Exploring Thermalization from Tipicallity's viewpoint in the XY model

An alternative mechanism for thermalization

Jose Alejandro Montaña Cortes

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

Exploring Thermalization from Tipicallity's viewpoint in the XY model

An alternative mechanism for thermalization

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

"*one can prove that infinitely many more initial states evolve after a long time towards a more uniform distribution of states than to a less uniform one, and that even in that latter case, these states will become uniform after an even longer time*" (L. Boltzmann 1866) [1]

Abstract

I think the abstract is going to be the last thing i am going to write
Key Words: Typicality, Thermalization, Ultra-Orthogonality..

Acknowledgments

Content

Chapter 1

Quantum Thermodynamics And Canonical Typicality

1.1 Introduction

The foundations of Statistical Mechanics remain a debatable subject. Fundamental questions have emerged form these discussions regarding the role of probabilities, entropy[1] and the relevance of time averages and ensemble averages to individual physical systems. One of the most controversial issue is the validity of the postulate of equal a priori probability, postulate which can not be proved.

In this chapter we are going to discuss some ideas based on typicality addresed by several authors [2–4], who have abandoned the unprovable aforementioned postulate and replace it with a new viewpoint, which is uniquely quantum, and which does not rely on any ignorance probabilities in the description of the state. Instead, is supported, and can be proved, by means of the entanglement between a physical system and it environment.

This chapter will be divided in three parts, first we are going to introduce the postulate of equal prior probability, discussing its commencement in the foundations of statistical mechanics, the idea of an ensemble as a collection of identical systems will be introduced and the postulate of equal prior probability will be translated to a particular version of ensemble, the microcanonical ensemble. Also a quantum version of the above-mentioned postulate will be exhibited in terms of the random phase postulate [5] and the derivation of the canonical ensemble for a weakly interacting system.

The second part will be dedicated to understand entanglement and therefore we present the phenomenon of canonical typicality, where thermalisation emerges as a consequence of typicality. These ideas are taken from *S.Popescu, A. J. short and A. Winter* [4, 6] work which we follow very close[2]. Quantitative arguments will be provided and explained with its

[1] As it is well known, probabilities and entropy are both referred as measures of ignorance

[2] However ,it is important to stress that an independet work done by *Goldstein et. al.* [3] discuss a similar issues addressed by *S.Popescu et. al.* in [4, 6].

respective mathematical tools needed to comprehend the ideas in which typicality is built on.
To conclude, we use all the previous results in the second part to present in the third part an approach studied by *Linden et. al.* [7] in which rather than a kinematic insight, they address the dynamical aspects of thermalisation, explaining a mechanism in which systems far from equilibrium would reach a generic state[3] [7]. Furthermore, we will talk about the main subject of this work, an idea which intuition is built on the ideas of typicality to show why in the case when a state can be decomposed as a linear superposition of states that have approximately the same energy, typicality will still holds and the random phase postulate could be replace with something we called we call super-orthogonality over partial traces, and which can be understood in terms of error an error correction scheme. therefore, it give us an a mathematical proof of why thermalisation will still occurs in the case when we have states that have close eigenenergies.

1.2 Postulate of Equal a Prior Probability

Statistical physics have been historically inspired in phenomenological thermodynamics. In fact quantities such as temperature, pressure, heat and entropy emerged as concepts to describe empirical and macroscopical principles. the fact that we could describe a system by understanding how its microscopic constituents behave via some effective laws, known as *the law of thermodynamics [8]*[4]. A well known result form these studies was the *the second law of thermodynamics* which let us know what kind of natural phenomena could occur.
It was not until Boltzmann, Gibbs and Maxwell that statistical physics became known as a fundamental theory, providing bridge between experimental thermodynamics and the theoretical notions of the atomic world. However, this approach has been a subject of strong debates. Statistical physics differ from phenomenological thermodynamics by two main reasons. The first is the consideration that the behaviour of microscopic constituents could be described with classical dynamics, and the second is the introduction of probability and statistics to deduce the laws of thermodynamics.
We are going to introduce the main theoretical concepts proposed at the end of the 19th century, related to the description of the microscopic world[5].
Consider a mechanical system of N ($N \ggg 1$) particles subject to a time-independent potential V. Supposing every particle in the system has mass m and coordinates in its phase

[3]Here a generic state has to be understood as a typical state, meaning the one that corresponds to its canonical state.

[4]For this very section we follow the ideas provided in [8] which turned up to be a very interesting historical review of statistical mechanics and i sincerely recommend this text.

[5]An interesting fact is that, it was Maxwell who really marks the rebirth of kinetic theory. His ideas about the how in a system with a large number of particles could be described taking a frequency interpretation of the probability was the base for Boltzmann's work.

space $(\vec{q}_i, \vec{p}_i)$ we easily see that the state of the system is represented by a point in the phase space Γ $(\vec{q}, \vec{p}) = (\vec{q}_1, \ldots, \vec{q}_N, \vec{p}_1, \ldots, \vec{p}_N)$, and its time evolution follows the dynamics induced by its Hamiltonian. Moreover, imposing the restriction that the energy E is conserved, the evolution is confined to a set of Hamiltonians such that fulfill this restriction $(\Sigma_E = \{(\vec{q}, \vec{p}) \in \Gamma | H(\vec{q}, \vec{p}) = E\})$. At this moment is when probability comes to play an important role. In [1] Boltzmann proposed to interpret the probability associated to a particular state as the relative time spent by the system in that state[6]. This assumption relies on the hypothesis that the total time of a measurement is extremely long if compared with the intrinsic time scales of the evolution.
By Liouville's theorem, f is a constant function on all the admissible trajectories on Σ_E. Moreover, if we assume that the trajectory of a single point in phase space fills densely Σ_E (Ergodic hypothesis)

$$\rho(\vec{q}, \vec{p}) \propto \delta(H(\vec{q}, \vec{p}) - E). \tag{1.1}$$

based on this argument, Boltzmann was able to prove that thermal equilibrium could be described in terms of Maxwell's distribution of velocities.
In fact in both of his works [9], and [10], Boltzmann mentioned his dissatisfaction about the Ergodic hypothesis and slowly abandoned it. It seems that he only considered it an useful assumption for his general result

> For any system for which the hypothesis is true, its equilibrium state is characterized by[7],
>
> $$\rho_{\text{mc}}\left(p|\vec{q}_1, \ldots, \vec{q}_N\right) dp = \frac{1}{\sqrt{2m\pi}} \frac{\Gamma\left(\frac{3N}{2}\right)}{\Gamma\left(\frac{3N-1}{2}\right)} \frac{\left(E - U - \frac{p^2}{2m}\right)^{(3N-2)/2}}{(E-U)^{(3N-3)/2}} dp, \tag{1.2}$$
>
> from which an analogy to the Maxwell distribution may be recovered in the limit $N \to \infty$, regardless of any details of the inter-particle interactions, or indeed whether the system represented is a gas, fluid, solid or any other thermal body.

1.2.1 Postulate of Equal a Priori Probability and the Microcanonical Ensemble

Al alternative path was taken by Gibbs who made use of an alternative idea of probability in what he named as "ensembles". In his book [11] Gibbs introduce the probability as a distribution function on a collection of identical systems instead of introducing the probability

[6] Concretely the probability of the phase point lies in an infinitesimal region of Σ_E is given by

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

where $dpdq$ is the Lebesgue measure of Γ and f is a suitable function.

[7] This equation was first derived by Boltzmann in [1], and this is the first occasion where probability considerations are applied to the state of the mechanical system as whole, instead of its individual particles.

as an ingredient associated to the state of a single system, as Boltzmann proposed. Gibbs considered "ensembles" in his pioneering treatment of statistical mechanics. His ensembles were infinite sets of macroscopically identical systems, each represented by a point in his phase space (also called *microstates*) being compatible with a single macroscopic configuration (*macrostate*). It is worth stressing that this approach is not interested in following the temporal evolution of a single microscopic configuration, but rather is concerned about the distribution of all the available microscopic configurations, even if the *microstate* associated, is entirely fictitious.
More precisely an ensemble is introduced as a probability density function on the phase space Γ, such that the average number of *microstates* in a region R of Γ is given by $\int_R \rho(\vec{q},\vec{p})d\vec{q}d\vec{p}$. Furthermore, the expectation value of an observable $f : \Gamma \to \mathbb{R}$ is the average of f over Γ

$$\langle f \rangle = \int_\Gamma f((\vec{q},\vec{p})\rho(\vec{q},\vec{p})d\vec{q}d\vec{p}. \tag{1.3}$$

In this approach we take the condition that the ensembles must be stationary[8] , to derive the microcanonical ensemble, which is the only compatible with the conservation of energy is such that

$$\rho_{\mu c}(\vec{q},\vec{p}) = \frac{1}{|\Sigma_E|}\delta(H(\vec{q},\vec{p}) - E), \tag{1.4}$$

with $|\Sigma_E|$ the measure of the set Σ_E.
As mentioned before Classical statistical mechanics is built on the postulate of Equal Probability. This result can be interpreted as follows:

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

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

1.2.2 The Ergodic Hypothesis

We dedicate this special section to discuss a little more about the ergodic hypothesis and how Boltzmann could have understood this to give a formal argument of why it is needed that independent of the initial distribution of the system, this has to evolve towards a uniform distribution (equal a prior probability). First we will show some previous concepts to fully understand this ideas.
The ergodicity provides us an equality between ensemble and time averages of observables.

[8]This condition reads as

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

Each measurement of an observable f at time t_0 takes a certain time to accomplish. In this period of time the observable f samples different values so that the effectively quantity is the time average

$$\frac{1}{t}\int_0^t f\left(T_s\left(\vec{q}_0,\vec{p}_0\right)\right)\mathrm{d}s, \tag{1.5}$$

with $(\vec{q}_0,\vec{p}_0)$ is the initial microstate (at $t=0$), and $\{T_s\}_{s\in\mathbb{R}}$ is the Hamiltonian flow generated by H[9]. Thus we are sampling f trajectories whose initial points are $(\vec{q}_0,\vec{p}_0)$. Furthermore, due to the difference in scales of time between the measurement and the intrinsic time of the microscopical constituents, it is possible to consider the limit

$$f^*\left(\vec{q}_0,\vec{p}_0\right)=\lim_{t\to\infty}\frac{1}{t}\int_0^t f\left(T_s\left(q_0,p_0\right)\right)\mathrm{d}s. \tag{1.6}$$

The problem about the ergodic hypothesis is to determine when the average ensemble average equals the time averages. Of course the general answer to this question is that this equality does not always holds, since $f^*(\vec{q}_0,\vec{p}_0)$ depends on the initial conditions whereas $\langle f\rangle$ does not. nevertheless these differences, the cases when the equality between time and ensemble averages holds, we say that the ergodic hypothesis is valid.

Perhaps the most well-known, interpretation to this problem is the one provided by the Ehrenfests [12]. In essence, they suggest that Boltzmann somehow relied on the ergodic hypothesis in his argument, even though he never explicitly mentioned it.

It is indeed evident that if the ergodic hypothesis holds, a state will spend time in the various regions of the energy hypersurface in phase space in proportion to their volume. In other words, during the evolution of the system along its trajectory, regions with small volume, corresponding to highly non-uniform distributions of state, are visited only sporadically, and regions with larger volume, corresponding to more uniform distributions of state, more often. Therefore, the latter could also make plausible the idea that if a system starts out from a very small region (a very improbable state), it will display a tendency to evolve towards the overwhelmingly larger equilibrium state.

Thus, as is mention in [12], it is suggested that Boltzmann relied on the ergodic hypothesis in order to equate time averages and phase averages, or in other words, to equate two meanings of probability (relative time and relative volume in phase space.) There is however no evidence that Boltzmann ever followed this line of reasoning neither in the 1870s, nor later. He simply never gave any justification for equivocating time and particle averages, or phase averages, at all. Presumably, he thought nothing much depended on this issue and that it was a matter of taste [10].

[9]Here H is the Hamiltonian function describing the system of N particles subject to a time-independet potential V

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

[10]For further readings on ergodic hypothesis i sincerely recommend [8, 13–16]

1.2.3 The Quantum Postulate of Equal a Priori Probability

The latter discussion is completely classical. Since our world is quantum, a quantum formulation of statistical mechanics is needed. Instead of probability distributions that live in phase space, one should consider density matrices, which encode the whole physical content of the system. Therefore, the way we write the postulate of equal a priori probability has to be adapted to the formalism of quantum mechanics.
In quantum mechanics a system is described in a Hilbert space $\mathcal{H}$ and its evolution is generated by a Hamiltonian operator $\hat{H}$. By fixing the energy to belong to an energy shell around the value E ($[E, E+\delta]$), with $\delta \ll E$, but δ large enough so that the shell contains many energy eigenvalues of $\hat{H}$. Rather than considering a phase space region Γ_E, as we did in the classical case, we consider a subspace spanned by all eigenvectors with energy eigenvalues belonging to the energy shell ($\mathcal{H}_R = \mathcal{H}_{[E,E+\delta]}$).
The postulate of equal prior probability affirms that all the states compatible with the energy E are equiprobable[11]. However, in the quantum case this is not enough. In fact, these states must be in an incoherent superposition. This what is know as *the random phase postulate* and its a contribution to the foundations of statistical mechanics which is purely quantum.

The idea here is to assume that the coupling between the system and its environment is sufficiently weak such that the energy of the system is found with in a macroscopically narrow range $[E, E+\delta]$ containing many possible states of the system. Transitions between this energy are mediated by ΔH. All states within energy range which can be connected by range which can be connected by ΔH are considered accessible. We assume that the environment is sufficiently complex, its states so numerous, and its transitions so rapid that phase relationships between different states of the system can not be maintained over microscopically long time intervals. Formally the idea is then that the Fourier coefficient of a vector state in $\mathcal{H}_R$ should have equal probability and completely random phases, due to the unavoidable interactions between the environment and the system [17].
Using these two postulates we conclude that the state of the universe is described in terms of a projector of the Hamiltonian $\hat{H}$ in the constrained Hilbert space $\mathcal{H}_R$. We define $\mathcal{E}_R$, the equiprobable state of the universe corresponding to the restriction R by

$$\mathcal{E}_R = \frac{\mathbb{I}_R}{d_R}, \tag{1.7}$$

where $\mathbb{I}_R$ is the identity (projection) operator on $\mathcal{H}_R$, and d_R is the dimension of $\mathcal{H}_R$. $\mathcal{E}_R$ is the maximally mixed state in $\mathcal{H}_R$, in which each pure state has equal probability. This corresponds to the standard intuition of assigning equal a priori probabilities to all states of the universe consistent with the constraints.

[11] In the quantum mechanical frame this reads as every eigenstate belonging to $\mathcal{H}_R$

1.2.4 The canonical Ensemble

In this part we will show that under some quite general assumptions it is possible to show that if the universe is in the equiprobable state, every small part of it (system) will be in a canonical state Ω_S characterised by a Boltzmann distribution among its eigenstates at a given temperate β^{-1}, this is

$$\Omega_S \propto e^{-\beta H_S} \tag{1.8}$$

To prove this we make a division of the universe in two parts. The system S and its environment E. So Hilbert space will then be described by $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$. Then the Hamiltonian of the universe will be written as

$$\hat{H} = \hat{H}_S \otimes \hat{\mathbb{I}}_E + \hat{\mathbb{I}}_S \otimes \hat{H}_E + \hat{H}_{\text{int}}, \tag{1.9}$$

where $\hat{H}_S$ and $\hat{H}_E$ act separately on the system and its environment respectively, $\hat{\mathbb{I}}_S$ and $\hat{\mathbb{I}}_E$ are the identity operators on $\mathcal{H}_S$ and $\mathcal{H}_E$, and $hat H_{\text{int}}$ describes the interaction between the system and its environment. Supposing the interaction is weak[12] and assuming the dimension of $\mathcal{H}_E$ is much larger than the dimension of $\mathcal{H}_S$ ($d_E = \dim \mathcal{H}_E \gg d_S = \dim \mathcal{H}_S$). We suppose the macroscopic energy E belongs to a small energy interval ($[E, E+\delta]$)[13]. Let $\mathbb{I}_R$ be the identity (projector) operator of $\hat{H}$ on the energy shell

$$\mathcal{H}_R = \mathcal{H}_{[E,E+\delta]}. \tag{1.10}$$

Assuming the universe is in the equiprobable state $\mathcal{E}_R$ given by (1.7). The state of the system S can be obtained by partial trace the state of the universe over the environment, that is,

$$\Omega_S = \text{Tr}_E \, \mathcal{R}. \tag{1.11}$$

From now on the main idea will be to show that Ω_S corresponds to the thermal state at a given temperature[14].
Let $\{|E_k\rangle\}_{k=1}^{d_E} \subset \mathcal{H}_E$ and $\{|\varepsilon_\alpha\rangle\}_{k=1}^{d_S} \subset \mathcal{H}_S$ be the energy basis of $\hat{H}_E$ and $\hat{H}_S$ respectively[15]. Since we use the hypothesis that the interaction is weak, we have that the equiprobable state

[12] This can be understood in terms of the Hamiltonians as

$$||\hat{H}_{\text{int}}|| \lll ||\hat{H}_S||, ||H_E||.$$

[13] As mentioned before here we suppose $\delta \lll E$ on a macroscopic scale but large enough to contain many eigenvalues on H_E

[14] The following prove was taken from [17]

[15] In other words this mean that the Hilbert $\mathcal{H}$ space associated with the universe has basis given by

$$\{|\varepsilon_\alpha\rangle \otimes |E_k\rangle : 1 \leq \alpha \leq d_S, 1 \leq k \leq d_E\}$$

$\mathcal{E}_R$ is written as[16].

$$\mathcal{E}_R = \frac{\mathbb{I}_R}{d_R} \approx \frac{1}{d_R} \sum_{\alpha,k} \chi_{[E,E+\delta]} (\varepsilon_\alpha + E_k) |\varepsilon_\alpha\rangle \langle\varepsilon_\alpha| \otimes |E_k\rangle \langle E_k| . \tag{1.12}$$

By tracing over the environment we get

$$\Omega_S = \mathrm{Tr}_E \, \mathcal{E}_R = \frac{1}{d_R} \sum_{\alpha,k} \chi_{[E,E+\delta]} (\varepsilon_\alpha + E_k) |\varepsilon_\alpha\rangle \langle\varepsilon_\alpha| = \frac{1}{d_R} \sum_{\alpha} d_\alpha^{(E)} |\varepsilon_\alpha\rangle \langle\varepsilon_\alpha| , \tag{1.13}$$

where

$$d_\alpha^{(E)} = \sum_k \chi_{[E,E+\delta]} (\varepsilon_\alpha + E_k) = \sum_k \chi_{[E-\varepsilon_\alpha,E-\varepsilon_\alpha+\delta]} (E_k) . \tag{1.14}$$

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

$$d_\alpha^{(E)} = \mathrm{Tr} \, \chi_{[E-\varepsilon_\alpha,E-\varepsilon_\alpha+\delta]} (H_B) = \dim \mathcal{H}^{(E)}_{[E-\varepsilon_\alpha,E-\varepsilon_\alpha+\delta]}, \tag{1.15}$$

where $\mathcal{H}^{(E)}_{[E_1,E_2]} \subset \mathcal{H}_E$ is the subspace generated by all eigenstates with energy in $[E_1, E_2]$. Thus $d_\alpha^{(E)}$ is a non-negative integer.
Defining the bath entropy at energy E as the logarithm of the number on energy levels in the bath,

$$S_E(E) = \ln \left(\dim \mathcal{H}^{(E)}_{[E,E+\delta]} \right) , \tag{1.16}$$

we deduce,

$$d_\alpha^{(E)} = \dim \mathcal{H}^{(E)}_{[E-\varepsilon_\alpha,E-\varepsilon_\alpha+\delta]} = \mathrm{e}^{S_E(E-\varepsilon_\alpha)}. \tag{1.17}$$

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

$$S_E (E - \varepsilon_\alpha) \approx S_E(E) - \frac{\mathrm{d}S_E(E)}{\mathrm{d}E} \varepsilon_\alpha \tag{1.18}$$

Thus, we get

$$\Omega_S = \frac{1}{d_R} \sum_{\alpha} d_\alpha^{(E)} |\varepsilon_\alpha\rangle \langle\varepsilon_\alpha| \approx \frac{1}{Z} \sum_{\alpha} \mathrm{e}^{-\beta\varepsilon_\alpha} |\varepsilon_\alpha\rangle \langle\varepsilon_\alpha| = \frac{1}{Z} \mathrm{e}^{-\beta \hat{H}_S}, \tag{1.19}$$

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

$$\beta = \frac{\mathrm{d}S_E(E)}{\mathrm{d}E}. \tag{1.20}$$

is the thermodynamical expression of the inverse temperature of the bath.

[16]Here the sums is over indexes k and α such that $\varepsilon_\alpha + E_k \in [E, E + \delta]$. Which is equivalent to say where the characteristic function does not vanish

1.3 Entanglement and Canonical Typicality

As mentioned before we dedicate this section to show how the postulate of equal a priori probability can be proved rather than postulated as we quoted before. The most interesting thing is that this postulate emerges from the very structure of quantum mechanics and it is a simply consequence of entanglement.

1.3.1 Entanglement and the Foundations of Statistical Mechanics

> When two systems, of which we know the states by their respective representatives, enter into temporary physical interaction due to known forces between them, and when after a time of mutual influence the systems separate again, then they can no longer be described in the same way as before, viz. by endowing each of them with a representative of its own (E. Schrödinger 1995 [18]).

To understand what Schrödinger meant by this, we can illustrate the situation by considering a composite system of two spins on the Hilbert space $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$, where $\mathcal{H}_S = \mathcal{H}_E = \mathbb{C}^2$. Denoting $\{|\uparrow\rangle, |\downarrow\rangle\}$, our computational basis of $\mathbb{C}^2$ [19–21]. If the bipartite system is described in terms of a factorized state, for example,

$$|\phi\rangle = |\uparrow\rangle_S \otimes |\uparrow\rangle_E, \tag{1.21}$$

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

$$\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle_S \otimes |\uparrow\rangle_E + |\downarrow\rangle_S \otimes |\downarrow\rangle_E\right) \tag{1.22}$$

then the state is entangled. In fact the density matrix associated with this system is

$$\rho_S = \mathrm{Tr}_E \left(\left|\Phi^+\right\rangle \left\langle\Phi^+\right|\right) = \frac{1}{2}\hat{\mathbb{I}}_2 \tag{1.23}$$

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

$$|\Phi_\alpha\rangle = \sqrt{\alpha}|\uparrow\rangle_S \otimes |\uparrow\rangle_E + \sqrt{1-\alpha}|\downarrow\rangle_S \otimes |\downarrow\rangle_E, \quad \alpha \in [0,1], \tag{1.24}$$

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

$$\rho_S^\alpha = \alpha|\uparrow\rangle\langle\uparrow| + (1-\alpha)|\downarrow\rangle\langle\downarrow|. \tag{1.25}$$

which is a state whose mixture depends on α.
Summarising, if the state of the composite system is factorised, the information on the whole

state and on every subsystem is completely accessible. On the other hand, if the global state is entangled, even though we have complete knowledge of the state in the universe, a prior only a partial knowledge of the subsystem can be obtained. Particularly, if the global state is maximally entangled ($\alpha = 1/2$) one has no information at all on the subsystem. Mathematically speaking, the information content of a state ρ is described by the von Neumann entropy [22].

$$S(\rho) = -\operatorname{tr}(\rho \ln \rho) = -\sum_k p_k \ln p_k, \tag{1.26}$$

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

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

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

$$S\left(\rho_S^\alpha\right) = -\alpha \log \alpha - (1-\alpha)\log(1-\alpha), \tag{1.27}$$

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

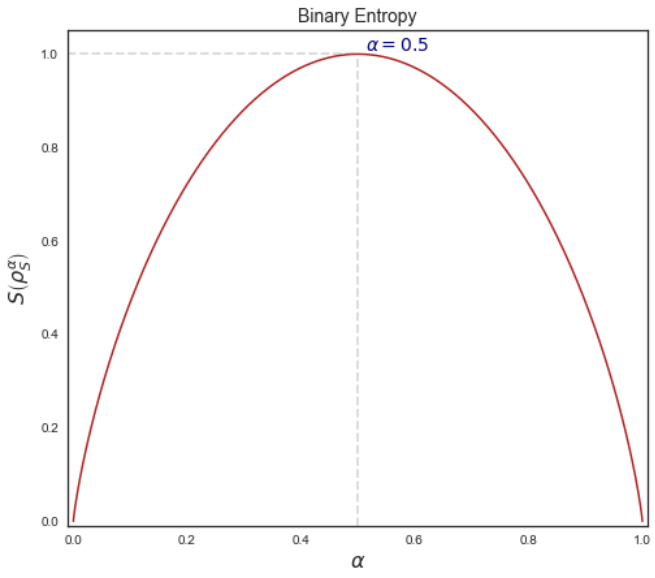

Figure 1.1: Binary Entropy as a function of binary outcome probability α.

and reaches its maximum for the maximally entangled Bell state ($\alpha = 1/2$). In this latter example the entropy is maximal and it corresponds to a complete ignorance on the subsystem

S. Observe that the expression given in (1.27) is nothing but the binary entropy function [23] which is equivalent to the Shannon entropy [24] of the probability vector $(\alpha, 1-\alpha)$.
Form the latter example it is easy to generalise it to a given pure state $|\Psi\rangle$ of a composite system $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$ with generic dimensions $d_S = \dim \mathcal{H}_S \leq d_E = \dim \mathcal{H}_E$, one gets that

$$0 \leq S(\rho_S) \leq \ln d_S. \tag{1.28}$$

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

$$|\Psi\rangle = \frac{1}{\sqrt{d_S}} \sum_{k=1}^{d_S} |u_k\rangle \otimes |v_k\rangle \,, \tag{1.29}$$

with $\{u_k\}, \{v_k\}$ are the orthogonal basis for each of system and environment. The von Neumann entropy is a measure of entanglement which leads to an objective lack of knowledge[17]. In fact, this objective "lack of knowledge" is related to the state of the system, because even if we knew everything about the state of the universe, every subsystem (small portion of the universe) could be mixed [26,27]. This is something astonishing since this lack of knowledge came form nothing but the true nature of quantum mechanics, and no randomness is introduced[18].
In the following part we will show that thermalisation appear as a generic property of pure states of the universe, meaning that for the overwhelming majority of them, the reduced state of the system is the canonical mixed state. As a conclusion we can state that the postulate of equal a prior probability, which refers to ensembles or time averages of states of the universe, and as such relies on a subjective lack of information, can be disregard and one can refer only to pure state of the universe. The lack of information which will give a canonical density matrix for the system is just a physical consequence of entanglement between the system and its environment.

1.3.2 Canonical Typicality

In this part we are going to show that the principle of equal prior probability, which can not be proved, should be replaced with the principle of *Canonical typicality*, which is based on individual states rather than ensembles or time averages, and most importantly, can be

[17] In general entropy provides a tool that can be used to quantify entanglement, although other entanglement measures exist [25] If the overall system is pure, the entropy of one subsystem can be used to measure its degree of entanglement with the other subsystems. In particular for bipartite pure states, the von Neumann entropy of reduced states is the unique measure of entanglement in the sense that it is the only function on the family of states that satisfies certain axioms required of an entanglement measure. Being this the reason why we only introduce this entropy.

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

proved.
Consider a large quantum mechanical system, "The Universe", which we decompose in two parts, the system S and the environment E[19]. Now, suppose the universe has to obey some global constraint R, which translates into the choice of a subspace of the total Hilbert space, say

$$\mathcal{H}_R \subset \mathcal{H}_S \otimes \mathcal{H}_E \tag{1.30}$$

As we mentioned before, the dimensions of $\mathcal{H}_S$, $\mathcal{H}_E$ and $\mathcal{H}_R$ will be denote by d_S, d_E and d_R respectively. In the standard approach to statistical mechanics, as we aforementioned, the restriction is imposed on the total energy. However, as *Popescu et. al.* mention in [4, 6], this restriction can be let to be completely arbitrary.
Using the definition of the equiprobable state in (1.7), we know that $\mathcal{E}_R$ is the maximally mixed state in $\mathcal{H}_R$, in which pure state has equal probability[20].
The *canonical state* of the system S is defined as the trace over the environment,

$$\Omega_S = \mathrm{Tr}_E\, \mathcal{E}_R. \tag{1.31}$$

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

$$\rho_S = \mathrm{tr}_B(|\phi\rangle\langle\phi|) \tag{1.32}$$

Here we ask ourselves, how different is ρ_S form the canonical state Ω_S. The answer to this is provided by *Popescu et. al.* in [4, 6], 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 $\mathcal{E}_R$.
Nevertheless, it is important to stress that Ω_S is not necessarily the thermal canonical state (1.19), but rather a generalised canonical state with respect to the arbitrary restriction R chosen[21].
As an illustrative example, we can interpret this result by representing our universe obeying the global constraint R as a map chart in which we see nothing but a vast ocean representing the equiprobable states in $\mathcal{H}_R$ and some islands representing those states who differ drastically from the canonical state. Particularly we represent with a boat a random picked state in the universe, as one can see, for the vast majority of random choices we have, this boat would land in a portion of this ocean.

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

[20] As Popescu mention in his paper, this corresponds to the standard intuition in classical mechanics of assigning equal a priori probabilities to all states of the universe consistent with the constraints [4, 6].

[21] The thermodynamic interpretation is recover is the restriction imposed coincide with the total energy. In that case we can state that almost every pure state $|\phi\rangle$ of the universe is such that the system S is approximately in the canonical thermal state $e^{-\beta \hat{H}_S}/Z$.

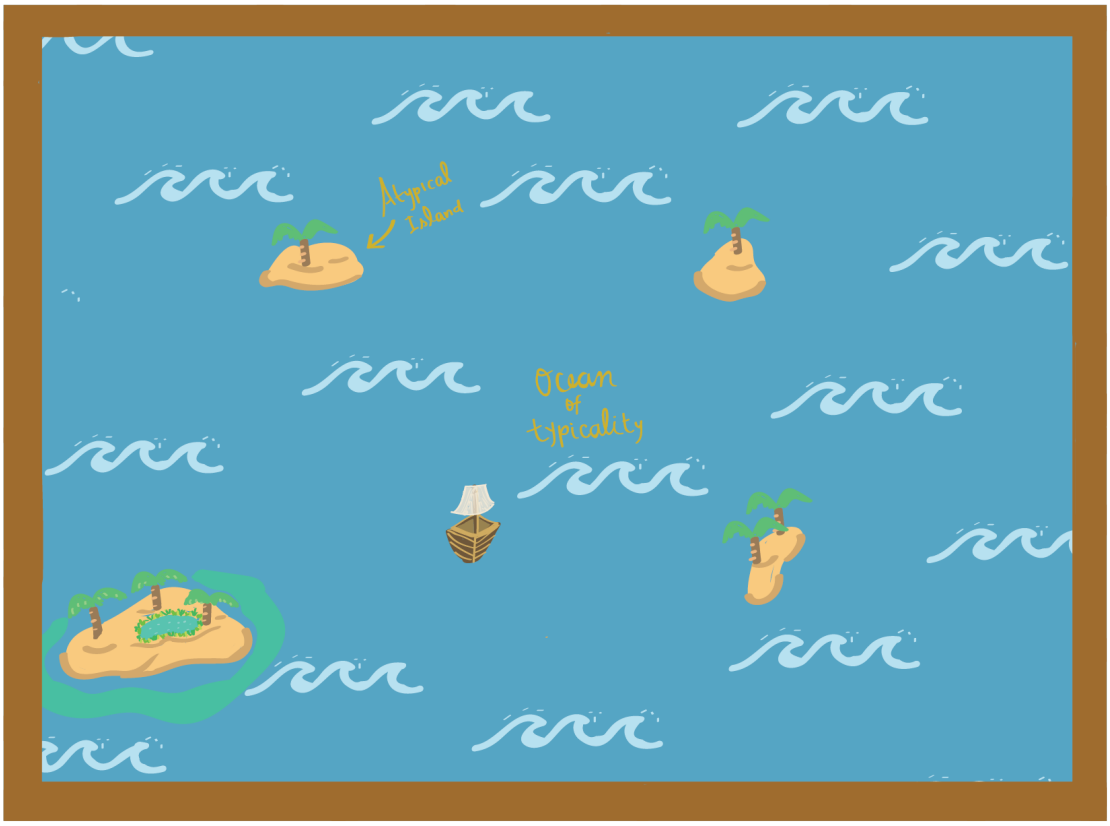

Figure 1.2: Illustration of the typicality in a map chart. The boat here refers to a random picked state in our universe. Here the islands refer to those states who radically differ from the Canonical state.

1.3.3 Quantitative Arguments behind Typicality

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

$$||\rho||_1 = \mathrm{Tr}\,|\rho| = \mathrm{Tr}\left(\sqrt{\rho^\dagger \rho}\right) \tag{1.33}$$

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

$$\|\rho\| = \sup_{\|M\|=1} |\,\mathrm{Tr}(\rho M)|. \tag{1.34}$$

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

$$|\mathrm{Tr}\,(\rho_S M) - \mathrm{Tr}\,(\Omega_S M)| \leq \|\rho_S - \Omega_S\|_1 \, \|M\|, \tag{1.35}$$

An important fact to stress is the choice of distance, one could wonder, why do not we use other operator distance such like Hilbert-Schmidt norm, which is defined by taking the root square of the trace, instead of the trace of the root square, as the trace norm is defined. The reason to do so is simply because the Hilbert-Schmidt distance in higher dimensions can tend to be very small even when the states have a disjoint support [28]. To clarify this, consider in $\mathbb{C}^{2d}$ the two states $\rho_1 = \mathbb{I}_1/d$ and $\rho_2 = (1 - \mathbb{I}_1)/d$, where $\mathbb{I}_1$ is a projector over d dimensions. An important fact over these states is that they have disjoint supports just as we mentioned before. So computing the distance between these two states give us

$$\|\rho_1 - \rho_2\|_1 = 2, \quad \|\rho_1 - \rho_2\|_2 = \sqrt{\frac{2}{d}}, \tag{1.36}$$

and then the trace norm give us a constant while the Hilbert-Schmidt distance decreases d increases $1/\sqrt{d}$.

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 the real sphere $(2d_R - 1)$-dimensional. Therefore, the states we are interested on, live over the surface of a sphere of d_R dimensions, thus by random sample one of this states we will have to sample with the measure $\sigma(\mathbb{S}^{2d_R-1})$, which is known as the Haar measure. From this definition, it is clear to see if the selected random state over the Haar measure, the average state of the universe in $\mathcal{H}_R$ is nothing but the equiprobable state, and then we get that average state of the system is nothing but the canonical state $\Omega_S = \langle\rho_S\rangle$.

Now with the notion of the distance we chose to work with, and the space where these pure states live in, we are ready to announce the general result in typicality.

Theorem of Canonical Typicality [4, 6].

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:

$$\mathrm{Prob}\,(\|\rho_S - \Omega_S\|_1 \geq \eta) \leq \eta', \tag{1.37}$$

where

$$\eta = \varepsilon + \sqrt{\frac{d_S}{d_E^{\mathrm{eff}}}}, \quad \eta' = 2\exp\left(-C d_R \varepsilon^2\right), \tag{1.38}$$

with

$$C = \frac{1}{18\pi^3}, \quad d_E^{\mathrm{eff}} = \frac{1}{\mathrm{Tr}\,\Omega_E^2} \geq \frac{d_R}{d_S}, \quad \Omega_E = \mathrm{Tr}_S\,\mathcal{E}_R \tag{1.39}$$

Form the latter result we notice that η and η' are small quantities, and then, the state will be close to the canonical state with high probability, whenever $d_E^{\text{eff}} \ggg d_S$ and $d_R\varepsilon^2 \ggg 1 \ggg \varepsilon$. The latter condition can be ensured when $d_R \ggg 1$.
So what our latter results holds 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, we will have always (almost) that every system, with small enough dimension, will be indistinguishable from the canonical state. Even more, *Popescu et. al.* get an expression to show how the fluctuations around the average behave. For this they get the next result

$$\langle \|\rho_S - \Omega_S\|_1 \rangle \leq \sqrt{\frac{d_S}{d_E^{\text{eff}}}} \leq \sqrt{\frac{d_S^2}{d_R}}, \tag{1.40}$$

which tells us, if the dimension of our accessible space d_R is large enough, compared to the dimension of the system, the fluctuations on the system are also very small. Thus these two results give us the quantitative arguments of typicality.
Up to this moment we have provided the qualitative as well quantitative arguments to show why typicality gives us an approach different than the equal a priori probability, that can not be proved, and replace it with a key quantum property, entanglement, which provides us a subjective lack of knowledge and more important that can be proved, to show that thermalisation turns up to be a generic property of pure states of the universe. However, despite this result explains very well the reason why by choosing randomly 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) under certain unitary evolution can reach its thermalisation. This means that the latter result is Kinematic rather than dynamical and therefore can not be used to explain how a thermalisation in a system can occur. The goal of the next section is to introduce a work done by *Linden et. al.* [7] in which a dynamical approach is done in order to explain how thermalisation can occur in a system reliant on a given unitary dynamic.

1.4 Evolution Towards Equilibrium.

As we above-mentioned, the latter result is only valid for a given time and state, meaning that this can not be used to study the how thermalisation occurs. Now we are interested in states that are atypical, in the sense that are going to be those states such that drastically differ from the canonical state. Even though, we might think that if typicality holds, most evolutions will quickly drive us from a state in which the system is not thermalised into one that is, and that system will remain in this states of "thermalisation" for most of its evolution. Nevertheless, there are some problems which are worth to mention about why this is much harder problem to solve. In the following we will show what could be named

as the roadmap of what is needed to be proved in order to show that a thermal state has thermalised.

- **Equilibration**: It is possible to affirm that a system will equilibrate if its states evolve towards a particular state, which can be in the more general case a mixed state, and remains in that state, or at least quite close to it, for every time[22]
- **Environment state independence**: The equilibrium state that the system reaches has to be independent of the initial state of the environment, this is, when the system reaches its state of equilibrium, this state should depend only on macroscopic parameters of the environment, like temperature or similar macroscopic parameters.
- **System state independence**: If the system is much smaller than the environment, the state of equilibrium should not depend of its initial state.

Having this in mind, it is possible to tackle these problems one by one. An interesting result is provided by *Linden et. al.* [7], where they manage to prove that with relatively full generality, equilibrium is an universal property of Quantum systems and even more that equilibrium state does not depend on the state of the environment.
To completely understand *Linden et. al.* work is necessary to first understand couple of definitions they provide in their paper [7].
Universe: Here we will refer always to a large quantum universe living in a Hilbert space $\mathcal{H}$. As previously, we are considering an universe that can be decomposed in two, in this decomposition we refer the system S to a small part of the Hilbert space and the rest we will call it the environment. Explicitly we decompose the Hilbert space of the universe as a tensor product of the Hilbert space of the system and the environment, $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$, where d_S and d_E the respective dimension of the system and the environment. Notice that here the environment nor the system have provided with any special property. Here the system could be a single particle or even a section of a lattice.
Hamiltonian: The evolution of the universe will be governed by a Hamiltonian given by

$$\hat{H} = \sum_k E_k \left|E_k\right\rangle \left\langle E_k\right| . \tag{1.41}$$

with $|E_k\rangle$ the eigenstate in the energy basis with energy E_k. The main assumption here is related with the possible values of energies this Hamiltonian can have. The only requirement needed the Hamiltonian to have non-degenerate energy gaps.
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

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

$k = \ell$, or $k = m$ and $\ell = n$. Which turns out to be the same condition of imposing that the energy levels have to be non-degenerate.
Notice that the restriction imposed to the Hamiltonian is a extremely natural constraint, this is because all Hamiltonian that lack of symmetries have non-degenerate energies, so we refer to a set of Hamiltonians with measure 1.
Notation: We will work here with pure states for the universe represented by $|\Psi(t)\rangle$ with a density matrix state given by $\rho(r) = |\Psi(t)\rangle \langle\Psi(t)|$. As above-mentioned, the state of the system is obtained by tracing out the environment at a time t, that is, $\rho_S(t) = \mathrm{Tr}_E\, \rho(t)$. Similarly we define the state of the environment as $\rho_E(t) = \mathrm{Tr}_S\, \rho(t)$.
We define a convenient quantity which is the transient state, or the time averaged state ω

$$\omega = \langle \rho(t) \rangle_t = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \rho(t) dt, \tag{1.42}$$

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

$$d^{\mathrm{eff}}(\rho) = \frac{1}{\mathrm{Tr}\,(\rho^2)}, \tag{1.43}$$

which is generally a better measurement of the effective dimension than the dimension of the support of ρ. This measure roughly tells us how many states contribute to the mixture[23], carrying the probabilistic weight of different states in the mixture, and is a continuous measure.
With the concepts aforementioned Linden et. al. [7] are able to mathematically prove

> *Every pure state of a quantum universe, composed by a large number of eigenstates of energy*[24] *such that evolves under an arbitrary Hamiltonian, is such that every system small enough will equilibrate.*

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

$$|\Psi(t)\rangle = \sum_k c_k e^{-iE_k t} |E_k\rangle \tag{1.44}$$

[23]Particularly, a mixture of n orthogonal states with equal probability has effective dimension of n
[24]The reason why we need the global state to have many eigenstates of energy is because by imposing this we can assure that there will be a large quantity of changes throughout the evolution of the system.

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

$$\rho(t) = \sum_{k,l} c_k c_l^* e^{-i(E_k - E_l)t} |E_k\rangle \langle E_l| , \tag{1.45}$$

which can be expanded and written as

$$\begin{aligned} \rho(t) &= \underbrace{\sum_n \|c_n\|^2 |E_n\rangle \langle E_n|}_{\omega} + \underbrace{\sum_{m \neq n} c_n c_m^* |E_n\rangle \langle E_m| e^{-it(E_n - E_m)}}_{\lambda(t)} \\ &= \omega + \lambda(t), \end{aligned} \tag{1.46}$$

Now in the case of non-degeneracy of the energy levels we have

$$\omega = \langle \rho(t) \rangle_t = \sum_k |c_k|^2 |E_k\rangle \langle E_k| , \tag{1.47}$$

leading us to

$$d^{\text{eff}}(\omega) = \frac{1}{\text{Tr}\,(\omega^2)} = \frac{1}{\sum_k |c_k|^4}. \tag{1.48}$$

Thus formally the statement above-mentioned can be mathematically written in terms of central quantity $D\left[\rho_S(t), \omega_S\right]$[25], the distance between $\rho_S(t)$, the state of the system at a time t, and its time average, $\omega_S = \langle \rho_S(t) \rangle_t$. The difference between $\rho_S(t)$ and ω_S in the energy eigenstates can be written as

$$\rho_S(t) - \omega_S = \sum_{m \neq n} c_m c_n^* e^{-i(E_m - E_n)t} \,\text{Tr}_E \, |E_m\rangle \langle E_n| . \tag{1.49}$$

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. Therefore, in order to characterise these fluctuations we will be interested in the time average of distance $\langle D\left[\rho_S(t), \omega_S\right] \rangle_t$, so the value this average takes will tell us about where the system is spending most of its time. In other words $\langle D\left[\rho_S(t), \omega_S\right] \rangle_t$ will be small when the system equilibrates to ω_S.
To be able to prove what is announced as the *Theorem 1* in [7] it is useful to relate the trace distance to the square of the Hilbert-Schmidt distance using a standard bound provided in [29]

$$D\left(\rho_1, \rho_2\right) = \frac{1}{2} \,\text{Tr}_S \sqrt{\left(\rho_1 - \rho_2\right)^2} \leq \frac{1}{2} \sqrt{d_S \,\text{Tr}_S \left(\rho_1 - \rho_2\right)^2}. \tag{1.50}$$

By using the concavity of the square-root function, we therefore have

$$\langle D\left[\rho_S(t), \omega_S\right] \rangle_t \leq \sqrt{d_S \left\langle \text{Tr}_S \left[\rho_S(t) - \omega_S\right]^2 \right\rangle_t}, \tag{1.51}$$

[25]Here has in the case of typicality we use the trace distance.

which will provide us the bound we need to proof the theorem. Now using (1.49) we write

$$\left\langle \mathrm{Tr}_{\mathcal{S}}\left[\rho_{\mathcal{S}}(t)-\omega_{\mathcal{S}}\right]^2\right\rangle_t = \sum_{m\neq n}\sum_{k\neq l}\mathcal{T}_{klmn}\,\mathrm{Tr}_{\mathcal{S}}\left(\mathrm{Tr}_E\left|E_k\right\rangle\left\langle E_l\right|\mathrm{Tr}_E\left|E_m\right\rangle\left\langle E_n\right|\right), \tag{1.52}$$

where $\mathcal{T}_{klmn} = c_k c_l^* c_m c_n^* e^{-i(E_k-E_l+E_m-E_n)t}$. Computing this time average taking into account that the Hamiltonian has non-degenerate energy gaps[26] we find that

$$\begin{aligned}
\left\langle \mathrm{Tr}_S\left[\rho_S(t)-\omega_S\right]^2\right\rangle_t &= \sum_{k\neq l}\left|c_k\right|^2\left|c_l\right|^2\mathrm{Tr}_S\left(\mathrm{Tr}_E\left|E_k\right\rangle\left\langle E_l\right|\mathrm{Tr}_E\left|E_l\right\rangle\left\langle E_k\right|\right)\\
&= \sum_{k\neq l}\left|c_k\right|^2\left|c_l\right|^2\sum_{ss'bb'}\left\langle sb|E_k\right\rangle\left\langle E_l|s'b\right\rangle\left\langle s'b'|E_l\right\rangle\left\langle E_k|sb'\right\rangle\\
&= \sum_{k\neq l}\left|c_k\right|^2\left|c_l\right|^2\sum_{ss'bb'}\left\langle sb|E_k\right\rangle\left\langle E_k|sb'\right\rangle\left\langle s'b'|E_l\right\rangle\left\langle E_l|s'b\right\rangle\\
&= \sum_{k\neq l}\left|c_k\right|^2\left|c_l\right|^2\mathrm{Tr}_E\left(\mathrm{Tr}_S\left|E_k\right\rangle\left\langle E_k\right|\mathrm{Tr}_S\left|E_l\right\rangle\left\langle E_l\right|\right)\\
&= \sum_{k\neq l}\mathrm{Tr}_E\left[\mathrm{Tr}_S\left(\left|c_k\right|^2\left|E_k\right\rangle\left\langle E_k\right|\right)\mathrm{Tr}_S\left(\left|c_l\right|^2\left|E_l\right\rangle\left\langle E_l\right|\right)\right]\\
&= \mathrm{Tr}_E\,\omega_E^2 - \sum_k\left|c_k\right|^4\mathrm{Tr}_S\left[\left(\mathrm{Tr}_E\left|E_k\right\rangle\left\langle E_k\right|\right)^2\right]\\
&\leq \mathrm{Tr}_E\,\omega_E^2,
\end{aligned} \tag{1.53}$$

where $\omega_E = \mathrm{Tr}_S\,\omega$. To obtain a further bound, we invoke weak sub-additivity of the Reényi entropy [30]

$$\mathrm{Tr}\left(\omega^2\right) \geq \frac{\mathrm{Tr}_E\left(\omega_E^2\right)}{\mathrm{rank}\left(\rho_S\right)} \geq \frac{\mathrm{Tr}_E\left(\omega_E^2\right)}{d_S}, \tag{1.54}$$

and therefore combining (1.51), (1.53) and (1.54) we get

$$\left\langle D\left[\rho_S(t),\omega_S\right]\right\rangle_t \leq \frac{1}{2}\sqrt{d_S\,\mathrm{Tr}_E\left(\omega_E^2\right)} \leq \frac{1}{2}\sqrt{d_S^2\,\mathrm{Tr}\left(\omega^2\right)}, \tag{1.55}$$

which by taking the definition of effective dimension, we get the main result shown in [7]

$$\left\langle D\left[\rho_S(t),\omega_S\right]\right\rangle_t \leq \frac{1}{2}\sqrt{\frac{d_S}{d^{\mathrm{eff}}\left(\omega_E\right)}} \leq \frac{1}{2}\sqrt{\frac{d_S^2}{d^{\mathrm{eff}}(\omega)}}. \tag{1.56}$$

As we can see the result obtained by Linden et. al. tell us that the vast majority of quantum systems 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[27].

[26] This condition is reflected in the evaluation by considering the terms where $k \neq \ell$ and $m \neq n$, leading to $m = \ell$ and $k = n$ are the only terms that contribute.

[27] Notice that the last result is not necessarily considering that the state of equilibration will coincide with the canonical state, this result is very general. Exponential bounds can be achieve if we consider the energy eigenvalues of the Hamiltonian to have no rational dependencies, this restriction leads to exponential bound in $d^{\mathrm{eff}}(\omega)$ (check Appendix C in [7]) .

To illustrate this result we will picture an universe as the one described before in which the islands refer to the systems regions of the universe such that are atypical and the ocean the region which represent the region where the system equilibrates. As shown in the figure 1.3 we can start from a region in the island (out of the equilibrium) and going out to the ocean (region of equilibrium), as is shown we jump from the island to the ocean so that in average we spend most of the time out of the island. Even though, this results tell us that apparently

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

the quantum property of systems is to be close to its equilibrium state, it does not take into consideration how fast it is getting to this state. Since the time averages are taken, we can say that the term of λ goes to zero in trace norm, that is $\langle \|\lambda\|_1 \rangle_t \to 0$, this can be assured since the *Random phase postulate* described above, holds. However, it is easy to see that if we take a state which its energy eigenvalues are very close to each others, the time it will take to equilibrate is exponentially large. It is not of our purpose to measure how fast a system presenting but instead we present an alternative to the *Random phase postulate*, for the special case in which the energy eigenvalues are very close to each other.
To better explain our idea consider an universe $|\Psi\rangle$ which is decomposed on two states that are very close to each other (meaning that live on the shell of a defined energy), for example a state of the universe belonging to an shell of a defined energy, the energy eigenstates $|E_n\rangle$

and $|E_m\rangle$, the state of the universe can be written as

$$\begin{aligned}\rho &= |\Psi\rangle\langle\Psi| \\ &= \|c_n\|^2 |E_n\rangle\langle E_n| + \|c_m\|^2 |E_m\rangle\langle E_m| \\ &+ c_n c_m^* |E_n\rangle\langle E_m| + c_m c_n^* |E_m\rangle\langle E_n|,\end{aligned} \tag{1.57}$$

if we take the partial trace of the equation (1.57) this will lead us to the state of the system,

$$\begin{aligned}\rho_S &= \mathrm{Tr}_E\,\rho = \mathrm{Tr}_E\,|\Psi\rangle\langle\Psi| \\ &= \|c_n\|^2 \mathrm{Tr}_E\,|E_n\rangle\langle E_n| + \|c_m\|^2 \mathrm{Tr}_E\,|E_m\rangle\langle E_m| \\ &+ c_n c_m^* \mathrm{Tr}_E\,|E_n\rangle\langle E_m| + c_m c_n^* \mathrm{Tr}_E\,|E_m\rangle\langle E_n|.\end{aligned} \tag{1.58}$$

Something we notice from (1.58) is that since we consider the state to have its eigenvalues of energy to be quite close to each other, by using typicality we get that $\mathrm{Tr}_E\,|E_n\rangle\langle E_n| = \mathrm{Tr}_E\,|E_m\rangle\langle E_m| = \Omega(E)$, with $\Omega(E)$ the canonical state. Therefore the equation (1.58) can be written as

$$\rho_S = \Omega(E) + c_n c_m^* \mathrm{Tr}_E\,|E_n\rangle\langle E_m| + c_m c_n^* \mathrm{Tr}_E\,|E_m\rangle\langle E_n|. \tag{1.59}$$

Since the canonical state is time independent we see that the last terms in (1.59) should cancel in order to be consistent with the result of typicality discussed before. The only way those terms vanish, is by a property that we named call ultra-orthogonality, which refers to an apparent orthogonality of states with respect to the partial trace, that it

$$\mathrm{Tr}_E\,|E_n\rangle\langle E_m| = \mathrm{Tr}_E\,|E_m\rangle\langle E_n| \approx 0 \tag{1.60}$$

Considering this, we state that the mechanism which makes the systems to equilibrate emerge naturally as a consequence of typicality.

From these ideas we decided to explore this super-orthogonality in a physical system such that can be solved analytically (or numerically). Take into account that the latter requirement is needed to have a complete knowledge of the state of our universe, and be able to then take a small portion of it to study this reduced state. Thus in order to do that, it is necessary to be able to compute with relative easiness the diagonalization of our universe as well as the reduced states of it in an efficient way[28]. In the following sections we will provide the background required to understand the choice of the system we made as well as a detailed calculations which provide a full characterization of this system and its reduced states. More over we show a way to generalise this result to any fermionic system and how this Ultra orthogonality is related with a minimal distance code.

[28]We emphasise that it has to be efficient because the trend of exponential growth inherit by Hilbert spaces turns out to be a challenging problem when working with large systems.

Chapter 2

Fermionic States, its characterisation and the connection with Coding theory.

Up to this point we provided a detailed explanation of how in quantum mechanical description, reaching equilibrium emerges as a consequence of the structure of quantum mechanics. Even more, we stated that in the case we deal with states which its eigenenergies are close to each other, typicality will also provide us an answer of how this kind of states will reach equilibrium. As we would like to illustrate this phenomena occur, we focus our study in the fermionic case. In this Chapter we are Going to provide a background to understand how fermionic states are usually treated and why we choose to work with them. More specifically, we provide the overview of solvable fermionic systems, its connection to Majorana fermions and Gaussian states, the formalism of Grassmann for anticommuting variables, and the link between all the formalism for fermions and coding theory.

2.1 Overview

In many areas of physics one has to has to deal with solving quantum many body problems, which is often a computationally difficult if not impossible task. However, the cases which can be annalitically solved are very well known, and some some assumptions have to be taken into account. In spite of this considerations it has been found that a wide class of complicated Hamiltonians with many-body interactions can be often be mapped onto Hamiltonians that are quadratic in annihilation and creation operators and have the generic form [31]

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

where i, j run from 1 to the number of modes in the system (N) and $\hat{a}_i$, $\hat{a}_i^\dagger$ are Fermionic annihilation and creation operators which satisfy the canonical anti-commutation relations (CAR) [32]

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

A convenience when working with these kind of Hamiltonians is that can be diagonalized via a Bogoliubov- Valantin transformations transformation (i.e., canonical transformations), which maps Fermionic creation and annihilation operators on the creation and annihilation operators of non-interacting quasi-particles [33, 34]. Explicitly the transformation looks like

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

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

Many relevant physics models are diagonalizable via a Bogoliubov-Valantin transformations, some examples are the Hubbard model, BCS theory of superconductivity in the mean field or Hartree-Fock approximation, and certain solvable spin-chain models (After a Jordan-Wigner transformation) [32]. As we will later explain, an important feature about the class of Hamiltonians described by (2.1), if not the most important for the purpose of this project, is that not only the ground state (quasi-particle vacuum) but all eigenstates describing a excitation in a set of quasi-particles, belong to the class of so-called Fermionic Gaussian states [31]. What is important about this class of states is that are fully characterized by second order correlations, because all the higher moments factorize. This result is very well known and is known as Wick theorem.

2.2 Majorana Fermions

Majorara fermions are fermions such that they are their own antiparticle. In the frame of condensed matter Majorana Fermions (quasi-particles) can be interpreted as a superposition of a electron state and a hole [35].

The formalism of Majorana fermions or Majorana modes of the system (with N modes) can be introduce with the operators

$$\hat{c}_{2j-1} = \hat{a}_j^\dagger + \hat{a}_j, \quad \hat{c}_{2j} = (-i)\left(a_j^\dagger - a_j\right). \tag{2.4}$$

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

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

and

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

In which case its canonical anti-commutation relations (CAR) take the form

$$\{\hat{c}_k, \hat{c}_l\} = 2\delta_{kl}. \tag{2.5}$$

The anti-commutation relations is seen to be a consequence of an $\mathbb{R}^{2N}$ Clifford algebra[2]. Transformation in between Fermionic and Majorana operators is achieved by matrix of the block form

$$\Omega = \begin{pmatrix} \mathbb{I} & \mathbb{I} \\ i\mathbb{I} & -i\mathbb{I} \end{pmatrix}, \tag{2.6}$$

and then the map from Fermionic ($\vec{\hat{a}}^T = (\hat{a}_1, \ldots, \hat{a}_1^\dagger, \ldots)$) and Majorana ($\vec{\hat{c}}^T = (\hat{c}_1, \ldots, \hat{c}_1^\dagger, \ldots)$) operators is written as $\Omega\vec{\hat{a}} = \vec{\hat{c}}$.
By changing from Fermionic operators to Majorana operators is possible, and convenient, to define de Fermionic covariance matrix which as we mentioned before, fully characterise Gaussians states[3].

2.3 Fermionic Covariance matrix

A system of N fermion modes, described by a set of creation and annihilation operators $\hat{a}^\dagger, \hat{a}$ and satisfying the canonical anti-commutations relations in (2.2), is Gaussian if for such system any state ρ can be written as [37]

$$\rho = \bigotimes_{k=1}^{N} \tilde{\rho}_k, \quad \tilde{\rho}_k = \frac{1}{2}\left(1 - \lambda_i \left[\tilde{a}_i^\dagger, \tilde{a}_i\right]\right), \tag{2.7}$$

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

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

with $\hat{c} = (\hat{c}_1, \hat{c}_2, \ldots, \hat{c}_{2N})$, the vector of Majorana operators (2.4), Z a normalization constant and G real anti-symmetric $2N \times 2N$ matrix.
Since G is a skew-symmetric matrix, it can always be brought to the block diagonal form

$$OGO^T = \bigoplus_{i=1}^{N} \begin{pmatrix} 0 & -\beta_j \\ \beta_j & 0 \end{pmatrix} \quad \text{with} \quad O \in \mathrm{SO(2N)}, \tag{2.9}$$

[2]By inspection of (2.5) we see that any linear transformation of the form $\tilde{\gamma}_\alpha = O_{\alpha\beta}\gamma_\beta$ where $O \in SO(2N)$, the special orthogonal group in $2N$ dimensions

[3]In comparison to its boson counterpart the fermion Gaussian states have the property that correlation functions for the creation/annihilation operators are completely determined by the two-point functions according to Wick's theorem [36], and moreover, since this property is extensible to correlation function pertaining to a reduced subset of the modes, it follows that any partial (reduced) density matrix obtained from ρ remains Gaussian.

by a special orthogonal matrix $O \in SO(2N)$ where the β_j are called the Williamson eigenvalues of the matrix G. From equation (2.8) it is clear that Gaussian states have an interpretation as thermal (Gibbs) states corresponding to a Hamiltonian of the form

$$\hat{H} = \frac{i}{4}\hat{c}^T G \hat{c} = \frac{i}{4}\sum_{k>l} G_{kl}\left[\hat{c}_k, \hat{c}_l\right]. \tag{2.10}$$

and (2.9) shows that every Gaussian state has a normal mode decomposition in terms of N single mode "thermal states" of the form (2.7) ($\sim \exp(-\beta \hat{a}^\dagger \hat{a})$) [38]. From this is clear, that the state can be fully determined by the expectation values of quadratic operators ($\hat{a}_i^{(\dagger)}\hat{a}_j^{(\dagger)}$ and $\hat{a}_i^\dagger \hat{a}_j$). So collecting these expectation values in a real and skew-symmetric covariance matrix Γ which is defined via

$$\Gamma_{kl} = \frac{i}{2}\operatorname{tr}\left(\rho\left[c_k, c_l\right]\right). \tag{2.11}$$

We will be able to bring this anti-symmetric matrix to its block diagonal form, via a canonical transformation.

$$\tilde{\Gamma} = O\Gamma O^T = \bigoplus_{i=1}^{M} \begin{pmatrix} 0 & \lambda_i \\ -\lambda_i & 0 \end{pmatrix}, \tag{2.12}$$

with $\lambda_i \geq 0$ the Williamson eigenvalues.
In terms of the creation/annihilation operators obtained from the transformed $\tilde{c} = O\hat{c}$, the Gaussian state ρ takes the form (2.8) with $\tilde{\Gamma}$ as its Fermionic covariance matrix. It is easy to see that the relation between G and Γ is given by $\lambda_i = \tanh(\beta_i/2)$, for $i = 1, 2, \ldots, N$ [38].
The equivalence between the special orthogonal group in $2N$ dimensions ($SO(2N)$) and the Fermionic Gaussian states, drives to an interesting property about states describing multi-particles excitations. If $|vacuum\rangle$ is the ground state of some Hamiltonian, with annihilation operators $\hat{a}_i$ in a given quasi-particle basis, then $\hat{a}_i^\dagger |vacuum\rangle = \hat{c}_{2i} |vacuum\rangle$. Meaning that any multi-particle state of this kind is obtained from some transformation (that preserves the canonical anti-commutation relations) of the ground state $|vacuum\rangle$, remains 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 as well as the extension of this property to a reduced subset of the modes, is quite important since a big part of this work is focused on the study of the Fermionic covariance matrix of the XY model.
Up to this point we have talked about some generalities about the quadratic Hamiltonians and how these can be brought to its diagonal form to be analytically solved. However, one may wonder how is this connected to some observables, if the Fermionic covariance matrix has something to do in the observables. The answer to these questions will be boarded in the next section in which we will provide the formalism needed to compute observables over anti-commuting variables as well as the differences we will have between bosons and fermions.

2.4 Grassmann Approach

Fermions are one kind of fundamental particles which can be used to represent quantum information, being one posibility to do quantum computation. From Knill, Laflamme and Milburn [39] work, we know that it is possible to provide an universal set of operations for quantum computation by just using passive linear objects. Right after this discovery, Terhal, DiVincenzo [40] and knill [41] provided a description of the computational capability for Fermionic linear optics (FLO)[4]. These results show that FLO is not a promising way to do quantum computation, since it can be efficiently simulated by classical means, on the other hand, this provides a very interesting tool to study general properties quantum channels for quantum communication.
When working with fermions one could think, weather or not is possible to use the same formalisms than the ones used when working with bosons[5], and as expected many features between them are shared. More specifically, in the case of FLO, it shares most of the features with a limited version of photon linear optics, where mostly Gaussian states appears [39][6].
It is worth mentioning that some of the results mentioned in here are quite general and can be used even beyond the scope of our work.

2.4.1 Fermionic Linear Optics

In order provide a more specific description of FLO, we start by defining N abstract Fermionic modes which can be defined by creation $\hat{a}_j$ and annihilation $\hat{a}_j^\dagger$, $j = 1, \ldots, N$, satisfying (2.2). As mentioned above, any state of the system can be generated from the vacuum state $|vacuum\rangle$ of the Fock basis, which we define as

$$|N_1, \ldots, N_N\rangle = \left(\hat{a}_1^\dagger\right)^{N_1} \cdots \left(\hat{a}_n^\dagger\right)^{N_n} |vacuum\rangle\,, \tag{2.13}$$

with $N_i \in \{0, 1\}$. We have then that any sequence $\{N_i\}$ can be understood as a Fermionic state which comes from a superposition of the Fock basis.
When working with FLO, quadratic Hamiltonians govern the dynamics, in which terms associated with individual energy modes, tunneling and bulk dynamics are included. The structure of these Hamiltonians is what allow us conveniently work with operators which generate the *Clifford* algebra ($\mathcal{C}_{2N}$), described in (2.5). So we will have that any arbitrary

[4]Whenever we are referring to this term, it basically has in consideration a system consisting of non-interacting electrons in a controllable external potential and a detector that measures projectively occupation numbers of single-electron modes.

[5]By this we mean that it could be also written in its correspond operators of annihilation and creation $\hat{a}^\dagger$,$\hat{a}$ [42]

[6]the main reason for this is that a set of states that can be achieved by FLO operations starting from the Fock vacuum is a set of Fermionic Gaussian states, just as the case of its bosonic counterpart.

operator $X \in \mathcal{C}_{2N}$ will be written as a polynomial in the operators $\{\hat{c}_a\}$ as

$$X = \alpha \hat{I} + \sum_{p=1}^{2n} \sum_{1 \leq a_1 < \cdots < a_p \leq 2n} \alpha_{a_1,\ldots,a_p} \hat{c}_{a_1} \cdots \hat{c}_{a_p} \tag{2.14}$$

with $\alpha = 2^{-n} \operatorname{Tr}(X)$. Particularly for the Hamiltonian operator we have that

$$H = \frac{i}{4} \sum_{a,b=1}^{2N} H_{ab} \hat{c}_a \hat{c}_b, \tag{2.15}$$

with $\{H_{ab}\}$ an anti-symmetric $2N \times 2N$ matrix.
We may think that when working with fermions could be the same as working with bosons by changing commutators with anti-commutators and changing to a fine algebra. However, this is not the case, as we will show in the next section observables have to be computed by using the corresponding calculus for fermionic systems.

2.4.2 Grassmann Calculus

There are a lot of results from the earlier seventies which describe techniques to study systems with infinite number of fermionic modes modes [43], nonetheless, the formalism we are going to present here is a specific case where a finite number of modes is considered, and consist in a customization of the Lagrangian representation for infinite number of Fermionic modes, presented in [44]. Consider $\theta_1, \ldots, \theta_N$ span a N-dimensional complex linear space $\mathbb{C}^N$. From the anti-commutation relation we get that for $\theta_1, \ldots, \theta_N$,

$$\theta_a^2 = 0 \quad \text{and} \quad \theta_a \theta_b + \theta_b \theta_a = 0, \tag{2.16}$$

so it is clear that the most general function we can built over the Grassmann algebra with complex coefficients $\mathcal{G}_N$ is a polynomial of θ's

$$f(\theta) = \alpha + \sum_{p=1}^{n} \sum_{1 \leq a_1 < \cdots < a_p \leq N} \alpha_{a_1,\ldots,a_p} \theta_{a_1} \cdots \theta_{a_p}, \tag{2.17}$$

where the coefficients α_* are complex numbers. A polynomial $f(\theta)$ will be called *even* if it involves only even powers of θ, nonetheless something to stress is that even elements constitute the center of the Grassmann algebra.
It is possible to differentiate functions of Grassmann variables. A partial derivative over θ_a is a linear operator from $\mathcal{G}_N$ to $\mathcal{G}_N$, which is defined by

$$\frac{\partial}{\partial \theta_a} 1 = 0, \quad \frac{\partial}{\partial \theta_a} \theta_b = \delta_{ab}, \tag{2.18}$$

and Liebniz's rule

$$\frac{\partial}{\partial \theta_a} \left(\theta_b f(\theta) \right) = \delta_{ab} f(\theta) - \theta_b \frac{\partial}{\partial \theta_a} f(\theta). \tag{2.19}$$

The equality (2.19) implies that

$$\left\{\frac{\partial}{\partial\theta_a},\frac{\partial}{\partial\theta_b}\right\}=0 \tag{2.20}$$

Since a derivative $\frac{\partial}{\partial\theta_a}f(\theta)$ does not depend upon variable θ_a, it is sometimes convenient to think about differentiation as a linear operator mapping $\mathcal{G}_n$ into $\mathcal{G}_{n-1}$. Such operator is called integration and is denoted as [45] [7]

$$\int d\theta_a \equiv \frac{\partial}{\partial\theta_a} : \mathcal{G}_n \to \mathcal{G}_{n-1}. \tag{2.21}$$

From equation (2.20) we have that

$$\int \mathrm{D}\theta \frac{\partial}{\partial\theta_a} f(\theta) = 0, \tag{2.22}$$

this equation combined with (2.19) give us the anti-commuting version of integration by parts.
With these properties we show the main equation that will help us calculate what is need for the purpose of this work.
For a vectors of Grassmann variables $\vec{\theta}$, $\vec{\eta}$ and complex anti-symmetric matrix M we have

$$\int \mathrm{D}\theta \exp\left(\frac{i}{2}\theta^T M\theta\right) = i^n \operatorname{Pf}(M), \tag{2.23}$$

and

$$\int \mathrm{D}\theta \exp\left(\eta^T\theta + \frac{i}{2}\theta^T M\theta\right) = i^n \operatorname{Pf}(M)\cdot\exp\left(-\frac{i}{2}\eta^T M^{-1}\eta\right). \tag{2.24}$$

In these formulas $\operatorname{Pf}(N)$ is the Pfaffian of a complex antisymmetric matrix N defined as anti-symmetrized product $\mathcal{A}(N_{1,2}N_{3,4}\cdots N_{2n-1,2n})$ that is

$$\operatorname{Pf}(N) = \frac{1}{2^n n!}\sum_{\sigma\in S_{2n}} \operatorname{sgn}(\sigma) N_{\sigma_1,\sigma_2}\cdots N_{\sigma_{2n-1}\sigma_{2n}} \tag{2.25}$$

2.4.3 Gaussian States

For this part we are going to introduce in a formal way the states we have been talking about in at the beginning of this chapter. Informally we can say that any Gaussian operator can be

[7] Here we will use a compact notation

$$\int \mathrm{D}\theta \equiv \int d\theta_n \cdots \int d\theta_2 \int d\theta_1,$$

and the order is chosen such that

$$\int D\theta\ \theta_1\cdots\theta_n = 1$$

represented as an exponent of another operator which is quadratic in creation/annihilation operators. When considering an operator $X \in \mathcal{C}_{2N}$ It is natural to assign a polynomial $\omega(X,\theta) \in \mathcal{G}_{2N}$ of $2N$ Grassmann variables defined by.

$$\omega\left(\hat{c}_p\hat{c}_q\cdots\hat{c}_r,\theta\right) = \theta_p\theta_q\cdots\theta_r, \quad \omega(\hat{\mathbb{I}},\theta) = 1, \tag{2.26}$$

where this definition comes from the general functions in the Grassmann algebra described in (2.17).
To illustrate what how this map can be done, consider as an example the projector $\hat{a}_1\hat{a}_1^\dagger$, which its action is nothing but to project into a state where the first mode is empy. For this precise case we have that the map to Grassmann variables will look as

$$\omega(\hat{a}_1\hat{a}_1^\dagger,\theta) = \frac{1}{2}\left(\mathbb{I} + i\theta_1\theta_2\right) = \frac{1}{2}\exp\left(i\theta_1\theta_2\right), \tag{2.27}$$

where all exponents were defined by a Taylor series.

For any two operators $X, Y \in \mathcal{C}_{2N}$ one can compute the trace by just using a simple formula,

$$\mathrm{Tr}(XY) = (-2)^N \int D\theta D\mu e^{\theta^T\mu}\omega(X,\theta)\omega(Y,\mu), \tag{2.28}$$

which can be directly verified. It is also easy to see that canonical transformations of an operator $X \in \mathcal{C}_{2N}$ are equivalent to an orthogonal change of basis in the space of the Grassmann variables, this is

$$\omega\left(VXV^\dagger,\theta\right) = \omega(X,\eta), \quad \eta_a = \sum_{b=1}^{2n} R_{ab}\theta_b. \tag{2.29}$$

So we will define a Gaussian states of N Fermionic modes as via its density operator $\rho \in \mathcal{C}_{2N}$. We say that a state is Gaussian in $\mathcal{C}_{2N}$ *iff* its Grassmann representation is as well Gaussian

$$\omega(\rho,\theta) = \frac{1}{2^n}\exp\left(\frac{i}{2}\theta^T M\theta\right), \tag{2.30}$$

for some $2N \times 2N$ antisymmetric matrix M. The matrix M is defined as

$$M_{ab} = \frac{i}{2}\mathrm{Tr}\left(\rho\left[\hat{c}_a,\hat{c}_b\right]\right) = \begin{cases} \mathrm{Tr}\left(\rho i\hat{c}_a\hat{c}_b\right) \text{ for } a \neq b \\ 0 \text{ for } a = b \end{cases}, \tag{2.31}$$

which is nothing but the covariance Fermionic matrix defined in (2.11), meaning that it can be brought to its Williamson form to get its eigenvalues.

Therefore the connection between the Grassmann formalism for Gaussian states and the Fermionic covariance matrix described at the beginning of the chapter, is easily understood

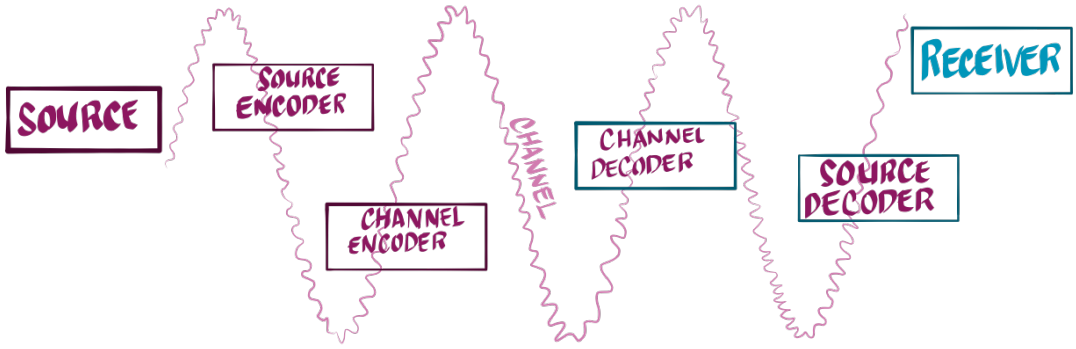

Figure 2.1: Representation of the scheme of communication. In the image the noise in the channel is represented by the noise wavy connection between the parts in the communication.

by simply assigning the correspondent covariant matrix to the states we work with. Nonetheless, one could be wondering how is this formalism connect to the error correcting code theory. In the next section we will provide a little historical background of how this theory was done and how one can think states of fermions as binary codes. More specifically we will provide the necessary background to link the theory of correcting errors with Fermionic states which are constrained over a shell of energy in the Hilbert Space.

2.5 Error correcting Code Theory

In this section we will provide some concepts and definitions in order to understand how the Coding theory can be linked to Fermionic states.
In 1948, Claude Shannon presents his extraordinary work named "A Mathematical Theory of Communication" [24] in which he provided a precise measure of the information content of a random variable in terms of its entropy. His work is divided in two parts the noiseless coding theorem and the noisy channel theorem. For the purpose of our needs we are going to focus only in the second part of his work, which states that a reliable communication is possible if we use schemes such that its rate is less than the capacity of the channel. Even though he never provided an idea of how this schemes could be found, his work is considered one of the most relevant discovery of the century.

Here we will consider codes in communication scenario, as the one showed in figure 2.1,

meaning that there will be a sender who wants to send k message symbols over a noisy channel and there will be a receiver who has to correct possible errors over the sent code to fully interpret it. The sender will first encodes the k message symbols into n symbols. The receiver then tries to recover the original k message symbols. thus, encoding is the process of adding redundancy and decoding is the process of removing errors and the communication can only be done over the channel. 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.

Definition 2.5.1 ***Code:*** *A code of block C length n over an alphabet Σ is a subset of Σ^n. If $|\Sigma| = q$, we say that C is a q-ary code.*

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

Definition 2.5.2 ***Dimension of a code:*** *Given a code $C \subset \Sigma^n$, its dimension is given by*

$$k \stackrel{\text{def}}{=} \log_q |C|, \tag{2.32}$$

An interesting fact about defining the dimension of the code in this way is that implicitly it is telling us that when working with codes exponential growth will be always taken into account.
We have to provide here a way to measure the amount of redundancy in a given message.

Definition 2.5.3 ***Rate of a code:*** *The rate of a code with dimension k and block length n is given by*

$$R \stackrel{def}{=} \frac{k}{n}, \tag{2.33}$$

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

$$R(C) = \frac{\log |C|}{n \log |\Sigma|}. \tag{2.34}$$

Definition 2.5.4 ***Hamming distance:*** *The Hamming distance between two strings x and y of the same length over a finite alphabet Σ, denoted $\Delta(x, y)$, is defined as the number of positions at which the two strings differ, i.e, $\Delta(x, y) = |\{i | x_i \neq y_i\}|$. The fractional Hamming distance or relative distance between $x, y \in \Sigma^n$ is given by $\delta(x, y) = \frac{\Delta(x,y)}{n}$.*

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

Definition 2.5.5 ***Hamming weight:*** *The Hamming weight of a string x over alphabet Σ is defined as the number of non-zero symbols in the string. More formally, the Hamming weight of a string $\mathcal{W}(x) = |\{i|x_i \neq 0\}|$. Note that $\mathcal{W}(x-y) = \Delta(x,y)$.*

Given a string $x \in \Sigma^N$, the Hamming ball of radius r around x is the set $\{y \in \Delta^n | \Delta(x,y) \leq r\}$.
The minimum distance, or simply distance, of a code C, denoted $\Delta(C)$, is defined to be the minimum Hamming distance between two distinct code words of C. That is

Definition 2.5.6 ***Minimum distance:*** *The minimum distance, or simply distance, of a code C, denoted $\Delta(C)$, is defined to be the minimum Hamming distance between two distinct code words of C. That is*

$$\Delta(C) = \min_{\substack{c_1,c_2 \in C \\ c_1 \neq c_2}} \Delta(c_1, c_2). \tag{2.35}$$

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

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

Definition 2.5.7 ***Notation:*** *A q-ary code of block length N and dimension k will be referred to as an $[N,k]_q$ code. Further, if the code has minimum distance d, it will be referred to as an $[N,k,d]_q$ code. When the alphabet size q is clear from the context, or not very relevant to the discussion, we omit the subscript.*

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

Definition 2.5.8 ***Family Codes:*** *Let $q \geq 2$. let $\{n_i\}_{i\geq 1}$ be and increasing sequence of block lengths and suppose there exists sequences $\{k_i\}_{i\geq 1}$ and $\{d_i\}_{i\geq 1}$ such that for all $i \geq 1$ there exist an $[n_i,k_i,d_i]_q$ code C_i. then the sequence $C = \{C_i\}_{i\geq 1}$ is a family of codes. The rate of C is defined as*

$$R(C) = \lim_{i\to\infty} \left\{ \frac{k_i}{n_i} \right\}, \tag{2.36}$$

and the relative distance of C is defined as

$$\delta(C) = \lim_{i\to\infty} \left\{ \frac{d_i}{n_i} \right\} \tag{2.37}$$

Chapter 3

The One Dimensional XY Model.

Bibliography

[1] Ludwig Boltzmann. *Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie: (vorgelegt in der Sitzung am 8. Februar 1866)*. Staatsdruckerei. Google-Books-ID: 8Hc5AAAAcAAJ. iv, 3

[2] 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

[3] Sheldon Goldstein, Joel L. Lebowitz, Roderich Tumulka, and Nino Zanghì. Canonical typicality. 96(5):050403. 1

[4] Sandu Popescu, Anthony J. Short, and Andreas Winter. Entanglement and the foundations of statistical mechanics. 2(11):754–758. 1, 12, 14

[5] L. D. Landau and E. M. Lifshitz. *Statistical Physics: Volume 5*. Butterworth-Heinemann, 3 edition edition. 1

[6] Sandu Popescu, Anthony J. Short, and Andreas Winter. The foundations of statistical mechanics from entanglement: Individual states vs. averages. 1, 12, 14

[7] Noah Linden, Sandu Popescu, Anthony J. Short, and Andreas Winter. Quantum mechanical evolution towards thermal equilibrium. 79(6):061103. 2, 15, 16, 17, 18, 19

[8] J. Uffink. *Handbook for Philsophy of Physics 924-1074*. Elsevier, Amsterdam. 2, 5

[9] Ludwig Boltzmann. Zur priorität der auffindung der beziehung zwischen dem zweiten hauptsatze der mechanischen wärmetheorie und dem principe der kleinsten wirkung. 219(6):211–230. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.18712190606. 3

[10] Ludwig Eduard Boltzmann. *Einige allgemeine sätze über Wärmegleichgewicht*. K. Akad. der Wissensch. 3

[11] Josiah Willard Gibbs. *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*. C. Scribner's Sons. 3

[12] Paul 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]. 5

[13] César R. de Oliveira and Thiago Werlang. Ergodic hypothesis in classical statistical mechanics. 29(2):189–201. Publisher: Sociedade Brasileira de Física. 5

[14] Larry Engelhardt, Marie Lopez del Puerto, and Norman Chonacky. Simple and synergistic ways to understand the boltzmann distribution function. 83(9):787–793. Publisher: American Association of Physics Teachers. 5

[15] Giovanni Gallavotti. *Statistical Mechanics: A Short Treatise*. Theoretical and Mathematical Physics. Springer-Verlag. 5

[16] Sheldon Goldstein. Boltzmann's approach to statistical mechanics. In Jean Bricmont, Giancarlo Ghirardi, Detlef Dürr, Francesco Petruccione, Maria Carla Galavotti, and Nino Zanghi, editors, *Chance in Physics: Foundations and Perspectives*, Lecture Notes in Physics, pages 39–54. Springer. 5

[17] Kerson Huang. *Statistical Mechanics, 2nd Edition*. John Wiley & Sons, 2 edition edition. 6, 7

[18] E. Schrödinger. Discussion of probability relations between separated systems. 31(4):555–563. Publisher: Cambridge University Press. 9

[19] Artur Ekert, Patrick Hayden, and Hitoshi Inamori. Basic concepts in quantum computation. version: 1. 9

[20] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press. Google-Books-ID: 65FqEKQOfP8C. 9

[21] 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. 9

[22] John Von Neumann. *Mathematical foundations of quantum mechanics*. Princeton, N.J. : Princeton University Press. 10

[23] David J. C. MacKay. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 1 edition edition. 11

[24] C. E. Shannon. A mathematical theory of communication. 27(3):379–423. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1538-7305.1948.tb01338.x. 11, 30

[25] Martin B. Plenio and S. Virmani. An introduction to entanglement measures. 11

[26] J. Gemmer and G. Mahler. Distribution of local entropy in the hilbert space of bi-partite quantum systems: origin of jaynes' principle. 31(2):249–257. 11

[27] Don N. Page. Average entropy of a subsystem. 71(9):1291–1294. 11

[28] Paolo Facchi and Giancarlo Garnero. Quantum thermodynamics and canonical typicality. 14(8):1740001. 14

[29] C.A. Fuchs and J. van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. 45(4):1216–1227. Conference Name: IEEE Transactions on Information Theory. 18

[30] Giancarlo Camilo, Gabriel T. Landi, and Sebas Eliëns. On the strong subadditivity of the r\'enyi entropies for bosonic and fermionic gaussian states. 99(4):045155. 19

[31] Alonso Botero and Benni Reznik. BCS-like modewise entanglement of fermion gaussian states. 331(1):39 – 44. 22, 23

[32] Eduardo Fradkin. *Field Theories Of Condensed Matter Systems Vol#82 On-demand Printing Of #52253*. CRC Press, 1 edition edition. 23

[33] F. A. Berezin. *The method of second quantization*. Number v. 24 in Pure and applied physics. Academic Press. 23

[34] N. N. Bogoljubov. On a new method in the theory of superconductivity. 7(6):794–805. 23

[35] Martin Leijnse and Karsten Flensberg. Introduction to topological superconductivity and majorana fermions. 27(12):124003. 23

[36] B. Westwanski and A. Pawlikowski. On a general statistical wick theorem. 43(2):201–202. 24

[37] Siew-Ann Cheong and Christopher L. Henley. Many-body density matrices for free fermions. 24

[38] Christina V. Kraus, Michael M. Wolf, J. Ignacio Cirac, and Geza Giedke. Pairing in fermionic systems: A quantum information perspective. 79(1):012306. 25

[39] E. Knill, R. Laflamme, and G. J. Milburn. A scheme for efficient quantum computation with linear optics. 409(6816):46–52. Number: 6816 Publisher: Nature Publishing Group. 26

[40] Barbara M. Terhal and David P. DiVincenzo. Classical simulation of noninteracting-fermion quantum circuits. 65(3):032325. 26

[41] E. Knill. Fermionic linear optics and matchgates. 26

[42] 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. 26

[43] D. E. Soper. Construction of the functional-integral representation for fermion green's functions. 18:12. Institution: Institute of Theoretical Science, University of Oregon, Eugene, Oregon 97403. 27

[44] Sergey Bravyi. Classical capacity of fermionic product channels. version: 1. 27

[45] Sergey Bravyi. Lagrangian representation for fermionic linear optics. 28

[46] Luca Guido Molinari. Notes on wick's theorem in many-body theory.

[47] Ingo Peschel. Calculation of reduced density matrices from correlation functions. 36(14):L205–L208.

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

[49] Philippe Francesco, Pierre Mathieu, and David Sénéchal. *Conformal Field Theory*. Graduate Texts in Contemporary Physics. Springer-Verlag.