

Università degli Studi di Genova

Facoltà di Scienze Matematiche, Fisiche e Naturali
Anno Accademico 2007/2008

Tesi di Dottorato in Fisica

Typicality in Quantum Statistical Physics

Candidato:
Christian Mastrodonato

Relatore:
Prof. Pierantonio Zanghì
(Università di Genova)

Correlatore:
Prof. Roderick Tumulka
(Rutgers University)

A Francesca

Introduction

The principle that an event with very high probability is *morally certain* is known since the beginning of the history of probability. Bernoulli was the first to mention it in his “Ars Conjectandi” published in 1713 [1, 24]. In this book Bernoulli also proved his celebrated theorem: in a sufficiently long sequence of independent trials of an event, there is a very high probability that the frequency with which the event happens will be close to its probability. Bernoulli explained that we can treat the very high probability as *moral certainty* and so use the frequency of the event as an estimate of its probability. This consideration gives to the concept of *moral certainty* a great importance, since it actually links the probability theory to the reality.

During the nineteenth century Cournot [7, 24] put the concept of *moral certainty* in the context of geometric probability that was born in that period. Recalling the weak law of large numbers in this context: *it is physically impossible for the frequency of an event in a long sequence of trials to differ substantially from the event’s probability* ([7], pages 57 and 106).

From this perspective, in the second half of the nineteenth century, the principle that an event with a vanishingly small probability will not happen took on a real role in physics, most saliently in Ludwig Boltzmann’s statistical understanding of the second law of thermodynamics [5]. As Boltzmann explained in the 1870s, dissipative processes are irreversible because the probability of a state with entropy far from the maximum is vanishingly small. Therefore a macroscopic system tends to increase its entropy with time approaching in the end to equilibrium, which is the macroscopic configuration with maximum entropy associated. To denote this kind of behavior Boltzmann introduced the concept of “typicality”, which can be therefore regarded as the physical counterpart of the concept of “moral certainty”.

From its introduction by Boltzmann, the concept of typicality has played a fundamental role during the development of classical statistical mechanics. The aim of this thesis is to present some possible applications of this concept for the quantum extension of statistical mechanics. We shall see, indeed, that it is possible to define some “typical” properties also for macroscopic quantum systems.

This work starts from presenting a theorem on the asymptotic behavior of unitary matrices [17]. This theorem states that the joint distribution of the upper left (or, in fact, any) $k \times k$ entries of a random unitary $n \times n$ matrix, multiplied for a normalization factor \sqrt{n} , in the limit $n \rightarrow \infty$ converges in probability (or, in other words, *typically*) to a $k \times k$ matrix whose entries are independent complex gaussian random variables. This result has been first established by Petz and Réffy in [19], who showed the convergence of the joint distribution of the eigenvalues of a $k \times k$ submatrix of a unitary matrix to the corresponding distribution for a $k \times k$ Gaussian matrix. Here it is presented an alternative elementary proof based on the geometric properties of Gaussian random variables.

This theorem is particularly interesting since, in general, unitary matrices play an important role in quantum statistical mechanics, as we shall see in the rest of this work. In particular, this theorem is a crucial ingredient to provide a rigorous proof (in this work only sketched) for a statement proven (but not rigorously) in [11], which states that, and formulates the precise sense in which, wave functions indeed have *typically* a thermal equilibrium distribution, and this distribution is a GAP measure [12].

The family of GAP measures is a family of probability measures on Hilbert spaces, introduced in [11] as a candidate for the thermal equilibrium distribution of the wave function of a quantum system, in a roughly similar way as the canonical Gibbs distribution on the phase space of a classical system is the thermal equilibrium distribution of the phase point. There is one GAP measure for every density matrix ρ , denoted $GAP(\rho)$. The measures relevant to thermal equilibrium are those associated with canonical density matrices.

In more detail, the result sketched here asserts that, for a system coupled to a heat bath and a *typical* wave function of system and bath in a given small energy range, the conditional wave function of the system has

a distribution near the GAP measure associated with the canonical density matrix of the system; where the conditional wave function can be regarded as the formalization of the collapsed wave function. Finally, to complete the framework, it will be studied the existence, uniqueness and continuity of the GAP measure, also in infinite-dimensional spaces.

In the final part of this thesis it is treated the problem of the typical approach to equilibrium of a macroscopic quantum system [13]. This problem is studied starting from the well know analysis of von Neumann [28], who defined the macroscopic observable for a quantum system as a set of commuting operators, and the macrostates as the eigenspaces associated to this set. In terms of this framework it is posed the problem of convergence to equilibrium, defining a sufficient condition on the stationary states (and, therefore, on the Hamiltonian) of the system which guarantees the approach to equilibrium regardless of the initial state. It will be show, in fact, that, to have convergence to equilibrium, it is sufficient that all the stationary states of the system have probability near one to be into the equilibrium macrostate.

This condition is studied in the case of a simple system composed by two coupled subsystem, and, within this case, it is defined in terms of the unitary matrix which connects the exact Hamiltonian eigenvectors basis with the “uncoupled” Hamiltonian eigenvectors basis. By neglecting the interaction potential on macroscopic scale, the subsystems Hamiltonian can be regarded as the set of macroscopic observables of the system, and the “uncoupled” Hamiltonian eigenvectors forms an eigenbasis for this set. Therefore, this setup can be generalized to every macroscopic system, by considering the unitary matrix which connects the exact Hamiltonian eigenvectors basis with the eigenvectors basis of the macroscopic observables. Within this generalized framework it will be shown that, under suitable condition, for *typical* Hamiltonians a macroscopic quantum system spends the overwhelming majority of time instants in equilibrium.

Summary

Chapter 1 will be concerned with giving an alternative, elementary proof of a theorem first established by Petz and Réffy in [19], concerning the joint distribution of the upper left $k \times k$ entries of a random unitary $n \times n$ matrix

in the limit $n \rightarrow \infty$. This proof it will be useful in Chapter 2 and furthermore it will show a “typical” behavior for unitary matrices.

Chapter 2 will be based on the characteristic of the GAP measure. The main goal of this Chapter is to provide a partial proof for a statement sketched in [11], which states that wave functions have a thermal equilibrium distribution, and this distribution is a GAP measure.

Chapter 3 will deal with the approach to equilibrium for macroscopic quantum systems. We shall show that such a system reaches equilibrium for a “typical” Hamiltonian, regardless of the initial state.

Contents

Introduction	ii
1 Asymptotics for Large Haar-Distributed Unitary Matrices	1
1.1 Statement of the Theorem	2
1.2 First Part of the Proof: Construction of U_{ij}	3
1.3 Second Part of the Proof: Probable Geometry	4
1.4 Third Part of the Proof: Estimates	5
1.5 Final Remark	8
2 Measure of Typicality in Equilibrium Quantum Statistical Mechanics	9
2.1 An introduction to the problem	9
2.2 Review of Relevant Definitions	11
2.2.1 Definition of the GAP Measure	12
2.2.2 Conditional Wave Function	13
2.3 Typicality of the GAP Measure for a Subsystem of a Large System	14
2.3.1 Setup	14
2.3.2 Typicality Theorem	15
2.3.3 Proof of Lemma 2.3.2	17
2.3.4 Random Basis	20
2.4 The GAP Measure as the Thermal Equilibrium Distribution .	21
2.5 Further Results on GAP Measure	23
2.5.1 Continuity in Finite Dimension	23
2.5.2 The GAP Measure in Infinite Dimension	26

2.6	Final Remark	29
3	Typical Non-Equilibrium Behavior of Macroscopic Quantum Systems	30
3.1	The Problem of Convergence to Equilibrium	30
3.2	Setting the Stage of Convergence to Equilibrium in a Concrete Case	33
3.3	A General Framework for Convergence to Equilibrium	36
	3.3.1 Proof of Lemma 3.3.1	38
3.4	The Main Theorem on Convergence to Equilibrium	42
3.5	A Simple Counter-Example to Convergence to Equilibrium	43
3.6	Comparing different approaches to Convergence to Equilibrium	45
	3.6.1 An historical overview	45
	3.6.2 Summary of our results	48
4	Conclusions	49

Chapter 1

Asymptotics for Large Haar-Distributed Unitary Matrices

In this chapter it is shown an alternative, elementary proof of a theorem first established by Petz and Réffy in [19], concerning the joint distribution of the upper left $k \times k$ entries of a random unitary $n \times n$ matrix in the limit $n \rightarrow \infty$ and formulated as Theorem 1 below. This work has been done in joint work with Roderich Tumulka of Rutgers University [17]. The theorem treated in this chapter is of particular interest in quantum statistical mechanics, where one often studies the behavior of a small system (corresponding to dimension k) coupled to a heat bath—a much larger system corresponding to dimension n . Specifically, Theorem 1 can be used for studying the distribution of the conditional wave function of a system coupled to a heat bath in the relevant limit (in which the size of the heat bath tends to infinity). As it will be shown in the Chapter 2, this distribution should typically converge, as a consequence of Theorem 1, to the so-called “GAP” measure [11], which can thus be regarded as the thermal equilibrium distribution of the conditional wave function.

1.1 Statement of the Theorem

We fix some notation and terminology. Let \mathbb{P} denote probability and \mathbb{E} expectation, $U(n)$ the group of unitary $n \times n$ matrices, and $\text{Haar}(U(n))$ the (normalized) *Haar measure* on this group, representing the “uniform” probability distribution over $U(n)$. We write (a_{ij}) for the matrix with entries a_{ij} . The relevant notion of convergence of probability distributions is *weak convergence*, also known as “convergence in distribution” of the random variables [2, Sec. 25]. By a complex Gaussian random variable G with mean 0 and variance σ^2 we mean $G = X + iY$, where X and Y are independent real Gaussian random variables with means $\mathbb{E}X = 0$ and $\mathbb{E}Y = 0$ and variances $\mathbb{E}X^2 = \sigma^2/2$ and $\mathbb{E}Y^2 = \sigma^2/2$.

Theorem 1.1.1. *If (U_{ij}) is $\text{Haar}(U(n))$ distributed, then the upper left (or, in fact, any) $k \times k$ submatrix, multiplied by a normalization factor \sqrt{n} , converges in distribution, as $n \rightarrow \infty$, to a random $k \times k$ matrix (G_{ij}) whose entries G_{ij} are independent complex Gaussian random variables with mean 0 and variance $\mathbb{E}|G_{ij}|^2 = 1$.*

To understand the factor \sqrt{n} , note that a column of a unitary $n \times n$ matrix is a unit vector, and thus a single entry should be of order $1/\sqrt{n}$. A random $k \times k$ matrix such as (G_{ij}) , consisting of independent complex Gaussian variables with mean 0 and variance 1, is also called “ \sqrt{k} times a standard non-selfadjoint Gaussian matrix.”

Theorem 1 is a generalization of the familiar fact that the first k entries of a random unit vector in \mathbb{R}^n (with uniform probability distribution over the unit sphere), multiplied by a normalization factor \sqrt{n} , converge in distribution to a vector whose k entries are independent real Gaussian random variables with mean 0 and variance 1.¹ This fact (with \mathbb{R}^n replaced by \mathbb{C}^n)

¹As a physical interpretation of this fact, consider N classical particles without interaction in a box $\Lambda \subseteq \mathbb{R}^3$; a given energy corresponds to a surface in phase space $\Lambda^N \times \mathbb{R}^{3N}$ given by $\Lambda^N \times \mathcal{S}$, where \mathcal{S} is the sphere of appropriate radius $\propto \sqrt{N}$ in momentum space \mathbb{R}^{3N} ; assuming a random phase point with micro-canonical distribution (i.e., uniform on $\Lambda^N \times \mathcal{S}$), the marginal distribution of the momentum of the first particle is, in the limit $N \rightarrow \infty$, Gaussian. This fact is part of the justification of Maxwell’s law of the Gaussian distribution of momenta.

is contained in Theorem 1 by specializing to just the first columns of the matrices (U_{ij}) and (G_{ij}) .

The proof of Petz and Réffy is based on the convergence of the joint distribution of the eigenvalues of a $k \times k$ submatrix of an unitary matrix to the corresponding distribution for a $k \times k$ Gaussian matrix. Our proof, in contrast, is based on the geometric properties of Gaussian random matrices. While it involves some more cumbersome estimates, it employs only elementary methods.

1.2 First Part of the Proof: Construction of U_{ij}

We write M_j for the j -th column of any $n \times n$ matrix (M_{ij}) and

$$\langle M_j | M_\ell \rangle = \sum_{i=1}^n M_{ij}^* M_{i\ell}, \quad \|M_j\|^2 = \sum_{i=1}^n |M_{ij}|^2. \quad (1.1)$$

For $i, j = 1, \dots, n$ let G_{ij} be i.i.d. complex Gaussian random variables with mean 0 and variance 1. To the n columns of the matrix (G_{ij}) apply the Gram–Schmidt orthonormalization procedure, and call the resulting matrix (U_{ij}) . That is,

$$U_{ij} = \frac{G_{ij} - \Delta_{ij}}{\|G_j - \Delta_j\|} \quad (1.2)$$

with

$$\Delta_{ij} = \sum_{\ell=1}^{j-1} \langle G_j | U_\ell \rangle U_{i\ell}. \quad (1.3)$$

The procedure fails if the columns of (G_{ij}) are linearly dependent, but this event has probability 0. Then, as also remarked in [19], (U_{ij}) is $\text{Haar}(U(n))$ distributed because its first column is uniformly distributed over the unit sphere in \mathbb{C}^n , the distribution of the second column conditional on the first column is uniform over the unit sphere in the orthogonal complement of the first column, ..., the distribution of the $j + 1$ -st column conditional on the first j columns is uniform over the unit sphere in the orthogonal complement of the first j columns—and this is exactly the Haar measure.

Our method of proof is to show that $|\sqrt{n}U_{ij} - G_{ij}|$ is in fact small if n is large. More precisely, we show that for every $\varepsilon > 0$,

$$\mathbb{P}\left(\sum_{i,j=1}^k |\sqrt{n}U_{ij} - G_{ij}| < \varepsilon\right) \rightarrow 1 \quad (1.4)$$

as $n \rightarrow \infty$. This is called convergence in probability, and to obtain the claim of the theorem we use the known fact [2, Theorem 25.2, p. 284] that convergence in probability implies weak convergence (of the joint distribution of $\sqrt{n}U_{ij}$ for $i, j = 1, \dots, k$), provided that all random variables are defined on the same probability space. Here, we can assume that for all $i, j \in \mathbb{N}$, the G_{ij} are defined on the same probability space.

1.3 Second Part of the Proof: Probable Geometry

The proof of (1.4) is based on the following observations:

- Any two different columns of (G_{ij}) tend to be nearly orthogonal.
- Every column of (G_{ij}) tends to have norm close to \sqrt{n} .
- The size of every single entry, $|G_{ij}|$, stays bounded as n grows.

These statements are to be understood in the sense that they are fulfilled with high probability for sufficiently large n . We now make them precise.

Fix a (small) $\delta > 0$. Choose $R > 0$ so large that

$$\mathbb{P}(|G_{ij}| < R) \geq 1 - \delta. \quad (1.5)$$

Define the following events corresponding to the three bullets above:

$$A_{j\ell}^n := \left\{ |\langle G_j | G_\ell \rangle| < \sqrt{\frac{n}{\delta}} \right\} \quad (1.6)$$

$$B_j^n := \left\{ \left| \frac{\|G_j\|^2}{n} - 1 \right| < \sqrt{\frac{2}{n\delta}} \right\} \quad (1.7)$$

$$C_{ij}^n := \{|G_{ij}| < R\} \quad (1.8)$$

for $i, j, \ell \leq k$. (C_{ij}^n actually does not depend on n , but never mind.) Each of these events has at least probability $1 - \delta$: $A_{j\ell}^n$ and B_j^n by Chebyshev's inequality and C_{ij}^n by (1.5). Thus, the event

$$D^n := \bigcap_{\substack{j,\ell=1 \\ j \neq \ell}}^k A_{j\ell}^n \cap \bigcap_{j=1}^k B_j^n \cap \bigcap_{i,j=1}^k C_{ij}^n \quad (1.9)$$

has at least probability $1 - 2k^2\delta$, as $2k^2$ is the number of intersecting sets.

We now show that for sufficiently large n , $D^n \subseteq E^n$, where

$$E^n := \left\{ \sum_{i,j=1}^k |\sqrt{n}U_{ij} - G_{ij}| < \varepsilon \right\} \quad (1.10)$$

is the event in the brackets of (1.4). Since δ was arbitrary, this fact implies (1.4). The remainder of the proof makes no reference to probabilities, but concerns only the inclusion $D^n \subseteq E^n$, which can be regarded as an inclusion between subsets of \mathbb{C}^{n^2} . Also $A_{j\ell}^n$, B_j^n , and C_{ij}^n will from now on be regarded as subsets of \mathbb{C}^{n^2} . (Now the upper index n in the notation C_{ij}^n becomes useful.) We thus regard G_{ij} as fixed numbers, and assume that the matrix (G_{ij}) lies in the set D^n . When we refer to "the condition $A_{j\ell}^n$ " we mean the condition that the $n \times n$ matrix (G_{im}) lies in the set $A_{j\ell}^n$.

We proceed to show, by induction over $j \in \{1, \dots, k\}$, that for sufficiently large n we have that for all $(G_{ij}) \in D^n$, and for all $i = 1, \dots, k$,

$$|\sqrt{n}U_{ij} - G_{ij}| < \frac{\varepsilon}{k^2} \quad (1.11)$$

and there are constants $C_1, \dots, C_k > 0$ such that for sufficiently large n

$$\|\sqrt{n}U_j - G_j\| < C_j. \quad (1.12)$$

From (1.11) we see that $(G_{ij}) \in E^n$, which is what we need to show. This induction is the contents of the next, and last, section.

1.4 Third Part of the Proof: Estimates

For $j = 1$, note that $U_1 = G_1/\|G_1\|$. By conditions B_1^n and C_{i1}^n ,

$$|\sqrt{n}U_{i1} - G_{i1}| = \left| \frac{\sqrt{n}}{\|G_1\|} - 1 \right| |G_{i1}| < \frac{2}{\sqrt{\delta n}} R < \frac{\varepsilon}{k^2} \quad (1.13)$$

for sufficiently large n . By condition B_1^n ,

$$\|\sqrt{n}U_1 - G_1\| = \left| \frac{\sqrt{n}}{\|G_1\|} - 1 \right| \|G_1\| < \frac{2}{\sqrt{\delta n}} 2\sqrt{n} = \frac{4}{\sqrt{\delta}} =: C_1. \quad (1.14)$$

We now collect four estimates. For $\ell < j$ and sufficiently large n we find

$$|\langle G_j | \sqrt{n}U_\ell \rangle| \leq \left| \langle G_j | \sqrt{n}U_\ell - G_\ell \rangle \right| + \left| \langle G_j | G_\ell \rangle \right| \leq \quad (1.15)$$

$$\leq \|G_j\| \|\sqrt{n}U_\ell - G_\ell\| + \sqrt{n/\delta} < 2\sqrt{n}C_\ell + \sqrt{n/\delta} =: C'_\ell \sqrt{n} \quad (1.16)$$

where we have used the Cauchy–Schwarz inequality, $A_{j\ell}^n$, B_j^n , and the induction hypothesis (1.12). As the next estimate, for $i \leq k$,

$$|\Delta_{ij}| \leq \sum_{\ell=1}^{j-1} \frac{|\langle G_j | \sqrt{n}U_\ell \rangle|}{\sqrt{n}} \frac{|\sqrt{n}U_{i\ell}|}{\sqrt{n}} \leq \quad (1.17)$$

$$\leq \sum_{\ell=1}^{j-1} C'_\ell \frac{1}{\sqrt{n}} \left(|\sqrt{n}U_{i\ell} - G_{i\ell}| + |G_{i\ell}| \right) < \quad (1.18)$$

$$< \left(\sum_{\ell=1}^{j-1} C'_\ell \right) \left(\frac{\varepsilon}{k^2} + R \right) \frac{1}{\sqrt{n}} =: \frac{C''_j}{\sqrt{n}} \quad (1.19)$$

using (1.16), the induction hypothesis (1.11), and $C_{i\ell}^n$. As the third estimate, for $j \leq k$

$$\|\Delta_j\|^2 = \sum_{i=1}^n |\Delta_{ij}|^2 \leq \frac{1}{n} \sum_{i=1}^n \sum_{\ell=1}^{j-1} |\langle G_j | \sqrt{n}U_\ell \rangle|^2 |U_{i\ell}|^2 < \quad (1.20)$$

$$< \frac{1}{n} \sum_{\ell=1}^{j-1} (C'_\ell)^2 n \|U_\ell\|^2 = \sum_{\ell=1}^{j-1} (C'_\ell)^2 =: C'''_j \quad (1.21)$$

using (1.16) and the fact that U_ℓ is a unit vector. As the last estimate,

$$\left| \frac{\sqrt{n}}{\|G_j - \Delta_j\|} - 1 \right| < \left(2\sqrt{\frac{2}{\delta}} + 2\sqrt{C'''_j} \right) \frac{1}{\sqrt{n}} =: \frac{\tilde{C}_j}{\sqrt{n}}, \quad (1.22)$$

which is easily obtained from $\|G_j\| - \|\Delta_j\| \leq \|G_j - \Delta_j\| \leq \|G_j\| + \|\Delta_j\|$ and (1.21) in the following way:

$$\frac{\sqrt{n}}{\|G_j - \Delta_j\|} - 1 \leq \frac{\sqrt{n}}{\|G_j\| - \|\Delta_j\|} - 1 < \quad (1.23)$$

$$< \frac{\sqrt{n}}{\|G_j\| - \sqrt{C_j'''}} - 1 = \frac{1}{\|G_j\|/\sqrt{n} - \sqrt{C_j'''}/n} - 1. \quad (1.24)$$

Since, using B_j^n , $\|G_j\|/\sqrt{n} > \sqrt{1 - \sqrt{2/n\delta}} > 1 - \sqrt{2/n\delta}$, and since

$$\frac{1}{1-x} < 1 + 2x \quad (1.25)$$

for sufficiently small $x > 0$, we obtain that, for sufficiently large n ,

$$\frac{\sqrt{n}}{\|G_j - \Delta_j\|} - 1 < \frac{1}{1 - \sqrt{2/n\delta} - \sqrt{C_j'''}/n} - 1 < 2\sqrt{\frac{2}{n\delta}} + 2\sqrt{\frac{C_j'''}{n}}. \quad (1.26)$$

Together with an (even narrower) lower bound obtained by similar arguments, this yields (1.22).

From these four estimates, the first induction claim (1.11) follows for $j \leq k$ because, for sufficiently large n ,

$$|\sqrt{n}U_{ij} - G_{ij}| = \left| \frac{\sqrt{n}}{\|G_j - \Delta_j\|} (G_{ij} - \Delta_{ij}) - G_{ij} \right| \leq \quad (1.27)$$

$$\leq \left| \frac{\sqrt{n}}{\|G_j - \Delta_j\|} - 1 \right| |G_{ij}| + \frac{\sqrt{n}}{\|G_j - \Delta_j\|} |\Delta_{ij}| < \quad (1.28)$$

$$< \frac{\tilde{C}_j}{\sqrt{n}} R + 2\frac{C_j''}{\sqrt{n}} < \frac{\varepsilon}{k^2}, \quad (1.29)$$

where we have used (1.22), C_{ij}^n , (1.22) with $\tilde{C}_j/\sqrt{n} < 1$, and (1.19). The second induction claim (1.12) follows from

$$\|\sqrt{n}U_j - G_j\| \leq \left| \frac{\sqrt{n}}{\|G_j - \Delta_j\|} - 1 \right| \|G_j\| + \frac{\sqrt{n}}{\|G_j - \Delta_j\|} \|\Delta_j\| < \quad (1.30)$$

$$< \frac{\tilde{C}_j}{\sqrt{n}} 2\sqrt{n} + 2\sqrt{C_j'''} = 2\tilde{C}_j + 2\sqrt{C_j'''} =: C_j. \quad (1.31)$$

This completes the proof.

1.5 Final Remark

The parenthesis in Theorem 1—the upper left (or, in fact, any) $k \times k$ submatrix—needs one further remark. In particular the meaning of “any” needs some further discussions. To select a $k \times k$ submatrix means to select k rows and k columns. This selection must be deterministic (i.e., non-random, or at least independent of the U_{ij}) but may depend on n . Indeed, if the selection depended on the U_{ij} , one could, for example, select those rows and columns where U_{ij} happens to be exceptionally close to zero, which would lead to a different asymptotic distribution. On the other hand, for a selection depending on n , Theorem 1 remains true: to see this, recall that for a compact group such as $U(n)$, the Haar measure is both left-invariant and right-invariant; as a consequence, $\text{Haar}(U(n))$ is invariant under any (non-random) permutation of either the rows or the columns, and thus all $k \times k$ submatrices have the same distribution.

Chapter 2

Measure of Typicality in Equilibrium Quantum Statistical Mechanics

The main goal of this chapter is to provide a partial proof for a statement sketched in [11], here called Theorem 2.4.1, which states that, and formulates the precise sense in which, wave functions indeed have a thermal equilibrium distribution, and this distribution is a GAP measure. The work of this chapter has been done jointly with Sheldon Goldstein, Joel Lebowitz and Roderich Tumulka of Rutgers University and Nino Zanghi of University of Genoa.

2.1 An introduction to the problem

The family of GAP measures is a family of probability measures on Hilbert spaces, introduced in [11] as a candidate for the thermal equilibrium distribution of the wave function of a quantum system, in a roughly similar way as the canonical Gibbs distribution on the phase space of a classical system is the thermal equilibrium distribution of the phase point.

There is one GAP measure for every density matrix ρ , denoted $GAP(\rho)$. The measures relevant to thermal equilibrium are those associated with canonical density matrices

$$\rho_\beta = \frac{1}{Z} e^{-\beta H}, \quad (2.1)$$

where $Z = \text{Tr} e^{-\beta H}$ is the normalization constant, β the inverse temperature and H the Hamiltonian. A detailed discussion of GAP measures can be found in [11]. See [26] for a study about the support of GAP measures, that is, about what $GAP(\rho_\beta)$ -distributed wave functions look like.

The main object of this chapter, Theorem 2.4.1, asserts that, for a system coupled to a heat bath and a typical wave function of system and bath in a given small energy range, the *conditional* wave function of the system has a distribution near a GAP measure (i.e., one that converges – in the sense of the Weak Law of Large Numbers – to a GAP measure in the thermodynamic limit in which the heat bath becomes infinite). This statement involves two concepts that deserve explanation:

- The concept of typical wave function means that the statement refers to the overwhelming majority of wave functions, while exceptional wave functions may exist. As already mentioned typicality is strictly related to the concept of convergence in probability.
- The concept of conditional wave function is a precise version of the collapsed wave function, which is normally what is meant when speaking of “the” wave function of a particle or system. We recall the definition of conditional wave function in Section 2.2.2 below.

Theorem 2.4.1 rests on two main ingredients, which can be formulated separately:

- (i) The typicality of the GAP measure, given the reduced density matrix. This fact was proven in Section 5.1.3 of [11] but not in a rigorous way, and the main steps for a rigorous proof are shown here as Theorem 2.3.1. In more detail, Theorem 2.3.1 asserts that the conditional wave function of a small subsystem of a big system is asymptotically $GAP(\rho)$ -distributed if the wave function of the big system is typical among those wave functions with reduced density matrix of the subsystem equal to ρ .

The proof of Theorem 2.3.1 is based on the result of Petz and Réffy [19] that for a random $m \times m$ unitary matrix with distribution given by the Haar measure on the unitary group $U(m)$, the upper left (or any

other) $k \times k$ submatrix converges in distribution, as $m \rightarrow \infty$ and after multiplying by a normalization factor \sqrt{m} , to a matrix of independent complex Gaussian random variables with mean 0 and variance 1. (To understand the factor \sqrt{m} , note that a column of a unitary $m \times m$ matrix is a unit vector, and thus a single entry should be of order $1/\sqrt{m}$.) An alternative, elementary proof of the Petz–Rényi theorem is given in [17] and presented in Chapter 1.

- (ii) Canonical typicality [14, 10, 20, 21], i.e., the fact that for a typical wave function from the microcanonical ensemble of a system plus an heat bath, the reduced density matrix of the system is (approximately) of the canonical form (2.1).

As a more general version of canonical typicality, we use, in the proof of Theorem 2.4.1, a beautiful theorem of Popescu, Short, and Winter [20]. The canonical form (2.1) arises in the limit of small interaction between the system and the heat bath. To achieve greater generality and allow any interaction, we follow [20] in defining the generalized canonical density matrix as the partial trace (over the heat bath) of the microcanonical density matrix for system and bath, which then includes (2.1) in the special case of no interaction.

2.2 Review of Relevant Definitions

We begin by introducing some notation. For any complex Hilbert space \mathcal{H} , let $\mathcal{D}(\mathcal{H})$ be the set of all density operators, i.e., of all positive operators with trace 1. On $\mathcal{D}(\mathcal{H})$ and on the trace class of \mathcal{H} , the norm we use is the *trace norm*

$$\|M\|_1 = \text{Tr}|M| = \text{Tr}\sqrt{M^*M}. \quad (2.2)$$

For the unit sphere in \mathcal{H} we write

$$\mathcal{S}(\mathcal{H}) = \{\psi \in \mathcal{H} : \|\psi\| = 1\}. \quad (2.3)$$

We write $\mu_n \Rightarrow \mu$ to denote that the sequences of measures μ_n (on a Hilbert space) converges weakly to μ . This means that $\mu_n(f) \rightarrow \mu(f)$ for every

bounded continuous function f such that $\mu(D(f)) = 0$, where $D(f)$ is the set of discontinuities of f and we have written

$$\mu(f) := \int \mu(d\psi) f(\psi). \quad (2.4)$$

As pointed out in [11] the density matrix ρ associated to a measure μ is its covariance matrix:

$$\rho = \int \mu(d\psi) |\psi\rangle\langle\psi| \quad (2.5)$$

2.2.1 Definition of the GAP Measure

The measure $GAP(\rho)$ on (the Borel σ -algebra of) $\mathcal{S}(\mathcal{H})$ is built starting from the measure $G(\rho)$, which is the Gaussian measure on \mathcal{H} with covariance matrix ρ . Let's assume $\mathcal{H} = \mathbb{C}^n$ (we will discuss in section 2.5.2 some aspects related to infinite-dimensional Hilbert spaces), and there $G(\rho)$ can be defined as follows: let S be the subspace of \mathcal{H} on which ρ is supported, i.e., its positive spectral subspace, or equivalently the orthogonal complement of its kernel, or equivalently its range; let $k = \dim S$ and ρ_+ the restriction of ρ to S ; then $G(\rho)$ is a measure on \mathcal{H} supported on S with the following density relative to the Lebesgue measure λ on S :

$$\frac{dG(\rho)}{d\lambda}(\psi) = \frac{1}{\pi^k \det \rho_+} \exp(-\langle\psi|\rho_+^{-1}|\psi\rangle). \quad (2.6)$$

Now we define the adjusted Gaussian measure $GA(\rho)$ on \mathcal{H} as:

$$GA(\rho)(d\psi) = \|\psi\|^2 G(\rho)(d\psi) \quad (2.7)$$

If Ψ^{GA} is a $GA(\rho)$ -distributed vector, then $GAP(\rho)$ is the distribution of the vectors projected on the unit sphere, namely:

$$\Psi^{GAP} := \frac{\Psi^{GA}}{\|\Psi^{GA}\|} \quad (2.8)$$

Like $G(\rho)$, but not like $GA(\rho)$, the covariance matrix of $GAP(\rho)$ is ρ .

More generally, one can define for any measure μ on \mathcal{H} the “adjust-and-project” procedure. We denote by $A\mu$ the adjusted measure

$$A\mu(d\psi) = \|\psi\|^2 \mu(d\psi) \quad (2.9)$$

and the projection on the unit sphere is defined as:

$$P : \mathcal{H} \setminus \{0\} \rightarrow \mathcal{S}(\mathcal{H}), \quad P(\psi) = \frac{\psi}{\|\psi\|}. \quad (2.10)$$

Then the adjusted-and-projected measure is $P_*(A\mu) = A\mu \circ P^{-1}$, where P_* denotes the action of P on measures, thus defining a mapping $P_* \circ A$ from the measures on \mathcal{H} with $\int \mu(d\psi) \|\psi\|^2 = 1$ to the probability measures on $\mathcal{S}(\mathcal{H})$.

2.2.2 Conditional Wave Function

Consider a composite system in a pure state $\psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$, and ask what might be meant by the wave function of system 1, i.e., the subsystem with Hilbert space \mathcal{H}_1 . For this we use the notion of *conditional wave function*, which was introduced in [8] for Bohmian mechanics, a theory of quantum mechanics with particle trajectories, as follows. The conditional wave function Ψ_1 of the system is defined as

$$\Psi_1(x) = \mathcal{N} \psi(x, Y) \quad (2.11)$$

for all x in the configuration space of system 1, where \mathcal{N} is a normalizing factor and Y is the actual configuration of system 2. We adopt here the more general definition of the conditional wave function given in [11]: Let $\{|b_j\rangle\}$ be a (generalized) orthonormal basis of \mathcal{H}_2 . For each choice of b_j , the (partial) scalar product $\langle b_j | \psi \rangle$, taken in \mathcal{H}_2 , is a vector belonging to \mathcal{H}_1 . Regarding $|b_j\rangle$ as random, we are led to consider the random vector $\Psi_1 \in \mathcal{H}_1$ given by

$$\Psi_1 = \mathcal{N} \langle b_J | \psi \rangle \quad (2.12)$$

where $\mathcal{N} = \mathcal{N}(\psi, b_J) = \|\langle b_J | \psi \rangle\|^{-1}$ is the normalizing factor and $|b_J\rangle$ is a random element of the basis $\{|b_j\rangle\}$, chosen with the quantum distribution

$$\mathbb{P}(b_J = b_j) = \|\langle b_j | \psi \rangle\|^2. \quad (2.13)$$

We refer to Ψ_1 as the conditional wave function of system 1. It is interesting to note that this can be regarded as a kind of collapsed wave function, in fact if we consider system 2 as the apparatus, $|b_j\rangle$ would be an outcome (i.e.

a measurement) of this apparatus and therefore Ψ_1 would be the state of system 1 after such a measurement result.

The distribution of Ψ_1 corresponding to (2.12) and (2.13) is given by the probability measure on $\mathcal{S}(\mathcal{H}_1)$,

$$\mu_1 = \mathbb{P}(\Psi_1 \in \cdot) = \sum_{b_j} \|\langle b_j | \psi \rangle\|^2 \delta_{\frac{\langle b_j | \psi \rangle}{\|\langle b_j | \psi \rangle\|^2}}, \quad (2.14)$$

where δ_ϕ denotes the Dirac “delta” measure (a point mass) concentrated at ϕ . While the density matrix ρ_{μ_1} associated with μ_1 always equals the reduced density matrix ρ_1^{red} of system 1, given by

$$\rho_1^{\text{red}} = \text{Tr}_2 |\psi\rangle\langle\psi| = \sum_j \langle b_j | \psi \rangle \langle \psi | b_j \rangle, \quad (2.15)$$

the measure μ_1 itself usually depends on the choice of the basis $\{b_j\}$. It turns out, nevertheless, as we show in Section 2.3, that the distribution of the conditional wave function $\mu_1(d\psi_1)$ is a universal function of ρ_1^{red} in the special case that system 2 is large and ψ is typical (with respect to the uniform distribution on all wave functions), namely $GAP(\rho_1^{\text{red}})$. Thus $GAP(\rho)$ has a distinguished, universal status among all probability measures on $\mathcal{S}(\mathcal{H})$ with density matrix ρ .

2.3 Typicality of the GAP Measure for a Subsystem of a Large System

2.3.1 Setup

Let $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, where \mathcal{H}_1 and \mathcal{H}_2 have respective dimension k and m , with $k < m < \infty$. For any given density matrix ρ_1 on \mathcal{H}_1 , let

$$\mathcal{R}^m(\rho_1) = \{\psi \in \mathcal{S}(\mathcal{H}) : \rho_1^{\text{red}}(\psi) = \rho_1\}, \quad (2.16)$$

where $\rho_1^{\text{red}}(\psi) = \text{Tr}_2 |\psi\rangle\langle\psi|$ is the reduced density matrix of system 1.

Our typicality theorem concerns typical wave functions in $\mathcal{R}^m(\rho_1)$, i.e., typical wave functions with fixed reduced density matrix. For this we need to define the measure $u_{\rho_1}^m$ that serves as the uniform distribution on $\mathcal{R}^m(\rho_1)$.

For this definition, we proceed as follows. According to the Schmidt decomposition, every $\psi \in \mathcal{H}$ can be written in the form

$$\psi = \sum_{i=1}^k c_i \chi_i \otimes \phi_i \quad (2.17)$$

where $\{\chi_i\}$ is an orthonormal basis in \mathcal{H}_1 , $\{\phi_i\}$ is an orthonormal system in \mathcal{H}_2 (i.e. a set of orthonormal vectors not necessarily complete) and the c_i are coefficients which can be assumed real and non-negative. Thus, the reduced density matrix of the system 1 is

$$\rho_1^{\text{red}} = \sum_i c_i^2 |\chi_i\rangle\langle\chi_i|. \quad (2.18)$$

As the reduced density matrix is given, $\rho_1^{\text{red}} = \rho_1$, the orthonormal basis $\{\chi_i\}$ (up to a phase) and the coefficients c_i are determined (when ρ_1 is nondegenerate) as the eigenvectors and the square-roots of the eigenvalues of ρ_1 . Besides, $\mathcal{R}^m(\rho_1)$ is in a natural one-to-one correspondence with the set $ONS(\mathcal{H}_2, k)$ of all orthonormal systems $\{\phi_i\}$ in \mathcal{H}_2 of cardinality k , provided that ρ_1 has no vanishing eigenvalues. (If some of the eigenvalues of ρ_1 vanish, the one-to-one correspondence is with $ONS(\mathcal{H}_2, k')$ where $k' = \dim \text{support}(\rho_1)$.) In the following we will assume for simplicity that ρ_1 has no vanishing eigenvalues, then the Haar measure on the unitary group of \mathcal{H}_2 defines the uniform distribution on the set of orthonormal basis of \mathcal{H}_2 , of which the uniform distribution on $ONS(\mathcal{H}_2, k)$ is a marginal, and thus defines the uniform distribution $u_{\rho_1}^m$ on $\mathcal{R}(\rho_1)$. (When ρ_1 is degenerate, $u_{\rho_1}^m$ does not depend upon how the eigenvectors χ_i of ρ_1 are chosen.)

2.3.2 Typicality Theorem

Theorem 2.3.1. *For any weak neighborhood \mathcal{N} of $GAP(\rho_1)$,*

$$u_{\rho_1}^m \left\{ \psi \in \mathcal{R}^m(\rho_1) : \mu_1^\psi \in \mathcal{N} \right\} \rightarrow 1 \quad \text{as } m \rightarrow \infty, \quad (2.19)$$

where μ_1^ψ is the distribution of the conditional wave function Ψ_1 of system 1 relative to the basis $\{b_1, \dots, b_m\}$, regardless of how the basis $\{b_j\}$ is chosen.

We note that this statement is equivalent to say

$$\mu_1^\psi \xrightarrow{u_{\rho_1}^m} GAP(\rho_1) \text{ as } m \rightarrow \infty \quad (2.20)$$

where $\xrightarrow{u_{\rho_1}^m}$ indicates (generalized) convergence in measure.

To obtain Theorem 2.3.1 we need the following lemma about the Gaussian measures $G(\rho_1)$.

Lemma 2.3.2. *For every $\varepsilon > 0$ and every bounded continuous function $f : \mathcal{H}_1 \rightarrow \mathbb{R}$,*

$$u_{\rho_1}^m \left\{ \psi \in \mathcal{R}^m(\rho_1) : |\bar{\mu}_1^\psi(f) - G(\rho_1)(f)| < \varepsilon \right\} \rightarrow 1 \text{ as } m \rightarrow \infty, \quad (2.21)$$

where $\bar{\mu}_1^\psi$ is the distribution of $\sqrt{m}\langle b_J | \psi \rangle \in \mathcal{H}_1$ (not normalized) with respect to the uniform distribution of $J \in \{1, \dots, m\}$, regardless of how the basis $\{b_j\}$ is chosen.

A trivial consequence of Lemma 2.3.2 is the corresponding statement of Theorem 2.3.1 about the Gaussian measures $G(\rho_1)$:

Lemma 2.3.2'. *For any weak neighborhood \mathcal{N} of $G(\rho_1)$,*

$$u_{\rho_1}^m \left\{ \psi \in \mathcal{R}^m(\rho_1) : \bar{\mu}_1^\psi \in \mathcal{N} \right\} \rightarrow 1 \text{ as } m \rightarrow \infty, \quad (2.22)$$

where $\bar{\mu}_1^\psi$ is the distribution of $\sqrt{m}\langle b_J | \psi \rangle \in \mathcal{H}_1$ (not normalized) with respect to the uniform distribution of $J \in \{1, \dots, m\}$, regardless of how the basis $\{b_j\}$ is chosen.

As for Theorem 2.3.1 this statement is equivalent to say

$$\bar{\mu}_1^\psi \xrightarrow{u_{\rho_1}^m} G(\rho_1) \text{ as } m \rightarrow \infty \quad (2.23)$$

where $\xrightarrow{u_{\rho_1}^m}$ indicates (generalized) convergence in measure.

A few comments about notation. In [11], the notation for the basis $\{b_1, \dots, b_m\}$ was $\{|1\rangle, \dots, |m\rangle\}$. For enumerating the basis, we will use the letter j , and thus write b_j ; in [11], the notation was q_2 for j (subscript 2 because it refers to \mathcal{H}_2). For a random choice of j , we write J ; the corresponding notation in [11] was Q_2 .

2.3.3 Proof of Lemma 2.3.2

We note that $\bar{\mu}_1^\psi$ is the average of m delta measures with equal weights,

$$\bar{\mu}_1^\psi = \frac{1}{m} \sum_{j=1}^m \delta_{\psi_1(j)}, \quad (2.24)$$

located at the points

$$\psi_1(j) = \sqrt{m} \langle b_j | \psi \rangle = \sum_{i=1}^k c_i \sqrt{m} \langle b_j | \phi_i \rangle \chi_i \quad (2.25)$$

Now we regard ψ as random with distribution u_{ρ_1} ; then the $\psi_1(j)$ are m random vectors, and $\bar{\mu}_1^\psi$ is their empirical distribution.

As noted above $\{\phi_i\}$ is a random orthonormal system distributed as a marginal of the Haar measure; in other words, the expansion coefficients $\langle b_j | \phi_i \rangle$ of

$$\phi_i = \sum_{j=1}^m \langle b_j | \phi_i \rangle b_j$$

form a $k \times m$ matrix that arises as the first k rows of a unitary $m \times m$ matrix distributed according to the Haar measure. It is therefore interesting to us how a submatrix of a Haar distributed unitary matrix is distributed. The relevant fact for us is the following theorem of Petz and Réffy [19, 17] which is also stated in Chapter 1 as Theorem 1.1.1: *If (U_{ij}) is Haar($U(m)$) distributed, then the joint distribution of the upper left (or, in fact, any) $k \times k$ submatrix, multiplied by the normalization factor \sqrt{m} , converges in distribution, as $m \rightarrow \infty$, to \sqrt{k} times a standard non-selfadjoint Gaussian matrix, i.e., converges to a random $k \times k$ matrix (G_{ij}) whose entries G_{ij} are independent complex Gaussian random variables with mean 0 and variance $\mathbb{E}|G_{ij}|^2 = 1$.*

We will use two consequences of this theorem. First, *if (U_{ij}) is Haar distributed in $U(m)$ then the first k entries of the first column, U_{i1} for $i = 1, \dots, k$, multiplied by a normalization factor \sqrt{m} , converge in distribution, as $m \rightarrow \infty$, to a vector of complex random variables G_i , $i = 1, \dots, k$, that are independent and have Gaussian distribution with mean 0 and variance $\mathbb{E}|G_i|^2 = 1$.* In particular, if $g : \mathbb{C}^k \rightarrow \mathbb{R}$ is any bounded continuous function

then

$$\mathbb{E}g(\sqrt{m}U_{11}, \dots, \sqrt{m}U_{k1}) \longrightarrow \mathbb{E}g(G_1, \dots, G_k) \quad \text{as } m \rightarrow \infty. \quad (2.26)$$

Second, if (U_{ij}) is Haar distributed in $U(m)$ then the first k entries of the first two columns, U_{i1} and U_{i2} for $i = 1, \dots, k$, multiplied by the normalization factor \sqrt{m} , converge in distribution, as $m \rightarrow \infty$, to two independent vectors of complex random variables G_{i1} and G_{i2} , $i = 1, \dots, k$, all of which are independent with Gaussian distribution with mean 0 and variance 1. In particular, if $g : \mathbb{C}^k \rightarrow \mathbb{R}$ is a bounded continuous function then

$$\begin{aligned} \mathbb{E}g(\sqrt{m}U_{11}, \dots, \sqrt{m}U_{k1}) g(\sqrt{m}U_{12}, \dots, \sqrt{m}U_{k2}) \longrightarrow \\ \mathbb{E}g(G_{11}, \dots, G_{k1}) g(G_{12}, \dots, G_{k2}) = \left(\mathbb{E}g(G_{11}, \dots, G_{k1}) \right)^2 \end{aligned} \quad (2.27)$$

as $m \rightarrow \infty$. We now proceed with the proof of Lemma 2.3.2.

Proof. We use the basis $\{\chi_i\}$ to identify \mathcal{H}_1 with \mathbb{C}^k . Let $f : \mathbb{C}^k \rightarrow \mathbb{R}$ be a bounded continuous function, and let $c_i \geq 0$ be as before (the square roots of the eigenvalues of ρ_1). Set

$$g(z_1, \dots, z_k) = f(c_1 z_1, \dots, c_k z_k). \quad (2.28)$$

Then g , too, is bounded and continuous. As we said above, the distribution of the $\langle b_j | \phi_i \rangle$ is the same as that of U_{ij} with $j \in \{1, \dots, m\}$, that is, the first k columns of a Haar distributed unitary $m \times m$ matrix. We thus write U_{ij} instead of $\langle b_j | \phi_i \rangle$, and obtain

$$\bar{\mu}_1^\psi(f) =: \bar{\mu}(f) = \frac{1}{m} \sum_{j=1}^m g(\sqrt{m}U_{1j}, \dots, \sqrt{m}U_{kj}). \quad (2.29)$$

For the expression to be estimated, we have by the triangle inequality that

$$\left| \bar{\mu}(f) - G(\rho_1)(f) \right| \leq \left| \bar{\mu}(f) - \mathbb{E}\bar{\mu}(f) \right| + \left| \mathbb{E}\bar{\mu}(f) - G(\rho_1)(f) \right| \quad (2.30)$$

Note that the second contribution in (2.30) is nonrandom.

We first check that $|\mathbb{E}\bar{\mu}(f) - G(\rho_1)(f)| < \varepsilon/2$ if m is sufficiently large: In our present notation,

$$G(\rho_1)(f) = \mathbb{E}f(c_1 G_1, \dots, c_k G_k) = \mathbb{E}g(G_1, \dots, G_k), \quad (2.31)$$

and, by (2.29),

$$\mathbb{E}\bar{\mu}(f) = \frac{1}{m} \sum_{j=1}^m \mathbb{E}g(\sqrt{m}U_{1j}, \dots, \sqrt{m}U_{kj}) = \mathbb{E}g(\sqrt{m}U_{11}, \dots, \sqrt{m}U_{k1}) \quad (2.32)$$

because each column of (U_{ij}) has the same distribution (because the columns are exchangeable because of the invariance of the Haar measure). By (2.26), (2.32) converges to (2.31) as $m \rightarrow \infty$.

Concerning the first contribution in (2.30), Chebyshev's inequality (see, e.g., [2, p. 65]) asserts that

$$\mathbb{P}\left(|\bar{\mu}(f) - \mathbb{E}\bar{\mu}(f)| > \varepsilon/2\right) < \frac{4}{\varepsilon^2} \text{var}(\bar{\mu}(f)), \quad (2.33)$$

where \mathbb{P} is the Haar measure for U_{ij} and $\text{var}(X)$ is the variance of X . Now it only remains to be shown that

$$\text{var}(\bar{\mu}(f)) \rightarrow 0 \quad \text{as } m \rightarrow \infty. \quad (2.34)$$

We find that

$$\begin{aligned} \text{var}(\bar{\mu}(f)) &= \mathbb{E}\bar{\mu}(f)^2 - (\mathbb{E}\bar{\mu}(f))^2 \\ &= \frac{1}{m^2} \sum_{j,j'=1}^m \mathbb{E}g(\sqrt{m}U_{1j}, \dots, \sqrt{m}U_{kj})g(\sqrt{m}U_{1j'}, \dots, \sqrt{m}U_{kj'}) - \\ &\quad - (\mathbb{E}\bar{\mu}(f))^2. \end{aligned} \quad (2.35)$$

Since the columns of (U_{ij}) are exchangeable, the joint distribution of any two different columns is the same as the joint distribution of the first two columns, so that all summands with $j \neq j'$ are equal (and all summands with $j = j'$ are equal), and we can write

$$\begin{aligned} \text{var}(\bar{\mu}(f)) &= \frac{m^2 - m}{m^2} \mathbb{E}g(\sqrt{m}U_{11}, \dots, \sqrt{m}U_{k1})g(\sqrt{m}U_{12}, \dots, \sqrt{m}U_{k2}) + \\ &\quad + \frac{m}{m^2} \mathbb{E}g(\sqrt{m}U_{11}, \dots, \sqrt{m}U_{k1})^2 - (\mathbb{E}\bar{\mu}(f))^2. \end{aligned} \quad (2.36)$$

As $m \rightarrow \infty$, these expressions converge in the following way: $(m^2 - m)/m^2 \rightarrow 1$, the subsequent expectation according to (2.27), $m/m^2 \rightarrow 0$, the subsequent expectation (by applying (2.26) to the bounded continuous function g^2 instead of g) to the finite number $\mathbb{E}g(G_1, \dots, G_k)^2 \leq \|g\|_\infty^2$, and

$(\mathbb{E}\bar{\mu}(f))^2 \rightarrow (G(\rho_1)(f))^2$ by (2.26). Thus,

$$\text{var}(\bar{\mu}(f)) \longrightarrow \left(\mathbb{E}g(G_{11}, \dots, G_{k1}) \right)^2 - G(\rho_1)(f)^2 = 0, \quad (2.37)$$

proving (2.34) and thus completing the proof of Lemma 2.3.2.

Proof of Theorem 2.3.1

We first check that $\bar{\mu}_1^\psi(\|\cdot\|^2) = 1$. We have that

$$\int_{\mathcal{H}} \bar{\mu}_1^\psi(d\psi_1) \|\psi_1\|^2 = \mathbb{E}(m \|\langle b_j | \psi \rangle\|^2) = \sum_{j=1}^m \|\langle b_j | \psi \rangle\|^2 = \|\psi\|^2 = 1. \quad (2.38)$$

Now we consider that: the mapping P_*A is continuous on the set of measures μ on \mathcal{H}_1 such that $\int \mu(d\psi) \|\psi\|^2 = 1$ by Lemma 2.5.2 and the weak continuity if P_* . The distribution of the conditional wave function is

$$\mu_1^\psi = P_*A(\bar{\mu}_1^\psi). \quad (2.39)$$

Then if μ_1^ψ is in a neighborhood \mathcal{N} of $GAP(\rho)$ we have that:

$$\left\{ \mu_1^\psi \in \mathcal{N} \right\} = \left\{ P_*A(\bar{\mu}_1^\psi) \in \mathcal{N} \right\} = \left\{ \bar{\mu}_1^\psi \in (P_*A)^{-1}\mathcal{N} \right\} = \left\{ \bar{\mu}_1^\psi \in \mathcal{N}' \right\} \quad (2.40)$$

where \mathcal{N}' is a neighborhood of $G(\rho_1)$, since $(P_*A)G(\rho_1) = GAP(\rho_1)$. By Lemma 2.3.2' this completes the proof. \square

2.3.4 Random Basis

As already explained in [11], instead of considering a typical wave function and a fixed basis one can consider a fixed wave function and a typical basis. Let $ONB(\mathcal{H}_2)$ be the set of all orthonormal bases of \mathcal{H}_2 .

Corollary 2.3.3. *For any weak neighborhood \mathcal{N} of $GAP(\rho_1)$,*

$$\nu \left\{ (b_1, \dots, b_m) \in ONB(\mathcal{H}_2) : \mu_1^b \in \mathcal{N} \right\} \rightarrow 1 \text{ as } m \rightarrow \infty, \quad (2.41)$$

where ν is the uniform probability measure on $ONB(\mathcal{H}_2)$, corresponding to the Haar measure on $U(m)$, regardless of how $\psi \in \mathcal{R}^m(\rho_1)$ is chosen.

Proof. Note that for any unitary U on \mathcal{H}_2

$$\langle U^{-1}b_j|\psi\rangle = \langle b_j|I_1 \otimes U\psi\rangle. \quad (2.42)$$

From this fact it follows that when $\psi \in \mathcal{R}^m(\rho_1)$ is $u_{\rho_1}^m$ -distributed the distribution of the conditional wave function μ_1^ψ actually doesn't depend on the choice of the basis $\{b_j\}$. It similarly follows from (2.4) (and the fact that the Haar measure is invariant under $U \mapsto U^{-1}$) that for fixed $\psi \in \mathcal{R}^m(\rho_1)$, $U\psi$ is $u_{\rho_1}^m$ -distributed. Thus the statement is equivalent to Theorem 2.3.1. \square

2.4 The GAP Measure as the Thermal Equilibrium Distribution

Here it will be stated the fundamental theorem of this chapter. This theorem has been already anticipated in [11] and there were shown the fundamental ingredients of its proof. In the previous part of the chapter there are rigorously stated all the necessary steps to arrive to the proof of this statement. Unfortunately the proof of the following theorem is simply sketched, since Theorem 2.3.1 and Corollary are valid only for those ψ such that $\rho_1 = \text{Tr}_2|\psi\rangle\langle\psi|$ is fixed, but a varying ρ_1 is a necessary condition for

Theorem 2.4.1. *Suppose we have a given Hilbert space \mathcal{H}_1 of dimension k , and for every positive integer m a Hilbert space \mathcal{H}_2 of dimension m , and a subspace \mathcal{H}_R of $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ such that*

$$d_R := \dim \mathcal{H}_R \rightarrow \infty \quad \text{as } m \rightarrow \infty. \quad (2.43)$$

Let ρ_R be $1/d_R$ times the projection to \mathcal{H}_R , $u_R = \text{GAP}(\rho_R) = u_{\mathcal{S}(\mathcal{H}_R)}$, and $\Omega_m := \text{Tr}_2 \rho_R$. Suppose that

$$\Omega_m \rightarrow \Omega \quad \text{as } m \rightarrow \infty. \quad (2.44)$$

Then for every bounded continuous function $f : \mathcal{S}(\mathcal{H}_1) \rightarrow \mathbb{R}$

$$u_R \times \nu \left\{ (\psi, b) \in \mathcal{S}(\mathcal{H}) \times \text{ONB}(\mathcal{H}_2) : |\mu_1^{\psi, b}(f) - \text{GAP}(\Omega)(f)| < \varepsilon \right\} \rightarrow 1 \quad (2.45)$$

as $m \rightarrow \infty$.

To understand the importance of this theorem let's consider an important example: choose for every m a Hamiltonian H on \mathcal{H} and a suitable energy E_m , such as corresponding to a fixed energy per particle or per volume, and let \mathcal{H}_R be the microcanonical subspace, i.e., the spectral subspace of H associated with the interval $[E_m, E_m + \delta]$. It is a standard fact that when the interaction energy is negligibly small then Ω is of the exponential form $Z^{-1} \exp(-\beta H)$ with $Z = \text{Tr} \exp(-\beta H)$ for suitable $\beta > 0$, i.e. the canonical density matrix ρ_β . Then by Theorem 2.4.1 in this "special case" of negligible interaction we have that

$$\mu_1^{\psi, b} \xrightarrow{u_R \times \nu} \text{GAP}(\rho_\beta) \quad (2.46)$$

This fact leads to consider $\text{GAP}(\rho_\beta)$ as the thermal equilibrium distribution of the wave functions of the system 1.

Proof of Theorem 2.4.1. Sketched. In the case that Corollary 2.3.3 is valid also for varying ρ_1 , we would obtain that for an arbitrary probability measure λ on $\mathcal{S}(\mathcal{H})$, depending on m in an arbitrary way,

$$(\lambda \times \nu)(A) \rightarrow 1 \quad (2.47)$$

as $m \rightarrow \infty$ for

$$A := \left\{ (\psi, b) \in \mathcal{S}(\mathcal{H}) \times \text{ONB}(\mathcal{H}_2) : |\mu_1^\psi(f) - \text{GAP}(\rho_1^\psi)(f)| < \varepsilon \right\} \quad (2.48)$$

where $\rho_1^\psi = \text{Tr}_2 |\psi\rangle\langle\psi|$. Fix f and ε . Theorem 2.5.1 provides us with $r > 0$ so that for every ψ with $\|\rho_1^\psi - \Omega\|_1 < r$, we have that

$$|\text{GAP}(\rho_1^\psi)(f) - \text{GAP}(\Omega)(f)| < \varepsilon. \quad (2.49)$$

Let $B := \{\psi \in \mathcal{S}(\mathcal{H}) : \|\rho_1^\psi - \Omega\|_1 < r\} \times \text{ONB}(\mathcal{H}_2)$. Thus, for every pair $(\psi, b) \in A \cap B$, we have that

$$\begin{aligned} |\mu_1^\psi(f) - \text{GAP}(\Omega)(f)| &\leq |\mu_1^\psi(f) - \text{GAP}(\rho_1^\psi)(f)| + |\text{GAP}(\rho_1^\psi)(f) - \text{GAP}(\Omega)(f)| \\ &< 2\varepsilon. \end{aligned} \quad (2.50)$$

Therefore, $A \cap B$ is contained in the set mentioned in Theorem 2.4.1 (with ε replaced by 2ε).

It remains to show that for $\lambda = u_R$, $(\lambda \times \nu)(A \cap B) \rightarrow 1$, which follows if $(\lambda \times \nu)(A) \rightarrow 1$ and $(\lambda \times \nu)(B) \rightarrow 1$. The first statement was (2.47) above, and the second follows from

$$u_R\{\psi \in \mathcal{S}(\mathcal{H}) : \|\rho_1^\psi - \Omega\|_1 < r\} \rightarrow 1, \quad (2.51)$$

which we now derive from Theorem 1 of [20]. It asserts, in a somewhat specialized and simplified form that suffices for our purposes: *Given a Hilbert space \mathcal{H}_1 of dimension k , another Hilbert space \mathcal{H}_2 of finite dimension and a subspace $\mathcal{H}_R \subseteq \mathcal{H}_1 \otimes \mathcal{H}_2$ of dimension d_R . Let ρ_R again be $1/d_R$ times the projection to \mathcal{H}_R . Then for every $\epsilon > 0$,*

$$u_R \left\{ \psi \in \mathcal{S}(\mathcal{H}) : \left\| \text{Tr}_2 |\psi\rangle\langle\psi| - \text{Tr}_2 \rho_R \right\|_1 \geq \epsilon + \frac{k}{\sqrt{d_R}} \right\} \leq 2 \exp(-C d_R \epsilon^2), \quad (2.52)$$

where C is a positive constant.

By the assumptions of Theorem 2.4.1, in our case $d_R \rightarrow \infty$. Thus, provided m is so large that $k/\sqrt{d_R} < r/4$ and setting $\epsilon = r/4$, we obtain that

$$u_R\{\psi \in \mathcal{S}(\mathcal{H}) : \|\rho_1^\psi - \Omega_m\|_1 < r/2\} \geq 1 - 2 \exp(-C d_R r^2/4) \rightarrow 1. \quad (2.53)$$

Since $\Omega_m \rightarrow \Omega$, $\|\Omega_m - \Omega\|_1 < r/2$ for sufficiently large m , and therefore the last equation implies (2.51). This completes the proof of Theorem 2.4.1.

2.5 Further Results on GAP Measure

2.5.1 Continuity in Finite Dimension

Here we show that (in finite dimension) $GAP(\rho)$ is continuous as a function of ρ . (A corresponding theorem for infinite dimension is established in section 2.5.2.)

Theorem 2.5.1. *$\rho \rightarrow GAP(\rho)$ is continuous as a map from $\mathcal{D}(\mathcal{H})$ to the set of the probability measures on $\mathcal{S}(\mathcal{H})$: suppose $\rho_n \in \mathcal{D}(\mathbb{C}^k)$ for every $n \in \mathbb{N}$, then if $\rho_n \rightarrow \rho$, $GAP(\rho_n) \Rightarrow GAP(\rho)$.*

To show this, we first establish the continuity of the mapping A defined in the previous section. For any topological space X let $\mathcal{B}(X)$ denote its Borel σ -algebra.

Lemma 2.5.2. *A is continuous on the set of probability measures μ on \mathcal{H} such that $\int \mu(d\psi) \|\psi\|^2 = 1$: suppose that for every $n \in \mathbb{N}$, μ_n is a probability measures on $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ such that $\int \mu_n(d\psi) \|\psi\|^2 = 1$. If $\mu_n \Rightarrow \mu$ with $\int \mu(d\psi) \|\psi\|^2 = 1$ $A\mu_n \Rightarrow A\mu$.*

Proof. Fix $\varepsilon > 0$ and an arbitrary bounded, continuous function $f : \mathcal{H} \rightarrow \mathbb{R}$. Set $g(\psi) = \|\psi\|^2$. Since, by hypothesis, $\mu(g) = 1$ in the notation (2.4), there exists $R > 0$ so large that

$$\mu(1_{B_R}g) = \int_{B_R} \mu(d\psi) \|\psi\|^2 > 1 - \frac{\varepsilon}{6\|f\|_\infty}, \quad (2.54)$$

where 1_{B_R} denotes the indicator function of the set $B_R = \{\psi \in \mathcal{H} : \|\psi\| < R\}$. Let the ‘‘cut-off function’’ $\chi : \mathcal{H} \rightarrow \mathbb{R}$ be any continuous function such that $\chi = 1$ on B_R , $\chi = 0$ outside B_{2R} , and $0 \leq \chi \leq 1$. Because χg is a bounded continuous function, and because $\mu_n \Rightarrow \mu$, we have that $\mu_n(\chi g) \rightarrow \mu(\chi g)$, that is, there is an $n_1 \in \mathbb{N}$ such that for all $n > n_1$,

$$|\mu_n(\chi g) - \mu(\chi g)| < \frac{\varepsilon}{6\|f\|_\infty}. \quad (2.55)$$

Therefore, using that $\chi \geq 1_{B_R}$,

$$\mu_n(\chi g) > \mu(\chi g) - \frac{\varepsilon}{6\|f\|_\infty} \geq \mu(1_{B_R}g) - \frac{\varepsilon}{6\|f\|_\infty} > 1 - \frac{\varepsilon}{3\|f\|_\infty}, \quad (2.56)$$

and thus that

$$\left| \mu_n(f(1 - \chi)g) \right| \leq \mu_n(|f(1 - \chi)g|) \leq \|f\|_\infty \mu_n((1 - \chi)g) = \quad (2.57)$$

$$= \|f\|_\infty (\mu_n(g) - \mu_n(\chi g)) < \|f\|_\infty \left(1 - 1 + \frac{\varepsilon}{3\|f\|_\infty}\right) = \varepsilon/3. \quad (2.58)$$

Likewise,

$$\left| \mu(f(1 - \chi)g) \right| \leq \mu(|f(1 - \chi)g|) \leq \|f\|_\infty \mu((1 - \chi)g) \leq \quad (2.59)$$

$$\leq \|f\|_\infty (\mu(g) - \mu(1_{B_R}g)) < \|f\|_\infty \left(1 - 1 + \frac{\varepsilon}{6\|f\|_\infty}\right) = \varepsilon/6 < \varepsilon/3. \quad (2.60)$$

Because $f\chi g$ is a bounded continuous function, and because $\mu_n \Rightarrow \mu$, we have that $\mu_n(f\chi g) \rightarrow \mu(f\chi g)$, that is, there is an $n_2 \in \mathbb{N}$ such that for all $n > n_2$,

$$|\mu_n(f\chi g) - \mu(f\chi g)| < \varepsilon/3. \quad (2.61)$$

Thus,

$$|A\mu_n(f) - A\mu(f)| = |\mu_n(fg) - \mu(fg)| \leq \quad (2.62)$$

$$\leq |\mu_n(f\chi g) - \mu(f\chi g)| + |\mu_n(f(1-\chi)g) - \mu(f(1-\chi)g)| < \varepsilon. \quad (2.63)$$

□

We remark that the hypothesis $\int \mu(d\psi)\|\psi\|^2 = 1$ cannot be dropped, that is, does not follow from $\int \mu_n(d\psi)\|\psi\|^2 = 1$. An example is $\mu_n = (1 - 1/n)\delta_0 + (1/n)\delta_{\psi_n}$, where δ_ϕ means the Dirac delta measure at ϕ and ψ_n is any vector with $\|\psi_n\|^2 = n$; then μ_n is a probability measure with $\int \mu_n(d\psi)\|\psi\|^2 = 1$ but $\mu_n \Rightarrow \delta_0$, which has $\int \delta_0(d\psi)\|\psi\|^2 = 0$.

The following lemma is the analog of Theorem 2.5.1 for $G(\rho)$.

Lemma 2.5.3. $\rho \rightarrow G(\rho)$ is continuous as a map from $\mathcal{D}(\mathcal{H})$ to the set of the probability measures on $\mathcal{S}(\mathcal{H})$: suppose $\rho_n \in \mathcal{D}(\mathbb{C}^k)$ for every $n \in \mathbb{N}$, then if $\rho_n \rightarrow \rho$, $G(\rho_n) \Rightarrow G(\rho)$.

Proof. We use characteristic functions; as usual, the characteristic function $\hat{\mu} : \mathbb{R}^{2k} \rightarrow \mathbb{C}$ of a probability measure μ on \mathbb{R}^{2k} is defined by

$$\hat{\mu}(t_1, \dots, t_{2k}) = \int \mu(dx_1 \cdots dx_{2k}) \exp\left(i \sum_{j=1}^{2k} x_j t_j\right), \quad (2.64)$$

or, in our notation on $\mathcal{H} = \mathbb{C}^k$,

$$\hat{\mu}(\phi) = \int \mu(d\psi) \exp\left(i \operatorname{Re}\langle \psi | \phi \rangle\right), \quad (2.65)$$

where Re denotes the real part. We write $\mu_n = G(\rho_n)$ and $\mu = G(\rho)$; their characteristic functions are:

$$\hat{\mu}_n(\psi) = \exp(-\langle \psi | \rho_n | \psi \rangle), \quad \hat{\mu}(\psi) = \exp(-\langle \psi | \rho | \psi \rangle). \quad (2.66)$$

If $\rho_n \rightarrow \rho$ then $\langle \psi | \rho_n | \psi \rangle \rightarrow \langle \psi | \rho | \psi \rangle$ for every ψ and thus $\hat{\mu}_n \rightarrow \hat{\mu}$ pointwise. Since (e.g., [2]), *pointwise convergence of the characteristic functions is equivalent (in finite dimension) to weak convergence of the associated measures*, it follows that $G(\rho_n) \Rightarrow G(\rho)$, which is what we wanted to show. □

Proof of Theorem 2.5.1. Suppose $\rho_n \rightarrow \rho$, we have that $GAP(\rho_n) = P_*A(G(\rho_n))$, that $\int \|\psi\|^2 G(\rho) = 1$ and that $(AG(\rho))(0) = 0$. Since $\psi \mapsto P\psi$ is continuous for $\psi \neq 0$, Theorem 2.5.1 follows from Lemma 2.5.2 and Lemma 2.5.3. \square

2.5.2 The GAP Measure in Infinite Dimension

Although in the other sections we have considered GAP measures only on finite-dimensional spaces, for the sake of completeness we discuss in this section the rigorous definition of GAP measures on infinite-dimensional Hilbert spaces. Let \mathcal{H} be a separable complex Hilbert space, and for any topological space X let $\mathcal{B}(X)$ denotes its Borel σ -algebra.

Existence and Uniqueness

Definition. A probability measure μ on $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ is a *Gaussian measure* if for every $\psi \in \mathcal{H}$, the random variable $\mathcal{H} \ni \omega \mapsto \langle \psi | \omega \rangle \in \mathbb{C}$ has Gaussian distribution in \mathbb{C} .

The existence and uniqueness of a Gaussian measure with given mean and covariance operator is guaranteed by the following theorem due to Prohorov [22] (cited from [15, p. 29]): *Every Gaussian measure on \mathcal{H} is uniquely determined by its mean, which is a point in \mathcal{H} , and its covariance operator, which is a positive trace-class operator on \mathcal{H} . Conversely, for every point ψ_0 in \mathcal{H} and every positive trace-class operator ρ on \mathcal{H} , there is a Gaussian measure μ with mean ψ_0 and covariance ρ . Its characteristic function is*

$$\hat{\mu}(\psi) = \exp(i\operatorname{Re}\langle \psi_0 | \psi \rangle - \langle \psi | \rho | \psi \rangle), \quad (2.67)$$

where Re denotes the real part. We write $G(\rho)$ for the unique Gaussian measure with mean 0 and covariance ρ .

Now the adjust-and-project procedure can be applied in the same way as in finite dimension, and we obtain:

Corollary 2.5.4. *For every positive operator ρ with $\operatorname{Tr} \rho = 1$ the measure $GAP(\rho)$ exists on $\mathcal{B}(\mathcal{S}(\mathcal{H}))$ and is uniquely defined.*

We think it is helpful to see the following alternative, more direct proof of existence and uniqueness of the Gaussian measure with given covariance ρ , which does not make use of characteristic functionals. Given $\rho = \sum p_n |n\rangle\langle n|$, the desired random vector Ψ^G has coefficients $Z_n = \langle \Psi^G | n \rangle$ in the eigenbasis of ρ which should be independent Gaussian random variables with mean 0 and variances p_n . By the Kolmogorov extension theorem [2], there exists such a random sequence $(Z_n)_n$. The following lemma asserts that the σ -algebras fit in the right way:

Lemma 2.5.5. *Let $\mathbb{C}^{\mathbb{N}}$ be the vector space of all complex sequences, and \mathcal{C} the σ -algebra generated by the cylinder sets. Then $\mathcal{B}(\ell^2) \subseteq \mathcal{C}$. (Here, ℓ^2 is understood, on the one hand, as a Hilbert space, and on the other hand as a subset of $\mathbb{C}^{\mathbb{N}}$.)*

Proof. The Borel σ -algebra of a Polish space (and a separable Hilbert space is a Polish space) has the property that any countable family of Borel sets that separates points generates the full σ -algebra [29]. There is clearly a countable family of cylinder sets that separates points. \square

We now continue the construction of $G(\rho)$. The random sequence $Z = (Z_n)_n$ lies almost surely in ℓ^2 because the expectation of $\|Z\|_2^2$ is finite,

$$\mathbb{E} \sum_n |Z_n|^2 = \sum_n \mathbb{E} |Z_n|^2 = \sum_n p_n = \text{Tr } \rho = 1, \quad (2.68)$$

and therefore $\|Z\|_2^2$ is almost surely finite. We can thus form the random vector

$$\Psi^G := \sum_n Z_n |n\rangle \quad (2.69)$$

whose distribution is $G(\rho)$ because it is Gaussian and has covariance ρ .

Concerning uniqueness of $G(\rho)$ the only non-uniqueness arises from the choice of the orthonormal basis diagonalizing ρ . But since every eigenspace with eigenvalue different from zero is finite-dimensional, and since on finite-dimensional space the Gaussian measure is uniquely determined by its covariance, all eigenbases of ρ lead to the same distribution of Ψ^G .

Continuity

The following theorem expresses that $GAP(\rho)$ depends continuously on ρ , and thus forms an infinite-dimensional analog of Theorem 2.5.1.

Theorem 2.5.6. *If $\rho, \rho_n \in \mathcal{D}(\mathcal{H})$ for every $n \in \mathbb{N}$, \mathcal{H} is a separable Hilbert space, and $\|\rho_n - \rho\|_1 \rightarrow 0$ then $GAP(\rho_n) \Rightarrow GAP(\rho)$.*

The following lemma is the corresponding statement about $G(\rho)$.

Lemma 2.5.7. *If $\rho, \rho_n \in \mathcal{D}(\mathcal{H})$ for every $n \in \mathbb{N}$ and $\|\rho_n - \rho\|_1 \rightarrow 0$ then $G(\rho_n) \Rightarrow G(\rho)$.*

Proof. As in the proof Lemma 2.5.3, we use characteristic functions. As usual, the characteristic function $\hat{\mu}$ of a probability measure μ on \mathcal{H} is defined as

$$\hat{\mu}(\phi) = \int \mu(d\psi) \exp\left(i\operatorname{Re}\langle\psi|\phi\rangle\right). \quad (2.70)$$

We use Lemma VI.2.1 on p. 153 of [18], which implies that *if a sequence μ_n of measures on \mathcal{H} is conditionally compact and the characteristic functions $\hat{\mu}_n$ converge pointwise to $\hat{\mu}$, then $\mu_n \Rightarrow \mu$* . Here, $\mu = G(\rho)$ and $\mu_n = G(\rho_n)$.

The second condition, pointwise convergence of the characteristic functions, is the same as in the finite-dimensional case. Due to (2.67), the characteristic functions are given by the same formula (2.66) as in the finite-dimensional case. Now the pointwise convergence of $\hat{\mu}_n$ to $\hat{\mu}$ follows from that of $\langle\psi|\rho_n\psi\rangle$ to $\langle\psi|\rho\psi\rangle$, which follows from the convergence of ρ_n to ρ in the operator norm, which follows from the convergence in the trace norm.

What differs from the finite-dimensional case is the requirement of conditional compactness. According to Theorem VI.2.2 on p. 154 of [18], *the sequence μ_n is conditionally compact if*

$$\sup_{n \in \mathbb{N}} \sum_{i=k}^{\infty} \langle b_i | \rho_n b_i \rangle \rightarrow 0 \quad \text{as } k \rightarrow \infty, \quad (2.71)$$

where ρ_n is the covariance operator of μ_n and $\{b_i\}$ an orthonormal basis of \mathcal{H} . To see that this is the case, fix any orthonormal basis $\{b_i\}$ and note that for the operator $\Delta_n := |\rho_n - \rho|$ our hypothesis says $\operatorname{Tr} \Delta_n \rightarrow 0$ or, in other words, that for any $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that, for all $n > N$,

$$0 \leq \operatorname{Tr} \Delta_n < \varepsilon. \quad (2.72)$$

Since $\text{Tr } \Delta_n < \infty$ for all $n \in \mathbb{N}$, there is $K \in \mathbb{N}$ such that

$$\sum_{i=K}^{\infty} \langle b_i | \Delta_n b_i \rangle < \varepsilon \quad \text{for all } n \leq N. \quad (2.73)$$

Moreover,

$$0 \leq \sum_{i=k}^{\infty} \langle b_i | \rho_n b_i \rangle \leq \sum_{i=k}^{\infty} \langle b_i | \rho b_i \rangle + \sum_{i=k}^{\infty} \langle b_i | \Delta_n b_i \rangle, \quad (2.74)$$

where the first term on the right hand side is less than ε if k is large enough because $\text{Tr } \rho < \infty$, and the second term is less than ε for all $n \in \mathbb{N}$ if $k > K$, either because of (2.73) (if $n \leq N$) or because of (2.72) (if $n > N$). \square

Proof of Theorem 2.5.6. The claim follows in the same way from Lemma 2.5.7 as Theorem 2.5.1 follows from Lemma 2.5.3. This is because Lemma 2.5.2 holds as well in infinite dimension, and the projection P to the unit sphere defined in (2.10) is continuous also in infinite dimension. \square

2.6 Final Remark

A final remark concerns the role of interaction (between the system and the heat bath) for obtaining the thermal equilibrium distribution. Interaction is relevant for creating typical wave functions, as it helps evolve atypical wave functions into typical ones. But once the composite of system and bath has a typical wave function, the presence or absence of interaction is irrelevant to the distribution of the conditional wave function. After all, the distribution of the conditional wave function at time t is determined by the total wave function at time t , and thus does not refer to the Hamiltonian. As some of our theorems assume a typical wave function by hypothesis, they need not, and do not, even mention interaction. Then again, interaction plays another role, namely for determining the system's canonical density matrix, which we take to be defined as the partial trace (over the bath) of the microcanonical density matrix of the composite. In the limit of negligible interaction, the canonical density matrix has the famous form $\rho = (1/Z)e^{-\beta H}$, while interaction makes the canonical density matrix deviate from this exponential form. This resolves the apparent paradox that we consider a system coupled to a heat bath, and then neglect the interaction.

Chapter 3

Typical Non-Equilibrium Behavior of Macroscopic Quantum Systems

After having considered a good candidate to be a typicality measure for quantum system in equilibrium, in this chapter it will be treated the approach to equilibrium for macroscopic quantum systems. Starting from a particular case we shall show that it is possible to find a formulation valid for a generic quantum system. Our result states that, under suitable conditions, and for a “typical” Hamiltonian, a quantum system reaches equilibrium regardless of the initial state. The work presented in this chapter it has been done jointly with Sheldon Goldstein, Joel Lebowitz and Roderich Tumulka of Rutgers University and Nino Zanghì of University of Genoa.

3.1 The Problem of Convergence to Equilibrium

In a macroscopic quantum system described by an Hilbert space \mathcal{H} , and whose dynamic is governed by an Hamiltonian H with eigenstates $|E_\alpha\rangle$, the macroscopic observables form a (complete) set of commuting operators

$$\{\widehat{M}_i\}_{i=1,\dots,N}, \quad (3.1)$$

according to the well known analysis of von Neumann[27, 28]. These observables have a common set of eigenspaces \mathcal{H}_ν identified by the eigenvalues of the operators \widehat{M}_i . Therefore such a set of operators generates naturally a linear decomposition of the Hilbert space

$$\mathcal{H} = \bigoplus_{\nu} \mathcal{H}_\nu, \quad (3.2)$$

where each \mathcal{H}_ν represents a macroscopic state of the system. In the following we shall denote by (\mathcal{H}_ν) orthogonal decomposition such as that given by (3.2).

Let's assume that the Hilbert space \mathcal{H} has finite dimensionality D . This is an adequate assumption in the case of a system with finite volume and with bounded energy, which is the case for a macroscopic quantum system. Let's furthermore consider the equilibrium macrostate \mathcal{H}_{eq} for which the fundamental property

$$D \approx d_{eq} \quad (3.3)$$

holds (we recall that the quantum entropy of a macrostate represented by \mathcal{H}_ν is given by $k_B \log d_\nu$, where d_ν is the dimension of \mathcal{H}_ν and k_B is the Boltzmann constant).

Most of the vectors $|\psi\rangle \in \mathcal{H}_{[E, E+\delta]}$ are practically in the equilibrium macrostate, the precise sense of "most" is expressed by $u_E(d\psi)$, the uniform measure on $\mathcal{H}_{[E, E+\delta]}$. In fact, since by (3.3)

$$\int \langle \psi | P_{eq} | \psi \rangle u_E(d\psi) = \frac{1}{D} \text{Tr} [P_{eq}] = \frac{d_{eq}}{D} \approx 1 \quad (3.4)$$

and the quantity $\langle \psi | P_{eq} | \psi \rangle$ is upper bounded by 1, we have for most $|\psi\rangle$

$$\langle \psi | P_{eq} | \psi \rangle \sim 1. \quad (3.5)$$

From a microscopic point of view the state of the system at time t is described as usual by a vector $|\psi(t)\rangle \in \mathcal{H}$. If the eigenstates $|E_\alpha\rangle$ of the Hamiltonian H are non-degenerate, $|\psi(t)\rangle$ can be expanded into the eigenstates of H ,

$$|\psi(t)\rangle = \sum_{\alpha=1}^D A_\alpha e^{-iE_\alpha t} |E_\alpha\rangle, \quad (3.6)$$

where $A_\alpha = \langle \psi(0) | E_\alpha \rangle$.

The probability to find the quantum system in the equilibrium macrostate at time t is

$$\langle \psi(t) | P_{eq} | \psi(t) \rangle = \sum_{\alpha, \beta} A_\alpha^* A_\beta \langle E_\alpha | P_{eq} | E_\beta \rangle e^{-i(E_\alpha - E_\beta)t}. \quad (3.7)$$

If, in addition to non-degeneracy of H , we assume the absence of resonances for its eigenvalues¹, it is easy to see that the long time average of (3.7) is

$$\overline{\langle \psi(t) | P_{eq} | \psi(t) \rangle} = \frac{1}{T} \int_0^T dt \langle \psi(t) | P_{eq} | \psi(t) \rangle = \sum_{\alpha=1}^D |A_\alpha|^2 \langle E_\alpha | P_{eq} | E_\alpha \rangle. \quad (3.8)$$

The main goal of the present work is to show the approach to equilibrium for a macroscopic quantum system. It is easy to see that a sufficient condition for this is

$$\overline{\langle \psi(t) | P_{eq} | \psi(t) \rangle} \approx 1. \quad (3.9)$$

In fact, since $\overline{\langle \psi(t) | P_{eq} | \psi(t) \rangle}$ is upper bounded by 1, having the long time average nearly one means that, whatever is the initial state, the system will reach the equilibrium macrostate in a short time, and it will remain in it for most of its subsequent evolution.

By expliciting the left hand side of (3.8), (3.9) becomes

$$\sum_{\alpha=1}^D |A_\alpha|^2 \langle E_\alpha | P_{eq} | E_\alpha \rangle \approx 1. \quad (3.10)$$

Clearly, this condition can be fulfilled by imposing ‘‘regularity’’ conditions on at least one of the two terms involved: $|A_\alpha|^2$ or $\langle E_\alpha | P_{eq} | E_\alpha \rangle$.

To demand a regularity condition on $|A_\alpha|^2$ amounts to a condition on the initial state of the system,

$$|\psi(0)\rangle = \sum_{\alpha} A_\alpha |E_\alpha\rangle. \quad (3.11)$$

On the other hand, a condition on $\langle E_\alpha | P_{eq} | E_\alpha \rangle$ amounts to a condition on the Hamiltonian (in addition to the two previous conditions that we have already assumed on H).

¹The absence of resonances means that the eigenvalues are not rationally related, namely $nE_\alpha \neq mE_\beta$ for all α, β and $n, m \in \mathbb{N}$.

Our goal is to show convergence to equilibrium for *any* initial state. Thus, we shall investigate which conditions on $\langle E_\alpha | P_{eq} | E_\alpha \rangle$ the system should satisfy. To start this investigation, let us consider the following situation. Suppose that the initial microstate (3.11) is such that the A_α are all equal to 0, except for a small number of them that we shall denote with A_{α^*} . If

$$\langle E_{\alpha^*} | P_{eq} | E_{\alpha^*} \rangle \sim 0, \quad (3.12)$$

where the $|E_{\alpha^*}\rangle$ are the eigenstates of H associated with A_{α^*} , then

$$\overline{\langle \psi(t) | P_{eq} | \psi(t) \rangle} = \sum_{\alpha^*} |A_{\alpha^*}|^2 \langle E_{\alpha^*} | P_{eq} | E_{\alpha^*} \rangle \sim 0. \quad (3.13)$$

As consequence the system will stay out of equilibrium for most of the time. Thus a sufficient condition for approach to equilibrium, is that

$$\langle E_\alpha | P_{eq} | E_\alpha \rangle \geq 1 - \delta \quad (3.14)$$

with $\delta \ll 1$. We stress that this condition is in general a condition on the Hamiltonian H .

3.2 Setting the Stage of Convergence to Equilibrium in a Concrete Case

To understand the problem of the approach to equilibrium for a macroscopic quantum system it is useful to start by focusing our attention on a very simple, but physically relevant example.

Let's consider a system composed of two identical subsystems, designated 1 and 2, with Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. The Hamiltonian of the total system is

$$H = H_1 + H_2 + \epsilon V, \quad (3.15)$$

where H_1 and H_2 are the Hamiltonians of subsystems 1 and 2 respectively, and ϵV is the interaction energy describing the coupling between the two subsystems. Under the assumptions of finite dimensions of the systems Hilbert spaces, we may regard V as a matrix with entries of order 1 and size of the coupling determined by ϵ .

Since the system is of macroscopic size and isolated, its state is in a shell of energy $[E, E + \delta]$, where E is the constant thermodynamic energy of the system and δ is small on a macroscopic scale, $\delta \ll E$, but large enough for the interval $[E, E + \delta]$ to contain very many eigenvalues. Let $\mathcal{H}_{[E, E + \delta]}$ be the subspace of \mathcal{H} associated with the spectral interval $[E, E + \delta]$ and D_E its dimensionality. From now on, when no ambiguity will arise, we shall drop the subscript E and simply denote by D the dimension of $\mathcal{H}_{[E, E + \delta]}$.

By neglecting on the macroscopic scale the interaction energy ϵV , one may consider as macroscopic observables (3.1) the Hamiltonians H_1 and H_2 . The eigenvalues of H_1 and H_2 are denoted by E_i^1 and E_j^2 , respectively, and $|E_i^1\rangle$ and $|E_j^2\rangle$ are their corresponding eigenvectors. Accordingly, $|E_{ij}\rangle^0 = |E_i^1\rangle \otimes |E_j^2\rangle$ is an eigenstate of the unperturbed Hamiltonian

$$H_0 = H_1 + H_2. \quad (3.16)$$

Clearly, since the system is isolated (and we are neglecting the contribution of the potential energy to the total energy), when the system is in the state $|E_i^1\rangle \otimes |E_j^2\rangle$ the total energy is

$$E \simeq E_i^1 + E_j^2.$$

In equilibrium the energies E_i^1 and E_j^2 of the subsystems should be such that $E_i^1 \simeq E_j^2 \simeq E/2$. Thus the equilibrium macrostate is associated with the following subspace of $\mathcal{H}_{[E, E + \delta]}$

$$\mathcal{H}_{eq}^E = \left\{ |\psi\rangle \in \mathcal{H}_{[E, E + \delta]} \mid |\psi\rangle = \sum_{ij} C_{ij} |E_i^1\rangle \otimes |E_j^2\rangle, E_i^1 \simeq E_j^2, E_i^1 + E_j^2 \simeq E \right\}. \quad (3.17)$$

In the following we shall call P_{eq} the projection operator on \mathcal{H}_{eq}^E .

Let's consider the system starting from a non-equilibrium initial state $|\psi(0)\rangle \in \mathcal{H}_{[E, E + \delta]}$ of the following form

$$|\psi(0)\rangle = |E_{in}^1\rangle \otimes |E_{in}^2\rangle, \quad (3.18)$$

where E_{in}^1 and E_{in}^2 are macroscopically different energies. Considering that the system is isolated, we have obviously that $E_{in}^1 + E_{in}^2 \simeq E$ where E is the energy of the shell. From an empirical point of view we can think of preparing

two similar systems with different energies, e.g., two metal blocks at different temperatures, and then to couple them.

The basic problem is to understand under what conditions the initial state (3.18) reaches the equilibrium macrostate, in the sense of (3.9) and (3.10).

Since the macroscopic observables we are considering are the energies of the two systems (with the initial microstate corresponding to a non-equilibrium macrostate of macroscopically different energies), it is convenient to write the eigenvectors $\{|E_\alpha\rangle\}$ of H in terms of the eigenvectors $\{|E_\beta\rangle^0\}$ of H_0 , i.e.,

$$|E_\alpha\rangle = \sum_{\beta=1}^D U_{\alpha\beta} |E_\beta\rangle^0. \quad (3.19)$$

From equation (3.19) it follows that the quantity $\langle E_\alpha | P_{eq} | E_\alpha \rangle$ can be written in terms of the unitary matrix U , namely,

$$\langle E_\alpha | P_{eq} | E_\alpha \rangle = \sum_{\beta=1}^D |U_{\alpha\beta}|^2 \langle E_\beta | P_{eq} | E_\beta \rangle^0 + \sum_{\beta \neq \beta'} U_{\alpha\beta'}^* U_{\alpha\beta} \langle E_{\beta'} | P_{eq} | E_\beta \rangle^0. \quad (3.20)$$

Considering that the definition (3.17) of the equilibrium macrostate \mathcal{H}_{eq} is given in terms of the unperturbed Hamiltonian eigenstates $\{|E_\beta\rangle^0\}$, clearly we have

$$\langle E_{\beta'} | P_{eq} | E_\beta \rangle^0 = 0 \text{ for } \beta' \neq \beta. \quad (3.21)$$

Therefore equation (3.20) becomes

$$\langle E_\alpha | P_{eq} | E_\alpha \rangle = \sum_{\beta=1}^D |U_{\alpha\beta}|^2 \langle E_\beta | P_{eq} | E_\beta \rangle^0 = \sum_{\beta=1}^{d_{eq}} |U_{\alpha\beta}|^2. \quad (3.22)$$

Recalling (3.9), (3.10) and (3.14), one sees that a sufficient condition for convergence to equilibrium is that

$$\langle E_\alpha | P_{eq} | E_\alpha \rangle = \sum_{\beta=1}^{d_{eq}} |U_{\alpha\beta}|^2 \geq 1 - \delta, \quad (3.23)$$

where $\delta \ll 1$.

Since the unitary matrix U connects the eigenbasis $\{|E_\alpha\rangle\}$ of $H = H_0 + \epsilon V$ with the eigenbasis $\{|E_\alpha\rangle^0\}$ of H_0 , condition (3.23) expresses how the two subsystems are coupled by ϵV , a property that one may call “goodness of the

coupling.” Then to prove convergence to equilibrium, one could show (3.23) for a specific model, i.e., for an explicit Hamiltonian H . Another approach is to establish that (3.23) is indeed a typical property.

The relevant notion for expressing typicality of the unitary matrix U in (3.23), is the Haar measure on the unitary group $\mathcal{U}(D)$, where D is the dimension of the energy shell $\mathcal{H}_{[E, E+\delta]}$. By normalizing this measure to 1, one obtains a probability measure on $\mathcal{U}(D)$ that we shall denote by \mathbb{P} . Thus, (3.23) is typical if the set A of unitary matrices in $\mathcal{U}(D)$ that do satisfy (3.23), for all α , has \mathbb{P} -measure near one, i.e.,

$$\mathbb{P}(A) = 1 - \delta, \quad (3.24)$$

where $\delta \ll 1$. Given that a typical U relates unitarily the eigenstates of H with those of H_0 , we may also regard \mathbb{P} as providing a notion of typical energy basis and, thus, of typical Hamiltonian.

If the interaction energy ϵV in (3.15) is not particularly contrived (and, of course, ϵ is different from 0 and not too small, see section 3.5), one expects that H , given by (3.15), is typical. Therefore to show (3.24), for this approach, is the main step for proving that the composite system thermalizes, i.e., that the two metal blocks do indeed reach the same final temperature.

3.3 A General Framework for Convergence to Equilibrium

The concrete case considered in the previous section can be generalized and one may arrive to a more general framework for studying the problem of convergence to equilibrium. One should simply observe that the energy H_0 , as macroscopic observable, does not play any special role in our analysis. Therefore, the ideas put forward at the end of the previous section can be extended to cover the case of an arbitrary set $\{\widehat{M}_i\}$ of N macroscopic commuting observables \widehat{M}_i , as we have already mentioned at the beginning of section 3.1.

One may write the eigenstates $\{|E_\alpha\rangle\}$ of Hamiltonian H in terms of the eigenvectors $\{|m_\beta\rangle\}$ of the macroscopic observables \widehat{M}_i , which is a fixed basis

for the macroscopic decomposition (\mathcal{H}_ν) associated with the commuting set $\{\widehat{M}_i\}$, namely:

$$|E_\alpha\rangle = \sum_{\beta} U_{\alpha\beta} |m_\beta\rangle \quad (3.25)$$

where U is a unitary matrix. Equation (3.25) is analogous to (3.19). Thus, in a completely analogous way one arrives at the equation corresponding to (3.22), not only for the equilibrium macrostate, but for a generic macrostate \mathcal{H}_ν , i.e.,

$$\langle E_\alpha | P_\nu | E_\alpha \rangle = \sum_{\beta=1}^D |U_{\alpha\beta}|^2 \langle E_\beta | P_\nu | E_\beta \rangle^0 = \sum_{\beta=1}^{d_\nu} |U_{\alpha\beta}|^2. \quad (3.26)$$

where now $U_{\alpha\beta}$ is given by (3.25).

Letting \mathbb{P} be, as before, the normalized probability measure on $\mathcal{U}(D)$, we have the following:

Lemma 3.3.1. *There exist positive numbers k and C such that for $0 < \epsilon < 2$ for any $\delta > 0$, there exists D_0 such that for all $D > D_0$, for all linear decomposition (\mathcal{H}_ν) and for all \mathcal{H}_ν such that*

$$d_\nu > C/\epsilon^2 \log(2kD/\delta) \quad (3.27)$$

is valid the following:

$$\mathbb{P}\left\{\bigcap_{\alpha=1}^D \left\{ \left| \langle E_\alpha | P_\nu | E_\alpha \rangle - \frac{d_\nu}{D} \right| < \epsilon \right\}\right\} > 1 - \delta. \quad (3.28)$$

The proof is given in section 3.3.1. We shall now provide some comments and remarks concerning the above Lemma.

- For the case of the equilibrium macrostate Lemma 3.3.1 and condition (3.3) immediately implies condition (3.14) and therefore the overwhelming majority of coupling are “good”.
- For general macrostates the meaning of Lemma 3.3.1 is that *the probability to find any stationary state in a macroscopic state is typically equivalent to the microcanonical probability of that specific macroscopic state.*

- Lemma 3.3.1 is not valid for a generic \mathcal{H}_ν , but only for those which satisfies the “macroscopicity” condition (3.27). In fact, if the subspace \mathcal{H}_ν is not great enough, equation (3.28) is no longer valid. The condition (3.27) is indeed reasonable, since it links the dimensionality of a macroscopic state to the dimensionality of the whole system. In general it could even be regarded as a necessary condition for a given subspace \mathcal{H}_ν to be considered a macrostate.
- Lemma 3.3.1 resembles to, in some way, the Schnirelman Theorem [6], which states that, in the semiclassical limit and with suitable hypothesis, the Wigner distribution corresponding to an eigenstate $|E_\alpha\rangle$ is the microcanonical measure. In fact, as we shall see in the proof, the property (3.28) is a sort of “coarse grained” uniformity of the modulus squared of $U_{\alpha\beta}$ appearing in (3.25). Like in the Schnirelman theorem we have a “coarse grained” (i.e. semiclassical) uniform distribution corresponding to these eigenstates. This analogy could be important to understand the nature of the Lemma 3.3.1, in a different context.

3.3.1 Proof of Lemma 3.3.1

Proof. Let’s start observing that for the probability on the left hand side of (3.28) is valid

$$\mathbb{P}\left\{\bigcap_{\alpha=1}^D\left\{\left|\langle E_\alpha|P_\nu|E_\alpha\rangle - \frac{d_\nu}{D}\right| < \epsilon\right\}\right\} = 1 - \mathbb{P}\left\{\bigcup_{\alpha=1}^D\left\{\left|\langle E_\alpha|P_\nu|E_\alpha\rangle - \frac{d_\nu}{D}\right| > \epsilon\right\}\right\}, \quad (3.29)$$

and considering that by (3.25)

$$\begin{aligned} \mathbb{P}\left\{\bigcup_{\alpha=1}^D\left\{\left|\langle E_\alpha|P_\nu|E_\alpha\rangle - \frac{d_\nu}{D}\right| > \epsilon\right\}\right\} &\leq D \mathbb{P}\left\{\left|\langle E_\alpha|P_\nu|E_\alpha\rangle - \frac{d_\nu}{D}\right| > \epsilon\right\} = \\ &= D \mathbb{P}\left\{\left|\sum_{\gamma=1}^{d_\nu} |U_{\alpha\gamma}|^2 - \frac{d_\nu}{D}\right| > \epsilon\right\}, \quad (3.30) \end{aligned}$$

we can conclude that for the left hand side of (3.28) we have

$$\mathbb{P}\left\{\bigcap_{\alpha=1}^D\left\{\left|\langle E_\alpha|P_\nu|E_\alpha\rangle - \frac{d_\nu}{D}\right| < \epsilon\right\}\right\} \geq 1 - D \mathbb{P}\left\{\left|\sum_{\gamma=1}^{d_\nu} |U_{\alpha\gamma}|^2 - \frac{d_\nu}{D}\right| > \epsilon\right\}. \quad (3.31)$$

Secondly we recall that any random unit vector \vec{U}_α , like the one present in the right hand side of (3.31)², can be regarded as

$$\vec{U}_\alpha = \frac{\vec{G}_\alpha}{\|\vec{G}_\alpha\|}, \quad (3.32)$$

where \vec{G}^α is a random gaussian vector with entries of mean 0. We consider for simplicity the element of \vec{G}^α as complex gaussian random variables with mean 0 and variance $1/D$, the same of the entries of the unitary vector \vec{U}_α . As consequence we can reduce the calculation of the probability on the left hand side of (3.28) in terms of gaussian probabilities, i.e.:

$$\mathbb{P}\left\{\bigcap_{\alpha=1}^D \left\{ \left| \langle E_\alpha | P_\nu | E_\alpha \rangle - \frac{d_\nu}{D} \right| < \epsilon \right\}\right\} \geq 1 - D \mathbb{P}\left\{ \left| \sum_{\gamma=1}^{d_\nu} \frac{|G_\gamma^\alpha|^2}{\|\vec{G}^\alpha\|^2} - \frac{d_\nu}{D} \right| > \epsilon \right\}. \quad (3.33)$$

To calculate the probability on the right hand side of (3.33) we need to introduce three different events:

$$A_D(\eta) := \left\{ \left| \|\vec{G}^\alpha\|^2 - 1 \right| < \eta \right\}, \quad (3.34)$$

$$B_D(\eta') := \left\{ \left| \sum_{\gamma=1}^{d_\nu} \frac{|G_\gamma^\alpha|^2}{\|\vec{G}^\alpha\|^2} - \frac{d_\nu}{D} \right| / \frac{d_\nu}{D} < \eta' \right\}, \quad (3.35)$$

$$C_D(\eta'') := \left\{ (1 - \eta'') \frac{d_\nu}{D} \leq \sum_{\gamma=1}^{d_\nu} |G_\gamma^\alpha|^2 \leq (1 + \eta'') \frac{d_\nu}{D} \right\}. \quad (3.36)$$

where clearly our target is to find a lower bound for

$$\mathbb{P}(B_D(\epsilon)) = 1 - \mathbb{P}\left\{ \left| \sum_{\gamma=1}^{d_\nu} \frac{|G_\gamma^\alpha|^2}{\|\vec{G}^\alpha\|^2} - \frac{d_\nu}{D} \right| > \epsilon \right\}. \quad (3.37)$$

If we consider the event $A_D(\eta)$, we see that implies

$$1 - \eta < \frac{1}{\|\vec{G}^\alpha\|^2} < 1 + 2\eta \quad \text{for} \quad \eta < \frac{1}{2}, \quad (3.38)$$

²More precisely it is a column of an unitary matrix U , and we are considering d_ν component of such a column.

and consequently $A_D(\eta) \cap C_D(\eta'')$ implies

$$\frac{d_\nu}{D}(1-\eta)(1-\eta'') < \frac{\sum_{\gamma=1}^{d_\nu} |G_\gamma^\alpha|^2}{\|G_\gamma^\alpha\|^2} < \frac{d_\nu}{D}(1+\eta)(1+2\eta''). \quad (3.39)$$

It is easy to see that $A_D(\eta) \cap C_D(\eta'') \subseteq B_D(2\eta + \eta'' + 2\eta\eta'')$, so if we choose $\eta = \eta'' = \epsilon/4$ we get

$$A_D(\epsilon/4) \cap C_D(\epsilon/4) \subseteq B_D(3/4\epsilon + \epsilon^2/8) \subseteq B_D(\epsilon) \quad \text{for } \epsilon < 2 \quad (3.40)$$

and therefore we have a lower bound for the probability $\mathbb{P}(B_D(\epsilon))$, i.e.:

$$\mathbb{P}(A_D(\epsilon/4)) \cdot \mathbb{P}(C_D(\epsilon/4)) \leq \mathbb{P}(B_D(\epsilon)) \quad \text{for } \epsilon < 2. \quad (3.41)$$

To find an estimate of $\mathbb{P}(A_D(\epsilon/4))$ and $\mathbb{P}(C_D(\epsilon/4))$ we are going to use the *Large Deviations Principle* [9]. It states that for a sequence of N i.i.d. random variable X_i it is valid the following

$$\mathbb{P}\left(\left|\sum_{i=1}^N \frac{X_N}{N} - E(X)\right| > \delta\right) \leq ke^{-NI(E(X)+\delta)}, \quad (3.42)$$

where k is a suitable constant and $I(x)$ is the *rate function* [9] associated to distribution of the random variable X .

In the case of A_D , considering the components $G_1^\alpha, \dots, G_D^\alpha$ of the vector \vec{G}^α as a sequence of D i.i.d gaussians, by (3.42) we get:

$$\begin{aligned} \mathbb{P}(\Omega \setminus A_D(\epsilon/4)) &= \mathbb{P}\left\{\left|\|\vec{G}^\alpha\|^2 - 1\right| > \epsilon/4\right\} = \\ &= \mathbb{P}\left\{\left|\sum_{\gamma=1}^D |G_\gamma^\alpha|^2 - D\frac{1}{D}\right| > \epsilon/4\right\} \leq ke^{-D I_D[(1+\epsilon/4)/D]}, \end{aligned} \quad (3.43)$$

where we recall that $E(|G_\gamma^\alpha|^2) = 1/D$ for all γ , and it is a simple exercise to see that the rate function associated to the random variable $|G_\gamma^\alpha|^2$ is

$$I_D(x) = \left[(Dx - 1) - \log(Dx) \right]. \quad (3.44)$$

Therefore $\mathbb{P}(\Omega \setminus A_D(\epsilon/4))$ becomes

$$\begin{aligned} \mathbb{P}\left\{\left|\|G^{\vec{\alpha}}\|^2 - 1\right| > \frac{\epsilon}{4}\right\} &\leq k \exp\left(-D\left[\frac{\epsilon}{4} - \log\left(1 + \frac{\epsilon}{4}\right)\right]\right) \leq \\ &\leq k \exp\left(-\frac{D\epsilon^2}{48}\right), \end{aligned} \quad (3.45)$$

from which follows

$$\mathbb{P}(A_D(\epsilon/4)) \geq 1 - k \exp\left(-\frac{D\epsilon^2}{48}\right), \quad (3.46)$$

where $k > 0$.

It is possible to use a completely analogous argument also for

$$\mathbb{P}(\Omega \setminus C_D(\epsilon/4)) = \mathbb{P}\left\{\left|\sum_{\gamma=1}^{d_\nu} |G_\gamma^\alpha|^2 - \frac{d_\nu}{D}\right|/\frac{d_\nu}{D} \geq \epsilon/4\right\}, \quad (3.47)$$

and so we can obtain

$$\mathbb{P}(C_D(\epsilon/4)) \geq 1 - k' \exp\left(-\frac{d_\nu \epsilon^2}{48}\right), \quad (3.48)$$

with $k' > 0$

We can now proceed to the final step of this proof. In fact by (3.41), (3.46) and (3.48) it follows that

$$\mathbb{P}(B_D(\epsilon)) \geq \left(1 - k \exp\left(-\frac{d_\nu \epsilon^2}{48}\right)\right) \left(1 - k' \exp\left(-\frac{D\epsilon^2}{48}\right)\right), \quad (3.49)$$

therefore by (3.33) and (3.37) we finally have

$$\begin{aligned} \mathbb{P}\left\{\bigcap_{\alpha=1}^D \left\{\left|\langle E_\alpha | P_\nu | E_\alpha \rangle - \frac{d_\nu}{D}\right| < \epsilon\right\}\right\} &\geq \\ &\geq 1 - D \left(1 - \left(1 - k \exp\left(-\frac{d_\nu \epsilon^2}{48}\right)\right) \left(1 - k' \exp\left(-\frac{D\epsilon^2}{48}\right)\right)\right) = \\ &= 1 - D \left(k \exp\left(-\frac{d_\nu \epsilon^2}{48}\right) + k' \exp\left(-\frac{D\epsilon^2}{48}\right) - k k' \exp\left(-\frac{(d_\nu + D)\epsilon^2}{48}\right)\right). \end{aligned} \quad (3.50)$$

Clearly, looking at the last two terms of the last line above, they approaches 0 when D tends to ∞ , then $\exists \delta > 0$ such that $\forall D > D_0$

$$D \left(k' \exp\left(-\frac{D\epsilon^2}{48}\right) - k k' \exp\left(-\frac{(d_\nu + D)\epsilon^2}{48}\right)\right) < \frac{\delta}{2}, \quad (3.51)$$

but for the first term inside the parenthesis we need to add a further condition, in fact $\exists \delta > 0$ such that $\forall D > D_0 > D'_0$

$$k \exp\left(-\frac{d_\nu \epsilon^2}{48} + \log D\right) \leq \frac{\delta}{2}, \quad (3.52)$$

only if

$$d_\nu \geq \frac{C}{\epsilon^2} \log\left(\frac{2kD}{\delta}\right), \quad (3.53)$$

where $C > 0$ is a suitable constant. In such a case we get finally

$$\mathbb{P}\left\{\bigcap_{\alpha=1}^D \left\{\left|\langle E_\alpha | P_\nu | E_\alpha \rangle - \frac{d_\nu}{D}\right| < \epsilon\right\}\right\} \geq 1 - \delta, \quad (3.54)$$

which completes the proof. \square

3.4 The Main Theorem on Convergence to Equilibrium

Since Lemma 3.3.1 gives an important information about the macroscopic probability of stationary states, it is interesting to see what are the consequences for the evolution in time of a generic $|\psi\rangle \in \mathcal{H}$. A very important, preliminary step, is the following:

Lemma 3.4.1. *There exist positive numbers k and C such that for $0 < \epsilon < 2$ and for any $\delta > 0$, there exists D_0 such that for all $D > D_0$, for all linear decomposition (\mathcal{H}_ν) , for any \mathcal{H}_ν such that $d_\nu > C/\epsilon^2 \log(2kD/\delta)$ and for all initial states $|\psi(0)\rangle$ is valid the following:*

$$\mathbb{P}\left\{\left|\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt \left(\langle \psi(t) | P_\nu | \psi(t) \rangle - \frac{d_\nu}{D}\right)\right| < \epsilon\right\} > 1 - \delta. \quad (3.55)$$

Proof. Given the time evolution of a quantum microstate when there is no degeneracy of the Hamiltonian eigenstates (3.6), we can write the time average of $\langle \psi(t) | P_\nu | \psi(t) \rangle$ as:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle \psi(t) | P_\nu | \psi(t) \rangle dt =$$

$$= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \sum_{\alpha, \beta} e^{-i(E_\alpha - E_\beta)t} A_\beta^* A_\alpha \langle E_\beta | P_\nu | E_\alpha \rangle$$

In the case of no-resonances for the Hamiltonian in the time average only the terms with $E_\alpha = E_\beta$ survive, i.e.:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle \psi(t) | P_\nu | \psi(t) \rangle dt = \sum_\alpha |A_\alpha|^2 \langle E_\alpha | P_\nu | E_\alpha \rangle$$

Considering that $\sum_\alpha |A_\alpha|^2 = 1$, clearly

$$\mathbb{P} \left\{ \bigcap_{\alpha=1}^D \left\{ \left| \langle E_\alpha | P_\nu | E_\alpha \rangle - \frac{d_\nu}{D} \right| < \epsilon \right\} \right\} \geq \mathbb{P} \left\{ \left| \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt \left(\langle \psi(t) | P_\nu | \psi(t) \rangle - \frac{d_\nu}{D} \right) \right| < \epsilon \right\} \quad (3.56)$$

and by Lemma 3.3.1 this inequality suffices to complete the proof. \square

The relevant aspect of this theorem is its validity for any initial microscopic state $|\psi(0)\rangle$, since $\langle E_\alpha | P_\nu | E_\alpha \rangle \approx d_\nu/D$ is true for all α and not only for most of them.

Since (3.3) implies straightforwardly (3.27) for a suitable choice of the constants C and k , Lemma 3.3.1 and Lemma 3.4.1 are valid for the equilibrium macrostate. Therefore we have the following:

Theorem 3.4.2. *For $0 < \epsilon < 2$ and for any $\delta > 0$ there exists D_0 such that for all $D > D_0$, for all initial states $|\psi(0)\rangle$ and for any linear decomposition (\mathcal{H}_ν) which contains \mathcal{H}_{eq} (defined by (3.3)) is valid the following:*

$$\mathbb{P} \left\{ \left| \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt \left(\langle \psi(t) | P_{eq} | \psi(t) \rangle - \frac{d_{eq}}{D} \right) \right| < \epsilon \right\} > 1 - \delta. \quad (3.57)$$

By considering that $\langle \psi(t) | P_{eq} | \psi(t) \rangle$ is upper bounded by 1, and recalling (3.3), the important consequence of Theorem 3.4.2 is that *for a typical Hamiltonian a quantum system spends in the equilibrium macrostate most of its time, regardless of the initial state.*

3.5 A Simple Counter-Example to Convergence to Equilibrium

We shall discuss here a simple example of atypical behavior in the concrete framework provided by Hamiltonian (3.15) and initial state (3.18).

Trivially in the case of $\epsilon = 0$ the subsystems Hamiltonians H_1 and H_2 commute with H , therefore the system does not reach equilibrium. Suppose now that $\epsilon \neq 0$ but very small respect to the energy levels spacing ΔE . Then also in this case it is easy to see that the system doesn't reach equilibrium, and we shall give now a simple explanation of this fact.

For a macroscopic system (that is with a huge number of particles) the ΔE is very small with respect to the total energy, so small that in many cases we can regard the energy distribution as continuous. Then the condition $\epsilon \ll \Delta E$ means a coupling so weak to be practically negligible, so that it becomes impossible for the system to approach equilibrium. More explicitly one may argue as follows.

Let $|\widehat{\psi}(0)\rangle$ be the eigenstate of H associated with the eigenvalue $E \simeq E_i^1 + E_j^2$ (the perturbed eigenstate), and $|\psi(0)\rangle$ be the eigenvector of H_0 associated with the eigenvalues $E = E_i^1 + E_j^2$ (the unperturbed eigenstate). A very weak assumption is continuity on ϵ of the function $f(\epsilon) = \||\psi(0)\rangle - |\widehat{\psi}(0)\rangle\|$ (no need to demand analyticity). Thus, for $\epsilon \ll \Delta E$ we have that $|\psi(0)\rangle \simeq |\widehat{\psi}(0)\rangle$, and this is also true when the system evolves, then we can write $|\psi(t)\rangle \simeq |\widehat{\psi}(0)\rangle$ for all t .

Let's now see that (3.14) is violated. Since the initial state is out of equilibrium, i.e., $\langle\psi(0)|P_{eq}|\psi(0)\rangle = 0$, by continuity we have

$$\langle\widehat{\psi}(0)|P_{eq}|\widehat{\psi}(0)\rangle \simeq \langle\psi(0)|P_{eq}|\psi(0)\rangle = 0. \quad (3.58)$$

Then also $|\widehat{\psi}(0)\rangle$ is out of equilibrium. Then, by the stationarity of $|\widehat{\psi}(0)\rangle$, i.e., $|\psi(t)\rangle \simeq |\widehat{\psi}(0)\rangle$, we have

$$\langle\psi(t)|P_{eq}|\psi(t)\rangle \simeq \langle\widehat{\psi}(0)|P_{eq}|\widehat{\psi}(0)\rangle \quad \text{for all } t \quad (3.59)$$

whence, since $|\psi(t)\rangle$ the system has to stay near $|\widehat{\psi}(0)\rangle$, we have

$$\langle\psi(t)|P_{eq}|\psi(t)\rangle \simeq 0 \quad \text{for all } t. \quad (3.60)$$

The foregoing provides a very simple example of "bad" couplings, but according to Lemma 3.3.1, the "bad" Hamiltonians, i.e. those with "bad" couplings, form a negligible set of all Hamiltonians (those characterized "good" couplings).

3.6 Comparing different approaches to Convergence to Equilibrium

To complete the framework it is interesting to see and compare the different approaches on the study of the convergence to equilibrium of macroscopic quantum system.

3.6.1 An historical overview

Our work follows the line traced by von Neumann in his article about the quantum H-Theorem [27] and continued by Bocchieri and Loinger [3, 4]. Their approach was based on studying the properties of macroscopic quantum system by averaging with respect to the Haar measure on the Unitary group $\mathcal{U}(D)$, since this group is isomorphic to the space of all the orthonormal basis on \mathcal{H} , as we shall explain next.

They noticed that given the fixed orthonormal basis $\{|m_1\rangle, \dots, |m_D\rangle\}$ of eigenstates of macrostates \widehat{M}_i , which is also a basis for

$$\mathcal{H} = \bigoplus_{\nu} \mathcal{H}_{\nu},$$

the probability to find a given state $|\psi\rangle$ in macrostate \mathcal{H}_{ν} is

$$\langle\langle \psi | P_{\nu} | \psi \rangle\rangle_U = \int_{\mathcal{U}(D)} dU \sum_{\alpha=1}^{d_{\nu}} \langle \psi | U | m_{\alpha} \rangle \langle m_{\alpha} | U^{-1} | \psi \rangle, \quad (3.61)$$

where P_{ν} is the projection on \mathcal{H}_{ν} and the summation is intended over all $|m_{\alpha}\rangle$ such that $\langle m_{\alpha} | P_{\nu} | m_{\alpha} \rangle = 1$. Their aim was to calculate the long time relative time frequency to be in the macrostate \mathcal{H}_{ν} , i.e.

$$\langle \psi(t) | P_{\nu} | \psi(t) \rangle = \frac{1}{T} \int_0^T dt \langle \psi(t) | P_{\nu} | \psi(t) \rangle = \sum_{\alpha=1}^D |A_{\alpha}|^2 \langle E_{\alpha} | P_{\nu} | E_{\alpha} \rangle. \quad (3.62)$$

To this end they used the formula (3.61) and other considerations that we shall now explain.

Von Neumann's observed that any "macroscopic basis" $\{U|m_1\rangle, \dots, U|m_D\rangle\}$, where $U \in \mathcal{U}(D)$, can be associated to a

so called “macroscopic observer” [27]. Indeed, by considering the right hand side of equation (3.62), von Neumann demanded a “regularity” condition on $\langle E_\alpha | P_{eq} | E_\alpha \rangle$, analogously to what we did (see equation (3.10) and below). However, instead of a condition on the Hamiltonian, and in particular on the stationary states, he assumed that the basis $\{|m_\alpha\rangle\}$ is typical.

Under this condition together with the hypothesis of no-degeneracy and no-resonances of the Hamiltonian H , von Neumann showed that for any H , for any initial state $|\psi(0)\rangle$, for most of macroscopic basis $\{|m_\alpha\rangle\}$, for all \mathcal{H}_ν and for most instants of time:

$$\langle \psi(t) | P_\nu | \psi(t) \rangle \approx \frac{d_\nu}{D}. \quad (3.63)$$

Where “most macroscopic basis $\{|m_\alpha\rangle\}$ ” is equivalent to “most linear decompositions (\mathcal{H}_ν)”.

It is convenient to introduce the symbol \mathbb{V} to denote “for most of” , in terms of which von Neumann’s result can be expressed as

$$\forall H \quad \forall \psi_0 \quad \mathbb{V}(\mathcal{H}_\nu) \quad \forall \nu \quad \mathbb{V}t : \langle \psi(t) | P_\nu | \psi(t) \rangle \approx \frac{d_\nu}{D} \quad (3.64)$$

The use of the term “for most of” in combination with the almost equality symbol \approx in (3.64) requires a bit of explanation: let $p(x, y, \epsilon)$ be a proposition, then the statement

$$\forall x \quad \mathbb{V}y : p(x, y, \epsilon)$$

means

$$\forall \epsilon > 0 \quad \forall \delta > 0 \quad \exists D_0 \text{ such that } \forall D > D_0 \quad \forall x \in S_D : \mu_D\{y : p(x, y, \epsilon)\} > 1 - \delta$$

where the domain S_D and the measure μ_D can possibly depend by D .

For the case of (3.64) we have $\mathbb{V}t$ with respect to the uniform measure on $[0, \infty)$; $\forall \nu$ in the set of all the subspaces \mathcal{H}_ν ; $\mathbb{V}(\mathcal{H}_\nu)$ with respect to the Haar measure on $\mathcal{U}(D)$; $\forall \psi_0$ in the whole Hilbert space \mathcal{H} and $\forall H$ in the set of all non-degenerate and non-resonant Hamiltonians. Finally

$$p(H, \psi_0, (\mathcal{H}_\nu), \nu, t, \epsilon) := |\langle \psi(t) | P_\nu | \psi(t) \rangle - \frac{d_\nu}{D}| < \epsilon$$

Bocchieri and Loinger in [3] obtained the result

$$\forall H \quad \forall \psi_0 \quad \forall t \quad \mathbb{V}(\mathcal{H}_\nu) \quad \forall \nu : \langle \psi(t) | P_\nu | \psi(t) \rangle \approx \frac{d_\nu}{D} \quad (3.65)$$

which implies (3.64). In addition, they criticized the notion of macroscopic observer introduced by von Neumann. Bocchieri and Loinger argued that the physical relevant result should be based on a “typical” initial state and not on a “typical” macroscopic observer. In accordance with this, they put conditions on the $|A_\alpha|^2$ in equation (3.62) and obtained also the following result:

$$\forall H \quad \forall(\mathcal{H}_\nu) \quad \forall \psi_0 \quad \forall t \quad \forall \nu : \quad \langle \psi(t) | P_\nu | \psi(t) \rangle \approx \frac{d_\nu}{D} \quad (3.66)$$

Where $\forall \psi_0$ is with respect to the Haar Measure on $\mathcal{U}(D)$.

A result related to (3.66), but obtained with a completely different approach has been recently presented by Reimann in [23]. Instead of working with unitary matrices Reimann focused his attention on the time evolution of the expectation values:

$$\text{Tr}\{\rho(t)A\} \quad (3.67)$$

where $\rho(t)$ is the density matrix of the system at time t and A is a generic operator. Making the very reasonable assumptions that the quantity

$$R := \left[\sum \rho_{nn}^2(0) \right]^{1/3} \leq \left[\max_n \rho_{nn}^2 \right]^{1/3} \quad (3.68)$$

is negligibly small and that the observable A has a “finite range”, i.e.

$$\Delta_A := \max_{\mathcal{H}_+} \langle \psi | A | \psi \rangle - \min_{\mathcal{H}_+} \langle \psi | A | \psi \rangle = a_{max} - a_{min} \quad (3.69)$$

where

$$\mathcal{H}_+ := \text{span}\{|E_n\rangle | \rho_{nn} > 0\}, \quad (3.70)$$

he showed that

$$\text{Prob}(|\text{Tr}\{\rho(t)A\} - \text{Tr}\{\rho_{eq}A\}| \geq R) \quad (3.71)$$

where the equilibrium density matrix is the mixed state

$$\rho_{eq} := \sum \rho_{nn}(0) |E_n\rangle \langle E_n| \quad (3.72)$$

This result is related to (3.66). In fact the assumption (3.68) is valid “for most of ψ_0 ”, which means, looking at the right hand side of equation (3.62), that (3.68) is a condition on the factors $|A_\alpha|^2$.

3.6.2 Summary of our results

By using the notation previously introduced, our Lemma 3.3.1, Lemma 3.4.1 and Theorem 3.4.2 can be stated as follows.

Lemma 3.3.1.

$$\forall(\mathcal{H}_\nu) \quad \forall H \quad \forall \alpha : \langle E_\alpha | P_\nu | E_\alpha \rangle \approx \frac{d_\nu}{D}$$

Lemma 3.4.1

$$\forall(\mathcal{H}_\nu) \quad \forall |\psi(0)\rangle \quad \forall \nu \in S_D \quad \forall H : \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle \psi(t) | P_\nu | \psi(t) \rangle dt \approx \frac{d_\nu}{D} \quad (3.73)$$

where $S_D := \{\nu \mid d_\nu > C/\epsilon^2 \log(2kD/\delta)\}$

Theorem 3.4.2

$$\forall(\mathcal{H}_\nu) \quad \forall |\psi(0)\rangle \quad \forall \nu \in S_D \quad \forall H \quad \forall t : \langle \psi(t) | P_{eq} | \psi(t) \rangle \approx \frac{d_{eq}}{D} \quad (3.74)$$

where $S_D := \{\nu \mid d_\nu > C/\epsilon^2 \log(2kD/\delta)\}$

Our final goal is to establish the result

$$\forall(\mathcal{H}_\nu) \quad \forall \psi_0 \quad \forall H \quad \forall t \quad \forall \nu : \langle \psi(t) | P_\nu | \psi(t) \rangle \approx \frac{d_\nu}{D} . \quad (3.75)$$

where $S_D := \{\nu \mid d_\nu > C/\epsilon^2 \log(2kD/\delta)\}$

However in this chapter we have shown (3.75) for the particular case of $\nu = eq$. This result implies the approach to equilibrium for any initial state and for a “typical” Hamiltonian.

Chapter 4

Conclusions

In Chapter 1 it is provided an alternative, elementary proof of a theorem first established by Petz and Réffy in [19], concerning the joint distribution of the upper left $k \times k$ entries of a random unitary $n \times n$ matrix in the limit $n \rightarrow \infty$. This proof it will be useful in Chapter 2 and furthermore it will show a “typical” behavior for unitary matrices.

In Chapter 2 is given a sketch for the rigorous mathematical foundation of the discussion of GAP measures in [11], in particular for justifying that the thermal equilibrium distributions of wave functions are GAP measures. The main result is a consequence of Theorem 2.4.1, according to which a system weakly coupled to a heat bath has (asymptotically) $GAP(\rho_\beta)$ -distributed conditional wave function, with $\rho_\beta = (1/Z) \exp(-\beta H)$, if the wave function of the composite is typical relative to the microcanonical distribution.

The key property that remain to be proven is an extended version of Theorems 2.3.1, which states a similar result but not for fixed $\rho_1 = \text{Tr}_2 |\psi\rangle\langle\psi|$. At the moment the most promising way to reach this goal is to extend the convergence inside Theorems 2.3.1 from pointwise to uniform.

It is also studied the rigorous existence, uniqueness, and continuity of $GAP(\rho)$, also on infinite-dimensional spaces.

Finally in Chapter 3 it is studied the problem of the approach to equilibrium for macroscopic quantum systems. Starting from a particular case it is shown that it is possible to find a formulation valid for a generic quantum system. The result proven here states that, under suitable conditions, and for

a “typical” Hamiltonian, a quantum system reaches equilibrium regardless of the initial state.

Acknowledgments. I would like to thank all the people I've collaborated with and helped me to arrive at this result. First of all my tutor Nino Zanghì who guided and supported me for more than 4 years. Then all my Rutgers friends, Shelly Goldstein, Joel Lebowitz and Roderich Tumulka; without them this thesis wouldn't exist. I must thank also Valia Allori and her husband Davide for their invaluable support during my American life. Finally I'm grateful to David Rouelle (IHES), Mikko Stenlund (Courant Institute), Tobias Kuna (Bielefeld), Ernesto de Vito (Genova) and Enrico Massa (Genova) for helpful discussions.

Bibliography

- [1] J. Bernoulli, *Ars Conjectandi*, Thurnisius, Basel, 1713. Part IV translated in English by O. Sheynin, 2005, downloadable from www.sheynin.de.
- [2] P. Billingsley: *Probability and Measure*. John Wiley & Sons, 1979.
- [3] P. Bocchieri, A. Loinger: *Ergodic Theorem in Quantum Mechanics*, Physical Review **111**, 668, 1958.
- [4] P. Bocchieri, A. Loinger: *Ergodic Foundation of Quantum Statistical Mechanics*, Physical Review **114**, 948, 1959.
- [5] L. Boltzmann: *Vorlesungen über Gastheorie*, 2 vol, Leipzig, Barth, 1896, 1898. Translated in English by S.G. Brush: *Lectures on Gas Theory*, Cambridge University Press, London, 1964.
- [6] Y. Colin de Verdière: *Ergodicité et fonctions propres du laplacien*, Comm. Math. Phys. **102**, 497–502, 1985.
- [7] A.-A. Cournot, *Exposition de la théorie des chances et des probabilités*, Hachette, Paris, 1843. Reprinted in 1984 as Volume I (B. Bru, ed.) of Cournot (1973–1984).
- [8] D. Dürr, S. Goldstein, N. Zanghì: *Quantum equilibrium and the origin of absolute uncertainty*, J. Statist. Phys. **67**, 843–907, 1992.
- [9] R.S. Ellis, *Entropy, Large Deviations and Statistical Mechanics*, Springer, New York, 1985.

- [10] J. Gemmer, G. Mahler, M. Michel: *Quantum Thermodynamics: Emergence of Thermodynamic Behavior within Composite Quantum Systems*. Lecture Notes in Physics **657**. Berlin: Springer, 2004.
- [11] S. Goldstein, J.L. Lebowitz, R. Tumulka, N. Zanghì: *On the Distribution of the Wave Function for Systems in Thermal Equilibrium*, J. Statist. Phys. **125**, 1193–1221, 2006. quant-ph/0309021
- [12] S. Goldstein, J.L. Lebowitz, C. Mastrodonato, R. Tumulka, N.Zanghì: *Typicality of the GAP Measure*, In preparation.
- [13] S. Goldstein, J.L. Lebowitz, C. Mastrodonato, R. Tumulka, N.Zanghì: *On the Approach to Equilibrium of Macroscopic Quantum Systems*, In preparation.
- [14] S. Goldstein, J.L. Lebowitz, R. Tumulka, N. Zanghì: *Canonical Typicality*, Phys. Rev. Lett. **96**, 050403, 2006. cond-mat/0511091
- [15] H.-H. Kuo: *Gaussian Measures in Banach Spaces*. Berlin: Springer, 1975.
- [16] J.L. Lebowitz: *Microscopic Origins of Irreversible Macroscopic Behavior: An Overview*, Physica A **263**, 516–527, 1999.
- [17] C. Mastrodonato, R. Tumulka: *Elementary Proof for Asymptotics of Large Haar-Distributed Unitary Matrices*, Journal of Mathematical Physics **82** Vol. 1 , 51–59, 2007.
- [18] K. R. Parthasarathy: *Probability Measures on Metric Spaces*. New York and London: Academic Press, 1967.
- [19] D. Petz, J. Réffy: *On asymptotics of large Haar distributed unitary matrices*, Periodica Mathematica Hungarica **49**, 103–117, 2004. math.PR/0310338
- [20] S. Popescu, A. J. Short, A. Winter: *The foundations of statistical mechanics from entanglement: Individual states vs. averages*, Preprint quant-ph/0511225, 2005.

- [21] S. Popescu, A. J. Short, A. Winter: *Entanglement and the foundation of statistical mechanics*, Nature Physics **21**(11), 754–758, 2006.
- [22] Yu. V. Prohorov: *Convergence of Random Processes and Limit Theorems in Probability Theory*, Theory of Probability and Its Applications **1**, 157–214, 1956.
- [23] P. Reimann: *Foundation of Statistical Mechanics under Experimentally Realistic Conditions*, Phys. Rev. Lett. **101**, 19043, 2008.
- [24] G. Shafer, V. Vork: *The Sources of Kolmogorov's Grundbegriffe*, Statistical Science **21**(1), 70–98, 2006
- [25] H. Tasaki: *From Quantum Dynamics to the Canonical Distribution: General Picture and a Rigorous Example*. Phys. Rev. Lett. **80**, 1373–1376, 1998.
- [26] R. Tumulka, N. Zanghì: *Smoothness of Wave Functions in Thermal Equilibrium*, Journal of Mathematical Physics **46**, 112104, 2005. math-ph/0509028
- [27] J. von Neumann: *Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik*, Z. Physik **57**, 30, 1929.
- [28] J. von Neumann: *Mathematical Foundation of Quantum Mechanics*, Princeton University Press, 1955. Translation of *Mathematische Grundlagen der Quantenmechanik*, Springer-Verlag, Berlin, 1932.
- [29] D. P. Williams: *Crossed Products of C^* -algebras*, Mathematical Surveys and Monograph, vol. **134**, American Mathematical Society, 2007.