

[69] R. G. Gallager, *Information Theory and Reliable Communication*. 44

[70] R. Gallager, **19**, 244. 44

[71] E. Knill, Fermionic linear optics and matchgates.

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

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

[74] E. Knill.

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

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

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

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

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

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

[81] B. Kožnjak, **51**, 42.

[82] J. C. Maxwell, *On Boltzmann's theorem on the average distribution of energy in a system of material points*, Printed at the University Press, 1879.

[83] G. Bierhalter, **24**, 195.

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

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

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

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

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

[89] E. Jaynes, Physical Review **106**, 620 (1957).

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

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

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

[93] 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.

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

[95] M. B. Plenio and S. Virmani.

[96] D. N. Page, **71**, 1291.

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

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

[99] S. Bravyi, version: 1.

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

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

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

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

[104] N. Merhav.

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

[106] 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.

[107] Markov chain monte carlo in practice.

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

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

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

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

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

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

[114] G. Szegö, **76**, 490.