Università degli Studi di Napoli “Federico II”

Scuola Politecnica e delle Scienze di Base
Area Didattica di Scienze Matematiche Fisiche e Naturali

Dipartimento di Fisica “Ettore Pancini”

Laurea Magistrale in Fisica

Quantum Measurement and Quantum Instruments

Relatori:
Dott. Paolo Aniello

Candidato:
Vincenzo Parisi
Matricola N94000477

A.A. 2018/2019

Contents

List of Symbols

$\mathcal{E}(\mathcal{H})$ Set of effects ($A \in \mathcal{L}_s(\mathcal{H})$ and $O \leq A \leq I$)

$\mathcal{L}(\mathcal{H})$ Bounded linear operators

$\mathcal{L}_s(\mathcal{H})$ Bounded selfadjoint operators ($A \in \mathcal{L}(\mathcal{H})$ and $A = A^*$)

$\mathcal{P}(\mathcal{H})$ Set of projections ($A \in \mathcal{L}_s(\mathcal{H})$ and $A = A^2$)

$\mathcal{S}(\mathcal{H})$ Set of states ($A \in \mathcal{T}_s(\mathcal{H})$, $A \geq O$ and $tr[A] = 1$)

$\mathcal{T}(\mathcal{H})$ Trace class operators

$\mathcal{T}_s(\mathcal{H})$ Selfadjoint trace class operators ($A \in \mathcal{T}(\mathcal{H}) \cap \mathcal{L}_s(\mathcal{H})$)

$\mathcal{U}(\mathcal{H})$ Set of unitary operators ($A \in \mathcal{L}(\mathcal{H})$ and $AA^* = A^*A = I$)

Introduction

When at the beginning of the twentieth century the microscopic world began to be investigated, it gradually became clear that it was necessary to adopt a theory different from classical mechanics to describe these processes.
With a work that lasted more than thirty years, a new scheme was introduced, known as quantum theory characterized by two crucial aspects. The first was the marked difference with classical mechanics, both in principles and in the laws themselves formulated with a different and more articulated mathematical language than that used in classical theories. The second was the incredible predictability and applicability of this theory, which over time was extended to phenomena concerning the fundamental interactions between elementary particles and the description of the different states and properties of condensed matter, to arrive at the most recent applications in the field of informatics (quantum computing) and information theory (quantum information). Although this theory proves to be strongly counterintuitive and presents complex conceptual and interpretative problems, the mathematical language with which it is described has been well developed and consolidated over time, reaching its complete form already with the works of Von Neumann [1] and Dirac [2].
Nowadays, there are generally two possible approaches to describe the mathematical structure of quantum mechanics. The first is the one developed by Von Neumann and founds quantum mechanics on Hilbert spaces. More precisely, quantum states are described by positive trace class trace one operators, while observables are represented by self-adjoint operators. For these operators (limited or not) it is possible to construct a projection valued measure PVM [3, 4], known as spectral measure, with respect to which they can be written as an integral on the spectrum. Then, the physical content of this scheme provides that the elementary propositions regarding

the possible measurements on the system are associated with projections, while the measurement outcomes are contained in the spectrum of the operator.
The second possible formulation is represented by the so-called algebraic formulation [5, 6]. In this case, the basic hypothesis is that the observables, both classical and quantum, constitute a C^*-algebra, while the states act as positive functionals on it. If the C^*-algebra is unital and Abelian, the Gelfand- Neumark theorem [7, 6] ensures that it is isometrically isomorphic to the algebra of continuous real-valued functions on the phase space vanishing at infinity. Then, Riesz's theorem [3] allows us to conclude that each state is uniquely associated with a Radon measure on the phase space. If the C^*-algebra is non-commutative, the so-called GNS construction [6, 7] shows how it is isomorphic to the C^*-algebra of bounded operators acting on a suitable Hilbert space, the states being represented by vectors of the Hilbert space itself. Eventually, the form of the algebra of the observables is established by the experiment, and it is, in particular, the Heisenberg uncertainty principle which confers the non-commutative character of the quantum one. This approach is undoubtedly powerful and elegant, since on the one hand, it frames the mathematical structure of quantum mechanics and classical mechanics in a single scheme, relegating the differences between the two schemes to the nature of C^*-algebra, and on the other, it can be extended to theories such as the quantum field theory on curved space-time. However, it should be noted that the great generality of the method requires that additional conditions must be imposed when restricting oneself to specific areas. In theories such as information theory and quantum mechanics in general, quantum states are σ-additive and are represented by density operators, which involve the need to select a precise class of states from those described by the algebraic approach.
In this thesis work, we will follow mostly an approach closer to that described by Von Neumann's model, introducing, where possible, a series of relatively new results which represent a generalization of the standard formalism. The aim will be to introduce two concepts, that of quantum instrument [8, 4] and quantum measurement model [8, 4, 9] which are of fundamental importance in the modern quantum measurement theory. In particular, we will discuss these concepts in the last chapter of this work.
Proceeding gradually, we first introduce the notion of quantum state, using the result of a theorem of fundamental importance for quantum mechanics, known as the Gleason theorem [3, 8]. This theorem was actually formulated by Gleason as an answer to Mackey's question if there were other rules other than Born's to calculate probabilities in quantum mechanics. What Gleason showed, is that under the hypotheses of associating projection operators with the possible measurement outcomes, and of considering a Hilbert space of dimension $d \geq 3$, every generalized

probability measure μ on the space of projections $\mathcal{P}(\mathcal{H})$ is uniquely associated with a positive trace class trace one operator $\rho_f \in \mathcal{S}(\mathcal{H})$ such that

$$\mu(P) = tr[\rho_f P],$$

Then, from one hand the Gleason theorem shows how the Born rule is the only possible one to calculate probabilities in quantum mechanics starting from the hypothesis that elementary events are associated with projections. On the other, it introduces the operators $\rho \in \mathcal{S}(\mathcal{H})$ that can be interpreted as quantum states.
As already mentioned, one of the fundamental ingredients of the mathematical formulation of quantum mechanics in Von Neumann's scheme is the association of projection operators with the elementary events characterizing a given measurement. However, observations made in quantum optics, as for example a photon detection measurement realized using a non-ideal photon detector i.e. a photon detector whose efficiency is not one, have shown how it is necessary to generalize this scheme by associating the elementary events of a measurement with positive unit bounded selfadjoint operators, known as effects [8, 4], rather than projections. Through a quantum effect E it is possible to calculate the probability associated with the corresponding elementary event as $tr[E\rho]$. Then, that quantum effects represent a generalization of projections, it follows from the fact that the latter are the extremal elements of the set of effects $\mathcal{E}(\mathcal{H})$. The introduction of the concept of quantum effect induces a generalization of the concept of observable. In fact, instead of defining a projection valued measure PVM, with respect to which selfadjoint operators are represented, it is possible to introduce a positive operator-valued measure (POVM), which acts as a map

$$\mathcal{M} \ni X \mapsto E(X) \in \mathcal{E}(\mathcal{H}),$$

$\mathcal{M}$ being the σ-algebra defined over the sample space M. POVMs therefore allow one to give a more general definition of observable since, as mentioned above, the projections operators are particular types of effects, thus showing that PVMs are a particular case of POVMs. Further, another generalization introduced by POVMs is that one can take as the space of the measurement outcomes an arbitrary metric space M instead of $\mathbb{R}$. Eventually, the generality of the definition of observable given in terms of POVMs allows to describe problems, such as that of the simultaneous measurement of observables that do not commute [8, 4, 10], which do not possess a quantitative description in the standard scheme of quantum mechanics.
Although states and observables represent the main objects in the description of quantum mechanics, and more generally of any physical theory, they do not exhaust the description of a quantum system. POVMs, in fact, quantitatively describe those processes that accept a state in input and produce a classical output given by the

probability distribution of the measurement results. Nonetheless, it is possible to think of transformations which, starting from an input state, produce another valid quantum state in output. A way in which similar processes can be described is through the concept of quantum operation [8, 4], which mathematically is described by a completely positive linear map $\mathscr{O}$ on $\mathcal{T}(\mathcal{H})$, trace non increasing. Quantum operations represent the most general transformations that can be performed on a physical system and, in fact, also the transformations induced by a measurement process can be described through them. Eventually, the trace non increasing property concerns the probabilistic nature of the processes described. As a particular case of quantum operation, one can consider the transformations that map states to states deterministically, i.e. with probability one. These maps are known as quantum channels [8, 4, 11] and are described by completely positive linear maps $\mathscr{C} : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})$ trace preserving, i.e. such that $tr[\mathscr{C}(\rho)] = 1$. Practical examples of quantum channels can be unitary transformations affecting a physical system, as well as the time evolution, for every fixed time t, of a quantum state. In particular, they represent a means through which it is possible to describe the interaction between an open quantum system and its environment [12], which is an essential element in the modern formulation of the measurement process in quantum mechanics.
POVMs, channels and quantum operations provide the necessary notions for the introduction of the concepts of quantum instruments and measurement models, which as anticipated, are the main topic covered in the following thesis work. Currently, there are, in fact, three possible descriptions of the measurement process in quantum mechanics:

- POVMs: The only information that one finds is the probability distribution of the possible measurement outcomes.

- Quantum Instruments: One finds the probability of the measurement outcomes as well as the output states conditioned by the measurement process. More precisely a quantum instrument $\mathscr{I}$ acts as a map that associates with every element X in the Borel σ-algebra $\mathcal{B}(\Omega)$ over the sample space Ω, a quantum operation $\mathscr{I}(X)$ on $\mathcal{T}(\mathcal{H})$, that reduces to a quantum channel when considering the whole space Ω.

- Measurement Model: It represents the most detailed description of the measurement process. The main idea behind a quantum measurement model is to treat the measurement process using the theory of open quantum systems [12]. More precisely, the system that is subject to the measurement process is seen as an open system that interacts with the measuring apparatus. Each measurement model $\mathcal{M}$ then consists of a quadruple $\mathcal{M} = \langle \mathcal{H}_P, \sigma, \mathscr{C}, E \rangle$, where E

is the pointer observable, i.e. the observable implemented on the measurement apparatus that reproduces the outcome probabilities of the initial observable considered in the measurement process. Then, $\mathcal{H}_P$ is the Hilbert space associated with point observable, σ is the initial state of the measurement apparatus, while $\mathscr{C}$ is a quantum channel describing the interaction between the measurement apparatus and the object system. Moreover, any measurement model uniquely induces a quantum instrument through which it is possible to gain information about the output states and the probabilities of the measurement outcomes.

Quantum instruments and Measurement models represent a powerful tool by which it is possible to face the measurement problem in quantum mechanics. Although they do not answer the questions relating to the nature of the interaction between the measuring apparatus and the measured object, they show as clearly and precisely as possible the different elements that come into play in a measurement process, thus favouring an ever deeper understanding.
Having described the main topics and the aims of the following thesis work, we now briefly describe its structure. The thesis is divided into four chapters:

- **chapter 1** *States and Effects*
  The first chapter deals with the concepts of quantum states and effects (sections 1.2, 1.3). In particular, we will see that the state space has a convex structure, the extremes of which are represented by pure states (proposition 1.2.1). Then, we investigate quantum effects, showing how they can be described by positive unit bounded selfadjoint operators (theorem 1.3.1). Eventually, we see that states act as generalized probability measures on effects (section 1.5). In this context, we first state and discuss the Gleason theorem (theorem 1.5.1), and after we present its generalization given by the Busch theorem (theorem 1.5.2).

- **chapter 2** *Quantum observables*
  The purpose of this chapter is to discuss the concept of positive operator valued measure (POVM). We start by defining POVMs (definition 2.2.1), passing then to give a series of examples in which this concept finds a concrete application. After we discuss the problem of the mixture of observables (section 2.4) by describing under what conditions the space of POVMs becomes a convex space. In this case, we will then characterize its extremal elements by using the so-called method of perturbations [13]. To clarify the relationships between selfadjoint operators and POVMs we state the Neumark's theorem (theorem 2.6.1), discussing its important mathematical and physical implications. We conclude the chapter by concisely introducing the concept of informational

completeness (section 2.7) of a set of POVMs, and its application to the problem of state reconstruction (section 2.8).

- **chapter 3** *Quantum Channels and Open Quantum Systems*
  This chapter introduces the central notions necessary for the arguments presented in the last chapter. First, we introduce quantum operations and channels (section 3.2), described as completely positive linear maps on $\mathcal{T}(\mathcal{H})$, characterized by the property of being trace non increasing and preserving respectively. We will then discuss in more precise terms the notion of complete positivity (section 3.3), starting from the more general case in which one considers linear maps $\mathfrak{L} : \mathcal{C} \to \mathcal{L}(\mathcal{H})$ from a *-algebra $\mathcal{C}$ into $\mathcal{L}(\mathcal{H})$, and stating in this context the Stinespring theorem (theorem 3.3.1) that gives a complete characterization of completely positive linear maps in the case $\mathcal{C}$ is a C^*-algebra. Then, we particularize this analysis to the case of quantum operations and channels, arriving eventually to state and prove a theorem known as Kraus theorem (theorem 3.3.5) which gives a very powerful representation of channels and operations. The last section of the chapter is left to the application of quantum channels to the theory of open quantum systems, obtaining in conclusion, under the Markovianity hypothesis, the equation for the generator of the dynamical semigroup eq.(3.5.19) of an open system.

- **chapter 4** *Measurement Models*
  The final chapter discusses the notions of quantum instrument and measurement model, which represent two key concepts in the modern theory of quantum measurement. The chapter begins by introducing the notion of a measurement model (definition 4.2.1) and by finding the quantum instruments uniquely induced by them (eq. 4.3.3). The Ozawa theorem (theorem 4.3.1), on the other hand, shows how every quantum instrument admits a measurement model that induces it. Eventually, it is observed that each quantum instrument uniquely induces an observable. The central part of the chapter is used to discuss three important applications represented by the notion of conditional state preparation (section 4.3), the proof that it is impossible to acquire information on a physical system without disturbing it, and the description of the so called Von Neumann measurement model (section 4.4). The last part of the chapter is devoted to the discussion of the role played by symmetries in measurement theory. First, we will state and prove, under suitable hypotheses, a theorem due to Wigner Araki and Yanase (theorem 4.5.1), which shows the limitations that a conservation law imposes on the possible measurements on a physical system. Then, we discuss the concepts of POVMs and quantum instruments covariant under the action of a locally

compact and unimodular group (section 4.5), observing, in particular, how the square integrability of the group representation is the necessary and sufficient hypothesis for their characterization (theorems 4.5.3, 4.5.5). We conclude by finding, as an important application, the POVMs covariant under the action of the Galilei group, thus obtaining the so called phase space measurements.

1 States and Effects

1.1 Introduction

States and observables play a central role in every physical theory. The main concept is in fact that states characterize the possible preparations in which a physical system can be found while the observables encode the possible measurements that can be performed on the system. What we will see [8] is that quantum states are described by normalized positive trace class operators, usually called density operators, while the possible outcomes of a physical experiment are associated with positive unit bounded selfadjoint operators known as effects [8, 4]. There are many ways through which quantum mechanical states can be introduced. A very general approach is represented by the GNS construction [7] in the contest of the algebraic formulation of quantum mechanics [6]. Starting from the hypotheses that the set of quantum observables constitutes a noncommutative C^*-algebra, and states act on them as positive linear functionals, one arrives at the conclusion that the observables can be represented by bounded selfadjoint operators acting on a suitable Hilbert space, while quantum states can then be represented by vectors in the Hilbert space. The generality of this approach lies in the observation that it also applies to the context of quantum field theory [5] and, moreover, a similar discussion allows one also to obtain the states of classical physics [6]. However, despite the great generality of the algebraic approach, we will introduce quantum states in a different way. The GNS construction introduces, in fact, more states than those used in contexts, such as quantum information theory, where one restricts to a class of states that respect the additional condition of σ-additivity, i.e. the so called completely additive states [3], which are represented by density operators. A direct way in which one can

introduce density operators is through a theorem due to Gleason and known as the Gleason's theorem [3]. This theorem shows that if the possible outcomes of a physical experiment are associated to orthogonal projections, and if the Hilbert space of the quantum system has dimension at last three, then every probability measure on a Hilbert space can be uniquely associated with a positive trace class operator, which can be interpreted as the density matrix of a quantum state, and the way through which the probabilities are calculated coincides exactly with the Born rule. The fundamental importance of this theorem in quantum mechanics is linked to the fact that it shows how the Born rule is the only rule through which quantum mechanical probabilities can be calculated and in fact, historically, it arose as an answer to the question posed by G. Mackey if there exist other ways to calculate probabilities in quantum mechanics different from the Born rule. However, we can consider this theorem also as an equivalent way through which one can introduce density operators in quantum mechanics starting from a few assumptions. As we have said, one of these is that the possible outcomes of an experiment are associated with projection operators. However, the new experimental evidence highlighted by the recent developments in the field of quantum optics has shown that there are cases in which the outcomes of an experiment must be associated with a more general class of unit bounded selfadjoint operators, known as effects. One can then ask what happens if effects are considered rather than projections operators in the Gleason's theorem. The surprising answer, as shown by Busch [14] and independently by Caves [15], is that not only the theorem remains true, but its validity extends to any dimension, thus including two level systems that play a central role in quantum optics and quantum information science in general.

Having discussed the main arguments of the following chapter, we briefly describe the order in which they will be presented: in section 1.2, we introduce quantum states as positive normalized trace class operators. We pass then to describe some aspects of the state space, focusing the attention on its convex structure and showing how each state can be decomposed as a convex combination of extremal states. In section 1.3, we introduce the concept of quantum effect and we show, moreover, how it can be realized in terms of a positive unit bounded selfadjoint operator. The last section will be devoted to the discussion of the Gleason theorem and the Busch theorem. First, we concisely introduce the concept of a state seen as a positive functional on the algebra of observables, starting from the classical case in which one associates a Radon probability measure to every state on phase space. At this point, passing to the quantum mechanical case, we will see, through the Gleason's theorem, how a probability measure on the Hilbert space of a quantum system can be uniquely associated with a density operator, i.e. to a quantum state. Finally, we will extend the Gleason theorem to any dimension, considering effects rather than

projections, thus stating the Busch theorem.

1.2 Physical States

The basic statistical framework

In this first section, the notion of state is described. In order to introduce this topic, we first observe that an experiment in quantum mechanics can be divided in two parts: the preparation and the measurement. This division is quite arbitrary and depends on the particular nature of the experiment we are describing [8, 4]. In particular, we assume that in every experiment there is a collection of possible preparations and a collection of possible measurements such that any preparation and any measurement can be combined obtaining in this way a probability distribution associated with the experimental outcomes [8, 4]. This constitutes the so-called basic statistical framework [8]. In principle, we observe that different preparations can lead to the same probability distribution in any chosen measurement. Thus, we can divide the set of preparations in an equivalence class of compatible preparations, that produce for every measurement the same probability distribution. These classes are called states of the system. Usually, we can think of a state as an ensemble of similarly prepared systems [4]. In the same way, we can introduce the notion of effect. In particular, we first define the concept of elementary event for a physical experiment. An elementary event is associated with only two possible outcomes: "yes" or "no" (an example of an elementary event could be "The recorded measurement outcome belongs to a subset $X \subset \Omega$" where Ω denotes the space of all possible numerical outcomes of the experiment). At this point, one can collect all the elementary events observed in various experiments that occur with the same probability for all states. This collection of equivalent events is called an effect. As we have said an effect that acts on a state, must produce a probability distribution. So with this idea in mind, we can see an effect as a function $\rho \mapsto E(\rho)$ from the set of states $\mathcal{S}$ to $[0, 1]$. In particular, the physical significance of $E(\rho)$ is that it represents the probability that the elementary event is realized [8, 4].

State space

One of the crucial aspects of quantum states is that they form a convex space. Intuitively, this feature can be justified in the following way: let us suppose that we perform an experiment in which we alternate randomly between two preparations procedure associated with two states ρ_1 and ρ_2, with probability λ and $1 - \lambda$, such that $0 \leq \lambda \leq 1$. In this way, we obtain a new preparation procedure, i.e. a new

state ρ that can be written as

$$\rho = \lambda\rho_1 + (1-\lambda)\rho_2. \tag{1.2.1}$$

From this one can then conclude that the state space is a set closed under convex combinations, i.e quantum states form a convex set. More precisely, we assume that the state space is the set

$$\mathcal{S}(\mathcal{H}) = \{\rho \in \mathcal{T}_s(\mathcal{H}) \mid \rho \geq O,\ tr[\rho] = 1\},$$

where the operators ρ are called density operators and $\mathcal{T}_s(\mathcal{H})$ denotes the set of self-adjoint trace class operators. We will give a rigorous justification for this assumption after we introduce the Gleason theorem in section 1.5.

Remark 1.2.1. Any measurement performed on the state 1.2.1 should give results that are consistent with the measurements performed on ρ_1 and ρ_2. This means that, for every effect E, we must have

$$E(\lambda\rho_1 + (1-\lambda)\rho_2) = \lambda E(\rho_1) + (1-\lambda)E(\rho_2), \tag{1.2.2}$$

thus implying that act as affine mappings on $\mathcal{S}(\mathcal{H})$.

Now we pass to describe some facts about this space. Firstly we note that $\mathcal{S}(\mathcal{H})$ is not only convex but also σ-convex. This means that if $\{\rho_i\}_{i=1}^{+\infty}$ is a sequence of states in $\mathcal{S}(\mathcal{H})$ and $\{\lambda_i\}_{i=0}^{+\infty}$ is a sequence of weights that satisfy the conditions $0 \leq \lambda_i \leq 1$ and $\sum_{i=1}^{+\infty} \lambda_i = 1$, then $\sum_{i=1}^{n} \lambda_i\rho_i$ is a Cauchy sequence in $\mathcal{T}(\mathcal{H})$. So it follows that

$$\lim_{n\to+\infty} \left(\sum_{i=1}^{n} \lambda_i\rho_i \right) = \sum_{i=1}^{+\infty} \lambda_i\rho_i \ \in\ \mathcal{T}(\mathcal{H})$$

with convergence in trace norm $\|\cdot\|_{tr}$, where we remember that such a norm is defined as $\|A\|_{tr} = tr[(AA^*)^{1/2}]$ for $A \in \mathcal{T}(\mathcal{H})$ [3]. In particular since

$$tr\Big[\sum_{i=1}^{+\infty} \lambda_i\rho_i\Big] = \sum_{i=1}^{+\infty} \lambda_i = 1$$

we also obtain that $\sum_{i=1}^{+\infty} \lambda_i\rho_i$ is an element of $\mathcal{S}(\mathcal{H})$. Moreover we observe that since $\mathcal{S}(\mathcal{H})$ is a subset of the space of trace class operators, using the spectral theorem we have the following [8]:

Theorem 1.2.1. *Every element $\rho \in \mathcal{S}(\mathcal{H})$ can be written as*

$$\rho = \sum_i \lambda_i P_i$$

where $\{\lambda_i\}$ is a finite or infinite sequence of positive numbers such that $\sum_i \lambda_i = 1$ and $\{P_i\}$ is a sequence of one dimensional orthogonal projections, i.e. of positive idempotent trace one operators that, using Dirac notation, can be written as $P_\phi = |\phi\rangle\langle\phi|$ for some unit vector $\phi \in \mathcal{H}$ and satisfying the condition $P_i P_j = \delta_{ji} P_j$.

Now in order to have a deeper understanding of the structure of the state space, we introduce the class of its extremal elements. In particular, an element σ of a certain convex set is called **extremal** if it can not be written as a convex combination of other elements of the set, i.e. that if we consider

$$\sigma = \lambda\sigma_1 + (1-\lambda)\sigma_2$$

for some $0 < \lambda < 1$ it must be $\sigma = \sigma_1 = \sigma_2$.

Remark 1.2.2. An element that is not extremal has uncountably many convex decompositions. We can see this in the following way. We consider a state $\rho = \lambda\rho_1 + (1-\lambda)\rho_2$ for $0 < \lambda < 1$ and $\rho_1 \neq \rho_2$. Let us define now a new state $\rho'_\alpha = \alpha\rho_1 + (1-\alpha)\rho_2$ for $\lambda < \alpha < 1$. So we have

$$\rho_1 = \frac{\rho'_\alpha}{\alpha} - \frac{1-\alpha}{\alpha}\rho_2$$

and so

$$\rho = \frac{\lambda}{\alpha}\rho'_\alpha + \rho_2\Big(1 - \frac{\lambda}{\alpha}\Big)$$

Since $\rho'_\alpha \neq \rho_2$ and since α is any number in the interval $[\lambda, 1]$ we have obtained, varying α, uncountably many different nontrivial convex decompositions of ρ.

As we will see nonextremal elements of $\mathcal{S}(\mathcal{H})$ can be written in terms of extremal elements. On the other hand, the extremal elements of $\mathcal{S}(\mathcal{H})$ are usually known as pure states. So we first give a

Definition 1.2.1. An extremal element in $\mathcal{S}(\mathcal{H})$ is called *pure state*. Any other element is called *mixed state*.

We can make a complete characterization of pure states through the following proposition [8]:

Proposition 1.2.1. *For $\rho \in \mathcal{S}(\mathcal{H})$ these conditions are equivalent:*

1) ρ is a pure state

2) ρ is a one-dimensional projection

Proof. 1) $\Rightarrow$ 2)
We suppose that ρ is a pure state. Then its canonical decomposition can contain only one term $\lambda_1 P_1$. Bat ρ must also satisfy the condition $tr[\rho] = 1$ and so we must have $\lambda_1 = 1$. Thus we conclude that $\rho = P_1$ and so it is a one dimensional projection.
1) $\Leftarrow$ 2)
Let $\rho = P_\phi$ be a one dimensional projection, for ϕ a unit vector in the Hilbert space $\mathcal{H}$. If ρ is not an extreme state there must exist two states ρ_1 and ρ_2 such that

$$\rho = \lambda\rho_1 + (1-\lambda)\rho_2, \tag{1.2.3}$$

for $0 < \lambda < 1$. Now we take a vector $\psi \neq \phi$, such that $P_\phi\psi = 0$. Since ρ_1 and ρ_2 are positive operators we have

$$0 = \langle\psi|P_\phi\psi\rangle = \lambda\langle\psi|\rho_1\psi\rangle + (1-\lambda)\langle\psi|\rho_2\psi\rangle \geq \lambda\langle\psi|\rho_1\psi\rangle \geq 0. \tag{1.2.4}$$

So we can conclude that $\langle\psi|\rho_1\psi\rangle = 0$ and so $\rho_1\psi = 0$. However if we consider

$$I - P_\phi = \lambda(I-\rho_1) + (1-\lambda)(I-\rho_2), \tag{1.2.5}$$

and we observe that $(I-P_\phi)\phi = 0$ is equivalent to $P_\phi\phi = \phi$, with a reasoning similar to that in eq. (1.2.4) we obtain also $\rho_1\phi = \phi$. So we observe that ρ_1 and P_ϕ act identically on all vectors $\psi \in \mathcal{H}$ and we can conclude that $\rho_1 = P_\phi$. From this, using eq.(1.2.3), we also obtain that $\rho_2 = P_\phi$, Thus concluding that ρ is a pure state.
□

So what we have seen is that, through Theorem 1.2.1, we can decompose every mixed state as a convex combination of pure states.

Remark 1.2.3. We can characterize the degree of purity of a state using a function known as Purity [8, 11] and defined as $\mathcal{P}(\rho) = tr[\rho^2] = \sum_i \lambda_i^2$. In particular this function satisfies the following properties

1) $\mathcal{P}$ is a convex map: $\mathcal{P}(\lambda\rho_1 + (1-\lambda)\rho_2) \leq \lambda\mathcal{P}(\rho_1) + (1-\lambda)\mathcal{P}(\rho_2)$

2) $\mathcal{P}$ is invariant under unitary transformations: $\mathcal{P}(U\rho U^*) = \mathcal{P}(\rho)$

3) $\mathcal{P}(\rho) = 1$ iff ρ is a pure state

Remark 1.2.4. For the set $\mathcal{S}(\mathcal{H}) \subset \mathcal{T}_s(\mathcal{H})$ we can also consider the elements that belong to the boundary. These elements are characterized by the fact that, for each $\epsilon > 0$, there exists an operator $\zeta \in \mathcal{T}_s(\mathcal{H})$ and $\zeta \notin \mathcal{S}(\mathcal{H})$ such that

$$\|\rho - \zeta\|_{tr} < \epsilon, \tag{1.2.6}$$

where $\| \|_{tr}$ denotes the trace norm.
Now we observe that all pure states are boundary points of $\mathcal{S}(\mathcal{H})$.

Proposition 1.2.2. *If a state ρ has eigenvalue* 0 *then it belongs to the boundary of $\mathcal{S}(\mathcal{H})$. All pure states are boundary points of $\mathcal{S}(\mathcal{H})$*

Proof. Suppose that we have a state ρ that has eigenvalue 0. Let $\phi \in \mathcal{H}$ be a corresponding normalized eigenvector. We can construct the following operator:

$$\zeta = \rho - \frac{1}{2}\epsilon|\phi\rangle\langle\phi| \tag{1.2.7}$$

We now observe that

$$\|\rho - \zeta\|_{tr} = \frac{1}{2}\epsilon\||\phi\rangle\langle\phi|\|_{tr} = \frac{1}{2}\epsilon \leq \epsilon. \tag{1.2.8}$$

So ζ is trace class and selfadjoint but not positive, since $\langle\phi|\zeta\phi\rangle = -\frac{1}{2}\epsilon \leq 0$. We conclude that $\zeta \notin \mathcal{S}(\mathcal{H})$ and so ρ is a boundary element. If we consider a projection $P \neq I \neq O$ it has eigenvalues 0 and 1. So we also conclude that all projections are boundary elements. □

It is important to observe that all projections are boundary elements of $\mathcal{S}(\mathcal{H})$, but there are elements in the boundary that are not projections. We can see this by making the following example. Consider an Hilbert space such that $dim\mathcal{H} \geq 3$. Let α_1, α_2 and α_3 be three orthogonal unit vectors. We can define a state ρ as

$$\rho = \lambda|\alpha_1\rangle\langle\alpha_1| + (1-\lambda)|\alpha_2\rangle\langle\alpha_2|, \tag{1.2.9}$$

where $0 < \lambda < 1$. The state ρ so obtained is a mixed state. However if we consider $\rho|\alpha_3\rangle$ we obtain

$$\rho|\alpha_3\rangle = \lambda|\alpha_1\rangle\langle\alpha_1|\alpha_3\rangle + (1-\lambda)|\alpha_2\rangle\langle\alpha_2|\alpha_3\rangle = 0. \tag{1.2.10}$$

Thus we conclude that since ρ has eigenvalue 0 it is a boundary element of $\mathcal{S}(\mathcal{H})$. However, it is not a pure state. We have considered an Hilbert space with $dim\mathcal{H} \geq 3$ because, in the following, we will observe an example, in dimension 2, in which we have a complete equivalence between the boundary of $\mathcal{S}(\mathcal{H})$ and the set of pure states.

Now we pass to another interesting question concerning mixed states. We have seen that every mixed state can be decomposed as a convex combination of at most countably many one dimensional projections $\rho = \sum_i \lambda_i P_i$ in which $P_i \neq P_j \; for \; i \neq j$, with the series converging in the trace norm. However, we note that in general ρ does not have a unique decomposition but admits, in fact, uncountably many decompositions. In order to see the non-uniqueness decomposition of mixed states, we make an example [4].

Example 1.2.1. In this example we consider a mixed state ρ of finite rank, i.e. that can be written as $\sum_{i=1}^{n} \lambda_i P_{\phi i}$, in which the $\{\lambda_i\}_i$ are (not necessarily different) positive numbers summing to one, and the one dimensional projections are associated with mutually orthogonal unit vectors $\phi_1, \ldots, \phi_n$. Now we consider the state $\tilde{\phi}_i = \sqrt{\lambda_i}\phi_i$, so that we can write $\rho = \sum_{i=1}^{n} |\tilde{\phi}_i\rangle\langle\tilde{\phi}_i|$. Now we consider a complex $m \times n$ matrix (c_{kl}),with $m \geq n$, such that all its columns $(c_{1i}, \ldots, c_{mi})$ for $i = 1, \ldots, n$, are mutually orthogonal unit vectors. If we define $\tilde{\psi}_k = \sum_{i=1}^{n} (c_{ki})\tilde{\phi}_i$ for $k = 1, \ldots, m$ and observe that $\sum_{i=1}^{n} \lambda_i |c_{ki}|^2 > 0$ and that $\sum_{k=1}^{m} \sum_{i=1}^{n} \lambda_i |c_{ki}|^2 = 1$, we can write $\rho = \sum_{k=1}^{m} |\tilde{\psi}_k\rangle\langle\tilde{\psi}_k|$. In this way we have obtained a new convex decomposition of the state ρ. ■

What we have seen in this example can be extended to a proposition that characterizes all the convex decompositions of a finite rank mixed state, the proof of which we omit (see [8] proposition 2.17).

Proposition 1.2.3. *If ρ is a mixed state that can be decomposed as*

$$\rho = \sum_{i=1}^{n} \lambda_i |\sigma_i\rangle\langle\sigma_i|, \tag{1.2.11}$$

then all its convex decompositions into pure states have the form

$$\rho = \sum_{k=1}^{m} q_k |\psi_k\rangle\langle\psi_k| \tag{1.2.12}$$

where the vectors $\psi_1, ..., \psi_m$ are orthogonal vectors in $\mathcal{H}$, and q_k for $k = 1, ..., m$ are defined through

$$\sqrt{\lambda_i}|\sigma_i\rangle = \sum_{k=1}^{m} u_{ik}\sqrt{q_k}|\psi_k\rangle \tag{1.2.13}$$

with the complex number u_{ik} satisfying

$$\sum_{i} u_{ik}\bar{u}_{ik'} = \delta_{kk'} \tag{1.2.14}$$

Remark 1.2.5. In the case of infinite but countable σ-convex mixtures, a complete characterization was given by Cassinelli [16].

We turn now our attention to the case of a finite-dimensional Hilbert space. We will see that, in this case, a state can be represented using the so called Bloch representation. In order to represent a state of $\mathcal{S}(\mathcal{H})$, we must find a suitable basis for this space. We begin by observing that if the Hilbert space $\mathcal{H}$ has dimension

n, then the space of trace class operators $\mathcal{T}(\mathcal{H})$ has complex dimension n^2, while the space $\mathcal{T}_s(\mathcal{H})$ has real dimension n^2. From this, we conclude that we can not find a basis for $\mathcal{T}_s(\mathcal{H})$ consisting of density operators. In fact if we consider pure states, since the orthogonality of pure states is equivalent to the orthogonality of the normalized vectors associated, and since there are at most n mutually orthogonal unit vectors in $\mathcal{H}$, we can not find a basis of n^2 pure states for $\mathcal{T}_s(\mathcal{H})$. If we consider mixed states we arrive at the same conclusion since the orthogonality of two mixed states is equivalent to the orthogonality of their supports and in $\mathcal{H}$ there are at most n mutually orthogonal subspaces [8].
Nevertheless we can construct a selfadjoint orthogonal basis $\{E_0, E_1, \dots, E_{n^2-1}\}$ for $\mathcal{T}_s(\mathcal{H})$, with respect to which we can represent the element of $\mathcal{S}(\mathcal{H})$, such that $E_0 = I$. If we impose the orthogonality of the basis elements with respect to the Hilbert-Schmidt inner product, i.e. the inner product defined as $\langle A_1|A_2\rangle_{HS} = tr[A_1^* A_2]$, for $A_1, A_2 \in \mathcal{T}(\mathcal{H})$, we observe that the operators $E_1, \dots, E_{n^2-1}$ are traceless. Moreover, we can also impose a normalization condition i.e. $\langle E_i|E_j\rangle_{HS} = n\delta_{ij}$. In this way, we can express a state as

$$\rho = \frac{1}{n}(I + \vec{r}\cdot\vec{E}) \tag{1.2.15}$$

where $\vec{E} = \{E_1, \dots, E_{n^2-1}\}$ and $\vec{r} = (tr[\rho E_1], \dots, tr[\rho E_{n^2-1}])$ is a real vector called Bloch vector.
In order to make a concrete realization, we now particularize our discussion to the case $n = 2$. In this case, we have a Hilbert space $\mathcal{H} \cong \mathbb{C}^2$ in which linear operators are represented by 2×2 matrices with complex entries. Thus we observe that the space of linear operators $\mathcal{L}(\mathcal{H})$ on $\mathcal{H}$ is isomorphic to $M_2(\mathbb{C})$ and is spanned by the Pauli matrices,

$$\sigma_0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.2.16}$$

Moreover we have also that $\mathcal{L}(\mathcal{H}) \cong \mathbb{C}^4$ through the correspondence

$$M_2(\mathbb{C}) \ni \ A = \sum_{i=0}^{3} c_i\sigma_i = a_0 I + \vec{a}\cdot\vec{\sigma} \ \longleftrightarrow (a_0, \vec{a}) \in \mathbb{C}^4 \tag{1.2.17}$$

The selfadjoint operators are instead represented by Hermitian 2×2 matrices, which form a 4 dimensional real vector space M_4, which is spanned by the Pauli matrices with now $(a_0, \vec{a}) \in \mathbb{R}^4$.
A generic state in this space can so be written as

$$\rho = \frac{1}{2}(I + \vec{a}\cdot\vec{\sigma}) \tag{1.2.18}$$

with $\vec{a} \in \mathbb{R}^3$. The eigenvalues of this operator are $\lambda_\pm = \frac{1}{2}(1 \pm \|\vec{a}\|)$. So the operator ρ is positive if and only if $\|\vec{a}\| \leq 1$ and we conclude that the Bloch vectors are contained in the unit sphere in $\mathbb{R}^3$, usually known as Bloch sphere. When $\|\vec{a}\| = 1$, the eigenvalues are equal to $\lambda_+ = 1$ and $\lambda_- = 0$. So we observe that with $\|\vec{a}\| = 1$ we have pure states and at the same time we are considering elements on the boundary of the Bloch sphere. Thus the boundary of the Bloch sphere consists of pure states only.

Example 1.2.2. As an example of what we have seen we consider a system of particles prepared in the mixed state

$$\rho = \frac{1}{3}|+\rangle\langle+| + \frac{1}{3}\left[\frac{1}{2}\Big(- |+\rangle + |-\rangle\Big)\Big(\langle-| - \langle+|\Big)\right] + \frac{1}{3}\left[\frac{1}{2}\Big(|+\rangle + i|-\rangle\Big)\Big(\langle+| - i\langle-|\Big)\right], \tag{1.2.19}$$

where $|\pm\rangle$ denotes the eigenvectors of the spin component along the z axis. So we have the following hermitian matrix:

$$\rho = \begin{pmatrix} \frac{2}{3} & \frac{1}{6}(-i-1) \\ \frac{1}{6}(i-1) & \frac{1}{3} \end{pmatrix}. \tag{1.2.20}$$

Finally we can find

$$r_x = tr[\sigma_x\rho] = -\frac{1}{3} \quad r_y = tr[\sigma_y\rho] = \frac{1}{3} \quad r_z = tr[\sigma_z\rho] = \frac{1}{3}. \tag{1.2.21}$$

So we have obtained a Bloch vector $\vec{r} = \left(- \frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. ■

We now conclude our brief discussion on the properties of states in quantum mechanics, describing the concept of superposition. We have seen that mixed states can be decomposed in terms of a mixture of pure states. However, there is something of similar also for pure states. First, we observe that pure states can be described also in terms of "rays". This concept is linked to the fact that if we consider a pure state P_ϕ, there exists a vector ϕ such that $P_\phi = |\phi\rangle\langle\phi|$. But if now, we consider another vector ψ, that differs from ϕ for a complex number of modulus one, then we obtain that $P_\psi = P_\phi$. So we can think of a pure state also as an equivalence class of vectors $[\phi]$ in which the equivalence relation is defined as

$$\phi \sim \psi \qquad \Longleftrightarrow \qquad \phi = z\psi \;\; for \;\; z \in \mathbb{C}, |z| = 1. \tag{1.2.22}$$

These equivalence classes are usually known as rays [8].
At this point, we can define the superposition of two linear independent unit vectors ψ,$\phi \in \mathcal{H}$. For two nonzero complex numbers a_1 and a_2 we can write

$$\sigma = \frac{1}{\|a_1\psi + a_2\phi\|}(a_1\psi + a_2\phi) \tag{1.2.23}$$

We now observe that σ is a unit vector and so $[\sigma]$, that represents a new pure state, is known as the superposition of the pure states $[\phi]$ and $[\psi]$.
If we consider the corresponding projection associated with the pure state $[\sigma]$, we have

$$P_\sigma = \frac{1}{\|a_1\psi + a_2\phi\|^2}(|a_1|^2 P_\psi + |a_2|^2 P_\phi + \bar{a}_1 a_2 |\phi\rangle\langle\psi| + a_1\bar{a}_2|\psi\rangle\langle\phi|). \tag{1.2.24}$$

So we see that we do not have a convex combination of P_ψ and P_ϕ because also cross terms appear. These terms are associated with interference, which is one of the characteristic elements of quantum mechanics.

1.3 Effects

In this section, we discuss the important notion of effect [8, 4]. As we have seen an effect can be described in terms of an equivalence class of compatible measures. When we act with an effect on a state we obtain a probability distribution associated with the experimental outcomes. This leads us to think that the effects are affine mappings from the set of states $\mathcal{S}(\mathcal{H})$, to $[0,1]$ (see eq.(1.2.2)). Now what we will see is that these maps can be associated with unit bounded trace class selfadjoint operators. We have the following [8]

Proposition 1.3.1. *Let E be an effect and so a mapping from $\mathcal{S}(\mathcal{H})$ to $[0,1]$. Then there exists a bounded selfadjoint operator $\hat{E}$ such that*

$$E(\rho) = tr[\rho\hat{E}] \qquad \forall\rho \in \mathcal{S}(\mathcal{H}) \tag{1.3.1}$$

Moreover the operator $\hat{E}$ satisfies $O \leq \hat{E} \leq I$.

Proof. In this proof [8] we prove that every effect can be extended to a continuous linear functional on $\mathcal{T}(\mathcal{H})$. In this way we can use the relation existing between the dual of $\mathcal{T}(\mathcal{H})$ and $\mathcal{L}(\mathcal{H})$,i.e. $\mathcal{T}(\mathcal{H})^* = \mathcal{L}(\mathcal{H})$.
So we start with an effect on $\mathcal{S}(\mathcal{H})$. Now we define a new functional such that

$$\tilde{E}(O) \coloneqq 0 \tag{1.3.2}$$

and

$$\tilde{E}(P) \coloneqq tr[P]E(tr[P]^{-1}P). \tag{1.3.3}$$

In this way, we have extended E such that it is defined on all positive trace class operators (not only trace class operators of trace one). We must now show that this extension is linear. So consider $a \geq 0$. We have

$$\tilde{E}(aP) = tr[aP]E(tr[aP]^{-1}aP) = a\tilde{E}(P). \tag{1.3.4}$$

If we consider two positive trace class operators P and R, we have also

$$\tilde{E}(P+R) = tr[P+R]E(tr[P+R]^{-1}(P+R)] = \tag{1.3.5}$$

$$tr[P+R]E\Big(\frac{tr[P]}{tr[P+R]}\frac{P}{tr[P]} + \frac{tr[R]}{tr[P+R]}\frac{R}{tr[R]}\Big) = \tag{1.3.6}$$

$$\tilde{E}(P) + \tilde{E}(R). \tag{1.3.7}$$

Now we want to extend $\tilde{E}$ to the whole $\mathcal{T}(\mathcal{H})$. In order to do this we observe that every trace class selfadjoint operator can be written as:

$$S = S^+ - S^-, \tag{1.3.8}$$

where

$$S^+ = \frac{1}{2}(|S| + S) \qquad S^- = \frac{1}{2}(|S| - S) \tag{1.3.9}$$

are the positive and the negative part respectively. If S is a trace class operator, then also S^+ and S^- are trace class operators [17]. We also observe that S^+ and S^- are positive operators. So we can consider the action of $\tilde{E}$ on these operators:

$$\tilde{E}(S) = \tilde{E}(S^+) - \tilde{E}(S^-), \tag{1.3.10}$$

for every $S \in \mathcal{T}_s(\mathcal{H})$, and where we have used the linearity of $\tilde{E}$.
Now we must prove that for two operators $P, R \in \mathcal{T}_s(\mathcal{H})$ we have

$$\tilde{E}(P+R) = \tilde{E}(P) + \tilde{E}(R). \tag{1.3.11}$$

i.e. that $\tilde{E}$ is a linear map on $\mathcal{T}_s(\mathcal{H})$. In order to show this we note that

$$P + R = (P+R)^+ - (P+R)^- \tag{1.3.12}$$

and also

$$P + R = P^+ - P^- + R^+ - R^-. \tag{1.3.13}$$

So we have

$$(P+R)^+ - (P+R)^- = P^+ - P^- + R^+ - R^-. \tag{1.3.14}$$

If we now apply $\tilde{E}$ on both sides of eq. (1.3.14), we obtain

$$\tilde{E}(P+R) = \tilde{E}(P) + \tilde{E}(R). \tag{1.3.15}$$

At this point, if we consider an operator $A \in \mathcal{T}(\mathcal{H})$, we can write it as $A = A_R + iA_I$, in which the A_R and A_I denote the real and the imaginary part of the operator A. In particular, we observe that these operators are selfadjoint and trace class (since

they are a combination of the operators A and A^*), hence we can extend $\tilde{E}$ to all trace class operators as

$$\tilde{E}(A) = \tilde{E}(A_R) + i\tilde{E}(A_I) \tag{1.3.16}$$

The mapping $\tilde{E}$ just obtained, is linear, as shown in eq.(1.3.15), but it is also bounded (and so continuous), since

$$|\tilde{E}(A)| = |\tilde{E}(A_R^+ - A_R^- + iA_I^+ - iA_I^-)| \leq \tag{1.3.17}$$

$$|\tilde{E}(A_R^+)| + |\tilde{E}(A_R^-)| + |\tilde{E}(A_I^+)| + |\tilde{E}(A_I^-)| \leq \tag{1.3.18}$$

$$tr[\tilde{E}(A_R^+)] + tr[\tilde{E}(A_R^-)] + tr[\tilde{E}(A_I^+)] + tr[\tilde{E}(A_I^-)] = \tag{1.3.19}$$

$$\frac{1}{2}tr[|A + A^*|] + \frac{1}{2}tr[|A - A^*|] \leq \|A\|_{tr}. \tag{1.3.20}$$

So we have extended $\tilde{E}$ to a linear operator on $\mathcal{T}(\mathcal{H})$. Now we make use of the following theorem [17]

Theorem 1.3.1. *Let $\mathcal{H}$ be a separable Hilbert space. For each bounded operator $A \in \mathcal{L}(\mathcal{H})$, consider the functional f_A on $\mathcal{T}(\mathcal{H})$ defined as*

$$f_A(T) \coloneqq tr[AT]. \tag{1.3.21}$$

Then the mapping $A \to f_A$ is a linear bijection from $\mathcal{L}(\mathcal{H})$ to $\mathcal{T}(\mathcal{H})^*$, such that $\|A\| = \|f_A\|$ for every $A \in \mathcal{L}(\mathcal{H})$.

So if we use the relation between the dual of $\mathcal{T}(\mathcal{H})$ and $\mathcal{L}(\mathcal{H})$ we conclude that there exists an operator $\hat{E} \in \mathcal{L}(\mathcal{H})$, such that

$$\tilde{E}(S) = tr[S\hat{E}]. \tag{1.3.22}$$

Moreover, if we consider the fact that by construction $\tilde{E}(S) \in \mathbb{R}$ for very $S \in \mathcal{T}_s(\mathcal{H})$ and that for an operator $A \in \mathcal{L}(\mathcal{H})$ it can be proved that $A = A^*$ if and only if $tr[AT] \in \mathbb{R}$, whenever the trace class operator T satisfies the condition $T = T^*$, then it follows that $\hat{E}$ must be a selfadjoint operator.
We must now prove that $O \leq \hat{E} \leq I$. The reason for this condition is that $tr[\rho E]$, being a probability, must belong to the interval $[0, 1]$. Firstly, if we consider the action of $\tilde{E}$ on a state ρ, we have $\tilde{E}(\rho) = tr[\rho\hat{E}]$. Now let us consider a pure state $\rho = |\psi\rangle\langle\psi|$ associated to a normalized vector $\psi \in \mathcal{H}$. Then we have

$$tr[|\psi\rangle\langle\psi|\hat{E}] = \langle\psi|\hat{E}\psi\rangle. \tag{1.3.23}$$

It follows that $tr[|\psi\rangle\langle\psi|\hat{E}] \geq 0$ for every $\rho \in \mathcal{S}(\mathcal{H})$ if and only if $\hat{E} \geq O$. But all other states are mixtures of pure states and so the condition $tr[\rho\hat{E}] \geq 0$ is satisfied

if and only if $\hat{E}$ is positive, i.e. $\hat{E} \geq O$.
Moreover, since $tr[\rho(I - \hat{E})] = 1 - tr[\rho\hat{E}]$, we conclude that $\hat{E} \leq I$. So in order to be verified the condition that $tr[\rho\hat{E}]$ is in $[0, 1]$, we must have that $O \leq \hat{E} \leq I$.
The last thing we must demonstrate is that if we consider two different selfadjoint trace class operators $\hat{E}_1$ and $\hat{E}_2$, such that $O \leq \hat{E}_1 \leq I, \quad O \leq \hat{E}_2 \leq I, \quad \hat{E}_1 \neq \hat{E}_2$, then $tr[\hat{E}_1\rho] \neq tr[\hat{E}_2\rho] \quad \forall\rho \in \mathcal{S}(\mathcal{H})$. This follows from the fact that if we consider a pure state $\rho = |\psi\rangle\langle\psi| \in \mathcal{S}(\mathcal{H})$ and if $tr[\rho\hat{E}_1] = tr[\rho\hat{E}_2]$, it must be $\langle\psi|\hat{E}_1\psi\rangle = \langle\psi|\hat{E}_2\psi\rangle$ for all vector $\psi \in \mathcal{H}$. So it must be $\hat{E}_1 = \hat{E}_2$. Therefore $\hat{E}_1 \neq \hat{E}_2$ implies that $tr[\rho\hat{E}_1] \neq tr[\rho\hat{E}_2]$. Since every mixed state can be written as a superposition of pure states, we conclude that this condition is true for every $\rho \in \mathcal{S}(\mathcal{H})$. □

Here and in the following we will indicate the set of effects as $\mathcal{E}(\mathcal{H})$

$$\mathcal{E}(\mathcal{H}) = \{E \in \mathcal{L}_s(\mathcal{H}) |\ O \leq E \leq I\}, \tag{1.3.24}$$

where we denote with $\mathcal{L}_s(\mathcal{H})$ the set of bounded selfadjoint operators. We note explicitly that the physical meaning of the quantity $tr[\hat{E}\rho]$, is that it represents the probability that the measurement event, represented by the effect E associated with the operator $\hat{E}$, occurs when the state of the system is described by ρ.

Example 1.3.1. As a first example of effects we can consider the identity effect I and the zero effect O. The identity effect assigns the probability 1 to any state ρ, i.e. $I(\rho) = 1 \ \ \forall\rho \in \mathcal{S}(\mathcal{H})$. The zero effect O assigns probability 0 to every state ρ: $O(\rho) = 0 \ \ \forall\rho \in \mathcal{S}(\mathcal{H})$. ■

Example 1.3.2. In the case of a two dimensional Hilbert space $\mathcal{H} \simeq \mathbb{C}^2$, the set of selfadjoint operators $\{I, \sigma_x, \sigma_y, \sigma_z\}$ constitutes a basis for $\mathcal{L}_s(\mathcal{H})$. Thus an operator $A \in \mathcal{L}_s(\mathcal{H})$ can be written as:

$$A = \frac{1}{2}(aI + \vec{b} \cdot \vec{\sigma}), \tag{1.3.25}$$

where $a \in \mathbb{R}$ and $\vec{b} \in \mathbb{R}^3$. The eigenvalues of this operator are

$$\lambda_\pm = \frac{1}{2}(a \pm \|\vec{b}\|). \tag{1.3.26}$$

Now from the condition $O \leq A \leq I$, we conclude that A is an effect if and only if $\lambda_+ \leq 1$ and $\lambda_- \geq 0$. Moreover A is a projection if and only if $\lambda_+ = 1$ and $\lambda_- = 0$, i.e. if and only if $a = \|\vec{b}\| = 1$. ■

We want now discuss some properties of the set $\mathcal{E}(\mathcal{H})$. We observe that this set is a convex set, that is for A and $B \in \mathcal{E}(\mathcal{H})$ and for $0 \leq \lambda \leq 1$ the operator

$\lambda A + (1-\lambda)B$ is an effect and so an element of $\mathcal{E}(\mathcal{H})$. Extremal elements of this set are effects A such that the condition $A = \lambda A_1 + (1-\lambda)A_2$ implies that $A = A_1 = A_2$. We now observe that the set of projections coincides with the set of extremal effects in $\mathcal{E}(\mathcal{H})$. We thus have the following [8]:

Proposition 1.3.2. *The set of extreme effects coincides with the set of projections.*

Proof. Let $P \in \mathcal{P}(\mathcal{H})$ be a projection. Suppose that we can write P as

$$P = \lambda A_1 + (1-\lambda)A_2, \tag{1.3.27}$$

for $0 \leq \lambda \leq 1$ and $A_1, A_2 \in \mathcal{E}(\mathcal{H})$. We also consider a vector $\phi \in \mathcal{H}$ such that $P\phi = 0$. Then

$$0 \leq \lambda \|A_1^{\frac{1}{2}}\phi\| = \lambda \langle \phi | A_1 \phi \rangle \leq \tag{1.3.28}$$

$$\lambda \langle \phi | A_1 \phi \rangle + (1-\lambda)\langle \phi | A_2 \phi \rangle = \langle \phi | P \phi \rangle = 0, \tag{1.3.29}$$

which implies that $A_1\phi = 0$. Now consider a vector $\psi \in \mathcal{H}$ such that $P\psi = \psi$ and so $(I-P)\psi = 0$. Since $I - P = \lambda(I - A_1) + (1-\lambda)(I - A_2)$, we obtain that $(I - A_1)\psi = 0$, i.e. $A_1\psi = \psi$. So, from the observation that every vector in $\mathcal{H}$ can be written as a sum of eigenvectors of P, we conclude that $P = A_1$. If $P = A_1$ it follows that $P = A_2$ and so that P is an extreme effect. On the other hand we must show that the projections are the only extreme elements on $\mathcal{E}(\mathcal{H})$. So let $A \in \mathcal{E}(\mathcal{H})$, such that A is not a projection. We have then $A \neq A^2$. Now we define two effects[1] : $E_1 = A^2$ and $E_2 = 2A - A^2 \neq A$. If we take the combination $\frac{1}{2}(E_1 + E_2)$ we observe that it is equal to A. Thus we conclude that every element that is not a projection, can be written as a non-trivial combination of effects. So the only extreme elements in $\mathcal{E}(\mathcal{H})$ are projections. □

Example 1.3.3. Consider the qubit effect that is defined as

$$A = \frac{1}{2}(I + \vec{a} \cdot \vec{\sigma}), \tag{1.3.30}$$

with $\vec{a} \in \mathbb{R}^3$ and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. We now show that this operator can be written as a convex combination of projections. Consider the term

$$\vec{a} \cdot \vec{\sigma} = \begin{pmatrix} a_z & a_x - ia_y \\ a_x + ia_y & -a_z \end{pmatrix} \tag{1.3.31}$$

[1]It can be shown [3] that for a bounded selfadjoint operator T such that $O \leq T \leq I$ the operator T^2 satisfies $O \leq T^2 \leq T$

This matrix has the eigenvalues $\lambda_\pm = \pm\|a\|$. Now if we indicate the eigenvectors of σ_z as $|\pm\rangle$, we can find the two eigenvectors of $\vec{a}\cdot\vec{\sigma}$ as

$$|\vec{a}\cdot\vec{\sigma},+\rangle = c|+\rangle + d|-\rangle$$
$$|\vec{a}\cdot\vec{\sigma},-\rangle = c'|+\rangle + d'|-\rangle,$$

where c and d are

$$c = -\frac{a_x - ia_y}{a_z - \|a\|}\sqrt{\frac{\|a\| - a_z}{2\|a\|}} \qquad d = \sqrt{\frac{\|a\| - a_z}{2\|a\|}}, \tag{1.3.32}$$

and c' and d' are

$$c' = -\frac{a_x - ia_y}{a_z + \|a\|}\sqrt{\frac{a_z + \|a\|}{2\|a\|}} \qquad d' = \sqrt{\frac{a_z + \|a\|}{2\|a\|}}. \tag{1.3.33}$$

In this way we can write

$$A = \frac{1}{2}(I+\vec{a}\cdot\vec{\sigma}) = \frac{1}{2}[|+\rangle\langle+| + |-\rangle\langle-| + \|a\|\ |\vec{a}\cdot\vec{\sigma},+\rangle\langle\vec{a}\cdot\vec{\sigma},+| - \|a\|\ |\vec{a}\cdot\vec{\sigma},-\rangle\langle\vec{a}\cdot\vec{\sigma},-|]$$

and so

$$A = \frac{1}{2}\big[(1 + \|a\|\ |c|^2 - \|a\|\ |c'|^2)|+\rangle\langle+| \ + \ (1 + |d|^2\ \|a\| - \|a\|\ |d'|^2)|-\rangle\langle-| + \\ (cd^* - c'd'^*)\|a\|\ |+\rangle\langle-| + (dc^* - d'c'^*)\|a\|\ |-\rangle\langle+|\big].$$

In this way we have written the qubit effect as a convex combination of the projections $\{|+\rangle\langle+|,\ |-\rangle\langle-|,\ |+\rangle\langle-|,\ |-\rangle\langle+|\}$. ■

1.4 Composite systems

Now we want to discuss some aspects of composite systems that will be useful in the following chapters. The problem we have to deal with is how to describe a system that is composed by the combination of two physical systems A and B. However we consider the case in which the two systems are distinguishable, and so they may represent two distinguishable particles, two different composite objects (e.g. two atoms or molecules), or two different degrees of freedom of the same object. These two systems will be called subsystems and we will formally denote the compound system as $A + B$. Another assumption we will make is that the two systems are statistically independent, which means that the two preparation procedures and the

possible measurements on the two systems A and B are independent.
So let us consider two systems A and B. These two systems are associated with two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, as well as the compound system $A+B$ is associated with the Hilbert space $\mathcal{H}_{AB}$. We now consider two effects $E_A \in \mathcal{E}(\mathcal{H}_A)$ and $E_B \in \mathcal{E}(\mathcal{H}_B)$ through which we can perform separate measurements on the two subsystems. Moreover, we must have an effect on the composite system $A+B$ that describes these possible measurements on the systems A and B. So we assume that there exists a mapping $\nu : \mathcal{E}(\mathcal{H}_A) \times \mathcal{E}(\mathcal{H}_B) \to \mathcal{E}(\mathcal{H}_{AB})$ such that $\nu(E_A, E_B)$ represents the separate measurement on the subsystems A and B in the system $A+B$. In the same way we must have a mapping λ from $\mathcal{S}(\mathcal{H}_A) \times \mathcal{S}(\mathcal{H}_B)$ in $\mathcal{S}(\mathcal{H}_{AB})$ such that $\lambda(\rho_A, \rho_B)$ describes the state of the two subsystems in the compound one. With the assumption of statistical independence of the two subsystems, we must have

$$tr[\lambda(\rho_A, \rho_B)\nu(E_A, E_B)] = tr[\rho_A E_A]tr[\rho_B E_B] \tag{1.4.1}$$

Up to this point, we have only stated the corresponding mathematical requirements associated with the request of statistical independence. Now it can be shown [8, 12] that if we consider the structure of tensor product, the condition (1.4.1) is satisfied. So, by defining $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$, $\lambda(\rho_A, \rho_B) = \rho_A \otimes \rho_B$ and $\nu(E_A, E_B) = E_A \otimes E_B$, we make a description of the compound system that agrees with the request of statistical independence. Now we want to see how we can recover the states ρ_A and ρ_B from the knowledge of the state ρ_{AB}. In order to do this we must use the notion of partial trace. So we premise the following [8]

Definition 1.4.1. The partial trace over the system A is defined as

$$tr_A : \mathcal{T}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathcal{T}(\mathcal{H}_B) \tag{1.4.2}$$

such that

$$tr_{\mathcal{H}_B}[tr_A[P]Q] = tr_{\mathcal{H}_{AB}}[P(I \otimes Q)], \tag{1.4.3}$$

where the $tr_{\mathcal{H}_B}$ denotes the trace on the Hilbert space $\mathcal{H}_B$, $tr_{\mathcal{H}_{AB}}$ denotes the trace in the Hilbert space $\mathcal{H}_{AB}$, $Q \in \mathcal{L}(\mathcal{H}_B)$ and $P \in \mathcal{L}(\mathcal{H}_A)$.

In the case the operator $P \in \mathcal{T}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is of the form $P = P_A \otimes P_B$, we have

$$tr[tr_A[P_A \otimes P_B]Q] \;=\; tr[(P_A \otimes P_B)(I \otimes Q)] \;=\; tr[P_A]tr[P_B Q] \tag{1.4.4}$$

where the operator $Q \in \mathcal{L}(\mathcal{H}_B)$, and it is understood that the traces are performed in the spaces in which the operators are defined. Since this is true for every $Q \in \mathcal{L}(\mathcal{H}_B)$ we can conclude that $tr_A[P_A \otimes P_B] = tr[P_A]P_B$ and similarly $tr_B[P_A \otimes P_B] = tr[P_B]P_A$.

Remark 1.4.1. If we consider an operator $E \in \mathcal{L}(\mathcal{H}_B)$ that is an effect, then the operator $I \otimes E$ represents a measurement that is performed only on the subsystem B, in the same way as $E \otimes I$, with $E \in \mathcal{L}(\mathcal{H}_A)$, represents a measurement performed only on subsystem A.

We can also show how the partial trace appears in coordinate, i.e. if we choose a basis in the Hilbert space.
It can be shown [17] that if $\{\alpha_i\}$ and $\{\beta_j\}$ are the bases respectively of the Hilbert space $\mathcal{H}_A$ and $\mathcal{H}_B$, we can obtain a basis for the whole Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ by forming the tensor product $\{\alpha_i \otimes \beta_j\}$ of the two bases. Now suppose we have an operator $S \in \mathcal{T}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Using the basis $\{\alpha_i \otimes \beta_j\}$ we can write it as

$$S = \sum_{i,j} \sum_{m,n} \langle \alpha_i \otimes \beta_j |\ S\alpha_m \otimes \beta_n \rangle\ |\alpha_i\rangle\langle\alpha_m| \otimes |\beta_j\rangle\langle\beta_n|. \tag{1.4.5}$$

Now the partial trace over A is such that

$$tr_A[|\alpha_i\rangle\langle\alpha_m| \otimes |\beta_j\rangle\langle\beta_n|] \ = \ tr_A[|\alpha_i\rangle\langle\alpha_m|]\ |\beta_j\rangle\langle\beta_n| \ = \ \delta_{im}\ |\beta_j\rangle\langle\beta_n|, \tag{1.4.6}$$

thus we obtain

$$tr_A[S] = \sum_{i,j,n} \langle \alpha_i \otimes \beta_j |\ S\alpha_i \otimes \beta_n \rangle\ |\beta_j\rangle\langle\beta_n|. \tag{1.4.7}$$

In a similar way we can also perform the trace tr_B over the system B.

Remark 1.4.2. Using (1.4.7) we can prove another useful relation that will be used in the following. If we consider a vector $\psi \in \mathcal{H}_B$ we have then

$$\langle \psi | tr_A[S] \psi \rangle \ = \ \sum_i \langle \alpha_i \otimes \psi | S\alpha_i \otimes \psi \rangle. \tag{1.4.8}$$

This follows from the fact that for $\psi \in \mathcal{H}_B$ we can write

$$|\psi\rangle = \sum_l \langle \beta_l | \psi \rangle\ |\beta_l\rangle, \tag{1.4.9}$$

such that, from a direct substitution of (1.4.9) in the left hand side of (1.4.8) and using (1.4.7) we obtain

$$\sum_{l,s} \sum_{i,j,n} \langle \alpha_i \otimes \beta_j |\ S\alpha_i \otimes \beta_n \rangle\ \langle \beta_l | \psi \rangle\ \langle \psi | \beta_s \rangle\ \langle \beta_j | \beta_s \rangle\ \langle \beta_n | \beta_l \rangle \ = \tag{1.4.10}$$

$$\sum_{l,s} \sum_{i,j,n} \langle \alpha_i \otimes \beta_j |\ S\alpha_i \otimes \beta_n \rangle\ \langle \beta_l | \psi \rangle\ \langle \psi | \beta_s \rangle\ \delta_{js}\delta_{nl} \ = \tag{1.4.11}$$

$$\sum_{i,j,n} \langle \alpha_i \otimes \beta_j |\ S\alpha_i \otimes \beta_n \rangle\ \langle \beta_n | \psi \rangle\ \langle \psi | \beta_j \rangle \ = \tag{1.4.12}$$

$$\sum_i \langle \alpha_i \otimes \psi |\ S\alpha_i \otimes \psi \rangle. \tag{1.4.13}$$

We now observe that the partial trace of a state is again a state. This is the content of the following proposition [8]:

Proposition 1.4.1. *If S is an operator in $\mathcal{T}(\mathcal{H}_A \otimes \mathcal{H}_B)$, then*

$$S \geq O \Longrightarrow tr_A[S] \geq O \ \ and \ \ tr_B[S] \geq O. \tag{1.4.14}$$

We have said that every effect of the form $E \otimes I$ with $E \in \mathcal{E}(\mathcal{H}_A)$ represents a measurement on the subsystem A only. On the other hand, from the definition of partial trace, we know that $tr[(E \otimes I)\rho] = tr[tr_B[\rho]E]$, and we have also seen, through proposition 1.4.1, that the partial trace of a state is again a state. From this, we can conclude that the state $tr_B[\rho]$ can be identified with the state of the subsystem A.
So what we observe is that if we start from a state $\rho \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the state $tr_A[\rho] = \rho_B$ represents the state of the subsystem B, in the same way as the state $tr_B[\rho] = \rho_A$ represents the state of the subsystem A.
Thus we conclude that starting from ρ_{AB}, we can recover the states of the two subsystems using the partial trace. However, this does not completely specify the relations between the marginals ρ_A, ρ_B and ρ_{AB}. In fact, if we start from the two marginals ρ_A and ρ_B, it can be shown [8] that in addition to the state $\rho_A \otimes \rho_B$, we can construct also other states in the total system $A+B$. So specifying the states ρ_A and ρ_B does not determine uniquely the state of the compound system. However, there is a case in which the reduced states specify the joint state completely. In fact we have

Proposition 1.4.2. *Let ρ_{AB} be a state in $\mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Suppose that $\rho_A = tr_B[\rho_{AB}]$ and $\rho_B = tr_A[\rho_{AB}]$ are pure states. Then the state ρ_{AB} is of the form $\rho_A \otimes \rho_B$*

Proof. We start from two states ρ_A and ρ_B such that $\rho_A = P_\psi$ with $\psi \in \mathcal{H}_A$ and $\rho_B = P_\phi$ with $\phi \in \mathcal{H}_B$. We must show that $\rho_{AB} = P_\psi \otimes P_\phi$. So we consider two orthonormal bases of the two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ that we denote respectively as $\{\alpha_i\}_i$ and $\{\beta_j\}_j$ such that $\alpha_1 = \psi$ and $\beta_1 = \phi$. Next we use (1.4.8) and so we obtain

$$\langle\psi|tr_B[\rho_{AB}]\psi\rangle = \langle\psi|\rho_A\psi\rangle = 1 = \sum_j \langle\alpha_1 \otimes \beta_j|\rho_{AB}\alpha_1 \otimes \beta_j\rangle \tag{1.4.15}$$

and similarly

$$\langle\phi|tr_A[\rho_{AB}]\phi\rangle = \langle\phi|\rho_B\phi\rangle = 1 = \sum_i \langle\alpha_i \otimes \beta_1|\rho_{AB}\alpha_i \otimes \beta_1\rangle. \tag{1.4.16}$$

However it is also true that

$$\langle \alpha_i \otimes \beta_j | \rho_{AB} \alpha_i \otimes \beta_j \rangle \geq O \tag{1.4.17}$$

and

$$\sum_{ij} \langle \alpha_i \otimes \beta_j | \rho_{AB} \alpha_i \otimes \beta_j \rangle = tr[\rho_{AB}] = 1. \tag{1.4.18}$$

From (1.4.15),(1.4.16) and (1.4.17),(1.4.18) it follows that

$$\langle \alpha_i \otimes \beta_j | \rho_{AB} \alpha_i \otimes \beta_j \rangle = \delta_{1j} \delta_{1i}. \tag{1.4.19}$$

Thus we can conclude that $\rho_{AB} \alpha_i \otimes \beta_j = 0$ if $i \neq 1$ and $j \neq 1$. If we now use the decomposition of ρ_{AB} in (1.4.5) we obtain

$$\langle \alpha_1 \otimes \beta_1 | \rho_{AB} \alpha_1 \otimes \beta_1 \rangle P_\psi \otimes P_\phi = \langle \psi \otimes \phi | \rho_{AB} \psi \otimes \phi \rangle P_\psi \otimes P_\phi. \tag{1.4.20}$$

From the observation that $tr[\rho_{AB}] = 1$ we conclude that $\rho_{AB} = P_\psi \otimes P_\phi$. □

1.5 States as Functionals

In this final subsection, we want to focus our attention on a particular aspect of states. As we have anticipated in the introduction of this chapter, we can naturally associate each state with a positive linear functional on the space of observables. Moreover, if the C^* algebra of observables is a commutative one, i.e. in the case of a classical system, through the Gelfand-Naimark theorem, and making use of the Riesz-Markov representation theorem, we can associate every state with a regular Borel measure on phase space. In order to clarify this point, we remember some notions about classical systems described on phase space. We know that the dynamic of a system can be described in an Hamiltonian form. In particular we consider a $2n + 1$ differentiable manifold, known as the phase space $\mathcal{S}$, in which locally we can chose a set of symplectic coordinates $t, q_1 \dots q_n, p_1 \dots p_n$. A physical system is described by a point on this space. Its dynamic is described in terms of Hamiltonian curves of the Hamiltonian H of the system, that are solutions of the following set of differential equations:

$$\frac{dq^k}{dt} = \frac{\partial H(t, q(t), p(t))}{\partial p_k} \quad k = 1 \dots n \tag{1.5.1}$$

$$\frac{dp_k}{dt} = -\frac{\partial H(t, q(t), p(t))}{\partial q^k} \quad k = 1 \dots n \tag{1.5.2}$$

Every integral curve determines, at fixed time t, once the initial conditions have been determined, a point $x(t)$, such that $(t, x(t))$ is the state of the system at time t.
However, there are also cases, such as in statistical mechanics, in which the state of the system is not known with certainty. In this case one usually use the concept of statistical ensemble, in which one considers a set of identical and independent copies of the system, whose states are distributed in the space $\mathcal{S}$ with a probability density given locally by a map $\rho = \rho(t, q(t), p(t))$. This function evolves in time according to the Liouville's equation

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{n} \left(\frac{\partial \rho}{\partial q^i} \frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial q^i} \frac{\partial \rho}{\partial p_i} \right) = 0. \tag{1.5.3}$$

$\rho(x, t)$ represents so the probability density that the physical system is in the state x at time t. In general, also $\rho(x, t)$ is considered as the state of the system at time t [3]. So we can distinguish between to possible states: the sharp state, that is described by a point on phase space $\mathcal{S}$, and the probabilistic state, that is described in terms of a probability density $\rho(x, t)$ on $\mathcal{S}$. In either case the state at time t can be seen as a Radon probability measure $\{\mu_t\}_{t\in\mathbb{R}}$. In particular for a sharp state we consider the measure $\mu_t = \delta_{x(t)}$ where $\delta_{x(t)}$ is the Dirac measure centred in the point x of phase space, i.e. the measure centred in $x \in X$ defined as

$$\delta_x(E) = 0 \;\; if \;\; x \notin E \quad and \quad \delta_x(E) = 1 \;\; if \;\; x \in E \quad \forall E \in \Sigma,$$

with (X, Σ) a measure space. For a probabilistic state we consider

$$\mu_t(E) = \int_E \rho(t, x) d\nu_t,$$

with $E \in \mathcal{B}(\mathcal{S})$, where $d\nu_t$ is the the Lebesgue measure locally written as $dq^1 \dots dq^n$ $dp_1 \dots dp_n$ and $\mathcal{B}(\mathcal{S})$ denotes the Borel σ-algebra of $\mathcal{S}$. In this way we see that we can associate a Radon measure with a physical state on phase space.
On the other hand, we know that observables are represented by functions on phase space, and so we can calculate the expectation value of an observable f on a given state ω as

$$\langle f \rangle_\omega = \int_E f d\mu_t, \tag{1.5.4}$$

where μ_t is the Radon measure associated with the state of the system. So in the case we are considering a sharp state, i.e. the Radon measure is a Dirac measure centered in a point x_0 in phase space, we obtain that

$$\langle f \rangle_\omega = \int_E f \delta_{x_0} = f(x_0), \tag{1.5.5}$$

if $x_0 \in E$, from which we recover the idea of a state as a point in phase space. In the same way for a probabilistic state, we obtain

$$\langle f \rangle_\omega = \int_E f d\mu_t = \int_E f \rho(t,x) d\nu_t. \tag{1.5.6}$$

However if we consider $\langle f \rangle_\omega \coloneqq \omega(f)$, i.e. the expectation value of f on ω as the value that the functional, denoted with the same letter ω, assumes on the function f, then equation (1.5.4) is nothing else that the Riesz-Markov representation theorem[17].

Remark 1.5.1. Through the measure μ_t just defined, we can assign a truth value to every proposition concerning the physical system. This is done in this way. First, let us consider a proposition $\bar{P}$ associated with our physical system. This proposition identifies a subspace P, of the phase space $\mathcal{S}$, that contains all the points $x(t)$ that make it true. Now let the state of our system be a sharp state, so that it is described by a Dirac measure $\mu_t = \delta_{x(t)}$. The proposition $\bar{P}$, at time t, is true for our system, if and only if the point $x(t)$ that describes the system is contained in the set P. So if we assign the value 0 to a false proposition at time t and the value 1 to a true one at the same time, we observe that the truth value of the proposition $\bar{P}$ can be obtained through $\mu_t(P)$. Similar considerations can also be made in the case the state of the system is described by a probabilistic state.

So what we have seen is that in classical mechanics there exist states that are "dispersion free", i.e. can assume only two possible values: 0 and 1. Moreover, this feature is a consequence [6] of the commutativity of the algebra on functions associated with classical observables. If we make similar considerations in quantum mechanics we will find that every quantum state can be associated with a generalized probability measure (Busch and Gleason theorem), but there not exist dispersion free states. This, in conclusion, is strictly linked to the non-commutativity of quantum mechanics, and so is what makes the crucial mathematical difference between quantum and classical mechanics.
So, we begin by observing how we can associate with every quantum state, i.e. with every density operator, a functional acting on the space of projectors, or said equivalently, a probability measure on the Hilbert space of the quantum system. In particular this is what is shown by the following theorem due to Gleason [18, 3]:

Theorem 1.5.1. *Let $\mathcal{H}$ be a separable Hilbert space of dimension $d \geq 3$. Then, for each probability measure μ on $\mathcal{P}(\mathcal{H})$, i.e. for each map $\mu : \mathcal{P}(\mathcal{H}) \to [0,1]$ satisfying the condition $\mu(I) = 1$ and $\mu(\sum_{i=1}^{+\infty} P_i) = \sum_{i=1}^{+\infty} \mu(P_i)$ for $\{P_i\}_i \subset \mathcal{P}(\mathcal{H})$ with $P_i P_j = O$ for $i \neq j$, there exists a unique operator $\rho_\mu \in \mathcal{S}(\mathcal{H})$ such that*

$$\mu(P) = tr[\rho_\mu P] \quad \forall P \in \mathcal{P}(\mathcal{H}) \tag{1.5.7}$$

Remark 1.5.2. A profound consequence of the Gleason's theorem is that it shows how the Born rule is the only possible rule for calculating probabilities in a theory founded on Hilbert space. In fact, historically the theorem arose as an answer to the question posed by Georege Mackey if one could calculate the probability in quantum mechanics in a way different from the Born rule. So Gleason proved that, if the dimension of the Hilbert space is at last three, the Born rule is the only way through which quantum probabilities can be computed. However, the theorem also say that in dimension two there are probability measures that do not correspond to quantum states.

So, what we see is that, as a consequence of this theorem, one can conclude that in quantum mechanics there not exist dispersion free states. In particular this follows immediately if we set $P = \frac{1}{2}\rho$ in (1.5.7) obtaining $\mu\left(\frac{1}{2}\rho\right) = \frac{1}{2}tr[\rho^2]$ and hence $0 < \mu\left(\frac{1}{2}\rho\right) < 1$.

A generalization of the Gleason's theorem can be obtained by introducing the concept of generalized probability measure on the set of effects. In this case, as we will see, every generalized probability measure is associated with a unique trace class operator, representing a quantum state, without any restrictions on the dimension of the Hilbert space, as is shown in the Busch theorem[14]. In order to define such a generalized probability measure, we must first observe how we can define a partial binary operation on the set of effects, that induces a partial order. Then, we must observe how we can associate with every state a mapping on the space of effects, that preserves the partial order induced by the partial binary operation. With these elements, we can define a mapping known as generalized probability measure. Finally, by considering for the set of effects the set $\mathcal{E}(\mathcal{H})$, we will state a theorem due to Busch that shows how each state can be seen as a generalized probability measure on the set of effects $\mathcal{E}(\mathcal{H})$.
Now we first start by noting that, from the fact that effects generate affine mappings on the set of states, we can define [8] a **partial binary operation** $\diamond$ on the set of effects. In particular let E_1,E_2 and E_3 be three effects. If for every state ρ the condition:

$$E_1(\rho) + E_2(\rho) = E_3(\rho) \tag{1.5.8}$$

is satisfied, then we can write

$$E_1 \diamond E_2 = E_3. \tag{1.5.9}$$

Since $E(\rho) \in [0, 1]$ for every state ρ, we observe that $E_1 \diamond E_2$ exists if and only if $E_1(\rho) + E_2(\rho) \leq 1$, for every state. From this condition, we see that the operation $\diamond$ is not defined for every pair of effects since if we consider for example the identity effect I we see that $I \diamond I$ is not defined. This also justifies the name partial operation.

Through the partial binary operation $\diamond$ we can also define the structure of **effect algebra** [8, 4]. More precisely, we have

Definition 1.5.1. A set $\mathcal{T}$ with two distinct elements O and I and equipped with a partial binary operation $\diamond$ is called an effect algebra, if the binary operation satisfies the following conditions:

1) if $E_1 \diamond E_2$ exists then $E_2 \diamond E_1$ also exists and $E_1 \diamond E_2 = E_2 \diamond E_1$

2) if $E_1 \diamond E_2$ and $(E_1 \diamond E_2) \diamond E_3$ exist then also $E_2 \diamond E_3$ and $E_1 \diamond (E_2 \diamond E_3)$ exist and $E_1 \diamond (E_2 \diamond E_3) = (E_1 \diamond E_2) \diamond E_3$

3) for every E there exists a unique $\bar{E}$ such that $E \diamond \bar{E} = I$

4) if $E \diamond I$ exists then $E = O$.

Remark 1.5.3. We observe that the set of effects is an effect algebra. In fact the O and I elements exist and coincide with the zero effect and the identity effect. If $E_1 \diamond E_2$ exists then $E_1(\rho) + E_2(\rho) = E_3(\rho) = E_2(\rho) + E_1(\rho)$ and so also $E_2 \diamond E_1$ exists and $E_1 \diamond E_2 = E_2 \diamond E_1$. If E_1, E_2, E_3 are three effects and if $E_1 \diamond E_2$ and $(E_1 \diamond E_2) \diamond E_3$ exist it means that $(E_1(\rho) + E_2(\rho)) + E_3(\rho)$ exists for every state ρ. But since $(E_1(\rho) + E_2(\rho)) + E_3(\rho) = E_1(\rho) + (E_2(\rho) + E_3(\rho))$ we have that also $E_1 \diamond (E_2 \diamond E_3)$ and $E_2 \diamond E_3$ exist. For every effect E, there exists a unique effect $\bar{E} = I - E$ such that $E \diamond \bar{E} = I$. Eventually if $E \diamond I$ exists, it must be $E(\rho) + I(\rho) \leq 1$ for all states ρ. But $I(\rho) = 1$ and from the positivity of effects we conclude that it must be $E = O$.

So we have seen that the set of effects equipped with the partial binary operation $\diamond$ constitutes an effect algebra. This partial operation also allows us to define a partial order on the set of effects. In fact, for two effects E_1 and E_2 we can say that $E_1 \leq E_2$ if there exists an effect $\tilde{E}_1$ such that $E_1 \diamond \tilde{E}_1 = E_2$. Through this partial order we can also treat infinite sequences by declaring that $E_1 \diamond E_2 \diamond E_3 \diamond \ldots$ is the last upper bound of the increasing sequence E_1, $E_1 \diamond E_2$, $E_1 \diamond E_2 \diamond E_3, \ldots$, if it exists. Now we observe that since the effects are affine mappings from the set of states to $[0, 1]$, we can associate with every state a functional on the set of effects through the formula [8]:

$$g_\rho(E) \coloneqq E(\rho), \tag{1.5.10}$$

such that $g_\rho(O) = O(\rho) = 0$ and $g_\rho(I) = I(\rho) = 1$. Moreover if E_1 and E_2 are effects and $E_1 \diamond E_2$ exists then we also have

$$g_\rho(E_1 \diamond E_2) = (E_1 \diamond E_2)(\rho) = E_1(\rho) + E_2(\rho) = g_\rho(E_1) + g_\rho(E_2). \tag{1.5.11}$$

We also see that the function thus defined preserves the order induced by the binary operation $\diamond$:

$$E_1 \leq E_2 \Rightarrow g_\rho(E_1) \leq g_\rho(E_2). \tag{1.5.12}$$

Once we have defined a partial order and we have associated with every state a mapping on the set of effects, we can now define the concept of generalized probability measure. We have the following:

Definition 1.5.2. A generalized probability measure is a mapping from the set of effects to the interval $[0, 1]$ such that $g(O) = 0$ and $g(I) = 1$ and

$$g(E_1 \diamond E_2 \diamond \dots) = g(E_1) + g(E_2) + \dots \tag{1.5.13}$$

whenever $E_1 \diamond E_2 \diamond \dots$ exists.

Now we make the hypothesis that the space of effects is the set $\mathcal{E}(\mathcal{H})$, and so the binary operation $\diamond$ can be identified with the addition of operators. Then, we can state the following theorem due to Busch [14]:

Theorem 1.5.2. *Let g be a generalized probability measure on $\mathcal{E}(\mathcal{H})$. Then there exists a unique operator $\rho_g \in \mathcal{S}(\mathcal{H})$ such that*

$$g(E) = tr[\rho_g E] \qquad \forall E \in \mathcal{E}(\mathcal{H}) \tag{1.5.14}$$

So from this theorem we conclude that, if we chose for the set of effects the set $\mathcal{E}(\mathcal{H})$, we can identify the states in $\mathcal{S}(\mathcal{H})$ as generalized probability measures on effects.

We can demand if there exist generalized probability measures in quantum mechanics that are dispersion free. A simple argument shows that it is not the case. In fact what we need is a state ρ such that $g_\rho(E) = 0 \ \ or \ \ 1$ for every effect in $\mathcal{E}(\mathcal{H})$. Let us consider two effects E_0 and E_1 such that $g_\rho(E_0) = 0$ and $g_\rho(E_1) = 1$. Then we observe that $g_\rho(\frac{1}{2}E_0 + \frac{1}{2}E_1) = \frac{1}{2}$. So we have shown that there is no dispersion free generalized probability measure.

Quantum Observables 2

2.1 introduction

Every measurement can be described in mathematical terms using the concept of observable, whose mathematical form depends on the nature, quantum or classical, of the theory described. In the classical case, observables are represented by continuous real-valued functions on the phase space vanishing at infinity [6], while in the quantum case, observables are described in terms of selfadjoint operators [3]. As is known [1, 2], every selfadjoint operator admits a spectral representation with respect to a uniquely associated spectral measure or projection valued measure (PVM). In particular, a projection valued measure acts as a map that associates with every element of the Borel σ-algebra over $\mathbb{R}$ a projection operator. In other terms, the spectral theorem represents a way in which one can associate with every outcome of an experiment, i.e. with every element in the spectrum of the observable, a suitable projection operator, through which one can compute quantum mechanical probabilities. However, there are cases in which the possible experimental measurement results are associated not with projection operators but with effects, i.e. unit bounded selfadjoint operators whose properties were discussed in the previous chapter. An example can be a photon-detection measurement realized using a non-ideal photon-detector, i.e. a detector whose efficiency is not one. In this case, the distribution probability of the revealed photons can be calculated [19, 10] using a set of operators that are not projections but effects. Similarly, the measurements performed using a photon beam splitter in which the beam splitter has transparency $\eta < 1$, are associated with effects rather than orthogonal projections [8]. So, what one can see is that there are experiments, particularly in the field of quantum optics,

which require a generalization of the concept of observable as a projection valued measure. The form and the characteristics of this generalization is the content of the following chapter. Preliminarily, one can generalize the concept of quantum observable by defining it as a map that associates to every measurement result, an effect. In order to acquire the case in which the observable has a continuous set of outcomes, it is possible to introduce the concept of positive operator valued measure, POVM [10, 8]. A POVM is a map that associates with each element of the σ-algebra of events, in particular with each element of the Borel σ-algebra over $\mathbb{R}$, an effect. That this is a generalization of the concept of PVM, it follows from the fact that projection operators are a particular type of effects, i.e. they are extreme effects [8].

An important aspect of the spectral theorem is that there is a one to one correspondence between observables and PVMs. So, one can ask whether such a correspondence also applies to the generalized case in which one has a POVM rather than a PVM. The answer to this question is a consequence of the Naumark dilation theorem [20, 21] that we also will state and discuss. In particular, what we will observe is that any maximal symmetric operator can be extended, on a larger Hilbert space, to a selfadjoint operator that admits a spectral decomposition. Then, from this, it can be shown [10] that every maximal symmetric operator is uniquely associated with a POVM. So, the conclusion at which one arrives is that similarly to the case of selfadjoint operators which are in one to one correspondence with PVMs, maximal symmetric operators possess a spectral representation with respect to a unique POVM. Moreover, the Neumark theorem has relevance also from the physical point of view [10]. In fact, the larger Hilbert space on which the extension is realized can always be considered as the tensor product between the Hilbert space of the physical system $\mathcal{H}$, and some Hilbert space $\mathcal{H}'$ which may represent the Hilbert space of an environment system or a measuring apparatus. So, starting from an observable represented by a POVM, the Neumark's theorem ensures the existence, in principle, of a measurement device for it [10].

Having outlined the purposes and contents of the following chapter we now briefly describe its structure: In sec. 2.2 and 2.3 we introduce the definition of POVM and discuss some physical examples in which this generalized concept of observable naturally arises. In section 2.4 we concentrate on the structure of the space of POVMs. First, we discuss the problem of mixtures of observables [8] and describe under what conditions the space of POVMs became a convex set. Then, we will characterize the extremal elements of this space using the method of perturbations [13]. Sections 2.5 and 2.6 are devoted to the relationships between selfadjoint operators and POVMs. In particular, we will state the Neumark theorem and discuss some of its important consequences. Sections 2.7 and 2.8 are left to the discussion of the

informational completeness of a set of POVMs, and its application to the problem of state reconstruction.

2.2 Observables

When we consider an experiment, it is in general characterized by a set of possible outcomes, each of which is associated with an effect E. This collection of effects is what, in a preliminary way, we will call [8] an observable. However this is not the most general definition of a quantum observable and, in what follows, we will observe that each observable is more properly described by a POVM (positive operator-valued measure).
In order to fix the ideas let us consider an experiment in which we have n possible outcomes. Every outcome is associated with an effect E_i for $i = 1, \ldots, n$. In the following, we will assume that the condition

$$tr[\rho E_1] + tr[\rho E_2] + \cdots + tr[\rho E_n] = 1 \tag{2.2.1}$$

is always respected, where ρ is the density operator associated with the physical system. This condition means that every time we perform a measurement on the system, an outcome associated with an effect in the set $\{E_1, \ldots, E_n\}$ is found. So we assume that the set of effects that describe the experiment is "complete" [4, 8].

Remark 2.2.1. If the set of effects is not complete, we can always add some effects of the form $E_{n+1} = I - \sum_{i=1}^{n} E_i$ such that the condition (2.2.1) is satisfied.

Taking into account that the quantity $tr[\rho E_i]$ is the probability that the outcome associated with the effect E_i is registered, we can conclude that measurement devices are described by a set of effects $\{E_1, \ldots, E_n\}$ that satisfy these conditions:

$$0 \leq tr[\rho E_i] \leq 1 \tag{2.2.2}$$

$\forall i = 1, \ldots, n$ and

$$\sum_{i=1}^{n} tr[\rho E_i] = 1. \tag{2.2.3}$$

Since this conditions must be true for every state we can conclude that

$$O \leq E_i \leq I \qquad \forall i = 1, \ldots, n, \tag{2.2.4}$$

and also that

$$\sum_{i=1}^{n} E_i = I. \tag{2.2.5}$$

Indeed, if we consider the case in which the state of the system is described by a pure state $|\phi\rangle\langle\phi|$, for a vector $\phi \in \mathcal{H}$, then the condition $0 \leq tr[\rho E_i] \leq 1$ translates into $0 \leq tr[|\phi\rangle\langle\phi|E_i] \leq 1 \Longrightarrow 0 \leq \langle\phi|E_i\phi\rangle \leq 1$ which is satisfied if and only if $O \leq E_i \leq I$. In a similar way $\sum_{i=1}^{n}\langle\phi|E_i\phi\rangle = 1$ is fulfilled if and only if $\sum_{i=1}^{n} E_i = I$. Since all states can be decomposed as a combination of pure states we can conclude that these conditions are true for a generic state ρ. In this way we have obtained that the set of effects associated with a measurement device must verify (2.2.4) and (2.2.5).

Remark 2.2.2. An important consequence of condition (2.2.5) is that if the effects considered are represented by projection operators, then they describe disjoint events if and only if they are orthogonal. In fact, in the case we consider two effects associated with two one dimensional projections P_ϕ, P_ψ, for $\psi, \phi \in \mathcal{H}$, and we impose the condition (2.2.5) for a quantum state $\rho = \frac{1}{2}(|\psi\rangle\langle\psi| + |\phi\rangle\langle\phi|)$, we obtain

$$tr[\rho P_\phi] + tr[\rho P_\psi] = \langle\psi|\rho\psi\rangle + \langle\phi|\rho\phi\rangle = 1 + |\langle\phi|\psi\rangle|^2 \geq 1, \tag{2.2.6}$$

which implies $\langle\phi|\psi\rangle = 0$, i.e. that the two one dimensional projections must be orthogonal. Moreover, if we consider two effects described by projection operators not necessarily one dimensional, we arrive at the same conclusion, i.e. that they must be orthogonal if they describe two disjoint events of a fixed observable. In fact from the condition (2.2.5) we have that two effects S and T must satisfy $S + T < I$. However, for two projections this condition is equivalent to the orthogonality [4, 3].

So in conclusion we can think of an observable, in a preliminary way, as a collection of effects $\{E_1, \ldots, E_n\}$ satisfying the conditions (2.2.4) and (2.2.5).
This definition of observable, however, must be generalized in order to take in consideration the case in which one has a measurement with a continuous set of outcomes, such as measurements of position and momentum. In particular, we must generalize the concept of observable as a collection of effects to that of POVM [8, 10]. In order to describe the concept of POVM, we remember the definition of probability measure on a σ-algebra. So let A be a set and $\mathcal{A}$ its σ-algebra, i.e. a non-empty collection of subsets of A closed under complements and countable unions; that is if $X \in \mathcal{A}$ then its complement $A \setminus X$ is still an element of $\mathcal{A}$, and any countable union of elements $\{X_j\}_{j\in\mathbb{N}}$ of $\mathcal{A}$, $\cup_{j\in\mathbb{N}} X_j$, is again an element of $\mathcal{A}$. The pair $(A, \mathcal{A})$ is usually known as measurable space [22], and any subset $X \in \mathcal{A}$ of the σ-algebra is called an event. A function $\mu : \mathcal{A} \longrightarrow [0, 1]$ is a probability measure if it satisfies the following conditions:

1) $\mu(\emptyset) = 0$

2) $\mu(A) = 1$

3) $\mu(\cup_j X_j) = \sum_j \mu(X_j)$ for any sequence of sets $\{X_j\} \subset \mathcal{A}$ such that $X_i \neq X_j$ for $i \neq j$.

In particular the quantity $\mu(X)$ is the probability for an event X to occur. Now we can define the notion of POVM:

Definition 2.2.1. A POVM (positive operator-valued measure) is a mapping $E : \mathcal{A} \longrightarrow \mathcal{E}(\mathcal{H})$ that satisfies

1) $E(\emptyset) = O$

2) $E(A) = I$

3) $E(\cup_j X_j) = \sum_j E(X_j)$ for any sequence of sets $\{X_j\} \subset \mathcal{A}$ such that $X_i \neq X_j$ for $i \neq j$

So a POVM is a map that associates with every subset X of a σ-algebra, an effect $E(X)$. In this way a POVM defines a mapping such that $X \longrightarrow \langle\phi|E(X)\phi\rangle = tr[|\phi\rangle\langle\phi|E(X)]$ is a probability measure. Since a state can be written as a combination of pure states we conclude that a mapping $E : \mathcal{A} \longrightarrow \mathcal{E}(\mathcal{H})$ is a POVM if and only if the map $X \mapsto tr[\rho E(X)]$ is a probability measure for every $\rho \in \mathcal{S}(\mathcal{H})$. We identify observables with POVMs.
Now we pass to the discussion of some experiments in which the notion of POVM arises naturally.

Example 2.2.1. In this example, we consider the Stern-Gerlach experiment [8, 10]. In particular we consider a beam of silver atoms, produced in a furnace, that is directed through an inhomogeneous magnetic field and subsequently impinges on a glass plate. The result of the experiment is that on the screen there appears a pair-of-lips shape, i.e. the original beam was split into two beams according to the possible values of the spin component in the direction of the magnetic field.
So let an atom carrying spin-$\frac{1}{2}$ be prepared in a spin state $|\phi\rangle = c_+|\phi_+\rangle + c_-|\phi_-\rangle$. We suppose that its center of mass is represented by a wave packet $|\psi\rangle$, so that the initial state of the atom is described by the vector $|\phi \otimes \psi\rangle$. The passage in the Stern-Gerlach apparatus produces a unitary evolution which couples the spin degrees of freedom with its translational motion. We obtain in exit the state $|\Phi\rangle = c_+|\phi_+ \otimes \psi_+\rangle + c_-|\phi_- \otimes \psi_-\rangle$ where the states $|\psi_\pm\rangle$ represent the wave packets deflected up or down. We now must describe the registration on the screen. This can be achieved by using the two projections P_+ and P_- (which constitute the so called screen observable), which describe the localization of the atom in the upper or in the lower half planes of the screen. The corresponding probabilities expressed with respect to the incoming spin state ϕ can be found to be

$$\langle\Phi|P_\pm \otimes I\Phi\rangle = |c_+|^2\langle\psi_+|P_\pm\psi_+\rangle + |c_-|^2\langle\psi_-|P_\pm\psi_-\rangle = \langle\phi|F_\pm\phi\rangle, \tag{2.2.7}$$

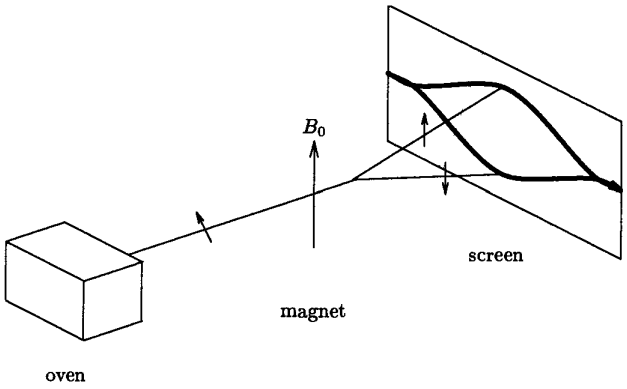

Figure 2.1: Stern-Gerlach experiment

where the operators F_+ and F_- correspond to two effects:

$$F_\pm \coloneqq \langle\psi_+|P_\pm\psi_+\rangle|\phi_+\rangle\langle\phi_+| + \langle\psi_-|P_\pm\psi_-\rangle|\phi_-\rangle\langle\phi_-|. \tag{2.2.8}$$

We observe that $F_+ + F_- = I$, but $F_\pm^2 \neq F_\pm$ and so they are not projections. So the set $\{F_+, F_-\}$ constitutes an observable known as unsharp spin observable [10]. If the center of the mass wave packets $|\psi_\pm\rangle$ are well separated and localized in the screen, i.e. if $\langle\psi_\pm|P_\pm\psi_\pm\rangle = 1$ and $\langle\psi_\pm|P_\mp\psi_\pm\rangle = 0$, then $F_\pm$ coincide with the usual projection operators $|\phi_+\rangle\langle\phi_+|$ and $|\phi_-\rangle\langle\phi_-|$. ■

Example 2.2.2. Another example of POVM emerges in a photodetection experiment using a nonideal photodetector [19]. This nonideality means that the detector does not detect all the photons that hit it. In particular, we assume that a photon hitting the detector is detected with probability η, with $0 \leq \eta \leq 1$. Now if the electromagnetic field of the light radiation is in the number state $|n\rangle$, the probability $p(k|n)$ that the detector will detect k photons is

$$\begin{aligned} p(k|n) &= \binom{n}{k}\eta^k(1-\eta)^{n-k} \qquad\qquad if \quad n \geq k \\ p(k|n) &= 0 \qquad\qquad if \quad n < k \end{aligned}$$

However we can obtain this probability by introducing a suitable observable. In particular there exists an effect $F(k)$ such that $p(k|n) = tr[F(k)|n\rangle\langle n|]$. The operator

$F(k)$ that reproduces the correct value of $p(k|n)$ is [19]

$$F(k) = \sum_{m=k}^{+\infty} \binom{m}{k} \eta^k (1-\eta)^{m-k} |m\rangle\langle m|. \tag{2.2.9}$$

We observe that $F(k) \geq O$ and it can be shown that $\sum_k F(k) = I$. However the effects $F(k)$ are not projections unless $\eta = 1$. So we can conclude that the $F(k)$ generate a POVM, known as the photon counting observable [8, 10]. ■

Remark 2.2.3. If we consider a POVM associated with an observable A, defined on a measurable space $(M, \mathcal{M})$ (also known as the outcome space of A), we have observed that for every $X \in \mathcal{M}$, we obtain, through the relation $tr[\rho E(X)]$, a probability measure. In this way, a POVM can also be seen as a collection of probability measures. In other terms, we can say that an observable A with outcome space $(M, \mathcal{M})$, defines a mapping $\theta_A : \mathcal{S}(\mathcal{H}) \longrightarrow \mathbb{P}(M)$ through

$$\theta_A(\rho) \coloneqq tr[\rho E(\cdot)] \tag{2.2.10}$$

such that for every state $\rho \in \mathcal{S}(\mathcal{H})$, it associates a probability measure on M (here we indicate with $\mathbb{P}(M)$ the set of probability measures on M). The map θ_A is usually called the statistical map associated with the observable A. We also note that, from the linearity of the trace, the statistical map is an affine mapping, i.e.

$$\theta_A(\lambda\rho_1 + (1-\lambda)\rho_2) = \lambda\theta_A(\rho_1) + (1-\lambda)\theta_A(\rho_2), \tag{2.2.11}$$

for every $\rho_1, \rho_2 \in \mathcal{S}(\mathcal{H})$ and for $0 \leq \lambda \leq 1$. Moreover it can be shown [8] the following

Proposition 2.2.1. *Let θ be an affine mapping from $\mathcal{S}(\mathcal{H})$ into $\mathbb{P}(M)$. Then there exists an observable A such that $\theta = \theta_A$.*

So we conclude that every affine mapping from $\mathcal{S}(\mathcal{H})$ to $\mathbb{P}(M)$ is of the form of eq.(2.2.10).

Example 2.2.3. As an example of statistical map [8], let us consider a qubit observable Q that consists of three effects $Q(l) = \frac{1}{2}(\beta_l I + \vec{r_l} \cdot \vec{\sigma})$, for $l = 1, 2, 3$, $\beta_l \in \mathbb{R}$ and $\vec{r_l} \in \mathbb{R}^3$ and $\|\vec{r_l}\| \leq \beta_l \leq 2 - \|\vec{r_l}\|$. The normalization condition (2.2.5) implies that $\sum_l \beta_l = 2$ and that $\sum_l \vec{r_l} = \vec{0}$. So we conclude that the vectors $\vec{r_l}$ are coplanar. Now if we consider the three effects $Q(l)$ given by:

$$Q(1) = \frac{1}{3}(I + \sigma_y)$$

$$Q(2) = \frac{1}{3}(I + \frac{\sqrt{3}}{2}\sigma_x - \frac{1}{2}\sigma_y)$$

$$Q(3) = \frac{1}{3}(I - \frac{\sqrt{3}}{2}\sigma_x - \frac{1}{2}\sigma_y),$$

the probability of observing the outcome l is given by $p_l = tr[Q(l)\rho]$. If the state ρ is expressed as $\rho = \frac{1}{2}(I + \vec{m}\cdot\vec{\sigma})$, we obtain

$$tr\Big[\frac{1}{2}(I + \vec{m}\cdot\vec{\sigma})\frac{1}{3}(I + \vec{r_l}\cdot\vec{\sigma})\Big] = \tag{2.2.12}$$

$$tr\Big[\frac{1}{6}(I + \vec{m}\cdot\vec{\sigma} + \vec{r_l}\cdot\vec{\sigma} + \vec{m}\cdot\vec{\sigma}\vec{r_l}\cdot\vec{\sigma})\Big] = \tag{2.2.13}$$

$$tr\Big[\frac{1}{6}(I + \vec{m}\cdot\vec{\sigma} + \vec{r_l}\cdot\vec{\sigma} + m_i r_{(l)j}(I\delta_{ij} + i\epsilon_{ijk}\sigma_k))\Big] = \tag{2.2.14}$$

$$\frac{1}{3}(1 + \vec{m}\cdot\vec{r_l}) \tag{2.2.15}$$

In this way the associated statistical map can be written as

$$\theta_Q : \rho \quad \mapsto \quad \vec{p} = \frac{1}{3}(1 + y, 1 + \frac{\sqrt{3}}{2}x - \frac{1}{2}y, 1 - \frac{\sqrt{3}}{2}x - \frac{1}{2}y), \tag{2.2.16}$$

where x and y are the components of the Bloch vector of the state ρ.

■

2.3 Real Observables

In the case the sample space M of the observable A is the set of real numbers $\mathbb{R}$, the observable is usually called a real observable. In this case the σ-algebra on $\mathbb{R}$ is represented by the Borel σ-algebra $\mathcal{B}(\mathbb{R})$, where we remember that given a topological space X, the Borel σ-algebra $\mathcal{B}(X)$ of X is the smallest σ-algebra containing all open sets of X. So we can make the following [10]

Definition 2.3.1. An observable A is real, or real valued, if the sample space of A is either $\mathbb{R}$ or a subset of $\mathbb{R}$.

Example 2.3.1. In the example of the Stern-Gerlach observable, we have seen that the outcome space is the set $\{+,-\}$. If however, we label the elements in the outcome space as $\{0,1\}$ we obtain a real observable.

For a real observable A, we can define the concepts of expectation value and variance. In particular, the expectation value of A is defined as

$$\langle A\rangle_\rho \coloneqq \int_{\mathbb{R}} x\ tr[\rho A(dx)], \tag{2.3.1}$$

where the integration is made with respect to the probability measure $tr[\rho A(\cdot)]$, while the variance $\Delta_\rho(A)$ is defined as

$$(\Delta_\rho(A))^2 \coloneqq \int_{\mathbb{R}} (x - \langle A\rangle_\rho)^2 tr[\rho A(dx)]. \tag{2.3.2}$$

In the case the sample space M is a countable set of real numbers, eq.(2.3.1) and (2.3.2) became respectively

$$\langle A\rangle_\rho = \sum_{x_i \in M} x_i \; tr[\rho A(x_i)], \tag{2.3.3}$$

$$(\Delta_\rho(A))^2 = \sum_{x_i \in M} (x_i - \langle A\rangle_\rho)^2 tr[\rho A(x_i)]. \tag{2.3.4}$$

Example 2.3.2. In this example, we consider the Polarization observable. If we consider an apparatus composed by a polarization filter, i.e. a filter that transmits only light polarized in a given direction, and a photodetector, then their action constitutes a measurement of the polarization of a photon. Assuming, as an experimentally verified fact, that the polarization property is associated with the two dimensional Hilbert space $\mathbb{C}^2$, we can consider the following POVM associated with such an apparatus:

$$A : \{0, 1\} \to \; \mathcal{E}(\mathbb{C}^2), \tag{2.3.5}$$

where 0 and 1 correspond respectively to no detection and detection of a photon. The effects corresponding to $A(0)$ and $A(1)$ are two projections:

$$A(0) = \frac{1}{2}(I - \vec{n}\cdot\vec{\sigma}), \qquad\qquad A(1) = \frac{1}{2}(I + \vec{n}\cdot\vec{\sigma}) \tag{2.3.6}$$

with $\vec{n}$ a unit vector in $\mathbb{R}^3$. Using eq.(2.3.3) and eq.(2.3.4), we can calculate the expectation value and the variance of A, assuming that the photon state is expressed as $\rho = \frac{1}{2}(I + \vec{m}\cdot\vec{\sigma})$. We find

$$\langle A\rangle_\rho = tr[\rho A(1)] = \frac{1}{2}(1 + \vec{n}\cdot\vec{m})$$

and

$$(\Delta_\rho(A))^2 = \langle A\rangle_\rho^2 tr[\rho A(0)] + (1 - \langle A\rangle_\rho)^2 tr[\rho A(1)] =$$
$$\frac{1}{8}(1 + \vec{n}\cdot\vec{m})^2(1 - \vec{n}\cdot\vec{m}) + \frac{1}{8}(1 - \vec{n}\cdot\vec{m})^2(1 + \vec{n}\cdot\vec{m}) =$$
$$\frac{1}{4}(1 + \vec{n}\cdot\vec{m})(1 - \vec{n}\cdot\vec{m}).$$

■

2.4 Mixture of observables and pure observables

Mixture of observables

We pass now to characterize the concept of mixture of observables [8]. We have seen that if we alternate the preparation procedure of a system while fixing the measurement, we obtain a new state that is the convex combination of the preparations. Similarly, we can fix the state of a system and alternate the measurements performed on it. In this way, we can describe a mixture of observables. However, in this case, the mixture of the observables depends on the particular nature of the outcome set of the observables considered, as well as the way they are combined. So we first consider the case in which we combine two observables A_1 and A_2, alternating them with probability η and $(1-\eta)$, and assuming that the outcome space M_1 and M_2 of the two observables have a finite number of outcomes and that the possible outcomes are distinguishable. So let M_1 be $M_1 = \{a_1, \ldots, a_n\}$ and $M_2 = \{a'_1, \ldots, a'_m\}$. We define now a new observable B such that

$$B = \eta A_1(X) + (1-\eta)A_2(X), \tag{2.4.1}$$

with $X \subseteq M_B$, M_B being the outcome space of the observable B. The set M_B is obtained by taking the union of M_1 and M_2, i.e. $M_B = M_1 \cup M_2$. The operators A_1 and A_2 are extended on M_B by writing $A_1(X) = A_1(X \cap M_1)$ and $A_2(X) = A_2(X \cap M_2)$ for $X \subseteq M_B$. In this way we have $A_1(a'_i) = O$, $A_2(a_i) = O$ and so

$$B(a_i) = \eta A_1(a_i), \qquad B(a'_i) = (1-\eta)A_2(a'_i). \tag{2.4.2}$$

We say that the observable B is a mixture of the two observables A_1 and A_2.
However we have also other cases in which we can combine two observables. In particular we can consider the case in which the outcome spaces M_1 and M_2 are not distinguishable. Then, if the common outcome space is $M = M_1 = M_2 = \{a_1, \ldots, a_n\}$, we can consider a new observable B defined as

$$B(a_i) = \eta A_1(a_i) + (1-\eta)A_2(a_i), \tag{2.4.3}$$

for all $a_i \in M$. What is changed is now the context in which we consider this combination. In particular the observable B describes the situation in which we perform a measurement of two observables A_1 and A_2 on a system and we can not determine whether the outcome registered pertained to the observable A_1 or to the observable A_2. In this case the parameter η, such that $0 \leq \eta \leq 1$, is the probability that the outcome a_i is due to a measurement of the observable A_1, as well as $(1-\eta)$ is the probability that it pertains to the observable A_2.

Extremal POVMs

In the case in which the observables mixed have the same outcome space, they constitute a subset, of all observables, closed under convex combinations, i.e. a convex subset. The convex structure of the space of POVMs is linked to the fact that a single measurement can be described by different measuring apparatuses. We can ask if there exist measurements that do not correspond to any random choice of measurement apparatuses, i.e. if there exist POVMs that can not be decomposed as a convex combination of other POVMs. So, similar to the case of states and effects, we can consider pure observables [4, 13, 23], i.e. observables that can not be written as a nontrivial convex combination of other distinct observables. However, we will see that differently from the case of states, extremal POVMs are not necessarily described by rank-one POVMs since there are higher rank POVMs that are indecomposable in the same way that there are rank-one POVMs that are not extremal. Nevertheless, in the case in which the POVM is described by a set of projections for every outcome in the outcome set, we have an extremal POVM, as is shown in the following

Proposition 2.4.1. *If the observable B with outcome space $(M, \mathcal{M})$ is a projection for every $x \in \mathcal{M}$, then it is a pure observable.*

Proof. suppose that $B(x)$ is a projection for every $x \in \mathcal{M}$, but that B is not pure. This means that there must exist two distinct observables A_1 and A_2 such that

$$B = \eta A_1 + (1 - \eta)A_2,$$

with $0 < \eta < 1$. So it must be true that there is at least one element $x \in \mathcal{M}$ such that $A_1(x) \neq A_2(x)$. For this x we have $B(x) = \eta A_1(x) + (1 - \eta)A_2(x)$. So we have obtained a nontrivial convex decomposition of $B(x)$, that by hypothesis is a projection. However this is not possible since projections are extreme elements in the set of effects. So it must be true that B is a pure observable. □

Nevertheless, the structure of the convex space of POVMs is much richer than that of the effects and states. Therefore in what follows we will try to describe some characteristics of the extremal elements of this convex space, using the method of perturbations [13], and subsequently, we will describe some aspects of its geometry [23]. In order to fix the ideas, we will consider discrete POVMs, i.e. POVMs that have a discrete outcome space. We denote the convex set of POVMs as $\mathscr{P}_N$, while $\Omega_N = \{a_1 \ldots, a_N\}$ will be the outcome space of the POVMs in $\mathscr{P}_N$.

Remark 2.4.1. In $\mathscr{P}_N$ there are also POVMs with outcome space $\Omega_M = \{a_1, \ldots, a_M\}$ with $M \leq N$. In this case, we assume that the effects associated to outcomes in

$\Omega_N \setminus \Omega_M$ are the null operator. In this way, we observe that $\mathscr{P}_M \subseteq \mathscr{P}_N \subseteq \mathscr{P}$ for $M \leq N$ and where $\mathscr{P}$ is the convex set of all POVMs with any (possibly infinite) discrete set of outcomes. Moreover, the extremal elements in $\mathscr{P}_M$, will be also extremal in $\mathscr{P}_N$, and in general in $\mathscr{P}$. In what follows we will consider POVMs with a finite discrete outcome space, such that every POVM in $\mathscr{P}_N$ is described by the set of effects $\{E_1, \ldots, E_N\}$, with $E_1, \ldots, E_N \in \mathcal{E}(\mathcal{H})$. In general, we can think a POVM in $\mathscr{P}_N$ as a vector of operators that we will indicate with $\mathbf{P}$.

At this point, in order to characterize extremal elements in $\mathscr{P}_N$, we introduce the method of perturbations [13]. According to this method an element $\mathbf{P}$ in $\mathscr{P}_N$ is not extremal if there exists a vector $\mathbf{D} = \{D_1, \ldots, D_N\}$, called a perturbation, consisting of N Hermitian operators, such that $\sum_{i=1}^{N} D_i = O$, and $\mathbf{P} \pm \epsilon \mathbf{D}$ is a POVM for some $\epsilon > 0$. This condition determines a necessary condition for a POVM to be extremal, i.e. a POVM is extremal if does not exist a non banal perturbation $\mathbf{D}$ for $\mathbf{P}$.

Example 2.4.1. Let us consider a two outcomes POVM $\mathbf{P} = (\frac{1}{2}|0\rangle\langle 0|, \frac{1}{2}|0\rangle\langle 0| + |1\rangle\langle 1|)$. Let us define the vector $\mathbf{D} = \frac{1}{2}(|0\rangle\langle 0|, -|0\rangle\langle 0|)$. Then the vectors $\mathbf{P}_\pm = \mathbf{P} \pm \mathbf{D}$ are two POVMs and we can write $\mathbf{P}$ as $\mathbf{P} = \frac{1}{2}\mathbf{P}_+ + \frac{1}{2}\mathbf{P}_-$, where $\mathbf{P}_+ = (|0\rangle\langle 0|, |1\rangle\langle 1|)$ and $\mathbf{P}_- = (O, I)$. So we observe that $\mathbf{P}$ is not an extremal element in $\mathscr{P}_2$ because it can be expressed as a combination of $\mathbf{P}_+$ and $\mathbf{P}_-$. ■

Remark 2.4.2. The condition that $\mathbf{P} \pm \epsilon \mathbf{D}$ is a POVM, and thus that $\mathbf{D}$ is a perturbation for $\mathbf{P} = \{E_1, \ldots, E_N\}$, implies that

$$E_i \pm \epsilon D_i \geq O \qquad\qquad \forall\, i = i, \ldots, N.$$

This condition can be equivalently expressed as $\epsilon |D_i| \leq E_i$, or in the same way, $supp(D_i) \subseteq supp(E_i)$, where *supp* denotes the support of the operator.

We can summarize all that we have seen up to this point in the following [13]

Theorem 2.4.1. *If a POVM $\mathbf{P} \in \mathscr{P}_N$ is extremal, then does not exist a non banal vector $\mathbf{D}$ of Hermitian operators satisfying the following conditions:*

$$\sum_{i=1}^{N} D_i = O \qquad supp(D_i) \subseteq supp(E_i) \qquad\qquad \forall i = 1, \ldots, N. \tag{2.4.4}$$

Remark 2.4.3. If we consider the eigenvectors $\{|e_n{}^{(i)}\rangle\}$ associated to E_i, such that $span(|e_n{}^{(i)}\rangle) = supp(E_i)$, then the condition in eq. (2.4.4) becomes

$$\sum_{i=1}^{N} \sum_{p,q=1}^{rank(E_i)} D^i_{pq} |e_p{}^{(i)}\rangle\langle e_q{}^{(i)}| = 0 \iff D^i_{pq} = 0 \qquad \forall i = 1, \ldots, N$$

$$and \;\; \forall p, q = 1, \ldots, rank(E_i).$$

Equivalently this condition means that the operators $|e_p{}^{(i)}\rangle\langle e_q{}^{(i)}|$ are linearly independent $\forall\, i = 1 \dots N$ and $\forall\, p, q = 1, \dots, rank(E_i)$.

From theorem (2.4.1) some useful corollaries follow [13]:

Corollary 2.4.1. *If $\sum_{i=1}^{N} dim[supp(E_i)]^2 > d^2$ for $d = dim\mathcal{H}$, then the POVM $\mathbf{P} = \{E_1, \dots, E_N\}$, is not extremal.*

So in this way, a POVM with more than d^2 outcomes is not extremal and can be written as a convex combination of POVMs with a number of outcomes less than d^2.

Corollary 2.4.2. *An extremal POVM with d^2 outcomes, with $d = dim\mathcal{H}$, is necessarily rank-one. Moreover, a rank-one POVM is extremal if and only if its effects E_i are linearly independent.*

Remark 2.4.4. For a POVM with a rank higher than 1 this condition became only necessary, i.e. if a POVM is extremal then the effects E_i are necessarily linearly independent.

Example 2.4.2. In this example [13], we want to find the extremal POVMs for a qubit system, i.e. for a two level system. From corollary (2.4.1), we know that an extremal POVM can have at most 4 outcomes. We observe that in addition to the identity, the other extremal POVMs must be described by one dimensional projectors. So the extremal elements $\mathbf{P} = \{E_i\}_{i=1}^{N}$ in $\mathscr{P}_N$ can be written, in the Bloch form, as

$$E_i = \lambda_i(I + \vec{n}_i \cdot \vec{\sigma}) \qquad for\ \ i = 1, \dots, N. \tag{2.4.5}$$

From the positivity and the normalization conditions for the POVM $\mathbf{P}$, it follows that

$$\lambda_i > 0 \qquad \sum_{i=1}^{N} \lambda_i = 1 \qquad \sum_{i=1}^{N} \lambda_i \vec{n}_i = O. \tag{2.4.6}$$

Now we will analyze the cases in which the outcome set is compound by two, three and four elements. In the case of an outcome set consisting of only two elements $0, 1$, the extremal POVM is $\mathbf{P} = \{|0\rangle\langle 0|, |1\rangle\langle 1|\}$, where $|0\rangle$ and $|1\rangle$ are the eigenvectors of $\vec{n}_i \cdot \vec{\sigma}$ associated with the eigenvalues 1 and -1 respectively. That this is the case, it follows from the following observations: from the condition $\sum_{i=1}^{2} \lambda_i = 1$ it follows that $\lambda_1 = \lambda_2 = \frac{1}{2}$, while from $\sum_{i=1}^{2} \lambda_i \vec{n}_i = 0$ we conclude that $\vec{n}_1 = -\vec{n}_2 \equiv \vec{n}$.
In the case of a POVM with 3 outcomes a necessary and sufficient condition for a POVM to be extremal is represented by

$$\sum_{i=1}^{3} \gamma_i \lambda_i (1 + \vec{n}_i \cdot \vec{\sigma}) = O \qquad \Longleftrightarrow \qquad \gamma_i = 0\ \forall i = 1, 2, 3, \tag{2.4.7}$$

or equivalently

$$\sum_{i=1}^{3}\gamma_i\lambda_i = 0, \qquad \sum_{i=1}^{3}\gamma_i\lambda_i\vec{n}_i = O \qquad \Longleftrightarrow \qquad \gamma_i = 0 \;\; for\; i = 1,2,3. \tag{2.4.8}$$

The third condition in eq. (2.4.6) implies that the $\lambda_i\vec{n}_i$ represent the edges of a triangle. Then the condition $\sum_{i=1}^{3}\gamma_i\lambda_i\vec{n}_i = O$ implies that $\gamma_i \equiv \gamma$ is independent of i. Finally from the condition $\sum_{i=1}^{3}\gamma_i\lambda_i$ it follows that $\gamma \equiv 0$. So we conclude that all three outcomes rank-one POVMs with pairwise non-proportional effects are extremal.
In the case of a POVM with an outcome space compound by four elements, it can be shown [13] that for a POVM written in Bloch form to be extremal, it must be true that the components $\{\vec{n_i}\}_{i=1}^{4}$ of the Bloch vector must not lie on a common plane. ■

The boundary of the convex set of POVMs

In this section, we want to analyze the boundary of $\mathscr{P}_N$ [13]. Intuitively we can make the following considerations: let us consider a polyhedron, and a point lying on same face of it. Then there exists a direction (for example the direction normal to the face) such that any shift of the point along that direction will bring it inside the convex set, while in the opposite direction it will bring the point outside the convex set. In mathematical terms we can express this in the following way: let S be a convex set and $p \in S$ one of its points. The point p lies in the boundary of S if and only if there exists another point $q \in S$ such that

$$p + \epsilon(q-p) \in S \qquad\qquad p - \epsilon(q-p) \notin S \qquad \forall \epsilon \in [0,1]. \tag{2.4.9}$$

With these ideas in mind we can characterize the POVMs lying in the boundary of $\mathscr{P}_N$ throught the following [13]

Theorem 2.4.2. *A POVM $\mathbf{P} \in \mathscr{P}_N$ is a boundary element of $\mathscr{P}_N$ if and only if there exists at least one effect E_f of $\mathbf{P}$ that has a non-trivial kernel.*

Proof. Let $\mathbf{P}$ and $\mathbf{Q}$ be two POVMs and for all $\epsilon \in [0,1]$ let us suppose that $\mathbf{P}+\epsilon\mathbf{D}$ is a POVM while $\mathbf{P}-\epsilon\mathbf{D}$ is not, where $\mathbf{D} \coloneqq \mathbf{Q}-\mathbf{P}$. Then it must be true that there exists an effect E_f such that $E_f - \epsilon D_f \ngeq O$. Therefore, one can find a $|\psi\rangle \in \mathcal{H}$ such that $\langle\psi|E_f\psi\rangle < \epsilon\langle\psi|D_f\psi\rangle \;\forall\epsilon \in [0,1]$. From the positivity of E_f we have that $\langle\psi|E_f\psi\rangle = 0$ and so $|\psi\rangle \in ker(E_f)$.
On contrary let us consider a POVM $\mathbf{P}$ with an effect E_f that has a non-trivial kernel, and let $|\phi\rangle \in ker(E_f)$. It must also exist an effect $E_m \neq E_f$ such that

$\langle\phi|E_m\phi\rangle > 0$ (such an effect must exist for the normalization condition of $\mathbf{P}$). Now let us define the operators $D_f = \eta|\phi\rangle\langle\phi|$, $D_m = -\eta|\phi\rangle\langle\phi|$ and $D_i = 0\ \forall i \neq e, m$, with η smaller than the minimum eigenvalue of E_m. In this way we can define the vector $\mathbf{D} = \{D_1, \ldots, D_f, \ldots, D_m \ldots, D_N\}$, such that $\mathbf{P} + \epsilon\mathbf{D}$ is POVM for all $\epsilon \in [0, 1]$, while $\mathbf{P} - \epsilon\mathbf{D}$ is not, since $E_f - \epsilon D_f$ is not positive definite as follows from the fact that $\langle\phi|(E_f - \epsilon D_f)\phi\rangle = -\eta\epsilon|\langle\phi|\phi\rangle|^2 < 0$. □

At this point, having described the conditions under which a POVM $\mathbf{P}$ lies in the boundary of $\mathcal{P}_N$, we pass now to discuss some geometrical aspects of it. For such a purpose we will consider $\mathscr{P}_N$ as a subset of an affine space, whose dimension equals the number of independent shifts that a point can do while remaining in the space. Intuitively we can make the following considerations: If we consider a cube and a point inside it, we can shift the point in three independent directions while remaining in the cube. However, if the point lies on a face of the cube, we have only two independent directions along which we can shift the point while remaining in the face. So we observe that if the number of independent shifts of a point is 3 then the point is inside the cube. On contrary if the number of independent directions is 2, then the point lies of a face. We can characterize the boundary of $\mathscr{P}_N$ in a similar way. As we have seen, a perturbation for a POVM $\mathbf{P}$ is a vector $\mathbf{D}$ of Hermitian operators such that

$$\sum_{i=1}^{N} \sum_{m,n=1}^{rank(E_i)} D^i_{nm} |{e_n}^{(i)}\rangle\langle{e_m}^{(i)}| = 0 \tag{2.4.10}$$

has a non-trivial solution, i.e. $|{e_n}^{(i)}\rangle\langle{e_m}^{(i)}|$ are linearly dependent for $i = 1, \ldots, N$ and $1 \leq n, m \leq rank(E_i)$. We now observe that the number of operators $|{e_n}^{(i)}\rangle\langle{e_m}^{(i)}|$ that we can form is

$$\sum_{i=1}^{N} rank(E_i)^2 \equiv Z(\mathbf{P}).$$

Moreover the number of independent operators of the form $|{e_n}^{(i)}\rangle\langle{e_m}^{(i)}|$ is

$$L(\mathbf{P}) \equiv dim[span(|{e_n}^{(i)}\rangle\langle{e_m}^{(i)}|)].$$

So we can conclude that the number of independent perturbations is $S(\mathbf{P}) = Z(\mathbf{P}) - L(\mathbf{P})$. If we consider that the dimension of the affine space in which $\mathscr{P}_N$ is embedded is $d^2(N-1)$[13], we can make the following characterization of the boundary of $\mathscr{P}_N$

Theorem 2.4.3. *a POVM $\mathbf{P}$ is an element of the boundary $\partial\mathscr{P}_N$ of $\mathscr{P}_N$, if and only if $S(\mathbf{P}) < d^2(N-1)$. Moreover $S(\mathbf{P})$ represents the dimension of the face in which $\mathbf{P}$ lies.*

2.5 Sharp observables

At this point, we turn our attention to the concept of PVM, also known as sharp observable. We have seen that a POVM is a mapping that associates with every subset in the outcome space, an effect. In the case in which the effect is represented by a projection, i.e. an extremal effect, we obtain a particular type of POVM that is said sharp observable. So we make the following definition:

Definition 2.5.1. An observable O with outcome space $(M, \mathcal{M})$ is a projection valued measure (PVM), iff $O(X)$ is a projection for every $X \in \mathcal{M}$. Such an observable is known as sharp observable.

Example 2.5.1. In the following example [8] we describe a class of PVMs, with a countable number of outcomes, associated with an orthonormal basis of an Hilbert space $\mathcal{H}$. So let us consider a d-dimensional Hilbert space, where $d < \infty$ or $d = \infty$, and an orthonormal basis $\{\psi_i\}_{i=1}^{d}$ for it. Then we construct the one dimensional projections $P(i) = |\psi_i\rangle\langle\psi_i|$ for $i = 1, \ldots, d$. As we now will see, the collection of these projections defines a PVM. In fact the probability of obtaining measurement outcome i for a state ρ is

$$tr[\rho P(i)] = \langle\psi_i|\rho\psi_i\rangle \tag{2.5.1}$$

and since $\{\psi_i\}_{i=1}^{d}$ is an orthonormal basis it is also true that the normalization condition holds:

$$\sum_{i=1}^{d} P(i) = \sum_{i=1}^{d} |\psi_i\rangle\langle\psi_i| = I. \tag{2.5.2}$$

So the mapping $i \mapsto P(i)$ defines a sharp observable associated with the sample space $M = \{1, \ldots, d\}$ ■

An interesting characteristic of the sharp observables is that their range consists of mutually commuting projections. In order to see this we must prove first the following:

Proposition 2.5.1. *If P is an observable with outcome space $(M, \mathcal{M})$, then the following conditions are equivalent:*

1) P is sharp

2) $P(X)P(Y)=P(X \cap Y)$ for every $X, Y \in \mathcal{M}$.

Proof. 1) $\Rightarrow$ 2)
Suppose that P is a sharp observable, i.e. $P(X)$ is a projection for every $X \in \mathcal{M}$. We now observe that in general, if we consider a POVM A, such that $X \subseteq Y$, for

X and Y two subsets in the sample space of A, then $A(X) \leq A(Y)$. That this is the case it follows from the observation that if $X \subset Y$ then $Y = X \cup (Y \setminus X)$ and so $A(Y) = A(X) + A(Y \setminus X)$, from which $A(X) \leq A(Y)$. So using this inequality, we have $P(X \cap Y) \leq P(X) \leq P(X \cup Y)$. Since all this operators are projections we also have that

$$P(X)P(X \cap Y) = P(X \cap Y) \tag{2.5.3}$$

and

$$P(X)P(X \cup Y) = P(X). \tag{2.5.4}$$

In fact if we consider two projections P_1 and P_2 such that $P_1 \geq P_2$ then for a vector $\phi \in \mathcal{H}$

$$\|P_1 P_2 \phi\| \leq \|P_1\| \|P_2 \phi\| = \|P_2 \phi\|. \tag{2.5.5}$$

But we have also that

$$\|P_1 P_2 \phi\|^2 = \langle P_1 P_2 \phi | P_1 P_2 \phi \rangle = \langle P_2 \phi | P_1 P_2 \phi \rangle \ \geq \ \langle P_2 \phi | P_2 P_2 \phi \rangle = \langle P_2 \phi | P_2 \phi \rangle = \|P_2 \phi\|^2.$$

So we have that $\|P_1 P_2 \phi\| = \|P_2 \phi\|$ from which it follows that $P_1 P_2 \phi = P_2 \phi$, i.e. $P_1 P_2 = P_2$.
Now using eqs. (2.5.3) ,(2.5.4) and the identity

$$P(X \cup Y) + P(X \cap Y) = P(X) + P(Y), \tag{2.5.6}$$

satisfied by every POVM P, multiplying by $P(X)$ from the left, we obtain

$$P(X)P(X \cup Y) + P(X)P(X \cap Y) = P(X)^2 + P(X)P(Y), \tag{2.5.7}$$

from which

$$P(X) + P(X \cap Y) = P(X) + P(X)P(Y). \tag{2.5.8}$$

1) $\Leftarrow$ 2)
Let us suppose that eq.(2.5.8) is satisfied by P for every $X \in \mathcal{M}$. If we consider $Y = \neg X \coloneqq M \setminus X$ we have

$$P(X \cup \neg X) = P(\emptyset) = P(X)P(\neg X) = O. \tag{2.5.9}$$

So we have obtained that

$$P(X)P(\neg X) = P(X)(I - P(X)) = P(X) - P(X)^2 = O. \tag{2.5.10}$$

Hence it follows that $P(X)$ is a projection for every $X \in \mathcal{M}$, i.e. P is a PVM. $\square$

Using this proposition we can now prove that the projections associated with a PVM are mutually commuting.

Proposition 2.5.2. *Let P be a sharp observable with outcome space $(M, \mathcal{M})$. The range of P consists of mutually commuting projections.*

Proof. Since P is a PVM we have that, for every $X \in \mathcal{M}$, $P(X)$ is a projection. Now if we consider $P(X)$ and $P(Y)$ for $X, Y \in \mathcal{M}$, using eq.(2.5.8), it follows that

$$P(X)P(Y) = P(X \cap Y) = P(Y \cap X) = P(Y)P(X). \tag{2.5.11}$$

So we have shown that $P(X)P(Y) = P(Y)P(X)$. Since this is true for every $X, Y \in \mathcal{M}$ we obtain the desired result. □

Remark 2.5.1. It can be shown that in the case of a finite dimensional Hilbert space $\mathcal{H}$, all sharp observables are necessarily discrete.

At this point we make an example of a sharp observable that has uncountably many outcomes.

Example 2.5.2. Let Q be the observable, with outcome space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, defined by

$$Q(X)\phi(x) = \chi_X(x)\phi(x), \tag{2.5.12}$$

where $\phi(x) \in L^2(\mathbb{R})$ and χ_X is the characteristic function of the set X. $Q(X)$ defines a PVM for which the quantity

$$tr[Q(X)\rho] = \langle \sigma | Q(X)\sigma \rangle = \int_X |\sigma(x)|^2 dx, \tag{2.5.13}$$

represents the probability that a particle in the state ρ is localized within X, where we have assumed that ρ is a pure state, i.e. $\rho = |\sigma\rangle\langle\sigma|$. From the definition of $Q(X)$ it also follows that

$$Q(X)Q(Y) = Q(X \cup Y) = Q(Y)Q(X) \tag{2.5.14}$$

and so $Q(X)$ is a sharp observable. It is known as **canonical position observable** ■

2.6 Selfadjoint Operators

We have seen that a sharp observable is an observable that associates a projection with every element in the outcome space. Using the spectral theorem we will now show that we can associate with every PVM a selfadjoint operator (bounded or

unbounded). In this way, we see that the observables usually used in quantum mechanics are just a particular class of the more general concept of POVM.
So let us start from the case of a finite-dimensional Hilbert space. If we consider a selfadjoint operator, we know that it can be diagonalized. This means that for an operator $\mathbf{B} \in \mathcal{L}(\mathcal{H})$, we can write $\mathbf{B}$ as

$$\mathbf{B} = \sum_{i=1}^{n} \eta_i P_{\eta_i},$$

where η_i are the eigenvalues of $\mathbf{B}$ while P_{η_i} denotes the projection on the linear subspace $\mathcal{H}_{\eta_i} \subset \mathcal{H}$ spanned by the eigenvectors associated to the eigenvalue η_i. Now we can define a PVM B in the following way: We consider as outcome space the pair $(M, \mathcal{M})$, where $M = \{\eta_1, \eta_2, \ldots, \eta_n\}$ and $\mathcal{M}$ the associated σ-algebra. The projection valued measure B is defined as $B(\eta_i) = P_{\eta_i}$. So we have associated a PVM with a selfadjoint operator. In the same way it is also possible to associate with every sharp discrete real observable B a selfadjoint operator $\mathbf{B}$ such that

$$\mathbf{B} = \sum_{i=1}^{n} x_i B(x_i),$$

where $\Omega = \{x_1, \ldots, x_n\}$ constitutes the sample space of the observable B.
We now turn our attention to the case of an infinite-dimensional Hilbert space. In this case, the spectral theorem for selfadjoint operators says that for each bounded or unbounded selfadjoint operator $\mathbf{B}$ there exists a unique projection valued measure $\mu_{\mathbf{B}}$ on the Borel space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$$\mathbf{B} = \int_{\mathbb{R}} x \mu_{\mathbf{B}}(dx).$$

In particular for every vector ϕ in the domain on $\mathbf{B}$ we have

$$\langle \phi | \mathbf{B} \phi \rangle = \int_{\mathbb{R}} x \langle \phi | \mu_{\mathbf{B}}(dx) \phi \rangle.$$

So we observe that each PVM B on $\mathbb{R}$, determines a unique selfadjoint operator $\mathbf{B}$ (bounded or unbounded).

Example 2.6.1. In this example we consider the canonical momentum observable P that is connected to the canonical position observable Q through the Fourier transform $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ (here we are considering the unitary extension of the Fourier transform to $L^2(\mathbb{R})$). In particular we can express $P(X)$ as $P(X) = \mathcal{F}^{-1} Q(X) \mathcal{F}$. If we denote the Fourier transform of $\phi \in L^2(\mathbb{R})$ as $\hat{\phi}$, we can write

$$\langle \phi | P(X) \phi \rangle = \int_X |\hat{\phi}|^2 dx. \tag{2.6.1}$$

We can also consider the corresponding unbounded selfadjoint operator $\mathbf{P}$, known as momentum operator, as $\mathbf{P} = \mathcal{F}^{-1}\mathbf{Q}\mathcal{F}$, where $\mathbf{Q}$ is the position operator associated with the position observable [8]. It can be shown that the action of the momentum operator on Schwartz functions ψ is as follows [24]:

$$\mathcal{P}\psi(x) = -i\frac{\partial}{\partial x}\psi(x). \tag{2.6.2}$$

■

Remark 2.6.1. As we have seen, selfadjoint operators give an alternative description for real sharp observables. In particular, this means that we can calculate the mean value and the variance of a sharp observables B using the associated selfadjoint operator $\mathbf{B}$:

$$\langle B\rangle_\rho = tr[\rho\mathbf{B}] = \langle\mathbf{B}\rangle_\rho, \tag{2.6.3}$$

where ρ is the state of the system. In a similar way, taking into account that

$$\langle\phi|\mathbf{B}^2\phi\rangle = \int_{\mathbb{R}} x^2\langle\phi|\mu_{\mathbf{B}}(dx)\phi\rangle, \tag{2.6.4}$$

we can express the variance of the observable B as

$$(\Delta_\rho(B))^2 = \langle\mathbf{B}^2\rangle_\rho - \langle\mathbf{B}\rangle_\rho^2. \tag{2.6.5}$$

As follows by the spectral theorem, there is a one to one correspondence between selfadjoint operators and PVMs. However one can ask if there is a similar correspondence also in the case of POVMs, i.e. if there is a class of operators that is in one to one correspondence with POVMs. The answer to this question arises as a consequence of a well-known theorem due to Neumark, which investigates on the relationships between POVMs and PVMs. For this reason, in what follows, we will state and discuss some of the consequences of the Neumark's dilation theorem.

Neumark's theorem

What the Neumark's theorem [10, 20] shows is that we can always pass from a POVM to a PVM. In particular, Neumark has demonstrated that it is possible to obtain a POVM on a Hilbert space $\mathcal{H}$ by means of a PVM on a Hilbert space $\mathcal{H}'$ containing $\mathcal{H}$ as a subspace.

Theorem 2.6.1. *(Neumark)*
Let M be a topological space, $\mathcal{M}$ be the associated σ-algebra, and let $\mathcal{H}$ be an Hilbert space. If $B : \mathcal{M} \to \mathcal{E}(\mathcal{H})$ is a POVM, then there exist an Hilbert space $\mathcal{H}'$, an isometry $U : \mathcal{H} \to \mathcal{H}'$ and a PVM $P : \mathcal{M} \to \mathcal{P}(\mathcal{H}')$ such that $B(X) = U^*P(X)U$ for every $X \in \mathcal{M}$.

So, as we see, starting from a POVM we can obtain a PVM provided we pass in a suitable Hilbert space containing the Hilbert space $\mathcal{H}$ on which the POVM acts. Usually one refers to the **spectral dilation** for the projection valued measure associated with it [10]. Moreover, there exists a minimal dilation which is unique up to a unitary isomorphism, minimality being defined in the sense that the Hilbert space $\mathcal{H}'$ is the smallest Hilbert space containing $\mathcal{H}$ [10, 20]. So we observe that using the Neumark theorem, we can, in some terms, reduce the problem of a POVM to that of a standard PVM. In fact, through the theorem we have

$$tr_{\mathcal{H}}[\rho B(X)] = tr_{\mathcal{H}'}[U^*\rho U P(X)], \tag{2.6.6}$$

i.e. a measurement of the generalized observable $B(X)$ in the state ρ can be replaced by a measurement of the standard observable $P(X)$ in the state $U^*\rho U$.
However, we must observe that, in general, the PVM P associated with the POVM B will not have a direct physical interpenetration. Nevertheless, it is possible to construct dilations by identifying $\mathcal{H}'$ as a tensor product $\mathcal{H} \otimes \mathcal{H}_0$ i.e. as a tensor product between the Hilbert space $\mathcal{H}$ associated with the physical system, and the Hilbert space $\mathcal{H}_0$ which may represent the Hilbert space of an environment system or a measuring apparatus. In this way the Neumark theorem can be restated in the following way [10]:

Proposition 2.6.1. *For every POVM $B : \mathcal{M} \to \mathcal{E}(\mathcal{H})$ there exists an Hilbert space $\mathcal{H}_0$, a state $\rho_0 \in \mathcal{S}(\mathcal{H}_0)$ and a PVM $P : \mathcal{M} \ni X \mapsto P(X) \in \mathcal{P}(\mathcal{H} \otimes \mathcal{H}_0)$ such that*

$$tr_{\mathcal{H}\otimes\mathcal{H}_0}[\rho \otimes \rho_0 P(X)] = tr_{\mathcal{H}}[\rho B(X)], \tag{2.6.7}$$

for any $\rho \in \mathcal{S}(\mathcal{H})$, and for every $X \in \mathcal{M}$.

This observation gives physical relevance to the Neumark theorem since the Neumark extension of a POVM is realized by means of coupling the object system to some probe system. At the same time it ensures the existence of a measurement for any observable.
According to the spectral theorem the selfadjoint operators are in one to one correspondence with projection valued measures. In a similar way there are certain symmetric operators that determine a unique POVM on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. More precisely, every maximal symmetric operator determines a POVM uniquely [10]. We remember that a symmetric operator A is an operator such that $\langle\phi|A\psi\rangle = \langle A\phi|\psi\rangle$ for all $\phi, \psi \in D(A)$, where $D(A)$ is the domain of the operator A, i.e. an operator is symmetric if A^* is an extension of A. Moreover, a symmetric operator is said to be maximal if it do not possess any symmetric extension within the Hilbert space on which it is defined. In particular, as we know, a symmetric operator is not necessary

selfadjoint. There are cases in physics where the operators used are symmetric but not selfadjoint [10]. In this case, we do not have a spectral resolution of these operators and so there is no probability distribution, whose existence is assured by the spectral theorem. Then one wonders if a symmetric operator admits a selfadjoint extension such that for the extended operator there exists a spectral representation. Now we will observe [10] that Neumark's theorem allows an affirmative answer to this question. In fact, Neumark's theorem guarantees that any maximal symmetric operator A can be extended to a selfadjoint operator A' acting on a Hilbert space $\mathcal{H}'$ containing $\mathcal{H}$. In this way using the spectral decomposition for A' in $\mathcal{H}'$ one obtains an analogue of the spectral theorem for symmetric operators. So for any $\xi \in \mathcal{H}$ and $\zeta \in D(A)$ one has

$$\langle \xi | A\zeta \rangle = \langle \xi | A'\zeta \rangle = \int x d\langle \xi | PE^{A'}(x)\zeta \rangle, \tag{2.6.8}$$

where $E^{A'}$ is the projection valued measure associated with A' while P denotes the projection of $\mathcal{H}'$ onto $\mathcal{H}$. In this way defining $PE^{A'}(X)$ as $B(X)$ for every $X \in \mathcal{B}(\mathcal{R})$, one obtains a POVM in $\mathcal{H}$ such that

$$\langle \xi | A\zeta \rangle = \int_{-\infty}^{+\infty} x d\langle \xi | B(x)\zeta \rangle. \tag{2.6.9}$$

So, what one sees is that starting from a symmetric operator one can extend it in order to obtain a selfadjoint operator or a maximal symmetric operator. In the case one has a maximal symmetric operator the Neumark theorem assures that it can be extended to a larger Hilbert space on which it becomes a selfadjoint operator. Moreover, it can be shown [20] that for every maximal symmetric operator there exists exactly one POVM that fulfils eq.(2.6.9). From this, one arrives at the conclusion that one can uniquely associate a POVM with every maximal symmetric operator.

2.7 Informational Completeness and State Tomography

One of the main purpose of a quantum measurement is the reconstruction of the state of the system under investigation. However if we consider the map

$$\mathcal{S}(\mathcal{H}) \ni \rho \mapsto tr[\rho A(X)] \in [0, 1],$$

we observe that there is no effect $A(X)$ for which it is injective and so invertible. This means that no single measurement result can, without prior knowledge on the

state, determine the state of the system. Nevertheless we can consider a subset $\mathcal{C} \subset \mathcal{E}(\mathcal{H})$ of effects, that can be associated with the range of a single observable or with the union of such ranges of a set of observables, for which the associated measurement outcome probabilities separate the states, i.e. such that for any two states ρ_1,ρ_2, it is true that $\rho_1 = \rho_2$ if and only if $tr[\rho_1 A] = tr[\rho_2 A]$ for all $A \in \mathcal{C}$. One usually says that such a set of observables is informationally complete [4, 8, 25]. If we consider an informationally complete set of observables we can try to reconstruct the state of the system knowing the totality of the measurement outcome probabilities associated with the observables in the set. The existence of such set of observables lies at the basis of the state estimation problem, in particular in the field of quantum tomography. So in the following, we will describe more in deep these concepts, specifying the notion of informationally completeness, giving some sufficient and necessary conditions for an observable to be informationally complete, and finally describing some techniques used in the state reconstruction.

Informational completeness

As we have observed in remark 2.2.3, we can associate with every POVM a probability measure. In particular we said that $\theta_A(\rho)$ represents the probability distribution of the outcomes of A in the state ρ. At this point we give the following [8]

Definition 2.7.1. A collection of observables $\{A, B, \dots\}$ is informationally complete if

$$\begin{cases} \theta_A(\rho_1) = \theta_A(\rho_2) \\ \theta_B(\rho_1) = \theta_B(\rho_2) \\ \qquad \vdots \end{cases} \implies \rho_1 = \rho_2, \tag{2.7.1}$$

$\forall \rho_1, \rho_2 \in \mathcal{S}(\mathcal{H})$.

So we see that with an informationally complete set of observables, for two states ρ_1, ρ_2 to be different, we must wave that at last one observable in the collection gives different probability distributions for ρ_1 and ρ_2. Finally, we observe that also a single observable can be informationally complete. In this case, it must be true that

$$\theta_A(\rho_1) = \theta_A(\rho_2) \qquad \Longrightarrow \qquad \rho_1 = \rho_2,$$

$\forall \rho_1, \rho_2 \in \mathcal{S}(\mathcal{H})$.

Example 2.7.1. In this example [8] we consider the informationally completeness of the qubit observables. As we know, we can represent a qubit state using the Bloch

representation associating in this way a Bloch vector with every state. If we consider a two outcomes sharp observable A, the effects $A(1)$ and $A(0)$ can be written in Bloch form with a unit Bloch vector $\vec{a}$. Now we observe that $tr[\rho A(1)] = \frac{1}{2}(1 + \vec{r}\cdot\vec{a})$ and so $\vec{r}\cdot\vec{a} = 2tr[A(1)\rho] - 1$. In order to find the vector $\vec{r}$, and so the state ρ, we need to know the projection of $\vec{r}$ on three independent directions, i.e. we need to know $\vec{r}\cdot\vec{a}$, $\vec{r}\cdot\vec{b}$ and $\vec{r}\cdot\vec{c}$. So we see that the sharp observables associated with the three vectors $\vec{a}, \vec{b}, \vec{c}$ determine an informationally complete set of observables. ■

Example 2.7.2. In this example we consider the set of operators $\{\mathbf{Q}, \mathbf{P}\}$, where $\mathbf{Q}$ and $\mathbf{P}$ are the position and the momentum operators. As we will see this set is not informationally complete. This problem is usually known as the Pauli problem. The informational incompleteness of the momentum-position pair appears evident in the following case: let as consider a function $\phi_{a,b}(x) = \left(\frac{2a}{\pi}\right)^{\frac{1}{4}} e^{-(a+ib)x^2}$, with $a, b \in \mathbb{R}$ and $a > 0$. Now we can calculate the momentum and the position distributions in this state as

$$|\phi_{a,b}(x)|^2 = \left(\frac{2a}{\pi}\right)^{\frac{1}{2}} e^{-2ax^2}, \qquad |\hat{\phi}_{a,b}(p)|^2 = \left(\frac{a}{2\pi(a^2+b^2)}\right)^{\frac{1}{2}} e^{-\frac{ap^2}{2(a^2+b^2)}}, \tag{2.7.2}$$

where $\hat{\phi}_{a,b}(p)$ denotes the Fourier transform of $\phi_{a,b}(x)$. If we consider now the states $\rho_1 = |\phi_{a,b}\rangle\langle\phi_{a,b}|$ and $\rho_2 = |\phi_{a,-b}\rangle\langle\phi_{a,-b}|$ we observe that $\rho_1 \neq \rho_2$ but from eq.(2.7.2) we obtain that the position and the momentum distributions are the same for the two states, i.e. $Q_{\rho_1} = Q_{\rho_2}$ and $P_{\rho_1} = P_{\rho_2}$. ■

Now we demonstrate [8] a necessary condition for an observable to be informationally complete.

Proposition 2.7.1. *If A is an informationally complete observable, then its outcome set $\mathcal{M}$ has at last d^2 elements, where $d = dim(\mathcal{H}) < \infty$.*

Proof. Let A be an observable and let $\mathcal{M} = \{a_1, \ldots, a_n\}$ be its outcome space, with $n < d^2$. The space $\mathcal{T}_s(\mathcal{H})$ of selfadjoint operators is d^2-dimensional and so there exists an operator $S \neq O$ such that $tr[A(a_i)S] = 0$, $\forall a_i \in \mathcal{M}$. Moreover since $\sum_{a_i\in\mathcal{M}} A(a_i) = I$ we also observe that $tr[S] = 0$. Now we can define the new operator

$$\rho' = \frac{1}{d}\left(I + \frac{S}{\|S\|}\right). \tag{2.7.3}$$

The operator ρ' is positive and $tr[\rho'] = 1$, so we can conclude that ρ' is a state. At this point we observe that $tr[A(a_i)\rho'] = tr[\frac{1}{d}IA(a_i)]$, $\forall a_i \in \mathcal{M}$. So the observable A can not distinguish between ρ' and the state $\frac{1}{d}I$, proving that A is not informationally complete. □

Another useful property of an informationally complete observable A is that the elements in its range span the space of selfadjoint operators, thus allowing the estimation of any ensemble average using the same fixed apparatus. In particular, this is the content of the following proposition [8]

Proposition 2.7.2. *An observable A is informationally complete if and only if every selfadjoint operator on a finite dimensional Hilbert space can be written as a real linear combination of the elements belonging to the range of A.*

Proof. Let A be an observable on an Hilbert space $\mathcal{H}$ and let $\mathcal{A}$ be the collection of all operators obtained by taking all real linear combinations of the elements in $ran(A)$, where ran denotes the range of A. $\mathcal{A}$ is a linear subspace in $\mathcal{L}_s(\mathcal{H})$, and so we can consider the space $\mathcal{A}^\perp$, i.e. the orthogonal complement of $\mathcal{A}$ with respect to the Hilbert-Schmidt inner product:

$$\mathcal{A}^\perp \coloneqq \{S \in \mathcal{L}_s(\mathcal{H}) \mid tr[SA(X)] = 0 \; \forall X\} \tag{2.7.4}$$

Now since $I \in ran(A)$ we observe that $tr[S] = 0 \; \forall S \in \mathcal{A}^\perp$, implying that if $tr[A\rho_1] = tr[A\rho_2]$ i.e. $\theta_A(\rho_1) = \theta_A(\rho_2)$, it must be $\rho_1 - \rho_2 \in \mathcal{A}^\perp$. Now we can have two possibilities: $\mathcal{A}^\perp = \{0\}$ or $\mathcal{A}^\perp \neq \{0\}$. In the first case the condition $\theta_A(\rho_1) = \theta_A(\rho_2)$ implies that $\rho_1 - \rho_2 = 0$ and so $\rho_1 = \rho_2$. Then we observe that A is informationally complete. In the second case, it must exists an nonzero operator $B \in \mathcal{A}^\perp$. Now we can split B as $B = B^+ - B^-$ where $B^\pm$ denotes its positive and negative parts respectively. Since $tr[B] = 0$ we also have $tr[B^+] = tr[B^-] \equiv b$, and so we can define $\rho^+ = \frac{B^+}{b}$ and $\rho^- = \frac{B^-}{b}$, which represent two states such that $\rho^+ - \rho^- = \frac{B}{b} \in \mathcal{A}^\perp$. From this it follows that $\theta_A(\rho^+) = \theta_A(\rho^-)$, but since $\rho^+ \neq \rho^-$ we can conclude that A is not informationally complete. □

2.8 State reconstruction

As we have previously said, if we consider an informationally complete set of POVMs, we can, through measurements of the observables in the set, reconstruct the quantum state. The set of procedures that lead to the reconstruction of the state of a quantum system is usually known as quantum tomography. Even if the problem of state reconstruction traced back to the early days of quantum mechanics, and tomographic procedures were known since 1957, when Fano [26] introduced the concept of **quorum** to indicate the set of observables sufficient for a complete determination of the density matrix, it is only with the pioneering experiments by Raymer's in the domains of quantum optics that quantum tomography began to have experimental relevance. In quantum optics, if fact, using a balanced homodyne detector one can

measure all possible linear combinations of position and momentum of a harmonic oscillator representing a single mode of the electromagnetic field. As was observed by Vogel and Risken the collection of probability distributions achieved by homodyne detection is just the Radon transform of the Wigner function W [27]. In this way by Radon transform inversion one can obtain W, and then from W the matrix elements of the density operator. Moreover using group theory [28], tomographic methods are extended from the harmonic oscillator to an arbitrary quantum system. However, an important observation must be done at this point. In order to have an exact reconstruction of a quantum state, a sequence of repeated measurements must be performed on the system. This induces two problems: first, if we make measurements on a single system we can not reconstruct the quantum state due to the perturbation induced on the system by the measurement process. Second, the true probability of an event is known only in the limit of an infinite number of measurements. The first problem can be solved considering a set of N identical copies of the system, but since N is always finite, we can never fix the second problem. However, what we can try to do is to reconstruct a state ρ of the system that reproduces experimental probabilities as faithfully as possible. This can be obtained through the maximization of the likelihood functional $\mathcal{L}(\rho)$, which quantifies the degree of belief in the hypothesis that for a particular set of measured probabilities, the system was prepared in the quantum state ρ. In the following, we will observe an important application of the theory of state tomography in the state reconstruction problem of multiple qubits systems. First, we see how we can reconstruct a qubit system in the case of an ideal experiment in which one can perform infinitely many measurements on the system. Subsequently, we describe the problem of state reconstruction in the case of real experiments, i.e. with a finite number of measurements, using the Maximum-Likelihood method.

Qubit Tomogaphy

We begin by describing the exact tomography, i.e. the reconstruction of a qubit state under the assumption that we can perform infinitely many exact measurements on the system. In particular, we will first observe the tomographic procedure in the case of a single qubit state and after we will make the generalization in the case of multiple, i.e. n-dimensional, qubit state. So starting from the single qubit case, we have already seen that a qubit state can be written as $\frac{1}{2}(I + \vec{r} \cdot \vec{\sigma})$ where $\vec{r} = (r_x, r_y, r_z)$ is the so called Bloch vector, and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. The components of the Bloch vector can be expressed as $r_i = tr[\rho\sigma_i] \;\; i = 1, 2, 3$. Physically every two level system, such as photons, spin-$\frac{1}{2}$ particles and two level atoms, can be represented in this way. What we can see is that, in the ideal

case of exact measurements performed infinitely many times, we can reconstruct a qubit state. In order to do this we must see what is the physical meaning of the components of the Bloch vector. In particular, we observe that if we denote with $P_{|\psi\rangle}$ the probability to make a measurement in the state $|\psi\rangle$, we have, in the case of a two level system spanned by base vectors $\{|0\rangle, |1\rangle\}$, that

$$\begin{aligned} r_x &= P_{\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)} - P_{\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)} \\ r_y &= P_{\frac{1}{\sqrt{2}}(|0\rangle+i|1\rangle)} - P_{\frac{1}{\sqrt{2}}(|0\rangle-i|1\rangle)} \\ r_z &= P_{|0\rangle} - P_{|1\rangle}. \end{aligned}$$

So we can conclude that if we know these probabilities, we can reconstruct the state of the system.

Remark 2.8.1. In the particular case of a photon we have six possible polarization states in which the photon can be found. They are the vertical polarization $|V\rangle$, the horizontal polarization $|H\rangle$, the diagonal and antidiagonal polarizations $|D\rangle = \frac{|H\rangle+|V\rangle}{\sqrt{2}}$ and $|A\rangle = \frac{|H\rangle-|V\rangle}{\sqrt{2}}$, and the right and left circular polarizations $|R\rangle = \frac{|H\rangle+i|V\rangle}{\sqrt{2}}$ and $|L\rangle = \frac{|H\rangle-i|V\rangle}{\sqrt{2}}$. From this we observe that by making measurement of D/A, R/L and H/V polarizations we can find the components of the Bloch vector and so we can reconstruct the state.

In the case of a single qubit state, the tomographic procedures can be pictorially seen using the Bloch sphere. In fact let us consider the state [27]

$$\rho = \begin{pmatrix} \frac{5}{8} & -\frac{i}{\sqrt{2}} \\ \frac{i}{\sqrt{2}} & \frac{3}{8} \end{pmatrix}.$$

We can write this state in the form

$$\rho = \frac{1}{2}\left(I + \frac{1}{\sqrt{2}}\sigma_2 + \frac{1}{4}\sigma_3\right),$$

from which we deduce that the components of the Bloch vector are

$$r_x = 0 \qquad r_y = \frac{1}{\sqrt{2}} \qquad r_z = \frac{1}{4}.$$

However if we start from the measurement of the following set of projectors

$$P = \left\{|H\rangle\langle H|, \frac{1}{2}(I + |H\rangle\langle V| + |V\rangle\langle H|), \frac{1}{2}[I + i(|V\rangle\langle H| - |H\rangle\langle V|)]\right\}$$

i.e.

$$P = \Big\{ |H\rangle\langle H|, |D\rangle\langle D|, |R\rangle\langle R| \Big\},$$

then we observe that a measurement of the effect $|R\rangle\langle R|$ forces the unknown state in the $z = \frac{1}{\sqrt{2}}$ plane in the Bloch sphere. A measurement of the effect $|D\rangle\langle D|$ further constrains the state to the $y = 0$ plane, while the final measurement of $|H\rangle\langle H|$ pinpoints the state on a line parallel and directly above the x axis. Pictorially we have the following situation [27]

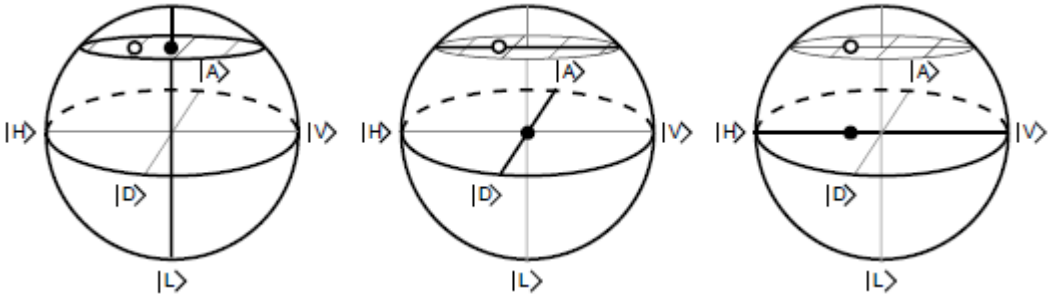

Figure 2.2: A sequence of measurements along the right-circular, diagonal and horizontal axes. The state of the system is represented by the open circle while the black dots correspond to the projection of the unknown state onto the measurement axis. The first measurement isolates the unknown state to a plane while the second and the third to a line and a point respectively.

Remark 2.8.2. The order of the measurements of the three effects just described is irrelevant, i.e. we obtain always the same result: the first measurement selects a plan, the second a line and the third a point.

So from what we have seen we observe that in the case of a single qubit systems we need of three suitable measurements in order to reconstruct the state.

Remark 2.8.3. Experimentally an arbitrary polarization measurement, in the case the qubit considered is associated to a photon, can be realized using a quarter-waveplate, a half waveplate and a polarizing beam splitter. In particular, a waveplate implements unitary operations, and in the Bloch sphere pictures, it acts as a rotation along an axis lying within the linear polarization plane (the equator). The magnitude of this rotation is equal to the waveplate's retardance [27] (90 deg for quarter waveplates and 180 deg for half waveplates.

We now pass to the case of multiple qubit systems. A state can be expressed in a way that is a direct generalization of the one qubit case:

$$\rho = \frac{1}{2^n}\left(I + \sum_{i_1,i_2,\ldots,i_n=1}^{3} r_{i_1,i_2,\ldots,i_n} \sigma_{i_1} \otimes \sigma_{i_2} \otimes \cdots \otimes \sigma_{i_n} \right),$$

where I is the identity on a $4^n - 1$ dimensional Hilbert space and where $r_{i_1,i_2,\ldots,i_n}$ are the components of the Bloch vector associated to the state such that

$$r_{i_1,i_2,\ldots,i_n} = tr[(\sigma_{i_1} \otimes \sigma_{i_2} \otimes \cdots \otimes \sigma_{i_n})\rho].$$

Also in this case we can obtain the components of the Bloch vector through a suitable set of measurements. We observe in fact that in general we can express a component of the Bloch vector as [27]

$$r_{i_1,i_2,\ldots,i_n} = P_{|\psi_1\rangle\otimes|\psi_2\rangle\otimes\cdots\otimes|\psi_n\rangle} \pm P_{|\psi_1\rangle\otimes|\psi_2^\perp\rangle\otimes\cdots\otimes|\psi_n\rangle} \pm \cdots \pm P_{|\psi_1^\perp\rangle\otimes|\psi_2^\perp\rangle\otimes\cdots\otimes|\psi_n^\perp\rangle},$$

where as in the single qubit case, $P_{|\psi_1\rangle\otimes|\psi_2\rangle\otimes\cdots\otimes|\psi_n\rangle}$ is the probability that a measurement on the system will find it in the state $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle$, while the signs on the last line depend on the parity of the number of orthogonal terms $\psi^\perp$, and on the number $i_j = 0, \;\; j = 1, \ldots, n$ of indices equal to zero.
Since we can associate to every component of the Bloch vector an experimental measurement, we are able to reconstruct a multi qubit state, and in the particular case of photon systems we need of $4^n - 1$ measurement, using 2^n detectors, in order to find all the components of the Bloch vector. In particular, if we consider a two qubit state we have the following relationship between experimental measurements and the components of the Bloch vector [27]:

$$\begin{aligned}
r_{1,1} =& P_{|DD\rangle} - P_{|DA\rangle} - P_{|AD\rangle} + P_{|AA\rangle} \\
r_{1,2} =& P_{|DR\rangle} - P_{|DL\rangle} - P_{|AR\rangle} + P_{|AL\rangle} \\
r_{1,3} =& P_{|DH\rangle} - P_{|DV\rangle} - P_{|AH\rangle} + P_{|AV\rangle} \\
r_{2,1} =& P_{|RD\rangle} - P_{|RA\rangle} - P_{|LD\rangle} + P_{|LA\rangle} \\
r_{2,2} =& P_{|RR\rangle} - P_{|RL\rangle} - P_{|LR\rangle} + P_{|LL\rangle} \\
r_{2,3} =& P_{|RH\rangle} - P_{|RV\rangle} - P_{|LH\rangle} + P_{|LV\rangle} \\
r_{3,1} =& P_{|HD\rangle} - P_{|HA\rangle} - P_{|VD\rangle} + P_{|VA\rangle} \\
r_{3,2} =& P_{|HR\rangle} - P_{|HL\rangle} - P_{|VR\rangle} + P_{|VL\rangle}
\end{aligned}$$

$$\begin{aligned}
r_{3,3} =& P_{|HH\rangle} - P_{|HV\rangle} - P_{|VH\rangle} + P_{|VV\rangle} \\
r_{0,1} =& P_{|DD\rangle} - P_{|DA\rangle} + P_{|AD\rangle} - P_{|AA\rangle} \\
r_{0,2} =& P_{|RR\rangle} - P_{|LR\rangle} + P_{|RL\rangle} - P_{|LL\rangle} \\
r_{0,3} =& P_{|HH\rangle} - P_{|HV\rangle} + P_{|VH\rangle} - P_{|VV\rangle} \\
r_{1,0} =& P_{|DD\rangle} + P_{|DA\rangle} - P_{|AD\rangle} - P_{|AA\rangle} \\
r_{2,0} =& P_{|RR\rangle} + P_{|LR\rangle} - P_{|RL\rangle} - P_{|LL\rangle} \\
r_{3,0} =& P_{|HH\rangle} + P_{|HV\rangle} - P_{|VH\rangle} - P_{|VV\rangle},
\end{aligned}$$

where as usual $|AB\rangle = |A\rangle \otimes |B\rangle$, and the vector $|H\rangle$, $|V\rangle$, $|D\rangle$, $|A\rangle$, $|R\rangle$ and $|L\rangle$ are the vectors defined in remark 2.8.1.

Remark 2.8.4. We observe that the components $r_{0,1}$, $r_{0,2}$, $r_{0,3}$, $r_{1,0}$, $r_{2,0}$, $r_{3,0}$, i.e. the components in which one of the two indices is zero, have expressions that, except for a sign, are similar to the respective elements in which the two indices are different from zero. This is not incidental and in fact it is true also in more general multi qubit states. The importance of this observation is that it can be shown [27] that only 3^n measurements are really necessary for the identification of the components of the Bloch vector, thus allowing in this way a drastic reduction of experimental measurements.

Example 2.8.1. Let us suppose that an experimental measurements of photon polarization produce the following probability: $P_{|AA\rangle} = P_{|DD\rangle} = P_{|HH\rangle} = P_{|VV\rangle} = P_{|LR\rangle} = P_{|RL\rangle} = \frac{1}{3}$, $P_{|DA\rangle} = P_{|AD\rangle} = P_{|RR\rangle} = P_{|LL\rangle} = P_{|HV\rangle} = P_{|VH\rangle} = \frac{1}{6}$, while all other possible combinations have probability $\frac{1}{4}$ to be found. In this way we obtain that the only components of the Bloch vector are:

$$r_{1,1} = \frac{1}{3} \qquad r_{2,2} = -\frac{1}{3} \qquad r_{3,3} = \frac{1}{3}.$$

From this it follows that the state of the system is

$$\rho = \frac{1}{4}\left(I + \frac{1}{3}\sigma_1 \otimes \sigma_1 - \frac{1}{3}\sigma_2 \otimes \sigma_2 + \frac{1}{3}\sigma_3 \otimes \sigma_3\right).$$

■

Maximum-Likelihood Methods in Quantum Mechanics

We now pass to the case in which we consider a real experiment. As we have seen in the previous paragraph, the main ingredient in the state reconstruction problem is the knowledge of the probability distribution associated with the experimental outcomes. However, in a real experiment, where only a finite number of measurements are possible, the only data we have access to are the relative frequencies f_i, which sample the principally unknowable probabilities p_i [27]. If we try to reconstruct the quantum state starting from that frequencies, and if the number of data is small, we can obtain operators that could not represent a quantum state since, for example, they may be not positive definite. So we must find other ways in which we can reconstruct a physical state using the experimental data. One way is to use the Maximum Likelihood method ML. The scope of this method is the maximization of a functional $\mathcal{L}(\rho)$, known as the likelihood functional [27], whose function

is to quantify the degree of belief in the hypothesis that for a particular data set the system was prepared in the state ρ. In this way, the likelihood method selects the state for which the likelihood attains its maximum value on the convex set of density matrices. In what follows, we define the functional $\mathcal{L}(\rho)$ in the case the measurement procedure is represented by a PVM. However similar considerations can be done also in the more general case where POVMs are considered. So if we consider an experiment, described by the PVM $P = \{|x_i\rangle\langle x_i|\}_{i=1}^{N}$, then we define [27] $\mathcal{L}(\rho)$ as

$$\mathcal{L}(\rho) \coloneqq \prod_i \langle x_i|\rho|x_i\rangle^{n_i}, \tag{2.8.1}$$

where n_i denotes the rate of registering a particular outcome i. Now in order to find the condition corresponding to the maximization of $\mathcal{L}(\rho)$, we make use of the Jensen inequality [29] between the geometric and arithmetic averages:

$$\prod_i \left(\frac{x_i}{a_i}\right)^{f_i} \leq \sum_i f_i \frac{x_i}{a_i}, \tag{2.8.2}$$

with $x_i \geq 0$, $a_i > 0$ and $f_i = \frac{n_i}{N}$ play the role of frequencies. This inequality may be easily adopted for the maximization of $\mathcal{L}(\rho)$:

$$(\mathcal{L}(\rho))^{\frac{1}{N}} = \prod_i (\langle x_i|\rho|x_i\rangle)^{f_i} = \prod_i a_i^{f_i} \left(\frac{\langle x_i|\rho|x_i\rangle}{a_i}\right)^{f_i} \leq \prod_i a_i^{f_i} \sum_i \frac{f_i}{a_i} \langle x_i|\rho|x_i\rangle =$$
$$\prod_i a_i^{f_i} tr[\rho R(\vec{x},\vec{a})] \leq \lambda(\vec{x},\vec{a}) \prod_i a_i^{f_i}, \tag{2.8.3}$$

where R is the positive semi-definite operator defined as

$$R(\vec{x},\vec{a}) = \sum_i \frac{f_i}{a_i} |x_i\rangle\langle x_i|,$$

$\vec{x} = (x_i, \ldots, x_N)$ are the possible outcomes in the measurement, and $\lambda(\vec{x},\vec{a})$ represents the greatest eigenvalue of the operator R [27]. The role of the parameters $\vec{a} = (a_1, \ldots, a_N)$ will be clarified later. At this point we want to observe, indeed, under what conditions the equality signs will be achieved in the last inequality in equation (2.8.3). In particular, we observe that fixing the parameters a_i, if the extremal state ρ_e that maximizes $\mathcal{L}(\rho)$ have its support in the subspace corresponding to the maximal eigenvalue λ, i.e. if the following equation is satisfied

$$R(\vec{x},\vec{a})\rho_e = \lambda(\vec{x},\vec{a})\rho_e, \tag{2.8.4}$$

then the equality sign appear in eq.(2.8.3). Moreover we note [27] that the equality sign is achieved also in the case in which we set the parameters a_i equal to $a_i = \langle x_i|\rho_e|x_i\rangle$. From these observations we can conclude that the extremal equation for the ML density operator is

$$R\rho = \rho, \tag{2.8.5}$$

where now the operator

$$R = \sum_i \frac{f_i}{\langle x_i|\rho_e|x_i\rangle}|x_i\rangle\langle x_i|$$

is state dependent. The eigenvalue $\lambda(\vec{x}, \vec{a})$, by imposing the normalization condition, must be set equal to 1.

Remark 2.8.5. We can generalize eq.(2.8.5) also in the case we consider a measurement described by a POVM $A = \{E_i\}_{i=1}^N$. In this case eq.(2.8.1) became

$$\mathcal{L}(\rho) \coloneqq \prod_i (tr[\rho E_i])^{n_i}, \tag{2.8.6}$$

while the operator R changes in

$$R = \sum_i \frac{f_i}{tr[\rho E_i]} E_i. \tag{2.8.7}$$

Remark 2.8.6. The extremal condition eq.(2.8.5) can be alternatively rephrased as

$$\sum_i \frac{f_i}{\langle x_i|\rho_e|x_i\rangle}|x_i\rangle\langle x_i| = I_\rho, \tag{2.8.8}$$

where I_ρ represents the identity operator defined on the support of the extremal density operator. In the same way eq.(2.8.7) is equivalent to

$$\sum_i \frac{f_i}{tr[\rho E_i]} E_i = I_\rho \tag{2.8.9}$$

Remark 2.8.7. Another way in which we can obtain the equation (2.8.5) is through the variational method. In particular, we observe that we can think of the likelihood as a sort of statistical distance $D(f_i, p_i)$ between the theoretical probability distribution and the experimental frequencies found. In this way through its minimization, we find the density operator ρ_e that generates probabilities p_i lying as close to the observed frequencies f_i as possible. A statistical distance $D(f_i, p_i)$ that satisfies similar conditions is the Kullback-Leibler divergence [27] defined as

$$D(f_i, p_i) = -\sum_i f_i ln(p_i). \tag{2.8.10}$$

The condition for ρ to be the maximum likely state is stated in terms of a stationary condition for the following functional:

$$G = \sum_i f_i ln(tr[\rho E_i]) - \lambda tr[\rho] = \sum_i f_i ln(tr[A^* A E_i]) - \lambda tr[A^* A], \tag{2.8.11}$$

where the Lagrange multiplier λ must be determined by imposing the normalization condition, and we have placed $\rho = A^* A$ in order to guarantee the positivity condition. In this way varying A to $A + \delta A$, the value of G will change in

$$\delta G = \sum_i \frac{f_i}{tr[\rho E_i]} tr[E_i A^* \delta A] - \lambda tr[A^* \delta A]. \tag{2.8.12}$$

Imposing that $\delta G = 0$ we obtain finally

$$\sum_i \frac{f_i}{tr[E_i \rho]} E_i A^* = \lambda A^*. \tag{2.8.13}$$

If we now multiply by A from the right side we obtain

$$\sum_i \frac{f_i}{tr[E_i \rho]} E_i \rho = \lambda \rho, \tag{2.8.14}$$

i.e.

$$R\rho = \lambda \rho, \tag{2.8.15}$$

that by imposing the normalization condition finally became

$$R\rho = \rho \tag{2.8.16}$$

that has the same form of eq.(2.8.5).

So what we have seen is that if we consider a set of informationally complete POVMs we can always reconstruct the physical state of a quantum system. However if we consider the case of real experiments we must also observe that, in order to reconstruct the quantum state, we must consider some optimization methods, such as the maximum likelihood method, from which one obtains eq.(2.8.5), such that, using numerical methods [27], it is possible to find the state that best reproduces the experimental values.

Quantum Channels and Open Quantum Systems

3

3.1 Introduction

In the previous chapters, we have seen that an experiment can be divided into two parts: the preparation and the measurement. In the preparation part, no input is required and a quantum output is obtained, while in the measurement part, a quantum input is accepted and a classical output is obtained, i.e. the measurement outcomes distribution. Until now, however, we have not described how a physical system changes. What we will describe in this chapter is the concept of quantum channel [8, 11]. A quantum channel is a map that accepts a quantum state in input and produces a quantum state at the output. Practical examples of quantum channels are symmetry transformations and the time evolution of a quantum system at a certain fixed time t. Moreover, also the dynamics of an open quantum system can be described in terms of quantum channels [12, 30]. From a mathematical point of view, a quantum channel can be seen as a completely positive trace preserving linear map on $\mathcal{T}(\mathcal{H})$. The linearity property is required such that every quantum channel can not change the statistical indistinguishability of different convex decompositions of a quantum state, while the trace preserving property means that the process associated with the transmission through a quantum channel is deterministic, i.e. it happens with probability one. Complete positivity assures that the output state is another valid quantum state, even in the case where the input state is a state of a composite system. A very convenient mathematical representation of quantum channels is the so called operator sum form or Kraus form [8, 4], in which the conditions listed above are manifestly satisfied. For this reason, we will state and discuss a theorem, according to which every quantum channel can be decomposed as

a sum of suitable operators known as Kraus operators and conversely, every linear map on $\mathcal{T}(\mathcal{H})$ that can be written in the Kraus form is a quantum channel. We will see a concrete use of the Kraus form of quantum channels in the study of the dynamics of an open quantum system. The mathematical implications, as well as the proof of this theorem, are linked to a very general result due to Stinespring in the field of dilation theory [4]. The Stinespring theorem shows how every completely positive linear map from a C^*-algebra $\mathcal{A}$ to $\mathcal{L}(\mathcal{H})$ can be written as a composition of a bounded map $V : \mathcal{H} \to \mathcal{K}$ and a *unital**-representation $\rho : \mathcal{A} \to \mathcal{L}(\mathcal{K})$, with $\mathcal{K}$ a suitable Hilbert space. Then, the Kraus theorem follows from the Stinespring theorem if one considers for the C^* algebra $\mathcal{A}$ the set of bounded linear operators $\mathcal{L}(\mathcal{H})$. This theorem, therefore, represents a useful tool in the characterization of completely positive maps and, in fact, it is not a case that the complete positivity condition for a quantum channel becomes evident in the Kraus form for it.
We now present the order in which these topics are discussed in the chapter: in section 3.2 we introduce the concept of quantum channel and describe the mathematical aspects of such a map. In particular, we discuss a very useful class of channels, known as unitary channels [8], that will be used later in the context of open quantum systems. Further, we observe that the set of quantum channels constitutes a semigroup under the binary operation of composition of functions, which becomes a group when considering unitary channels only. Section 3.3 is devoted to the discussion of the Stinespring theorem as well as the associated Kraus theorem. In sec. 3.4 we briefly discuss the matrix representation and the χ-matrix representation, which will be essential in the applications of quantum channels to open quantum systems. The last section is left to a relevant application of the concepts developed in the chapter to the theory of open systems. First, we introduce and discuss the general theory of open systems, focusing on the physical hypotheses underlying the Markovian approximation [12]. Then, we show [12] how the dynamics of an open system can be described in terms of quantum channels. In particular, in the Markovian approximation, we will obtain the so called Lindblad equation [31], i.e. the equation of motion for the density operator of an open quantum system.

3.2 Quantum Channels

In order to introduce the concept of quantum channel, we first introduce the slightly more general concept of quantum operation. Then a quantum channel will be a particular type of operation.
Briefly, an operation is the most general transformation that can be performed on a physical system, while a quantum channel describes transformations that map physical states into physical states deterministically, i.e. with probability one. From

this, we observe that particular conditions must be imposed on these maps. We begin by describing the operations.

Quantum Operations

The concept of quantum operation sums up all the possible transformations that can affect a quantum system. For example, the passage of a photon through an optical fibre or a polarizer is a practical example of a quantum operation [8]. Moreover, also the effects of a measurement process can be described in terms of a quantum operation acting on the system [32]. Every operation can be understood either as a part of the preparation or as a part of the measurement process. In the first case, it is described by a suitable map acting on the convex set of states while in the second it is understood as a map acting on bounded linear operators, i.e. as a map on $\mathcal{L}(\mathcal{H})$. For the present, we discuss quantum operations as maps on states, focusing on the mathematical requirements that must be imposed on them.
We start by observing that since an operation represents the most general transformation that can affect a physical system, a first assumption that can be made [8] is that it may destroy some fraction of the systems in the initial ensemble. This means that, we can consider an operation as a map that maps quantum states into subnormalized states, i.e. states for which the predicted probabilities associated with a quantum measurement sum up to a number less than or equal to one. If we denote with $\tilde{\mathcal{S}}(\mathcal{H})$ the set of subnormalized states, i.e. the set of trace class operators such that

$$\tilde{\mathcal{S}}(\mathcal{H}) = \{\rho \in \mathcal{T}(\mathcal{H}) \ : \rho \geq O, \ 0 \leq tr[\rho] \leq 1\},$$

then we can see an operation as a mapping $\mathscr{O} : \mathcal{S}(\mathcal{H}) \to \tilde{\mathcal{S}}(\mathcal{H})$.
Further, we require that quantum operations are convex-linear maps. This condition assures the preservation of convex combinations of states. In other terms, this means that through a quantum operation one can not distinguish between different convex decompositions of the same state ρ. So we require that

$$\mathscr{O}\left(\sum_i \eta_i \rho_i\right) = \sum_i \eta_i \mathscr{O}(\rho_i), \tag{3.2.1}$$

where $\rho_i \in \mathcal{S}(\mathcal{H})$, $0 \leq \eta_i \leq 1$ and $\sum_i \eta_i = 1$. Every convex-linear map that maps states into subnormalized states admits a unique linear extension (see [8] section 4.1.1) to the space of trace class operators $\mathcal{T}(\mathcal{H})$. So we can think, eventually, an operation as a linear mapping on $\mathcal{T}(\mathcal{H})$. Clearly, not all linear maps $\mathscr{O}$ on $\mathcal{T}(\mathcal{H})$ describe an operation and in fact $\mathscr{O}$ has to satisfy the following additional conditions [8]:

(i) $\mathscr{O}(\rho) \geq O$,

(ii) $tr[\mathscr{O}(\rho)] \leq 1$.

In particular, the first condition is necessary for $\mathscr{O}(\rho)$ to represent another valid quantum state, while the second condition means that the output state is an element of $\tilde{\mathcal{S}}(\mathcal{H})$. Further, the quantity $tr[\mathscr{O}(\rho)]$ can be interpreted (see the next chapter section 4.3) as the probability that the state ρ is transmitted through a device represented by the operation $\mathscr{O}$. Condition (ii) then says that an operation represents a probabilistic transformation affecting the system in the state ρ.

Remark 3.2.1. Since every positive trace class operator can be expressed as a scalar multiple of a density operator [8], we observe that the second requirement implies that $tr[\mathscr{O}(A)] \leq tr[A]$, for every positive trace class operator $A \in \mathcal{T}(\mathcal{H})$. In particular, we call a similar map $\mathscr{O}$ trace non increasing [8]. In the same way, the first condition implies that $\mathscr{O}(A) \geq O$ for all positive operators $A \in \mathcal{T}(\mathcal{H})$, and we say that $\mathscr{O}$ is positive.

As we will see, the operations and channels play an important role in the theory of open quantum systems. In this context, one usually considers the action of an operation (or a channel) on states of a composite system. Then, one sees that another condition must be imposed on the mathematical characterization of these maps. In fact, let us consider a composite system $A + B$ and suppose that we take a state of the form $\rho_A \otimes \rho_B$. If we have a quantum operation acting only on the system A, we can express it as $\mathscr{O}_A \otimes I_B$. In fact, when we consider the action of this map on $\rho_A \otimes \rho_B$ we obtain

$$\mathscr{O}_A \otimes I_B(\rho_A \otimes \rho_B) = \mathscr{O}_A(\rho_A) \otimes \rho_B. \tag{3.2.2}$$

However, it must be observed that even if $\mathscr{O}_A(\rho_A)$ is positive, as assured by the fact that $\mathscr{O}_A$ is an operation, in general it is not guaranteed that $\mathscr{O}_A(\rho_A) \otimes \rho_B$ is too. So according to the positivity condition of a quantum operation, i.e. that it maps positive operators to positive operators, we must impose that the map $\mathscr{O}_A \otimes I_B$ is positive for all the possible extensions I_B. When this is the case we say [8] that it is a completely positive map. Thus, we can make the following

Definition 3.2.1. A linear mapping $\mathscr{O}_A : \mathcal{T}(\mathcal{H}_A) \to \mathcal{T}(\mathcal{H}_A)$ is completely positive if the mapping $\mathscr{O}_A \otimes I_B$ on $\mathcal{T}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is positive for all finite dimensional extensions $\mathcal{H}_B$.

The condition of complete positivity is stronger than the condition of positivity since there are maps that are positive but not completely positive. An example of such a map is the partial transposition discussed in the following [8]:

Example 3.2.1. Let us consider the so called partial transposition. This operator is defined in the following way: for $\mathcal{H}$ a finite-dimensional Hilbert space, and $\{\psi_i\}_{i=1}^d$ an orthonormal basis for it, the partial transposition related to the basis $\{\psi_i\}_{i=1}^d$ is the map $\sigma : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})$ such that $\sigma(|\psi_i\rangle\langle\psi_j|) = |\psi_j\rangle\langle\psi_i|$. So if an operator T in $\mathcal{T}(\mathcal{H})$ is written in a matrix form $[T]_{ij}$ on the basis $\{|\psi_i\rangle\langle\psi_j|\}_{i,j}$, we observe that the transposition map acts on T as $\sigma([T]_{ij}) = [T]_{ji}$, i.e. it associates with the matrix $[T]_{ij}$ its transpose. Since the transposed matrix has the same eigenvalues of the original one, we conclude that the transposition is a positive mapping. However, we now show that it is not completely positive. In fact let us define the vector on $\mathcal{H} \otimes \mathcal{H}$

$$|\phi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |\psi_i\rangle \otimes |\psi_i\rangle.$$

Now, we consider the action of $\sigma_A \otimes I_B$ on $|\phi\rangle\langle\phi|$:

$$\sigma_A \otimes I_B(|\phi\rangle\langle\phi|) = \frac{1}{d} \sum_{i,j} \sigma_A \otimes I_B(|\psi_i\rangle\langle\psi_j| \otimes |\psi_i\rangle\langle\psi_j|) = \frac{1}{d} \sum_{i,j} |\psi_j\rangle\langle\psi_i| \otimes |\psi_i\rangle\langle\psi_j|.$$

However, this operator is not positive since

$$\frac{1}{d} \sum_{i,j} |\psi_j\rangle\langle\psi_i| \otimes |\psi_i\rangle\langle\psi_j|(|\psi_1\rangle \otimes |\psi_2\rangle - |\psi_2\rangle \otimes |\psi_1\rangle) = -(|\psi_1\rangle \otimes |\psi_2\rangle - |\psi_2\rangle \otimes |\psi_1\rangle).$$

■

On the complete positivity condition, we will say something more once we have introduced the concept of dual operation. For the moment, from what we have seen so far, we can define quantum operations in the following way:

Definition 3.2.2. A mapping $\mathscr{O}$ on $\mathcal{T}(\mathcal{H})$ is a quantum operation if it is

(i) linear,

(ii) completely positive,

(iii) trace non-increasing.

Now we introduce, as an example, an important class of quantum operations known as simple operations [8].

Example 3.2.2. Let us consider an operator $L \in \mathcal{L}(\mathcal{H})$. We introduce an operator $\mathscr{O}_L : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})$ defined as

$$\mathscr{O}_L(T) = LTL^*. \tag{3.2.3}$$

Then $\mathscr{O}_L$ is linear and positive. In order to conclude that it is a quantum operation we must check the complete positivity and the trace non-increasing conditions. Starting from the complete positivity, we begin by observing that for a vector $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$, it is true that

$$\langle\psi|((\mathscr{O}_L \otimes I_B)(T))\psi\rangle = \langle\psi|(L \otimes I_B)T(L^* \otimes I_B)\psi\rangle = \langle\tilde{\psi}|T\tilde{\psi}\rangle \geq 0, \tag{3.2.4}$$

where $\tilde{\psi} = (L^* \otimes I_B)\psi$, and where T is a positive trace class operator. So we observe that $(\mathscr{O}_L \otimes I_B)T$ must be positive, and since T is positive it must be true that $(\mathscr{O}_L \otimes I_B)$ is positive too. So we arrive at the conclusion that $\mathscr{O}$ is a completely positive map. Further, we consider

$$tr[\mathscr{O}_L(T)] = tr[LTL^*] = tr[LL^*T], \tag{3.2.5}$$

and so $\mathscr{O}_L$ is trace non-increasing if and only if $LL^* \leq I$. This condition is equivalent to $\|L\| \leq 1$ and so we conclude that any bounded operator $L \in \mathcal{L}(\mathcal{H})$ satisfying the condition $\|L\| \leq 1$ defines an operation. ■

Until now, we have seen quantum operations as maps acting on quantum states, i.e. we have described them as a part of the preparation process. However, we can consider operations also as a part of the measurement, i.e. as mappings on effects. The double role played by quantum operations is similar to the double picture used to describe processes in quantum mechanics, i.e. the Schrödinger and the Heisenberg picture. In particular, when we consider operations as maps acting on states we are considering the Schrödinger picture, while when we consider operations as a part of the measurement process we are in the Heisenberg picture [8, 4]. Operations in the Heisenberg picture are usually known as dual operations. In order to introduce such maps, we begin by remembering that the space of linear bounded operators $\mathcal{L}(\mathcal{H})$ is the dual space of $\mathcal{T}(\mathcal{H})$ [33]. So if we consider a linear map Φ on $\mathcal{T}(\mathcal{H})$, we have an associated linear operator Φ^* defined on the dual space $\mathcal{L}(\mathcal{H})$. The connection between these two maps is

$$tr[\Phi(T)E] = tr[T\Phi^*(E)], \tag{3.2.6}$$

for all $T \in \mathcal{T}(\mathcal{H})$ and all $E \in \mathcal{L}(\mathcal{H})$.
Then, we observe that for every quantum operation $\mathscr{O}$, it is possible to introduce the dual operation $\mathscr{O}^*$ that satisfies the following condition:

$$tr[\mathscr{O}(T)E] = tr[T\mathscr{O}^*(E)], \tag{3.2.7}$$

Moreover, since $(\mathscr{O} \otimes I)^* = (\mathscr{O}^* \otimes I)$ we observe that the complete positivity of $\mathscr{O}$ implies the complete positivity of $\mathscr{O}^*$. From this we can conclude that $\mathscr{O}^*$ is a

valid quantum operation acting on $\mathcal{L}(\mathcal{H})$. Further, from the identity $tr[\mathscr{O}(T)] = tr[\mathscr{O}(T)I] = tr[T\mathscr{O}(I)]$, and from the trace non-increasing property of $\mathscr{O}$, it follows that $\mathscr{O}^*$ satisfies the additional condition $\mathscr{O}^*(I) \leq I$. So, in the Heisenberg picture the effects rather then the states are transformed.
From now on, we will indicate with $\mathscr{O}^*$ the dual operation associated with $\mathscr{O}$.

Remark 3.2.2. The dual operations $\mathscr{O}^*$ are continuous linear maps with respect to the σ-weak topology on $\mathcal{L}(\mathcal{H})$. To clarify this point, we first introduce the concept of σ-weak topology on $\mathcal{L}(\mathcal{H})$.

Definition 3.2.3. The ultraweak or σ-weak topology of $\mathcal{L}(\mathcal{H})$ is the topology generated by the family of seminorms such that $A \mapsto |tr[AT]|$, for $A \in \mathcal{L}(\mathcal{H})$ and $T \in \mathcal{T}(\mathcal{H})$.

There is now a proposition [4] that clarifies the connections between continuous linear maps in the σ-weak topology and the existence of the dual of a linear map on $\mathcal{L}(\mathcal{H})$:

Proposition 3.2.1. *For a linear map $\Psi : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ the following conditions are equivalent:*

(i) Ψ is continuous with respect the σ-weak topology of $\mathcal{L}(\mathcal{H})$

(ii) there exists a bounded linear map $\Phi : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})$ such that $tr[A\Phi(T)] = tr[T\Psi(A)]$ for all $A \in \mathcal{L}(\mathcal{H})$, and $T \in \mathcal{T}(\mathcal{H})$. In particular, the map Φ is the dual map associated with Ψ.

If these conditions hold, we say that the map Ψ is normal

From this proposition, we can conclude that $\mathscr{O}^*$, the dual of an operation, is a normal map.

Example 3.2.3. Let $L \in \mathcal{L}(\mathcal{H})$ such that $\|L\| \leq I$, and let us consider the simple operation defined before in example (3.2.2). We want to determine the dual map associated with this operation. Applying the definition we find

$$tr[\mathscr{O}_L(T)E] = tr[LTL^*E] = tr[TL^*EL] = tr[T\mathscr{O}_L^*(E)]. \tag{3.2.8}$$

Thus we observe that $(\mathscr{O}_L)^* = \mathscr{O}_{L^*}$. ■

Quantum Channels

We have described quantum operations as mappings on $\mathcal{T}(\mathcal{H})$ that are linear, completely positive and trace nonincreasing. In particular, the last requirement concerns the possibility that a portion of the system can be lost after the action of a quantum operation. Moreover, we can consider a quantum operation as a probabilistic transformation that can be realized with probability $tr[\mathscr{O}(\rho)] \leq 1$. However, we can also describe transformations that act on $\mathcal{S}(\mathcal{H})$ and map states to states deterministically, i.e. with probability 1. We will call this type of maps quantum channels [8]. Physical examples of quantum channels can be symmetry transformations and the transient interaction of an open system with its environment. In particular, we will see that quantum channels play a central role in the dynamics of open systems. In this paragraph, we concentrate on their mathematical characteristics focusing on the differences with quantum operations previously introduced.
We start by noting that we must require linearity since a quantum channel must preserve the statistical indistinguishability of different convex decompositions of a given state. We have also to require the completely positive condition since otherwise, we can obtain unphysical states when we consider the action of a channel on a state of a composite system. The last condition, i.e. the trace non-preserving property must now be replaced by the condition of trace preservation, i.e. for a channel $\mathscr{C}$ it must be true that $tr[\mathscr{C}(\rho)] = 1$. From these observations, we can make the following

Definition 3.2.4. A mapping $\mathscr{C} : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})$ is a quantum channel if it is

(i) linear,

(ii) completely positive,

(iii) trace preserving, i.e. $tr[\mathscr{C}(\rho)] = 1$.

Remark 3.2.3. Since every trace class operator can be expressed as a linear combination of states, we conclude that the last property of a quantum channel, i.e. $tr[\mathscr{C}(\rho)] = 1$, can be extended to all trace class operators, i.e. $tr[\mathscr{C}(T)] = tr[T]$ for $T \in \mathcal{T}(\mathcal{H})$.

Also in the case of quantum channels, we can consider the Schrödinger and the Heisenberg picture, according to if we consider quantum channels as mappings on states or as mappings on effects respectively. Moreover, the relation between the two pictures is the same as in the case of quantum operations:

$$tr[\mathscr{C}(T)E] = tr[T\mathscr{C}^*(E)], \tag{3.2.9}$$

for $T \in \mathcal{T}(\mathcal{H})$ and $E \in \mathcal{E}(\mathcal{H})$. Further, from the identity

$$tr[\mathscr{C}(T)] = tr[T] = tr[T\mathscr{C}^*(I)],$$

we conclude that the trace preserving property implies that $\mathscr{C}^*$ is a unital map, i.e. that it satisfies $\mathscr{C}^*(I) = I$.

Remark 3.2.4. From now on, we will indicate a quantum channel with $\mathscr{C}$, and we use $\mathscr{C}^*$ for the dual channel associated to it.

Now, we make some [11] examples of quantum channels.

Example 3.2.4. We consider the mapping on $\mathcal{T}(\mathcal{H})$ given by $\mathscr{C}_U(S) \coloneqq USU^*$ for U a unitary operator on $\mathcal{H}$. That this map defines a channel it follows from the fact that it is linear, it is completely positive since for a vector $\psi \in \mathcal{H}_A \otimes \mathcal{H}_B$ we have

$$\langle\psi|(\mathscr{C}_U \otimes I_B)S\psi\rangle = \langle\psi|(U \otimes I_B)S(U^* \otimes I_B)\psi\rangle = \langle\tilde{\psi}|S\tilde{\psi}\rangle \geq 0, \tag{3.2.10}$$

where $\tilde{\psi} = (U^* \otimes I_B)\psi$, and if S is positive we obtain that also $\mathscr{C}_U \otimes I_B$ is positive implying that $\mathscr{C}_U$ is completely positive. Finally we observe that $tr[\mathscr{C}_U(T)] = tr[UTU^*] = tr[UU^*T] = tr[T]$ for $T \in \mathcal{T}(\mathcal{H})$, i.e. $\mathscr{C}_U$ is trace preserving. We can also determine the dual channel associated to $\mathscr{C}_U$, i.e. the channel in the Heisenberg picture. Using the defining equation (3.2.9) we obtain

$$tr[\mathscr{C}_U(S)E] = tr[USU^*E] = tr[SU^*EU] = tr[T\mathscr{C}_U^*(E)], \tag{3.2.11}$$

for $S \in \mathcal{T}(\mathcal{H})$ and $E \in \mathcal{E}(\mathcal{H})$. So we conclude that $(\mathscr{C}_U)^* = \mathscr{C}_{U^*}$. Every unitary transformation can be associated to a quantum channel that is usually known as unitary quantum channel . Moreover, it is also possible to see [8] that antiunitary transformations do not define channels. ■

Remark 3.2.5. If we consider simple operations, i.e. the operations acting on trace class operators as $\mathscr{O}_L(T) = LTL^*$ for $L \in \mathcal{L}(\mathcal{H})$, $\|L\| \leq 1$, we observe that if L satisfies the additional condition $LL^* = I$ we obtain a unitary quantum channel.

Example 3.2.5. In this example we describe a particular type of channel known as complete state space contraction [8]. It is defined in the following way:

$$\mathscr{E}_F(T) = \frac{tr[T]}{tr[F]}F \tag{3.2.12}$$

with F a fixed positive trace class operator. This map is linear and it can be shown that it is completely positive. Moreover $tr[\mathscr{E}_F(T)] = \frac{tr[T]}{tr[F]}tr[F] = tr[T]$, and so it is trace preserving. In particular, if we consider its action on a state ρ we obtain

$$\mathscr{E}_F(\rho) = \frac{1}{tr[F]}F. \tag{3.2.13}$$

So we observe that the whole state space is contracted into a single point represented by the state $\frac{1}{tr[F]}F$. We can also consider how this map appears in the Heisenberg representation:

$$tr[\mathscr{E}_F(S)E] = \frac{tr[S]}{tr[F]}tr[FE] = tr[S\mathscr{E}_F^*(E)], \tag{3.2.14}$$

from which we deduce that $\mathscr{E}_F^*(E) = \frac{tr[FE]}{tr[F]}I$. So we conclude that in the Heisenberg picture all effects are mapped into the one dimensional subspace spanned by the identity operator I. ■

Example 3.2.6. The example that we will now discuss plays a central role in the field of open quantum systems. So let's define the linear mapping $\mathscr{P}_\eta : \mathcal{T}(\mathcal{H}_A) \ni T \mapsto T \otimes \eta \in \mathcal{T}(\mathcal{H}_A \otimes \mathcal{H}_\eta)$, where $\mathcal{H}_\eta$ is the Hilbert space of an ancillary state η. We observe that this map is linear, it is trace preserving since

$$tr[\mathscr{P}_\eta(T)] = tr[T \otimes \eta] = tr[T], \tag{3.2.15}$$

and it is completely positive, as follows by observing that if we consider the action of $\mathscr{P}_\eta \otimes I_B$ on a state $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ we obtain $\rho_{AB} \otimes \eta \geq 0$, that implies the complete positivity of $\mathscr{P}_\eta$. We can also see how this channel appears in the Heisenberg picture using the fact that

$$tr[\mathscr{P}_\eta(T)E] = tr[(T \otimes \eta)E] = tr[(T \otimes I)(I \otimes \eta)E] = tr[Ttr_B[(I \otimes \eta)E] = tr[T\mathscr{P}_\eta^*(E)],$$

from which we conclude that $\mathscr{P}_\eta^*(E) = tr_B[(I \otimes \eta)E]$ for all $T \in \mathcal{T}(\mathcal{H})$ and $E \in \mathcal{E}(\mathcal{H})$. If we consider the action of this channel in the Schrödinger picture, and so as a map on states, we can consider its action as a tool through which one can describe the initial state of an open quantum system, under the hypothesis of initial statistical independence between the system and the reservoir. So, as we will see, this type of channel will be useful in the theory of open quantum systems in which one considers a system in interaction with an environment. We will say more on this in sec.(3.5) ■

Remark 3.2.6. We have seen that $\mathscr{P}_\eta^*$ is defined through the partial trace. It can be shown [8] that in general, if we consider the partial trace tr_B as a mapping $tr_B : \mathcal{T}(\mathcal{H}_A \otimes \mathcal{H}_B) \to \mathcal{T}(\mathcal{H}_A)$, then it is linear, completely positive and trace preserving, i.e. it defines a quantum channel. Further, we can also consider the Heisenberg picture $(tr_B)^*$ of such a channel, and it is easily shown, using eq.(3.2.9), that it acts on bounded linear operators as $(tr_B)^*(A) = A \otimes I$, for $A \in \mathcal{L}(\mathcal{H})$.

Remark 3.2.7. Since channels are functions, one can endow their set with the binary operation of composition. From a physical point of view, the composition

of two channels $\mathscr{C}_1$ and $\mathscr{C}_2$ corresponds to the sequential implementation of the transformations associated with $\mathscr{C}_1$ and $\mathscr{C}_2$. Since the composition of channels is associative, one sees that they constitute a semigroup. The reason for which they do not constitute a group is that not all channels are reversible, i.e. admit an inverse $\mathscr{C}^{-1}$ that is a channel and such that

$$\mathscr{C} \circ \mathscr{C}^{-1} = \mathscr{C}^{-1} \circ \mathscr{C} = \mathscr{I}, \tag{3.2.16}$$

where $\mathscr{I}$ denotes the identity channel, i.e. the channel that associates with each trace class operator itself. It can be shown [8] that a channel admits an inverse that is again a channel only in the case it is a unitary channel (example 3.2.4). Then, the subset of unitary channels constitutes a group, where the inverse of a channel $\mathscr{C}_U$ is $\mathscr{C}_{U^*}$ for U a unitary operator in $\mathcal{U}(\mathcal{H})$. Moreover, it can be shown [8] that the group of unitary channels is isomorphic to the quotient group $\mathcal{U}(\mathcal{H})/\mathbb{T}$, where $\mathbb{T}$ is the set of complex numbers of modulus one.

3.3 Kraus Decomposition Theorem

In this paragraph, we will state and prove the Kraus decomposition theorem. As we will see, this theorem allows to decompose every quantum channel in an operator sum form, i.e. in a suitable composition of bounded linear operators known as Kraus operators. This form for a quantum channel is very useful in applications and further, it provides a way in which the complete positivity property is evident. However, to prove this theorem, we must first discuss an important general result in the dilation theory known as the Stinespring's dilation theorem.
In order to state the Stinespring theorem, we restate in more precise terms the concept of complete positivity for a linear map, introduced in the last paragraph. In our discussion, we in particular follow [4].
Let $\mathcal{C}$ denote an arbitrary *-algebra. Let $M_n(\mathcal{C})$ be the linear space of $n \times n$ matrices with entries in $\mathcal{C}$, for $n \in \mathbb{N}$. $M_n(\mathcal{C})$ is a *-algebra if it is equipped with the natural product $(c_{ij})(a_{ij}) = \sum_{k=1}^{n} c_{ik}a_{kj}$ for $(c_{ij}), (a_{ij}) \in M_n(\mathcal{C})$, and the involution $(c_{ij})^* = (c_{ij}^*)$. We can now make the following

Definition 3.3.1. A linear map $\mathfrak{L} : \mathcal{C} \to \mathcal{L}(\mathcal{H})$ is said to be n-positive, for $n \in \mathbb{N}$, if the linear map $\mathfrak{L}_n : M_n(\mathcal{C}) \to M_n(\mathcal{L}(\mathcal{H}))$ defined by

$$\mathfrak{L}_n((c_{ij})) = (\mathfrak{L}_n(c_{ij})) \tag{3.3.1}$$

satisfies the condition $\mathfrak{L}_n(A^*A) \geq 0$ for all $A = (a_{ij}) \in M_n(\mathcal{C})$. If $\mathfrak{L}$ is n-positive for all $n \in \mathbb{N}$ we say that it is completely positive.

Proposition 3.3.1. *For a linear map $\mathfrak{L} : \mathcal{C} \to \mathcal{L}(\mathcal{H})$ and for any $n \in \mathbb{N}$ the following conditions are equivalent:*

(i) $\mathfrak{L}$ *is n-positive*

(ii) $\sum_{i=1}^{n}\sum_{j=1}^{n}\langle\zeta_i|\mathfrak{L}(a_i^* a_j)\zeta_j\rangle \geq 0$ *for all* $a_1, \ldots, a_n \in \mathcal{C}$ *and* $\zeta_1, \ldots, \zeta_n \in \mathcal{H}$.

So, we observe that in order to define the concept of complete positivity one must first define the weaker concept of n-positivity. The condition in (ii) in proposition (3.3.1) is useful in some contexts, as we will soon see. Further, we also observe that if we consider as the *-algebra the set of bounded linear operators $\mathcal{L}(\mathcal{H})$ we recover the condition of complete positivity for quantum channels in the Heisenberg picture. This follows from the fact that the *-algebra $M_n(\mathcal{A})$ can be identified with $M_n(\mathbb{C}) \otimes \mathcal{A}$. In this way, a linear map $\mathfrak{L}_n : M_n(\mathcal{A}) \to M_n(\mathcal{L}(\mathcal{H}))$ can be seen as the map $I_{M_n(\mathbb{C})} \otimes \mathfrak{L} : M_n(\mathbb{C}) \otimes \mathcal{A} \to M_n(\mathbb{C}) \otimes \mathcal{L}(\mathcal{H})$, where $I_{M_n(\mathbb{C})}$ denotes the identity matrix on $M_n(\mathbb{C})$. Then we see that the n-positivity condition translates into the positivity of $\mathfrak{L} \otimes I_n$. Thus, if we consider as the linear map $\mathfrak{L}$ the maps $\mathscr{C}^*$ or $\mathscr{O}^*$, i.e. the dual of a channel or an operation, we observe that the definition (3.3.1) is in accordance with the definition of complete positivity that we have used to characterize quantum channels and quantum operations in the first section.

In the Stinespring theorem, we will make use of the concept of unital *-representation of a *-algebra. For this reason, we introduce here this notion:

Definition 3.3.2. A unital *-homomorphism is a linear map $\pi : \mathcal{A} \to \mathcal{B}$, for $\mathcal{A}$ and $\mathcal{B}$ two *-algebras with unity, such that the following conditions hold:

(i) $\pi(I_{\mathcal{A}}) = I_{\mathcal{B}}$,

(ii) $\pi(ab) = \pi(a)\pi(b) \;\; \forall a, b \in \mathcal{A}$,

(iii) $\pi(a^*) = (\pi(a))^* \;\; \forall a \in \mathcal{A}$.

A unital*-homomorphism $\pi : \mathcal{A} \to \mathcal{L}(\mathcal{H})$, is called a unital*-representation of $\mathcal{A}$ in $\mathcal{H}$.

Remark 3.3.1. Every unital*-homomorphism π is a completely positive map (see [8] section 4.2.2).

Further, there is a very interesting result concerning completely positive linear maps that is closely related to the Stinesprng theorem, and that we present in the form of a proposition:

Proposition 3.3.2. *Let $\mathcal{C}$ be a *-algebra and $\mathfrak{L} : \mathcal{C} \to \mathcal{L}(\mathcal{H})$ be a linear map. Let $P : \mathcal{H} \to \mathcal{K}$ be a bounded linear map between the Hilbert space $\mathcal{H}$ and a suitable Hilbert space $\mathcal{K}$. Consider then a *-representation $\pi : \mathcal{C} \to \mathcal{L}(\mathcal{K})$ such that*

$$\mathfrak{L}(x) = P^*\pi(x)P \tag{3.3.2}$$

for all $x \in \mathcal{C}$. Then $\mathfrak{L}$ is completely positive.

Remark 3.3.2. The operator P^* in the proposition indicates the adjoint operator of P, i.e. the operator $P^* : \mathcal{K} \to \mathcal{H}$

Proof. For any $x_1, \dots, x_n \in \mathcal{C}$, $\zeta_1, \dots, \zeta_n \in \mathcal{H}$ and using (ii) in prop.(3.3.1), the following relations hold

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\langle\zeta_i|\mathfrak{L}(x_i^*x_j)\zeta_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\langle\zeta_i|P^*\pi(x_i^*x_j)P\zeta_j\rangle =$$
$$\sum_{i=1}^{n}\sum_{j=1}^{n}\langle\zeta_i|P^*\pi(x_i)^*\pi(x_j)P\zeta_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\langle\pi(x_i)P\zeta_i|\pi(x_j)P\zeta_j\rangle =$$
$$\left\|\sum_{i=1}^{n}P\pi(x_i)\zeta_i\right\|^2 \geq 0.$$

□

One can ask if also the converse of the above proposition is true, i.e. if any completely positive map can be decomposed as in eq.(3.3.2). The Stinespring theorem shows that, under certain conditions, this is in fact the case.

Theorem 3.3.1. *(Stinespring)*
Let $\mathcal{A}$ be a C^*-algebra with identity and $\mathfrak{L} : \mathcal{A} \to \mathcal{L}(\mathcal{H})$ a completely positive linear map. Then there exist a Hilbert space $\mathcal{K}$, a bounded linear map $P : \mathcal{H} \to \mathcal{K}$ and a unital *-representation $\pi : \mathcal{A} \to \mathcal{L}(\mathcal{K})$ such that

$$\mathfrak{L}(x) = P^*\pi(x)P, \tag{3.3.3}$$

for all $x \in \mathcal{A}$.

So, we observe that under the hypothesis that the *-algebra $\mathcal{A}$ is a C^*-algebra, the Stinespring theorem assures that any completely positive map can be put in the form of eq.(3.3.3). The triple $(\mathcal{K}, \pi, P)$ is said to be the Stinespring representation for $\mathfrak{L}$. It is said to be minimal if the linear combinations of the vectors $\pi(x)P\zeta$ for $x \in \mathcal{A}$ and $\zeta \in \mathcal{H}$ are dense in $\mathcal{K}$.

The proof of the Stinespring theorem can be derived from a more general version of this theorem in the bilinear dilation theory. We refer to the literature [4] for further details on these interesting topics. Here, we concentrate on the physical implications of this theorem. In particular, in order to apply the results of the Stinespring theorem to the theory of open quantum systems, we make the choice to consider as the C^* algebra $\mathcal{A}$ the C^*-algebra $\mathcal{L}(\mathcal{H})$. In this way, we can restate the Stinespring theorem in the following form [8]:

Theorem 3.3.2. *Let $\mathfrak{B} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ be a completely positive map. Then there exist an Hilbert space $\mathcal{K}$, a unital*-representation $\pi : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{K})$ and a bounded operator $P : \mathcal{H} \to \mathcal{K}$ such that*

$$\mathfrak{B}^*(A) = P^*\pi(A)P, \tag{3.3.4}$$

for every $A \in \mathcal{L}(\mathcal{H})$. If $\mathfrak{B}$ is unital, then $P^*P = I$.

An immediate consequence of this theorem is that it shows how every quantum channel can be expressed as the composition of the partial trace tr_B, the unitary channel, and the $\mathcal{P}_\eta$ channel discussed in example (3.2.6). In order to see this, we begin by observing that since $(tr_B)^*$ is a unital*-homomorphism [8], the Stinespring's theorem implies that every channel expressed in the Heisenberg picture is unitarily equivalent to $(tr_B)^*$. So, if $\mathscr{C}$ is a channel, using (3.3.4) we can express it in the Heisenberg form as

$$\mathscr{C}^*(T) = P^*(T \otimes I_B)P, \tag{3.3.5}$$

where $\mathcal{K} = \mathcal{H} \otimes \mathcal{H}_B$ and P satisfies the condition $PP^* = I$. In particular, for all vectors $\psi \in \mathcal{H}$ we observe that $P\psi$ is a vector in $\mathcal{H} \otimes \mathcal{H}_B$, and we can express it as $U(\psi \otimes \eta)$ for $\eta \in \mathcal{H}_B$ a fixed vector. U is a bounded operator and from the condition $PP^* = I$ it is easily shown [8] that it can be extended to a unitary operator on $\mathcal{H} \otimes \mathcal{H}_B$. Finally we observe that if $\{\zeta_i\}_i$ is a basis for $\mathcal{H}$ we have

$$\begin{aligned} tr[\mathscr{C}(\rho)A] = tr[\rho\mathscr{C}^*(A)] = tr[\rho P^*(A \otimes I)P] = \\ \sum_i \langle \zeta_i | \rho P^*(A \otimes I)P\zeta_i \rangle = \\ \sum_i \langle P\rho\zeta_i | (A \otimes I)P\zeta_i \rangle = \\ \sum_i \langle U(\rho\zeta_i \otimes \eta) | (A \otimes I)U(\zeta_i \otimes \eta) \rangle = \\ \sum_i \langle \zeta_i \otimes \eta | (\rho \otimes I)U^*(A \otimes I)U(\zeta_i \otimes \eta) \rangle = \\ tr[(\rho \otimes I)U^*(A \otimes I)U(I \otimes |\eta\rangle\langle\eta|)] = \\ tr[Atr_{\mathcal{K}}[U(\rho \otimes |\eta\rangle\langle\eta|)U^*], \end{aligned}$$

for every $\rho \in \mathcal{S}(\mathcal{H})$ and $A \in \mathcal{L}(\mathcal{H})$. So, from what we have observed, we can state the following corollary of the Stinespring's theorem:

Corollary 3.3.1. *If $\mathscr{C} : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})$ is a quantum channel, then there exist an Hilbert space $\mathcal{H}_B$, a pure state $\eta \in \mathcal{S}(\mathcal{H}_B)$ and a unitary operator U acting on $\mathcal{H} \otimes \mathcal{H}_B$ such that*

$$\mathscr{C}(\rho) = tr_B[U(\rho \otimes \eta)U^*]. \tag{3.3.6}$$

So we can conclude that with every channel we can associate a triple $\langle \mathcal{H}_B, \eta, U \rangle$ known as the dilation of the quantum channel, such that we can express it in the form (3.3.6).
Moreover, we also observe [4] that a channel is not associated with a unique dilation, i.e. different dilations can be associated with the same channel. When this is the case one says that the dilations considered are equivalent.

Example 3.3.1. In this example [8] we will find a dilation for the complete state space contraction channel introduced in example (3.2.5). In this case we set $F = I$, obtaining the channel $\mathscr{C}(A) = \frac{1}{2}tr[A]I$. Further, we consider the case in which $dim\mathcal{H} = 2$ and so the operators¡ $A \in \mathcal{T}(\mathcal{H})$ can be written in the usual Bloch form as $A = a_0 I + \vec{a} \cdot \vec{\sigma}$, with $\vec{a}$ the Bloch vector associated with the operator A. The action of the channel $\mathscr{C}$ on A is then $\mathscr{C}(A) = a_0 I$. We now consider a dilation $\langle \mathcal{H}_B, \eta, U \rangle$ in which $U = \sum_i \sigma_i \otimes |\psi_i\rangle\langle\psi_i|$ where ψ_i is an orthonormal basis for the four-dimensional Hilbert space $\mathcal{H}_B$, while we take $\eta = \frac{1}{4}I_B$. Then, we obtain for every operator $A \in \mathcal{T}(\mathcal{H})$

$$tr_B\left[U\left(A \otimes \frac{1}{4}I\right)U^*\right] = tr_B\left[\sum_{j=0}^{3} \sigma_j \otimes |\psi_j\rangle\langle\psi_j|\left(\sum_{i=0}^{3} a_i\sigma_i \otimes \frac{1}{4}I_B\right)\sum_{j=0}^{3} \sigma_j \otimes |\psi_j\rangle\langle\psi_j|\right] =$$

$$\frac{1}{4}tr_B\left[\sum_{i,j=0}^{3} \sigma_j a_i \sigma_i \sigma_j \otimes |\psi_j\rangle\langle\psi_j|\psi_j\rangle\langle\psi_j|\right] = \frac{1}{4}\sum_{i,j=0}^{3} a_i \sigma_j \sigma_i \sigma_j =$$

$$\frac{1}{4}\left(\sum_{i,j=0,i=j}^{3} a_i \sigma_i^2 \sigma_i + \sum_{i\neq j=0}^{3} a_i \sigma_j \sigma_i \sigma_j\right) = \frac{1}{4}(a_0 I + 3a_0 I + \vec{a} \cdot \vec{\sigma} - \vec{a} \cdot \vec{\sigma}) = a_0 I = \frac{1}{2}tr[A]I.$$

So we observe that $\langle \mathcal{H}_B, \eta, U \rangle$ constitutes a dilation for this particular type of channel. ■

Remark 3.3.3. We observe that the Stinespring theorem is in some sense a generalization of the Neumark theorem that we have discussed in the previous chapter. In particular, in the same hypotheses of the Stinespring theorem, we consider a particular C^*-algebra $\mathcal{B}$ that is defined in the following way. Let Ω be a set and $\mathcal{M}$ an algebra of its subsets. Then we consider the closed *-subalgebra $\mathcal{B}$ of the C^*-algebra

of all bounded complex functions $f : \Omega \to \mathbb{C}$ obtained as the closure of the space of the linear combinations of the characteristic functions of the sets in $\mathcal{M}$. We have then the following theorem [4]:

Theorem 3.3.3. *Let $\mathfrak{A} : \mathcal{B} \to \mathcal{L}(\mathcal{H})$ be a positive linear map. Then, there exist an Hilbert space $\mathcal{K}$, a bounded linear map $P : \mathcal{H} \to \mathcal{K}$ and a unital *-representation $\pi : \mathcal{B} \to \mathcal{L}(\mathcal{K})$ such that $\mathfrak{A}(x) = P^*\pi(x)P$ for all $x \in \mathcal{B}$*

The triple $(\mathcal{K}, \pi, P)$ is called the Neumark representation for $\mathfrak{A}$. It is said to be minimal if the linear combinations of the vectors $\pi(x)P\zeta$ for $x \in \mathcal{B}$ and $\zeta \in \mathcal{H}$ are dense in $\mathcal{K}$. The above theorem is linked to the Neumark theorem that we have discussed in chapter 2, through the following

Theorem 3.3.4. *Let $\mathfrak{A} : \mathcal{B} \to \mathcal{L}(\mathcal{H})$ be a positive linear map, with $\mathcal{B}$ defined in the same way as in theorem (3.3.3). Then*

(i) There is a minimal Neumark representation for $\mathfrak{A}$

(ii) If $\mathcal{M}$ is a σ-algebra and $\mathcal{M} \ni X \mapsto \mathfrak{A}(\chi_X)$ is a POVM on $\mathcal{M}$, where χ_X is the characteristic function on X, then in the minimal Naimark representation $(\mathcal{K}, \pi, P)$ for $\mathfrak{A}$ the map $\mathcal{M} \ni X \mapsto \pi(\chi_X)$ from $\mathcal{M}$ into $\mathcal{L}(\mathcal{K})$ is a PVM.

From this, one observes that the Neumark theorem is a special case of the more general Stinespring theorem.

At this point, having discussed some interesting aspects of the Stinespring theorem, we are ready to introduce a very useful consequence of it, i.e. the the Kraus decomposition theorem.

Theorem 3.3.5. *(Kraus)*
A linear mapping $\mathfrak{L} : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H})$ is a channel if and only if there exists a finite or infinite sequence of bounded operators $L_1, L_2, \ldots$ such that

$$\mathfrak{L}(T) = \sum_i L_i T L_i^*, \qquad \sum_i L_i^* L_i = I. \tag{3.3.7}$$

Proof. If $\mathfrak{L}(T) = \sum_i L_i T L_i^*$ and $\sum_i L_i^* L_i = I$ then $\mathfrak{L}$ satisfies the conditions of def.(3.2.4) and so it is a quantum channel. We now prove that the converse is also true. So let $\mathfrak{L}$ be a channel. Using the Stinespring theorem we have that

$$\mathfrak{L}(T) = tr_B[U(T \otimes |\psi_1\rangle\langle\psi_1|)U^*], \tag{3.3.8}$$

where we have considered the dilation $\langle \mathcal{H}_B, U, |\psi_1\rangle\langle\psi_1|\rangle$, for ψ_1 a fixed vector of the vector basis $\{\psi\}_{i=1}^n$ of $\mathcal{H}_B$. Now for every vector $\lambda, \zeta \in \mathcal{H}$ and for every pure state $|\eta\rangle\langle\eta|$ we have that

$$\begin{aligned}\langle \lambda \mathfrak{L}(|\eta\rangle\langle\eta|)\zeta\rangle = \langle\lambda|tr_B[U(|\eta\rangle\langle\eta| \otimes |\psi_1\rangle\langle\psi_1|)U^*]\zeta\rangle = \\ \sum_{i=1}^n \langle \lambda \otimes \psi_i|U(|\eta\rangle\langle\eta| \otimes |\psi_1\rangle\langle\psi_1|)U^*(\zeta \otimes \psi_i)\rangle = \\ \sum_{i=1}^n \langle \lambda \otimes \psi_i|U(\eta \otimes \psi_1)\rangle\langle \eta \otimes \psi_1|U^*(\zeta \otimes \psi_i)\rangle.\end{aligned}$$

Now we introduce a family of operators L_i defined for each i as

$$\langle \lambda|L_i\zeta\rangle = \langle \lambda \otimes \psi_i|U(\zeta \otimes \psi_1)\rangle, \tag{3.3.9}$$

for every vector $\lambda, \zeta, \in \mathcal{H}$. In particular, we observe that these operators are bounded since

$$|\langle\lambda|L_i\zeta\rangle| = |\langle \lambda \otimes \psi_i|U(\zeta \otimes \psi_1\rangle| \leq \|\lambda\|\|U\|\|\zeta\|. \tag{3.3.10}$$

So we have that

$$\sum_{i=1}^n \langle \lambda \otimes \psi_i|U(\eta \otimes \psi_1)\rangle\langle \eta \otimes \psi_1|U^*(\zeta \otimes \psi_i)\rangle = \sum_{i=1}^n \langle\lambda|L_i\eta\rangle\langle L_i\eta|\zeta\rangle, \tag{3.3.11}$$

and so

$$\mathfrak{L}(|\eta\rangle\langle\eta|) = \sum_{i=1}^n L_i|\eta\rangle\langle\eta|L_i^*. \tag{3.3.12}$$

Finally, since every state can be written as a convex combination of pure states we conclude that

$$\mathfrak{L}(\rho) = \sum_{i=1}^n L_i\rho L_i^*, \tag{3.3.13}$$

for every $\rho \in \mathcal{S}(\mathcal{H})$. □

Remark 3.3.4. In the case L_i are bounded operators and fulfil the condition $\sum_i L_i^* L_i < I$, the Kraus theorem is still true but now the eq.(3.3.13) describes an operation rather than a channel.

As an application of the Kraus theorem, we now prove a theorem due to Choi that provides a simple test of whether a linear map $\mathfrak{A} : \mathcal{L}(\mathbb{C}^n) \to \mathcal{L}(\mathbb{C}^{n'})$ is completely positive. Before state the Choi theorem, we remember that for a n-dimensional quantum system the operators can be identified with $n \times n$ complex matrices. Thus $M_n(\mathbb{C}) = \mathcal{L}(\mathbb{C}^n)$.

Theorem 3.3.6. *(Choi)*
Let $\Phi : M_n(\mathbb{C}) \to M_{n'}(\mathbb{C})$ be a positive linear mapping. Then, the following conditions are equivalent:

(i) Φ *is completely positive,*

(ii) $\Phi \otimes I_n$ *is a positive map,*

(iii) the "Choi matrix"

$$\Lambda_\Phi = \begin{pmatrix} \Phi(|\psi_1\rangle\langle\psi_1|) & \dots & \Phi(|\psi_1\rangle\langle\psi_n|) \\ \vdots & \ddots & \vdots \\ \Phi(|\psi_n\rangle\langle\psi_1|) & \dots & \Phi(|\psi_n\rangle\langle\psi_n|) \end{pmatrix} \tag{3.3.14}$$

is positive, where ψ_i are vectors of an orthonormal basis of the n-dimensional Hilbert space $\mathbb{C}^n$

Proof. $(i) \Rightarrow (ii)$: it follows from the definition of complete positivity. $(ii) \Rightarrow (iii)$: Assume that $\Phi \otimes I_n$ is positive. Then let us introduce the positive matrix M in $M_n(\mathbb{C}) \otimes M_n(\mathbb{C})$ defined as

$$M = \sum_{ij} |\psi_i \otimes \psi_i\rangle\langle\psi_j \otimes \psi_j|. \tag{3.3.15}$$

By the positivity of $\Phi \otimes I_n$ it follows that $(\Phi \otimes I_n)(M) = \Phi(|\psi_i\rangle\langle\psi_j|) \otimes |\psi_i\rangle\langle\psi_j|$ is also a positive matrix. But $\Phi(|\psi_i\rangle\langle\psi_j|) \otimes |\psi_i\rangle\langle\psi_j| = \Lambda_\Phi$ and so we conclude that the Choi matrix is positive. $(iii) \Rightarrow (i)$. Let us consider a set of $d \leq nn'$ eigenvectors ϕ_l of Λ_Φ that are linearly independent and form an orthogonal set in $\mathbb{C}^{n'} \otimes \mathbb{C}^n$. If we consider the tensor product as a direct sum $\mathbb{C}^{n'} \otimes \mathbb{C}^n = \mathbb{C}^{n'} \oplus \mathbb{C}^{n'} \oplus \cdots \oplus \mathbb{C}^{n'}$ for n times, we can introduce the projection operators $P_i : \mathbb{C}^{n'} \otimes \mathbb{C}^n \to \mathbb{C}^{n'}_i$, i.e. the projection onto the i th copy of $\mathbb{C}^{n'}$. Then we have

$$\Phi(|\psi_j\rangle\langle\psi_k|) = P_j \Lambda_\Phi P_k = \sum_l P_j |\phi_l\rangle\langle\phi_l| P_k = \sum_l |P_j\phi_l\rangle\langle P_k\phi_l|. \tag{3.3.16}$$

But now, we can define d operators $O_l : \mathbb{C}^n \to \mathbb{C}^{n'}$ such that $O_l\psi_k = P_k\phi_l$. In this way eq.(3.3.16) can be rewritten as

$$\Phi(|\psi_j\rangle\langle\psi_k|) = \sum_l |P_j\phi_l\rangle\langle P_k\phi_l| = \sum_l |O_l\psi_j\rangle\langle O_l\psi_k| = \sum_l O_l|\psi_j\rangle\langle\psi_k|O_l^*. \tag{3.3.17}$$

But from this it follows that for any $A \in \mathcal{L}(\mathbb{C}^n)$ one has that $\Phi(A) = \sum_l O_l A O_l^*$. Since the map Φ can be decomposed in Kraus form, the Kraus theorem assures that it is completely positive. □

3.4 Representations of quantum channels

The following paragraph is devoted to some useful representations of quantum channels. In particular, we will use the tools introduced here in the derivation of the quantum master equation, which is the dynamic equation of an open quantum system, as we will discuss in the next section.
Here, we present [8, 32] two types of representations known as the matrix representation and the χ-matrix representation, under the hypothesis that the Hilbert space of the system is finite dimensional.

Matrix Representation

Quantum channels act as linear maps on the space of trace class operators $\mathcal{T}(\mathcal{H})$ and so if we fix a basis for this space, we can represent channels as matrices. So we begin by fixing an orthogonal basis, with respect to the Hilbert-Schmidt inner product, for $\mathcal{T}(\mathcal{H})$ whose elements are $\{E_0, E_1, \dots, E_{d^2-1}\}$ where d is the dimension of the Hilbert space $\mathcal{H}$. Then, a general operator $A \in \mathcal{T}(\mathcal{H})$ can be represented as

$$A = \sum_i c_i E_i,$$

where the coefficients c_i are

$$c_i = \frac{1}{tr[E_i^* E_i]} tr[E_i^* A]. \tag{3.4.1}$$

We can now consider the action of a channel $\mathscr{C}$ on A obtaining

$$\mathscr{C}(A) = \mathscr{C}\left(\sum_i c_i E_i\right) = \sum_i \frac{1}{tr[E_i^* E_i]} tr[E_i^* \mathscr{C}(A)] E_i = \sum_{i,j} \frac{1}{tr[E_i^* E_i] tr[E_j^* E_j]} tr[E_i^* \mathscr{C}(E_j)] tr[E_j^* A] E_i = \sum_{i,j} \mathscr{C}_{ij} c_j E_i,$$

where we have defined the matrix $\mathscr{C}_{ij}$ as

$$\mathscr{C}_{ij} \coloneqq \frac{1}{tr[E_i^* E_i]} tr[E_i^* \mathscr{C}(E_j)]. \tag{3.4.2}$$

Remark 3.4.1. Clearly, in the matrix representation the composition of channels corresponds to the multiplication of the corresponding matrices.

Example 3.4.1. In this example [8], we consider the action of a channel on a state represented in Bloch form. We remember that the Bloch representation of a state of a d-dimensional quantum system can be obtained by fixing a basis of selfadjoint trace class operators $\{E_0, E_1, \ldots, E_{d^2-1}\}$ such that $E_0 = I$, while due to the orthonormality condition all other elements in the basis have a null trace. In this case, we have the following representation of a quantum state

$$\rho = \frac{1}{d}(I + \vec{r} \cdot \vec{E}), \tag{3.4.3}$$

with $\vec{r}$ the Bloch vector expressed as $r_i = tr[\rho E_i]$. Clearly, the vector $\vec{c}$ that we have previously defined is linked to the Bloch vector through $\vec{c} = \left(\frac{1}{d}, \frac{1}{d}\vec{r}\right)$. The action of a channel on the state ρ has the effect of changing the components of the vector $\vec{c}$. In particular, we observe that

$$c_0 \mapsto c_0' = \mathscr{C}_{00} c_0 = c_0, \tag{3.4.4}$$

where we have used the fact that $\mathscr{C}_{0j} = \frac{1}{d} tr[E_0^* \mathscr{C}(E_j)] = \frac{1}{d} tr[I\mathscr{C}(E_j)] = \frac{1}{d} tr[E_j] = \delta_{0j}$. Similarly the coefficients c_j are transformed in

$$c_j \mapsto c_j' = \sum_{i=1}^{d^2-1} (\mathscr{C}_{ji} c_i + \mathscr{C}_{i0} c_0). \tag{3.4.5}$$

So we can conclude that the effect of the channel $\mathscr{C}$ on $\vec{c}$ is such that:

$$c_0 \mapsto c_0' = \frac{1}{d} \qquad \frac{1}{d}\vec{r} \mapsto \frac{1}{d}(\mathbf{R}\vec{r} + \vec{r}_0), \tag{3.4.6}$$

where $\mathbf{R}$ is a $(d^2-1) \times (d^2-1)$ matrix, while r_0 is a d^2-1 dimensional vector whose components are $\mathscr{C}_{i0}$. ■

Remark 3.4.2. In this representation, there is no simple matrix property through which one can check the complete positivity condition of quantum channels.

The χ-matrix representation

If a channel is decomposed using Kraus operators, then it is simple to check the condition of complete positivity. So one can try to use the Kraus decomposition form of channels, and at the same time, the matrix representation described before. Then, what one obtains is the so called [8] χ-matrix representation.
The strategy is to use a set of Kraus operators and then to decompose them on a suitable operator basis trough which one obtains the matrix representation. So let

$\{K_n\}_n$ be a set o Kraus operators for the channel $\mathscr{C}$ and let $\{E_0, E_1 \ldots, E_{d^2-1}\}$ be a basis for the space of bounded operators $\mathcal{L}(\mathcal{H})$. Then any Kraus operator can be written as

$$K_n = \sum_i v_{ni} E_i, \tag{3.4.7}$$

where $v_{ni} = tr[E_i^* K_n]$. If we now consider a channel $\mathscr{C}$ acting on a state ρ we have

$$\mathscr{C}(\rho) = \sum_n K_n \rho K_n^* = \sum_{i,j} \sum_n v_{ni} v_{nj}^* E_i \rho E_j^* = \sum_{i,j} \chi_{ij} E_i \rho E_j^*, \tag{3.4.8}$$

where we have defined the χ matrix as

$$\chi \coloneqq \sum_n v_{ni} v_{nj}^*. \tag{3.4.9}$$

In this way we have obtained the χ-matrix representation of the channel $\mathscr{C}$.

Remark 3.4.3. the χ-matrix representation of a quantum channel $\mathscr{C}$ acts on a d^2-dimensional vector space. So we observe that for all d^2-dimensional complex vectors $\vec{c} = \{c_1, c_2, \ldots, c_{d^2}\}$ we have

$$\sum_{i,j} c_i^* \chi_{ij} c_j = \sum_n \sum_i c_i^* v_{ni} \sum_j v_{nj}^* c_j = \sum_n b_n b_n^* = \sum_n |b_n|^2 \geq 0, \tag{3.4.10}$$

and so the χ-matrix is a positive matrix. So what we observe is that in the χ-matrix representation the complete positivity condition translates into the positivity condition of the χ-matrix. However, in this representation, the composition of channels doesn't correspond to the matrix multiplication.

Remark 3.4.4. The relation between the matrix representation and the χ-matrix representation is

$$\mathscr{C}_{ij} = tr[E_i^* \mathscr{C}(E_j)] = \sum_{i,j,l,k} \chi_{lk} tr[E_i^* E_l E_j E_k^*]. \tag{3.4.11}$$

3.5 Dynamical Semigroups

This section is devoted to the application of the theory of quantum channels to open quantum systems [12, 30]. The interest in the study of open quantum systems is, in particular, justified by the observation that every realistic quantum system is to some extent open [32, 12]. Moreover, every system that is subjected to a measurement process interacts with a measurement apparatus. Thus it behaves like an open system, showing that open quantum systems are also relevant in measurement theory [8, 4]. We will discuss more on this point in the next chapter. Here we concentrate on the dynamical evolution of an open quantum system that, as we will see, is described by a quantum channel.

Open Quantum Systems

Open quantum systems are, by definition, quantum systems that interact with other quantum systems, usually called environments. In particular, the interaction leads to certain system-environment correlations such that the resulting state change of the open system can no longer, in general, be described in terms of unitary dynamics [12]. As we will see the dynamics of an open system can be described by a set of maps that are channels. The set of such maps, under certain hypotheses, constitutes a semigroup known as the quantum dynamical semigroup [12, 30].
We start by fixing some notations that will be used in the rest of this section. We use the term **reduced system** [12] to indicate the open system, while we refer to the surrounding system coupled with it as the **environment**. When the environment has an infinite number of degrees of freedom, as in the case for example of an electromagnetic field, the term **reservoir** is sometimes preferred [12]. We indicate with $\mathcal{H}_S$ and $\mathcal{H}_E$ the Hilbert spaces of the reduced system and the environment respectively. The term **compound system** indicates the total system composed by the reduced system and the environment. Its Hilbert space will be $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$.
The problem we want to study is the dynamics of the reduced system. The way in which one can approach this problem is to consider first the dynamics of the compound system. In fact, we can assume that the compound system is closed and so its dynamics can be described by unitary operators. However, in order to recover the state of the reduced system one must then use the partial trace over the degrees of freedom of the environment.
So, let $\rho(0)$ be the state of the compound system at the initial time $t = 0$. We assume that initially the environment and the reduced system are statistically independent [12] and so we can choose the initial state ρ in the form $\rho(0) = \rho_S(0) \otimes \rho_E$, where $\rho_S(0)$ is the initial state of the reduced system, while ρ_E is the state of the environment. If the unitary operator performing the evolution of the system is $U(t, 0)$, we obtain that the state of the compound system at time $t > 0$ is $U(t, 0)\rho(0)U(t, 0)^*$. At this point, we recover the final state of the reduced system performing the partial trace over the environment. So, the whole dynamics of the reduced system can be described by a channel $\mathscr{C}_O$ such that:

$$\rho_S(0) \mapsto \mathscr{C}_O(\rho_S(0)) \coloneqq tr_{\mathcal{H}_E}\{U(t, 0)(\rho_S(0) \otimes \rho_E)U(t, 0)^*\}. \tag{3.5.1}$$

That $\mathscr{C}_O$ is a channel it follows from the observation that it is the composition of a partial trace (remark 3.2.6), a unitary channel (example 3.2.4), and a $\mathcal{P}_{\rho_E}$ channel (example 3.2.6), all of which are channels. Moreover, this is also in accordance with the Stinespring theorem. So, we arrived at the conclusion that for any fixed time $t > 0$ and environment state ρ_E, the maps $\mathscr{C}_O(t)$ are channels. In particular, we use

the notation $\mathscr{C}_O(t)$ in order to emphasize that we are considering the channel $\mathscr{C}_O$ at a fixed instant of time t.

Remark 3.5.1. We consider channels for $t > 0$ because the open quantum systems describe, in general, irreversible processes [12].

Thus, the complete dynamics of the open system is known if the channels $\mathscr{C}_O(t)$ are known for every $t > 0$. The problem of finding such maps for every positive instant of time is simplified under the specific physical hypothesis of Markovian-type behaviour of the open system. In intuitive terms, a Markov process [12] is a stochastic process in which one can neglect memory effects, i.e. such that the probability of an event is independent of the whole history of the process. In the case of open quantum systems, this behaviour is justified under the hypothesis that the time scale over which the environment correlation functions decay is much smaller than the characteristic time scale of the reduced system evolution [12]. Then, in the case in which the open system shows a Markovian-type behaviour, it can be shown [30] that the set of channels $\mathscr{C}_O(t)$ constitutes a semigroup. This means that for every $t_1, t_2 > 0$ the following relation is always satisfied:

$$\mathscr{C}_O(t_1 + t_2) = \mathscr{C}_O(t_1)\mathscr{C}_O(t_2). \tag{3.5.2}$$

Moreover, it is possible to find a generator $\mathcal{G}$ for the semigroup such that every channel in it can be expressed as

$$\mathscr{C}_O(t) = exp(\mathcal{G}t). \tag{3.5.3}$$

Since $\rho_S(t) = \mathscr{C}_O(t)\rho_S(0)$, we obtain in this way an equation of motion for the reduced system density matrix

$$\frac{d}{dt}\rho_S(t) = \mathcal{G}\rho_S(t), \tag{3.5.4}$$

which is usually known as the Markovian quantum master equation or the Lindblad equation [12].

Remark 3.5.2. Here we are considering a generator for the dynamical semigroup that is time independent. However, the equation (3.5.4) can be generalized also to the more general case of a time dependent generator.

At this point, our problem is to find the most general form of a time independent generator. In particular, we concentrate on the simple case in which the Hilbert space $\mathcal{H}_S$ is finite dimensional. We start by rewriting the quantum channels $\mathscr{C}_O(t)$ in terms of the Kraus operators. In particular using the decomposition

$\rho_E = \sum_r \lambda_r |\phi_r\rangle\langle\phi_r|$, for $\{\phi_r\}_r$ a vector basis of the Hilbert space $\mathcal{H}_E$, we have

$$\begin{aligned}\mathscr{C}_O(t)\rho_S(0) = tr_{\mathcal{H}_E}\{U(t,0)(\rho_S(0)\otimes\rho_E)U(t,0)^*\} = \\ tr_{\mathcal{H}_E}\{U(t,0)(\rho_S(0)\otimes\sum_r \lambda_r|\phi_r\rangle\langle\phi_r|)U(t,0)^*\} = \\ \sum_{p,r}\langle\phi_p|U(t,0)(\rho_S(0)\otimes\sum_r \lambda_r|\phi_r\rangle\langle\phi_r|)U(t,0)^*|\phi_p\rangle = \\ \sum_{p,r} K_{pr}(t)\rho_S(0)K_{pr}(t)^*,\end{aligned}$$

where we have introduced the bounded operators

$$K_{pr}(t) = \sqrt{\lambda_r}\langle\phi_p|U(t,0)\phi_r\rangle. \tag{3.5.5}$$

These operators constitute a set of Kraus operators for the channel $\mathscr{C}_O(t)$, since the condition

$$\sum_{p,r} K_{pr}(t)K_{pr}(t)^* = I_S \tag{3.5.6}$$

is satisfied. Further, we can consider the χ-matrix representation for the Kraus operators. To do this we introduce a basis of operators $\{E_i\}_i$ $i = 1,\ldots,N^2$ where $N = dim\mathcal{H}_S$. The operators in the basis are orthonormal with respect to the Hilbert-Schmidt inner product, i.e. $tr[E_i^* E_j] = \delta_{ij}$. Moreover, in order to simplify the calculations, we choose [12] the operator E_{N^2} to be proportional to the identity operator, $E_{N^2} = \frac{1}{\sqrt{N}} I_S$. Every Kraus operator can now be expressed as

$$K_{pr}(t) = \sum_i E_i tr[E_i^* K_{pr}(t)]. \tag{3.5.7}$$

Then, via eq.(3.5.7) we obtain the χ-matrix representation of the quantum channel $\mathscr{C}_O(t)$:

$$\mathscr{C}_O(t)\rho_S(0) = \sum_{i,j=1}^{N^2} c_{ij}(t)E_i\rho_S(0)E_j^*, \tag{3.5.8}$$

where

$$c_{ij}(t) = \sum_{p,r} tr[E_i^* K_{pr}(t)]tr[K_{pr}(t)^* E_j]. \tag{3.5.9}$$

At this point, we return to the Lindblad equation (3.5.4) that we can rewrite in this way:

$$\frac{d}{dt}\rho_S(t) = \lim_{\epsilon\to 0}\frac{\rho_S(\epsilon+t)-\rho_S(t)}{\epsilon} =$$

$$\lim_{\epsilon\to 0}\frac{\mathscr{C}_O(\epsilon)\rho_S(t)-\rho_S(t)}{\epsilon}=\lim_{\epsilon\to 0}\frac{1}{\epsilon}\Big[\sum_{i,j=1}^{N^2}c_{ij}(\epsilon)E_i\rho_S(t)E_j^*\ -\rho_S(t)\Big]=$$

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\Big[\sum_{i=1}^{N^2}\Big(\frac{1}{\sqrt{N}}c_{iN^2}(\epsilon)E_i\rho_S(t)+\sum_{j=1}^{N^2-1}c_{ij}(\epsilon)E_i\rho_S(t)E_j^*\Big)-\rho_S(t)\Big]=$$

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\Big[\frac{1}{\sqrt{N}}\sum_{i=1}^{N^2}c_{iN^2}(\epsilon)E_i\rho_S(t)+\sum_{i=1}^{N^2}\sum_{j=1}^{N^2-1}c_{ij}(\epsilon)E_i\rho_S(t)E_j^*-\rho_S(t)\Big]=$$

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\Big[\frac{1}{N}C_{N^2N^2}(\epsilon)\rho_S(t)+\frac{1}{\sqrt{N}}\sum_{i=1}^{N^2-1}c_{iN^2}(\epsilon)E_i\rho_S(t)+\sum_{i=1}^{N^2}\sum_{j=1}^{N^2-1}c_{ij}(\epsilon)E_i\rho_S(t)E_j^*-\rho_S(t)\Big]=$$

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\Big[\frac{c_{N^2N^2}(\epsilon)-N}{N}\rho_S(t)+\frac{1}{\sqrt{N}}\sum_{i=1}^{N^2-1}c_{iN^2}(\epsilon)E_i\rho_S(t)+\frac{1}{\sqrt{N}}\sum_{j=1}^{N^2-1}c_{N^2j}(\epsilon)\rho_S(t)E_j^*$$
$$+\sum_{i,j=1}^{N^2-1}c_{ij}(\epsilon)E_i\rho_S(t)E_j^*\Big]=$$

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\Big[\frac{c_{N^2N^2}(\epsilon)-N}{N}\rho_S(t)+\frac{1}{\sqrt{N}}\sum_{i=1}^{N^2-1}\big(c_{iN^2}(\epsilon)E_i\rho_S(t)+c_{N^2i}(\epsilon)\rho_S(t)E_i^*\big)+$$
$$\sum_{i,j=1}^{N^2-1}c_{ij}(\epsilon)E_i\rho_S(t)E_j^*\Big]. \quad (3.5.10)$$

Now, by defining the following coefficients

$$a_{N^2N^2}=\lim_{\epsilon\to 0}\frac{c_{N^2N^2}(\epsilon)-N}{N},$$

$$a_{iN^2}=\lim_{\epsilon\to 0}\frac{c_{iN^2}(\epsilon)}{\epsilon},$$

$$a_{ij}=\lim_{\epsilon\to 0}\frac{c_{ij}(\epsilon)}{\epsilon},$$

we arrive at the following form of eq.(3.5.10)

$$\frac{a_{N^2N^2}}{N}\rho_S(t)+\frac{1}{\sqrt{N}}\sum_{i=1}^{N^2-1}\Big(a_{iN^2}E_i\rho_S(t)+a_{N^2i}\rho_S(t)E_i^*\Big)+\sum_{i,j=1}^{N^2-1}a_{ij}E_i\rho_S(t)E_j^*. \quad (3.5.11)$$

Further, we can introduce the operator M defined as

$$M \coloneqq \frac{1}{\sqrt{N}} \sum_{i=1}^{N^2-1} a_{iN^2} E_i$$

such that eq.(3.5.11) can be simplified in

$$\frac{a_{N^2N^2}}{N}\rho_S(t) + M\rho_S(t) + \rho_S(t)M^* + \sum_{i,j=1}^{N^2-1} a_{ij} E_i \rho_S(t) E_j^*. \tag{3.5.12}$$

For the bounded operator M it is possible to introduce the so called real and imaginary part of the operator [33], i.e. the selfadjoint operators M_R and M_I defined as

$$M_R = \frac{M + M^*}{2}, \tag{3.5.13}$$

and

$$M_I = \frac{M - M^*}{2i}, \tag{3.5.14}$$

such that $M = M_R + iM_I$. Thus, inserting eqs.(3.5.13) and (3.5.14) in eq.(3.5.12) one obtains

$$\frac{a_{N^2N^2}}{N} - i\left[\frac{M^* - M}{2i}, \rho_S(t)\right] + \rho_S(t)\frac{M + M^*}{2} + \frac{M + M^*}{2}\rho_S(t) + \sum_{i,j=1}^{N^2-1} a_{ij} E_i \rho_S(t) E_j^* =$$

$$-i\left[\frac{M^* - M}{2i}, \rho_S(t)\right] + \left\{\frac{M + M^*}{2} + \frac{1}{2N} a_{N^2N^2} I, \rho_S(t)\right\} + \sum_{i,j=1}^{N^2-1} a_{ij} E_i \rho_S(t) E_j^*.$$

By defining the Hermitian operator

$$H \coloneqq \frac{M^* - M}{2i} \tag{3.5.15}$$

and the operator

$$T \coloneqq \frac{M + M^*}{2} + \frac{1}{2N} a_{N^2N^2} I \tag{3.5.16}$$

we arrive at the following form of the genereto of the semigroup:

$$\mathcal{G}\rho_S(t) = -i\big[H, \rho_S(t)\big] + \big\{T, \rho_S(t)\big\} + \sum_{i,j=1}^{N^2-1} a_{ij} E_i \rho_S(t) E_j^*. \tag{3.5.17}$$

Moreover, using the trace preserving property of the channels, it must be true that $tr_{\mathcal{H}_S}[\mathcal{G}\rho_S(t)] = 0$, from which it follows that

$$tr_{\mathcal{H}_S}\left[\left(2T + \sum_{i,j=1}^{N^2-1} a_{ij}E_iE_j^*\right)\rho_S(t)\right] = 0,$$

and so

$$T = -\frac{1}{2}\sum_{i,j=1}^{N^2-1} a_{ij}E_i\rho_S(t)E_j^*.$$

In this way, eq.(3.5.17) can be written in the following concise form:

$$\mathcal{G}\rho_S(t) = -i\big[H, \rho_S(t)\big] + \sum_{i,j=1}^{N^2-1} a_{ij}\left(E_i\rho_S(t)E_j^* - \frac{1}{2}\big\{E_j^*E_i, \rho_S(t)\big\}\right). \tag{3.5.18}$$

Further, from the positivity of the coefficients a_{ij} [12], eq.(3.5.18) can be put in a diagonal form obtaining in this way the final form of the Lindblad equation:

$$\mathcal{G}\rho_S(t) = -i\big[H, \rho_S(t)\big] + \sum_{k=1}^{N^2-1} \zeta_k\left(L_k\rho_S(t)L_k^* - \frac{1}{2}L_k^*L_k\rho_S(t) - \frac{1}{2}\rho_S(t)L_k^*L_k\right) \tag{3.5.19}$$

where the ζ_k are the eigenvalues of the matrix a_{ij}, while the operators L_k, usually called Lindblad operators [12], are linked to the E_i by

$$E_i = \sum_{k=1}^{N^2-1} u_{ki}L_k, \tag{3.5.20}$$

the u_{ki} being the matrix coefficients of the unitary transformation that diagonalizes a_{ij}.
So, the conclusion at which one arrives is that the equation of motion for the reduced system density matrix can be written in the following form:

$$\frac{d}{dt}\rho_S(t) = \mathcal{G}\rho_S(t) = -i\big[H, \rho_S(t)\big] + \sum_{k=1}^{N^2-1} \zeta_k\left(L_k\rho_S(t)L_k^* - \frac{1}{2}L_k^*L_k\rho_S(t) - \frac{1}{2}\rho_S(t)L_k^*L_k\right). \tag{3.5.21}$$

The first term on the right hand side, represents the unitary evolution of the system, while the second is associated with the non-unitary contribution to the dynamics.

However, one must also note that the hermitian operator H does not correspond in general with the Hamiltonian of the open system since it may contain additional terms linked to the coupling of the system to its environment [12]. However, we must also observe that our derivation of eq.(3.5.21) is not a rigorous proof that eq.(3.5.19) is the most general form of the generator of a quantum dynamical semigroup for the case of a finite dimensional Hilbert space. The detailed proof that this is the case has been given by Gorini, Kossakowski and Sudarshan [34]. In the same time Lindblad [31] proved that eq.(3.5.19) is the most general form a bounded generator also in the case of a separable Hilbert space.
In the theory of open quantum systems, eq.(3.5.21) represents the equation of motion in the Schrödinger picture. However, it is also possible to write an equation of motion in the Heisenberg picture. In fact, the probability of every effect referred to the reduced system is obtained in the usual way as

$$tr_{\mathcal{H}_S}(E\rho_S(t)) = tr_{\mathcal{H}_S}(E\mathscr{C}_O(t)\rho_S(0)) = tr_{\mathcal{H}_S}((\mathscr{C}_O(t)^*E)\rho_S(0)). \tag{3.5.22}$$

The last term in eq.(3.5.22) is the definition for the Heisenberg picture of the quantum channel $\mathscr{C}_O(t)$. So one can obtain [12, 30], in a way similar to which we have obtained eq.(3.5.21), the equation of motion for an observable A in the Heisenberg picture:

$$\frac{d}{dt}A_H(t) = \mathcal{G}^*(t)A_H(t) = i[H, A_H(t)] + \sum_i \zeta_i\bigg(L_i^*A_H(t)L_i - \frac{1}{2}A_H(t)L_i^*L_i - \frac{1}{2}L_i^*L_iA_H(t)\bigg).$$

Remark 3.5.3. The way in which we have obtained the quantum master equation is not the only possible one. Another path through which eq.(3.5.21) can be derived is the so called microscopic derivation [12]. In this case, one starts by writing the Hamiltonian of the total system as:

$$H = H_S \otimes I_E + I_S \otimes H_E + H_I(t),$$

where H_S is the free Hamiltonian of the reduced system, H_E the free Hamiltonian of the environment while $H_I(t)$ is the interaction Hamiltonian. One then considers the interaction picture equation of motion:

$$\frac{d}{dt}\rho(t) = -i[H_I(t), \rho(t)], \tag{3.5.23}$$

where $\rho(t)$ is the state of the total system at time $t > 0$. Writing eq.(3.5.23) in integral form and performing a series of suitable approximations [12], one arrives

at an equation of motion for the reduced system that has the form of a quantum master equation. However, when one derives the master equation in this way, it appears that the coefficients ζ_i, previously introduced, are given in terms of certain correlation functions of the environment [12] and play the role of relaxation rates for the different decay modes of the open system.

Example 3.5.1. A typical field for the application of the theory of quantum Markovian master equations is the quantum optics. In this case, in fact, the Markovian condition is fully satisfied [12, 35]. The systems one considers are usually atoms or molecules in interaction with an electromagnetic field. In this case, the atom plays the role of the reduced system, while the electromagnetic field plays the role of the reservoir, i.e. an environment with an infinite number of degrees of freedom. However, the case we want to describe in the following example is the spontaneous emission of a two-level atom coupled to the vacuum [32]. In the spontaneous emission process for a two-level atom, the atom in a excited state emits a photon passing so to the ground state. Since the system under investigation is a two-dimensional system, we can consider a bi-dimensional Hilbert space $\mathcal{H}_S$ with basis vector $|g\rangle$ and $|e\rangle$. The vectors $|g\rangle$ and $|e\rangle$ denote, in particular, the ground and the excited state respectively. The free Hamiltonian of the atom can be described by $H_S = -\frac{\hbar\omega\sigma_z}{2}$, where $(\sigma_0, \vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3))$ constitutes a basis for $\mathcal{L}(\mathcal{H}_2)$, and where $-\hbar\omega$ is the energy difference between the excited and the ground state. The emission process can be described by the Lindblad operator $\sqrt{\gamma}|g\rangle\langle e|$ [12], where γ is the rate of the spontaneous emission. It is more convenient to resolve the interaction picture [24] master equation describing this process. Starting from eq.(3.5.21) one observes that in the interaction picture the term associated with the unitary dynamics disappear, thus leaving the following interaction picture master equation:

$$\frac{d}{dt}\rho_S(t) = \frac{\gamma}{2}\bigg(2\sigma_-\rho_S(t)\sigma_+ - \sigma_+\sigma_-\rho_S(t) - \rho_S(t)\sigma_+\sigma_-\bigg). \tag{3.5.24}$$

This equation can be solved by considering the Bloch representation form of the state $\rho_S(t)$. So by writing $\rho_S(t)$ as $\rho_S(t) = \frac{1}{2}\big(I + \frac{1}{2}\vec{\lambda}\cdot\vec{\sigma}\big)$ we arrive at the following equation:

$$\begin{pmatrix} \frac{d\lambda_z(t)}{dt} & \frac{d\lambda_x(t)}{dt} - i\frac{d\lambda_y(t)}{dt} \\ \frac{d\lambda_x(t)}{dt} + i\frac{d\lambda_y(t)}{dt} & -\frac{d\lambda_z(t)}{dt} \end{pmatrix} = \begin{pmatrix} 2\gamma(1-\lambda_z(t)) & (-\lambda_x(t) + i\lambda_y(t))\gamma \\ (-\lambda_x(t) - i\lambda_y(t))\gamma & -2\gamma(1-\lambda_z(t)) \end{pmatrix}$$

and so

$$\frac{d\lambda_x(t)}{dt} = -\gamma\lambda_x(t),$$
$$\frac{d\lambda_y(t)}{dt} = -\gamma\lambda_y(t),$$

$$\frac{d\lambda_z(t)}{dt} + 2\gamma\lambda_z(t) = 2\gamma.$$

The solution to these equations is easily found to be

$$\begin{aligned}\lambda_x(t) &= \lambda_x(0)e^{-\gamma t}\\ \lambda_y(t) &= \lambda_y(0)e^{-\gamma t}\\ \lambda_z(t) &= (\lambda_z(0) - 1)e^{-2\gamma t} + 1.\end{aligned}$$

The matrix elements $\rho_{11}(t)$ and $\rho_{22}(t)$ are the population of the exited and the ground state respectively, which show a decreasing exponential behaviour.
We can also write the Kraus operators for the quantum channel associated with this process (the density matrix is intended in the interaction picture [24]):

$$\rho_S(t) = \mathscr{C}_O(t)\rho_S(0) = E_0\rho(0)E_0^* + E_1\rho(0)E_1^* \tag{3.5.25}$$

with

$$E_0 = \begin{pmatrix}1 & 0\\ 0 & e^{-\gamma t}\end{pmatrix} \qquad E_1 = \begin{pmatrix}0 & \sqrt{1-e^{-2\gamma t}}\\ 0 & 0\end{pmatrix}. \tag{3.5.26}$$

■

Measurement Models and Quantum Instruments

4

4.1 Introduction

In this chapter, we describe some concepts that are of particular interest in the modern theory of quantum measurement. Every measurement process, admits three levels of description. The first level is the description provided by POVMs. In this case, the only information gained are the outcome probabilities of the measurement. The theory and some applications of these objects have been discussed in chapter 2. Further, a more complete level of description is the one offered by quantum instruments [8, 4]. A quantum instrument, in fact, allows one to know the outcome probabilities as well as the output states, i.e. the states conditioned by the measurement process. We will see, in particular, that this kind of object is described, mathematically, by a suitable linear map from a Borel σ-algebra $\mathcal{B}(\Omega)$ to the space of quantum operations on $\mathcal{T}(\mathcal{H})$. This, eventually, permits to see how every quantum instrument can work in two possible ways: as an operation, thus giving information about the measurement process, or as a quantum channel, thus providing a deterministic transmission of a quantum state. The last level of description, which is the most complete one, is represented by the so called quantum measurement model [8, 9]. In this case, in fact, the description of the measurement takes into consideration the interaction between the measurement apparatus and the measured system. In other terms, a measurement model describes the measurement process as an open quantum system, in which the measuring apparatus represents the environment, while the measured system is the open system. Clearly, these three levels of description are not independent. Indeed, every measurement model induces an instrument, and every instrument a POVM. Conversely, every POVM admits an entire class of compatible

instruments, and every instrument has a class of possible measurement models that induces it. This one to many relations is, eventually, what one expects [8] since fixed a POVM, different measurement settings can be employed to measure it.
Moreover, it is possible to restrict the possible measurements that can be performed on a system, by imposing some additional conditions. In particular, in this chapter, we will describe the constraints that a conserved quantity [8], or the action of a suitable symmetry group, impose on POVMs and quantum instruments. More precisely, given a conserved quantity, the Wigner-Araki-Yanase theorem [8], which we will discuss, shows how the only POVMs that can be measured on the system are those that commute with the conserved quantity. Similarly, by imposing that a POVM, or a quantum instrument, is covariant under the action of a group, in a way that we will describe, one obtains [36, 37, 38] a precisely defined class of POVMs or instruments that satisfy this condition.
After we have introduced the main subjects of the chapter, we pass to present the order in which they will be discussed: In section 4.2 we introduce the concept of a measurement model, while we discuss quantum instruments in section 4.3. In particular, we obtain a prototype form of a quantum instrument, by defining it as a suitable map induced by a measurement model. Then, via the Ozawa's theorem [9], we will see that every quantum instrument is induced by a measurement model. We conclude our general discussion about quantum instruments by describing two important implications: the notion of conditional state preparation, linked to the idea that an instrument can be used to prepare a state in a prefixed way, and the proof that it is impossible to gain some information on a physical system without disturbing it. Section 4.4 is devoted to present a particular measurement model, known as the Von Neumann measurement model [8], and the related class of instruments called Luders instruments [4, 8]. The last section is then left to the role played by conservation laws and covariance principles in a measurement process. About the first argument, we will state and prove, under simplified hypotheses, the fundamental theorem of Wigner, Araki, and Yanase. We pass then to discuss the covariance of POVMs and quantum instruments under the action of a group. In particular, under the hypotheses that the group is unimodular and locally compact, we will state two theorems that characterize the form of a covariant POVM and a covariant quantum instrument.

4.2 Measurement Models

In this section, we start by introducing the most complete description of a quantum measurement, leaving the next section to the description of the coarser level given by quantum instruments. In order to describe a measurement model [8, 4, 9], we begin

by discussing, in an informal way, how one can describe a measurement process. We start by an observable A and a system prepared in a state ρ. The measurement procedure requires that the measured system interacts with a measurement apparatus. The measurement apparatus is assumed to have, at least, a microscopic part, known as the probe system [4], which admits a quantum mechanical description. The measurement process, in this optics, is described as a transfer of information from the object system to the probe. Then one assumes [4, 8] that there is an observable, known as the pointer observable, into which the information about the measured observable is to be transcribed.
In conclusion, one sees that according to this model of the quantum measurement process, the measured system acts as an open quantum system that interacts with an environment represented by the probe system.
So, we can now restate in more precise terms what we said until now through the following [8]

Definition 4.2.1. Let Z be an observable, with outcome space $(\Omega, \mathcal{B}(\Omega))$, where we remember $\mathcal{B}(\Omega)$ denotes the Borel σ-algebra of sets of Ω. Suppose that we intend to measure the observable Z with Hilbert space $\mathcal{H}$ on a system in the state ρ. A measurement model $\mathcal{M}$ for Z is the quadruple $\mathcal{M} = \langle \mathcal{H}_P, \sigma, \mathscr{C}, E \rangle$ where

(i) $\mathcal{H}_P$ is the Hilbert space of the probe system,

(ii) σ is the initial state of the probe,

(iii) $\mathscr{C}$ is a quantum channel from $\mathcal{T}(\mathcal{H} \otimes \mathcal{H}_P)$ to $\mathcal{T}(\mathcal{H} \otimes \mathcal{H}_P)$ that describes the coupling between the measured system and the probe,

(iv) E is the pointer observable, i.e. the observable implemented on the probe that is assumed to have the same outcome space $(\Omega, \mathcal{B}(\Omega))$ of Z,

such that the probability reproducibility condition

$$tr_{\mathcal{H}}[\rho Z(X)] = tr_{\mathcal{H}\otimes\mathcal{H}_P}[\mathscr{C}(\rho \otimes \sigma)(I \otimes E(X))] \tag{4.2.1}$$

is satisfied for all $X \in \mathcal{B}(\Omega)$ and $\rho \in \mathcal{S}(\mathcal{H})$.

What we see is that in a quantum measurement model, one starts by considering the probe and the measured system to be statistically independent, such that the initial state of the compound system is $\rho \otimes \sigma$. Then, the interaction between the probe and the system is described by a quantum channel $\mathscr{C}$. Finally, the measurement is performed using the pointer observable acting on the probe only. The condition in (4.2.1) then simply means that the measurement of the pointer observable E leads to the same probabilities as if the measurement of Z was performed directly on the system.

Remark 4.2.1. The final state of the compound system, after the interaction, is described by $\mathscr{C}(\rho \otimes \sigma)$. From this, one can recover the final states of the probe and the measured object through the partial trace:

$$\sigma^f = tr_{\mathcal{H}}[\mathscr{C}(\rho \otimes \sigma)],$$

and

$$\rho^f = tr_{\mathcal{H}_P}[\mathscr{C}(\rho \otimes \sigma)].$$

Moreover a measurement model permits also to find the final form of the system state after the measurement is carried out. We will return more precisely on this point in the next section after we have introduced the concept of quantum instrument.

As a consequence of the definition, one sees that every measurement model induces a unique quantum observable acting on the Hilbert space $\mathcal{H}$, as is proved in the following [4]

Proposition 4.2.1. *Given a measurement model $\mathcal{M} = \langle \mathcal{H}_P, \sigma, \mathscr{C}, E \rangle$, the mapping*

$$\mathcal{B}(\Omega) \ni X \mapsto E_{\sigma^f}(X) \coloneqq tr_{\mathcal{H}_P}[\sigma^f E(X)] = tr_{\mathcal{H}\otimes\mathcal{H}_P}[\mathscr{C}(\rho \otimes \sigma)(I \otimes E(X))] \in [0,1]$$

is a probability measure for all $\rho \in \mathcal{S}(\mathcal{H})$ and $X \in \mathcal{B}(\Omega)$. In particular, it determines a unique observable $Z^{\mathcal{M}}$ acting on $\mathcal{H}$ and with outcome space $(\Omega, \mathcal{B}(\Omega))$, such that

$$tr[\rho Z^{\mathcal{M}}(X)] = E_{\sigma^f}(X). \tag{4.2.2}$$

So, when we say that a measurement model implements the observable Z what we are saying is that $Z = Z^{\mathcal{M}}$, i.e. that eq.(4.2.2) (or equivalently (4.2.1)) is fulfilled.

At this point, what we can ask is if, given an observable Z, there exists a measurement model that implements it. The answer to this question is the content of the following theorem [4, 10] that is a consequence of the Neumark theorem discussed in chapter 2 (see section 2.6)

Theorem 4.2.1. *For any observable Z of a quantum system S there exists a measurement model $\mathcal{M} = \langle \mathcal{H}_P, \sigma, \mathscr{C}, E \rangle$ such that $Z = Z^{\mathcal{M}}$, where $Z^{\mathcal{M}}$ is the observable induced by the measurement model. Moreover the pointer observable E can be chosen to be sharp and the initial probe state σ to be a pure state.*

Remark 4.2.2. A possible way [8] to describe the interaction between the probe and the measured system is the following: we fix an orthonormal basis $\{\psi_j\}_{j=1}^n$ for the Hilbert space system $\mathcal{H}$. After, we consider a fixed vector η_0 of the Hilbert space

$\mathcal{H}_P$ that represents the initial vector state of the probe. Then we can consider the following map

$$\psi_j \otimes \eta_0 \mapsto \psi_j \otimes \eta_j, \tag{4.2.3}$$

where η_j is a set of n vector states in $\mathcal{H}_P$. If we extend $\{\psi_j \otimes \eta_0\}$ and $\{\psi_j \otimes \eta_j\}$ to orthonormal bases of $\mathcal{H} \otimes \mathcal{H}_P$, this map has an extension to a unitary operator on $\mathcal{H} \otimes \mathcal{H}_P$. So if we consider this extension, we have obtained a possible tool through which one can describe the interaction of the system and the probe. We will see an application of this type of measurement interaction in the following example [8, 10], and later in section 4.4 in the context of Von Neumann's measurement model.

Example 4.2.1. As an example of a measurement model, we consider the Stern-Gerlach experiment discussed in chapter 2 (example 2.2.1). In this experiment, one considers a beam of spin-$\frac{1}{2}$ particle described by the tensor product Hilbert space $\mathbb{C}^2 \otimes L^2(\mathbb{R}^3)$. The particles pass through a magnetic field that couples the spin degrees of freedom to the spatial degrees of freedom. We assume that the initial spin state is a pure state and we can therefore write it as $\phi = a\phi_+ + b\phi_-$ with $a, b \in \mathbb{C}$, while $\{\phi_+, \phi_-\}$ defines an orthonormal basis for $\mathbb{C}^2 \equiv \mathcal{H}_S$. If we denote with $\xi \in L^2(\mathbb{R}^3)$ the initial spatial part of the state of the particle, we can write the initial compound state as $\phi \otimes \xi$. As we have said, the passage of the particles in the magnetic field couples the spatial degrees of freedom with the spin degrees of freedom and so we can describe this process through a channel $\mathscr{C}$ such that

$$\phi \otimes \xi \;\mapsto\; a\phi_+ \otimes \xi_+ + b\phi_- \otimes \xi_-, \tag{4.2.4}$$

where $\xi_\pm$ are fixed vectors in $L^2(\mathbb{R}^3)$ that describe the deflection of the particles up or down. Further, a plane orthogonal to the incoming beam collects the deflected particles. Then we can consider the plane as the probe system, while the pointer observable can be described in terms of projections $P_\pm$ in the lower or the upper half plane. So using eq.(4.2.1) we can find the probabilities $p(\pm)$ that the particles are deflected either up or down:

$$p(\pm) = tr_{\mathcal{H}_P}[tr_{\mathcal{H}_S}[\mathscr{C}(\phi \otimes \xi)]P_\pm] = tr_{\mathcal{H}_P}[(|a|^2|\xi_+\rangle\langle\xi_+| + |b|^2|\xi_-\rangle\langle\xi_-|)P_\pm] = |a|^2\langle\xi_+|P_\pm\xi_+\rangle + |b|^2\langle\xi_-|P_\pm\xi_-\rangle. \tag{4.2.5}$$

However we can also find the observable associated to this measurement model using again eq.(4.2.1). In particular we can introduce a discrete observable $S^{\mathcal{M}}$ that consists of two effects $S^{\mathcal{M}} = \{E(+), E(-)\}$. Then we find that

$$\begin{aligned} tr[E(\pm)|\phi\rangle\langle\phi|] = \\ tr[E(\pm)(|a|^2|\phi_+\rangle\langle\phi_+| + |b|^2|\phi_-\rangle\langle\phi_-| + a^*b|\phi_-\rangle\langle\phi_+| + b^*a|\phi_+\rangle\langle\phi_-|)] = \\ |a|^2\langle\phi_+|E(\pm)\phi_+\rangle + |b|^2\langle\phi_-|E(\pm)\phi_-\rangle + a^*b\langle\phi_+|E(\pm)\phi_-\rangle + b^*a\langle\phi_-|E(\pm)\phi_+\rangle. \end{aligned} \tag{4.2.6}$$

From the comparison of eq.(4.2.5) with eq.(4.2.6) we arrive at the conclusion that

$$E(\pm) = \langle \xi_+ | P_\pm \xi_+ \rangle |\phi_+\rangle\langle\phi_+| + \langle \xi_- | P_\pm \xi_- \rangle |\phi_-\rangle\langle\phi_-|. \tag{4.2.7}$$

The operator $S^{\mathcal{M}}$ coincides with the usual spin observable $S^{\vec{n}}$ if and only if $\langle \xi_\pm | P_\pm \xi_\pm \rangle = 1$ and $\langle \xi_\mp | P_\pm \xi_\mp \rangle = 0$. ■

4.3 Quantum Instruments

In this section, we discuss the notion of quantum instrument [8, 4]. As we have said, a quantum instrument is a map that permits the knowledge of the measurement outcome probabilities as well as the output states conditioned by the measurement process. As we will soon see, every quantum instrument is induced by a suitable measurement model. So let $\mathcal{M} = \langle \mathcal{H}_P, \sigma, \mathscr{C}, E \rangle$ be a measurement model associated with the observable Z with outcome space $(\Omega, \mathcal{B}(\Omega))$. Then, let us consider another observable R with outcome space $(\Omega', \mathcal{B}(\Omega'))$. For every $X \in \mathcal{B}(\Omega)$ and $Y \in \mathcal{B}(\Omega')$ we can consider the joint probability $p_\rho(Z \in X \;\&\; R \in Y)$ where ρ is the initial state on which Z is measured. Assuming that the measurement of the observable R is performed directly on the system (thus without considering a measurement model for it), we can express the joint probability as

$$p_\rho(Z \in X \;\&\; R \in Y) = tr_{\mathcal{H}\otimes\mathcal{H}_P}[\mathscr{C}(\rho \otimes \sigma)(R(Y) \otimes Z(X))]. \tag{4.3.1}$$

Using the definition of partial trace (see section 1.4), we can rewrite this quantity in the following way:

$$\begin{aligned} tr_{\mathcal{H}\otimes\mathcal{H}_P}[\mathscr{C}(\rho \otimes \sigma)(R(Y) \otimes Z(X))] = tr_{\mathcal{H}\otimes\mathcal{H}_P}[\mathscr{C}(\rho \otimes \sigma)(R(Y) \otimes I)(I \otimes Z(X)] = \\ tr_{\mathcal{H}}[R(Y) tr_{\mathcal{H}_p}[\mathscr{C}(\rho \otimes \sigma)(I \otimes Z(X))]], \end{aligned} \tag{4.3.2}$$

and by defining

$$\mathscr{I}_X^{\mathcal{M}}(\rho) \coloneqq tr_{\mathcal{H}_p}[\mathscr{C}(\rho \otimes \sigma)(I \otimes Z(X))], \tag{4.3.3}$$

we arrive at the following form of eq.(4.3.2)

$$p_\rho(Z \in X \;\&\; R \in Y) = tr_{\mathcal{H}}[\mathscr{I}_X^{\mathcal{M}}(\rho) R(Y)], \tag{4.3.4}$$

where $\mathcal{H}$ is the Hilbert space of the observable R. If now we consider the map $\rho \mapsto \mathscr{I}_X^{\mathcal{M}}(\rho)$ for fixed X, it can be shown (see [8] section 5.2.1) it has a unique extension to a linear map on the space of trace class operators $\mathcal{T}(\mathcal{H})$. From eq.(4.3.4) we observe that $\mathscr{I}_X^{\mathcal{M}}(\rho)$ describes the unnormalized conditional output state of the

measured system on which the measurement of $R(Y)$ is performed. Moreover if we consider $p_\rho(Z \in X, R \in \Omega')$ we are simply considering $p_\rho(Z \in X)$, and using eq.(4.3.4) we obtain

$$p_\rho(Z \in X) = tr[\mathscr{I}_X^{\mathcal{M}}(\rho)R(\Omega')] = tr[\mathscr{I}_X^{\mathcal{M}}(\rho)]. \tag{4.3.5}$$

So what we can conclude from this discussion is that the mapping $\mathscr{I}_X^{\mathcal{M}}$ represents a tool through which one can find the state $\mathscr{I}_X^{\mathcal{M}}(\rho)$ after the Z measurement, and at the same time the probability $tr[\mathscr{I}_X^{\mathcal{M}}(\rho)]$ of the measurement. This map, that as we have seen is induced by a measurement model, represents the prototype of an object, called quantum instrument, that we are now ready to introduce. In fact, abstracting from the previous discussion we can consider a map $\mathscr{I}_X^{\mathcal{M}}$ that satisfies, for an outcome space $(\Omega, \mathcal{B}(\Omega))$, the following conditions [8]:

(i) for each $X \in \mathcal{B}(\Omega)$ the map $\mathscr{I}_X^{\mathcal{M}}$ is a quantum operation,

(ii) $\mathscr{I}_\Omega^{\mathcal{M}}$ is a quantum channel and so $tr[\mathscr{I}_\Omega^{\mathcal{M}}(\rho)] = 1$ while $\mathscr{I}_\emptyset^{\mathcal{M}}(\rho) = O$ for all $\rho \in \mathcal{S}(\mathcal{H})$,

(iii) If $\{X_j\}_j$ is a sequence of mutually disjoint sets, then for all $\rho \in \mathcal{S}(\mathcal{H})$ the following condition is verified

$$tr[\mathscr{I}_{\bigcup_j X_j}^{\mathcal{M}}(\rho)] = \sum_j tr[\mathscr{I}_{X_j}^{\mathcal{M}}(\rho)]. \tag{4.3.6}$$

Definition 4.3.1. A mapping $\mathscr{I}_X^{\mathcal{M}}$ from an outcome space $(\Omega, \mathcal{B}(\Omega))$ to the set of operations on $\mathcal{T}(\mathcal{H})$ satisfying conditions (i)-(iii) is called a quantum instrument.

So, what we see is that a quantum instrument is a map that works in two ways: it can work as a quantum channel, thus allowing the transmission of a quantum state; it can work as an operation when one is interested in the output state conditioned by the measurement process.

Remark 4.3.1. The quantity $tr[\mathscr{I}_X^{\mathcal{M}}(\rho)]$ represents the probability that the observable Z takes values in the set $X \in \mathcal{B}(\Omega)$. Since $\mathscr{I}_X^{\mathcal{M}}$ for fixed X is an operation, we have obtained a justification of the interpretation of the quantity $tr[\mathscr{O}(\rho)]$ as the probability of the process associated with the quantum operation $\mathscr{O}$, as discussed in the previous chapter (see section 3.2).

In the case the set Ω of measurement results is countable, we write $\mathscr{I}_x^{\mathcal{M}}$ instead of $\mathscr{I}_{\{x\}}^{\mathcal{M}}$ for the related instrument. For the operation $\mathscr{I}_x^{\mathcal{M}}$, we can consider a decomposition in the Kraus form, such that $\mathscr{I}_x^{\mathcal{M}}(\rho) = \sum_i K_i \rho K_i^*$. If we consider the collection of all the Kraus operators associated with these decompositions, we obtain

[8] a Kraus decomposition for the channel $\mathscr{I}_\Omega^{\mathcal{M}}$. Conversely, if we start by a quantum channel $\mathscr{C}$, and consider a Kraus decomposition for it, i.e. $\mathscr{C}(\rho) = \sum_i K_i \rho K_i^*$, we can define $\mathscr{I}_i^{\mathcal{M}}(\rho) \coloneqq K_i \rho K_i^*$, obtaining in this way a countable instrument such that $\mathscr{I}_\Omega^{\mathcal{M}} = \mathscr{C}$.
However, we must also observe that although the Kraus decomposition theorem permits to obtain discrete instruments starting from quantum channels, the definition given above of quantum instruments implies that they need not be discrete while the Kraus decomposition is always countable.

Remark 4.3.2. There is another useful way in which a quantum instrument can be defined. We prefer to present here this alternative definition, because we will use it, later, when we discuss the concept of covariant quantum instrument (see section 4.5). Thus, let us introduce the set $\mathcal{M}(\Omega, \mathcal{T}(\mathcal{H}))$, that is the Banach space of $\mathcal{T}(\mathcal{H})$-valued Borel measures on Ω, i.e. the space of Borel measures $\mu : \mathcal{B}(\Omega) \to \mathcal{T}(\mathcal{H})$. Given an element μ in $\mathcal{M}(\Omega, \mathcal{T}(\mathcal{H}))$, the norm of μ is given by $\|\mu\| \coloneqq |\mu|(\Omega)$, where $|\mu|$ denotes the total variation of μ [3, 33]. We can then give the following alternative definition [37] of a quantum instrument:

Definition 4.3.2. A quantum instrument $\mathscr{I}$ with outcome space $(\Omega, \mathcal{B}(\Omega))$ is a linear map $\mathscr{I} : \mathcal{T}(\mathcal{H}) \to \mathcal{M}(\Omega, \mathcal{T}(\mathcal{H}))$ such that

i) for each $X \in \mathcal{B}(\Omega)$ the map $\mathscr{I}_X$ defined by

$$\mathscr{I}_X : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{H}), \qquad S \mapsto \mathscr{I}_X(S) \coloneqq (\mathscr{I}S)(X), \tag{4.3.7}$$

for all $S \in \mathcal{T}(\mathcal{H})$ is a completely positive map;

ii) If we consider the whole Ω it is true that

$$tr[(\mathscr{I}S)(\Omega)] = tr[S] = tr[\mathscr{I}_\Omega(S)], \tag{4.3.8}$$

for all $S \in \mathcal{T}(\mathcal{H})$.

So, as we see, condition *i*) implies that, for each fixed X, $\mathscr{I}_X$ is a quantum operation, while condition *ii*) that it is, indeed, a quantum channel if one considers the whole Ω.

From what we have said so far, it appears evident that a measurement model induces always a quantum instrument $\mathscr{I}^{\mathcal{M}}$ called the instrument induced by $\mathcal{M}$. However, we can also ask if the converse is true, i.e. if every instrument $\mathscr{I}$ admits a measurement model that induces it, i.e. a measurement model $\mathcal{M}$ such that $\mathscr{I} = \mathscr{I}^{\mathcal{M}}$. In this respect, a theorem due to Ozawa clarifies the relationship between the measurement models and the quantum instruments [9, 8],

Theorem 4.3.1. *(Ozawa)*
Let $\mathscr{I}$ be a quantum instrument. Then there exists a measurement model $\mathcal{M} = \langle \mathcal{H}_P, \sigma, \mathscr{C}, E \rangle$, such that $\mathscr{I} = \mathscr{I}^{\mathcal{M}}$. In particular it is always possible to choose $\mathcal{M}$ such that σ is a pure state, $\mathscr{C}$ is a unitary channel and E is a sharp observable.

So we see that it is not restrictive to think of a quantum instrument as a map induced by a measurement model, thus as the map defined in eq.(4.3.3).
At this point, we must make the following observation. As we have seen, every measurement model induces an observable $A^{\mathcal{M}}$ and an instrument $\mathscr{I}^{\mathcal{M}}$. Then what we will now see [8] is that also a quantum instrument induces an observable $A^{\mathscr{I}}$. In order to fix the ideas, let us consider a quantum instrument $\mathscr{I}$ defined on an outcome space $(\Omega, \mathcal{B}(\Omega))$. As we have seen, the quantity $tr[\mathscr{I}_X(\rho)]$ is, for every $X \in \mathcal{B}(\Omega)$, the probability of the measurement process described by the quantum instrument $\mathscr{I}$. However, $\mathscr{I}_X$ being an operation, for every fixed X, admits a dual operation (i.e. the Heisenberg picture) $\mathscr{I}_X^*$. In this way, we can consider

$$tr[\mathscr{I}_X(\rho)] = tr[\mathscr{I}_X^*(I)\rho], \tag{4.3.9}$$

i.e. $\mathscr{I}_X^*(I)$ defines an observable $A^{\mathscr{I}}(X)$ that has the outcome space $(\Omega, \mathcal{B}(\Omega))$. The observable $A^{\mathscr{I}}$ is said to be the observable induced by the quantum instrument $\mathscr{I}$. This result is what one expects since, as we have seen, every quantum instrument describes a measurement process on a physical system, and thus it implements the action of a suitable observable. In the same way, one can also ask if the converse is true, i.e. if given an observable A there exists a suitable instrument $\mathscr{I}$ that describes a certain way of measuring it. What one observes ([8], section 5.1.3) is that if such an instrument exists, it must satisfy a compatibility condition with the observable, i.e. it must be true that $A = A^{\mathscr{I}}$. When such a compatibility condition is satisfied, one says that the instrument $\mathscr{I}$ is "$A-$*compatible*". In general, an observable A admits infinitely many A-compatible instruments as the following example [8] shows.

Example 4.3.1. Let us consider an observable A and a fixed state η in the Hilbert space $\mathcal{H}$. We can introduce a kind of instrument, known as the trivial instrument, defined as

$$\mathscr{I}_X(\rho) \coloneqq tr[A(X)\rho]\eta. \tag{4.3.10}$$

This instrument defines an A-compatible instrument as follows from the observation that:

$$\begin{aligned} tr[A^{\mathscr{I}}(X)\rho] = tr[\mathscr{I}_X^*(I)\rho] = tr[\mathscr{I}_X(\rho)] \\ = tr[tr[A(X)\rho]\eta] = tr[A(X)\rho]tr[\eta] = tr[A(X)\rho], \end{aligned}$$

and since this condition is valid for every $\rho \in \mathcal{S}(\mathcal{H})$ we can conclude that the compatibility condition is satisfied and so that the instrument is A-compatible. Since

the vector η is arbitrary we can conclude that every observable admits infinitely many compatible instruments. ■

Another consideration that we must do is if the objects that we have defined so far "fit together" coherently. What we must verify is that the observable $A^{\mathscr{I}^{\mathcal{M}}}$, i.e. the observable induced by the instrument $\mathscr{I}^{\mathcal{M}}$, induced in turn by the measurement model $\mathcal{M}$, is compatible with the observable $A^{\mathcal{M}}$ induced by the measurement model. In particular, what we observe is that

$$tr_{\mathcal{H}}[\rho A^{\mathscr{I}^{\mathcal{M}}}(X)] = tr_{\mathcal{H}}[\mathscr{I}_X^{\mathcal{M}}(\rho)] = tr_{\mathcal{H}}[tr_{\mathcal{H}_P}[\mathscr{C}(\rho\otimes\sigma)(I\otimes E(X))]] = tr_{\mathcal{H}\otimes\mathcal{H}_P}[\mathscr{C}(\rho\otimes\sigma)(I\otimes E(X))] = tr[\rho A^{\mathcal{M}}(X)],$$

thus implying the consistency of the scheme described.

Remark 4.3.3. There is a crucial aspect in the relationship between measurement models, instruments and quantum observables that we must point out. Every measurement model induces a unique instrument, and every instrument a unique observable. Thus the relation is

$$\mathcal{M} \;\rightarrow\; \mathscr{I}^{\mathcal{M}} \qquad\qquad \mathscr{I} \;\rightarrow\; A^{\mathscr{I}}. \tag{4.3.11}$$

However if we reverse the order in eq.(4.3.11) what we obtain is

$$A \;\rightarrow\; [\mathscr{I}] \qquad\qquad \mathscr{I} \;\rightarrow\; [\mathcal{M}], \tag{4.3.12}$$

where $[\mathscr{I}]$ denotes a set of all instruments that are A-compatibles with A, while $[\mathcal{M}]$ denote a set of all measurement models that induce the instrument $\mathscr{I}$. Thus what we see is that while in eq.(4.3.11) the relation is one to one, the opposite direction is one to many. The physical significance of this circumstance is that fixed a measurement model, it describes the action of a unique observable, while fixed an observable, there are, in general, different measurement settings that can reproduce its action.

Conditional State Preparation

As an important application of the theory of quantum instruments, we discuss the concept of conditional state preparation [8]. In fact, a quantum instrument permits to know the state of a system after a measurement is carried out. Thus one can use the measurement process as a tool through with it is possible to obtain a suitable output state, i.e. as a state preparator.
To begin with, we describe in more precise terms the form of the conditional output

state. We have seen that if we have two observables Z and T, and if we measure first the observable Z and after the observable T, we can find the joint probability distribution of the values of Z and T as

$$p_\rho(Z \in X \;\&\; T \in Y) = tr[\mathscr{I}_X(\rho)T(Y)], \tag{4.3.13}$$

where ρ is the state on which the Z measurement is performed, while $\mathscr{I}_X$ is a Z-compatible instrument. From the joint probability we can obtain the conditional probability $p_\rho(T \in Y|Z \in X)$ as

$$p_\rho(T \in Y|Z \in X) = \frac{p_\rho(Z \in X \;\&\; T \in Y)}{p_\rho(Z \in X)} = \frac{tr[\mathscr{I}_X(\rho)T(Y)]}{tr[\mathscr{I}_X(\rho)]} = tr[\tilde{\rho}_X T(Y)],$$

where we have defined the state $\tilde{\rho}_X$ as

$$\tilde{\rho}_X \coloneqq \frac{\mathscr{I}_X(\rho)}{tr[\mathscr{I}_X(\rho)]}. \tag{4.3.14}$$

The state $\tilde{\rho}_X$, under the hypothesis that $p_\rho(Z \in X) \neq 0$, is called the conditional output state.

Example 4.3.2. As an example of conditional state preparation, we consider the case of a trivial instrument. We remember that the trivial instrument is defined as

$$\mathscr{I}_X(\rho) = tr[A(X)\rho]\eta. \tag{4.3.15}$$

From this we observe that

$$\tilde{\rho}_X = \frac{\mathscr{I}_X(\rho)}{tr[\mathscr{I}_X(\rho)]} = \eta. \tag{4.3.16}$$

In particular we also observe the reason for the name trivial instrument. In fact, every successive measurement following the trivial instrument is "trivialized", in the sense that the measurement

$$p_\rho(T \in Y|Z \in X) = tr[\tilde{\rho}_X T(Y)] = tr[\eta T(Y)] = p_\eta(T \in Y). \tag{4.3.17}$$

Thus the T measurement is the same as measuring the trivial observable $Y \mapsto tr[\eta T(Y)]I$ in the state ρ. ■

We are now ready to discuss the conditional state preparation. The idea is to fix a particular instrument $\mathscr{I}$ that induces a prefixed $\tilde{\rho}$. We start by considering an observable Z with a countable outcome set Ω. For each outcome $x \in \Omega$, we fix a state η_x and we introduce the Z-compatible instrument

$$\mathscr{I}_x(\rho) = tr[Z(X)\rho]\eta_x. \tag{4.3.18}$$

This instrument, called the conditional state preparator [8], induces a conditional output state $\tilde{\rho}_x = \eta_x$. Although the process associated with the instrument is probabilistic, the measurement outcome indicates which state is obtained.

Remark 4.3.4. The trivial instrument is a particular type of state preparator, since in this case the conditional output state is always the same.

We will now see two useful applications of conditional state preparation. The first is that through this notion it is possible to classify the compatible instruments of rank-1 observables, i.e. of discrete observables A such that each effect $A(x)$ is rank-1.
In particular, this is the content of the following proposition (for a proof see [8] proposition 5.11):

Proposition 4.3.1. *Let A be a rank-1 observable with outcome space $(\Omega, \mathcal{B}(\Omega))$. Then, all A-compatible instruments are of the form*

$$\mathscr{I}_x(\rho) = tr[A(x)\rho]\eta_x, \tag{4.3.19}$$

for all $x \in \mathcal{B}(\Omega)$ and for some state vectors $\eta_x \in \mathcal{H}$

Another important result that can be obtained using the conditional state preparation is that, as we will now show, it is impossible to gain information from a system without disturbing it. In particular, a crucial element in the proof of this "no information without disturbance" theorem is the linearity of quantum operations. The hypothesis [8] from which we start is that if we consider two observables Z and T, then the measurement of Z does not make any difference to the measurement outcome distribution of the subsequently measured observable T. In mathematical terms this condition translates in

$$p_\rho(T \in Y | Z \in X) = p_\rho(T \in Y), \tag{4.3.20}$$

for every $Y \in \mathcal{B}(\Omega')$, $X \in \mathcal{B}(\Omega)$ and $\rho \in \mathcal{S}(\mathcal{H})$.
Since T is arbitrary, eq.(4.3.20) implies that

$$\tilde{\rho} = \frac{\mathscr{I}_X(\rho)}{tr[\mathscr{I}_X(\rho)]} = \rho \Longrightarrow \mathscr{I}_X(\rho) = c_X(\rho)\rho, \tag{4.3.21}$$

where $c_X(\rho)$ is a nonnegative number depending on X and ρ, while $\mathscr{I}$ is a Z-compatible instrument. Owing to the linearity of the operation $\mathscr{I}_X$ we can show that $c_X(\rho)$ does not depend on ρ. In fact, for two states ρ_1 and ρ_2, we obtain

$$\mathscr{I}_X(\rho_1 + \rho_2) = c_X(\rho_1 + \rho_2)(\rho_1 + \rho_2), \tag{4.3.22}$$

and

$$\mathscr{I}_X(\rho_1+\rho_2)=\mathscr{I}_X(\rho_1)+\mathscr{I}_X(\rho_2)=c_X(\rho_1)\rho_1+c_X(\rho_2)\rho_2. \tag{4.3.23}$$

From this it follows that $c_X(\rho)\equiv c_X$ is independent of ρ. Now, returning to eq.(4.3.21) we obtain

$$\mathscr{I}_X(\rho)=c_X\rho, \tag{4.3.24}$$

and taking the trace of both sides we arrive to

$$tr[\mathscr{I}_X(\rho)]=tr[Z(X)\rho]=c_X. \tag{4.3.25}$$

If we take $\rho=|\psi\rangle\langle\psi|$, i.e. a pure state, eq.(4.3.25) implies that $Z(X)=c_XI$. Since every state can be decomposed as a mixture of pure states, we can conclude that the only way in which an observable does not disturb the system is if it is a trivial observable, i.e. an observable proportional to the identity operator.

Remark 4.3.5. One can think that condition eq.(4.3.20) is too strong and so one can consider the milder requirement [8]

$$p_\rho(T\in Y|Z\in\Omega)=p_\rho(T\in Y), \tag{4.3.26}$$

i.e. that the Z measurement does not disturb the system under the condition that we measure it, but we do not observe the outcome of the measurement. However, also in this case, it can be shown (see [8] section 5.2.2) that one arrives at the same conclusion as in eq.(4.3.25). The conclusion of the previous discussion is thus that every non trivial observable produces always a disturbance in the measured system if it provides some informations on it.

4.4 Von Neumann's Measurement Model

In this section, we present a particular type of measurement model for discrete sharp observables due to Von Neumann.
Let $\mathcal{H}$ be a Hilbert space and let $\{\xi_j\}_j$ be an orthonormal basis for it. We can consider the sharp observable V associated with such a basis, i.e. the sharp observable such that $V(j)=|\xi_j\rangle\langle\xi_j|$. Then, we fix a Hilbert space $\mathcal{H}_P$ that has the same dimension as $\mathcal{H}$ and endowed with the orthonormal basis $\{\eta_j\}_j$. We can introduce the sharp observable E defined as $E(j)=|\eta_j\rangle\langle\eta_j|$, which will represent the pointer observable of the measurement model. The last element that we must define is a suitable channel that describes the coupling between the system and the probe. In

particular, in the Von Neumann's measurement model the channel $\mathscr{C}$ acts in the following way on $\mathcal{T}(\mathcal{H}\otimes\mathcal{H}_P)$:

$$\mathscr{C}(\xi_j\otimes\eta_1)=\xi_j\otimes\eta_j, \tag{4.4.1}$$

where η_1 is a fixed vector in $\mathcal{H}_P$, i.e. the initial state of the probe system. What we will now show is that the quadruple $\mathcal{M}=\langle\mathcal{H}_P,\mathscr{C},\eta_1,E\rangle$ defines a measurement model, known as the Von Neumann measurement model. In order to show this, we must verify the probability compatibility condition (4.2.1). We start by considering an initial state vector for the measured system of the form $\rho=\sum_{i,k}c_i\bar{c}_k|\xi_i\rangle\langle\xi_k|$, such that the initial coupled state is $\rho\otimes|\eta_1\rangle\langle\eta_1|$. The action of the channel $\mathscr{C}$ on this state produces the state vector

$$\mathscr{C}(\rho\otimes|\eta_1\rangle\langle\eta_1|)=\sum_{i,k}c_i\bar{c}_k|\xi_i\rangle\langle\xi_k|\otimes|\eta_i\rangle\langle\eta_k|. \tag{4.4.2}$$

We are now ready to prove the condition given in (4.2.1). In fact on one hand we have

$$tr[V(j)\rho]=tr_{\mathcal{H}}\left[|\xi_j\rangle\langle\xi_j|\sum_{i,k}c_i\bar{c}_k|\xi_i\rangle\langle\xi_k|\right]=tr_{\mathcal{H}}\left[\sum_{i,k}c_i\bar{c}_k|\xi_i\rangle\langle\xi_j|\delta_{kj}\right]=$$
$$tr_{\mathcal{H}}\left[\sum_i c_i\bar{c}_j|\xi_i\rangle\langle\xi_j|\right]=|c_j|^2.$$

On the other hand we find

$$tr_{\mathcal{H}\otimes\mathcal{H}_P}[\mathscr{C}(\rho\otimes|\eta_1\rangle\langle\eta_1|)(I\otimes E(j))]=tr_{\mathcal{H}_P}\left[tr_{\mathcal{H}}\left[\sum_{i,k}c_i\bar{c}_k|\xi_i\rangle\langle\xi_k|\otimes|\eta_i\rangle\langle\eta_k|\right]E(j)\right]=$$
$$tr_{\mathcal{H}_P}\left[\left(\sum_i|c_i|^2|\eta_i\rangle\langle\eta_i|\right)|\eta_j\rangle\langle\eta_j|\right]=tr_{\mathcal{H}_P}[|c_j|^2|\eta_j\rangle\langle\eta_j|]=|c_j|^2,$$

from which we can conclude that $\mathcal{M}=\langle\mathcal{H}_P,\mathscr{C},\eta_1,E\rangle$ represents a valid measurement model for V.

Further, we can also find the instrument associated with this measurement model using the defining condition eq.(4.3.3). As we will see, it represents the prototype of a class of quantum instruments known as Luders instruments [8].

$$\mathscr{I}_l^{\mathcal{M}}(\rho)=tr_{\mathcal{H}_P}\left[\left(\sum_{i,k}c_i\bar{c}_k|\xi_i\rangle\langle\xi_k|\otimes|\eta_i\rangle\langle\eta_k|\right)(I\otimes|\eta_l\rangle\langle\eta_l|)\right]=$$
$$tr_{\mathcal{H}_P}\left[\sum_{i,k}c_i\bar{c}_k|\xi_i\rangle\langle\xi_k|\otimes|\eta_i\rangle\langle\eta_l|\delta_{kl}\right]=tr_{\mathcal{H}_P}\left[\sum_i c_i\bar{c}_l|\xi_i\rangle\langle\xi_l|\otimes|\eta_i\rangle\langle\eta_l|\right]=$$
$$\sum_i c_i\bar{c}_l|\xi_i\rangle\langle\xi_l|\delta_{il}=|c_l|^2|\xi_l\rangle\langle\xi_l|=tr[V(l)\rho]V(l).$$

Further, since $V(l)$ are one-dimensional projections, we can express the instrument associated with Von Neumann's measurement model as

$$\mathscr{I}_l^{\mathcal{M}}(\rho) = V(l)\rho V(l). \tag{4.4.3}$$

This instrument is known as the Luders instrument of V. More precisely, the quantum instrument in eq. (4.4.3) is only a particular type of Luders instrument, that, indeed, is defined in the following way [8]:

Definition 4.4.1. The Luders instrument $\mathscr{I}_x^L$ associated with a discrete observable T is defined as

$$\mathscr{I}_x^L(\rho) \coloneqq T(x)^{1/2}\rho T(x)^{1/2}. \tag{4.4.4}$$

We observe also that the Luders instrument thus defined is T-compatible. Indeed

$$tr[\mathscr{I}_x^L(\rho)] = tr[T(x)^{1/2}\rho T(x)^{1/2}] = tr[\rho T(x)], \tag{4.4.5}$$

from which we can conclude that this instrument is a T-compatible instrument.

Remark 4.4.1. If a system is in a pure state $|\xi_i\rangle\langle\xi_i|$, where $\{\xi_i\}_i$ is a basis for the Hilbert space $\mathcal{H}$, and A is the associated sharp observable $A(j) = |\xi_j\rangle\langle\xi_j|$, we observe that $A(i)|\xi_i\rangle\langle\xi_i|A(i) = |\xi_i\rangle\langle\xi_i|$. Then, the Luders instrument associated with A satisfies the condition

$$tr[\rho A(j)] = 1 \Rightarrow A(j)\rho A(j) = \rho, \tag{4.4.6}$$

for all j, and $\rho \in \mathcal{S}(\mathcal{H})$.
So what we observe is that if the probability of an outcome is 1, the corresponding state is not disturbed by the A-compatible Luders instrument. This feature is called ideality [8] and it can be shown that every instrument associated with a discrete sharp observable, which satisfies this condition, is of Luders type.

4.5 Covariant quantum instruments and covariant POVMs

In this last section, we will deal with the role played by symmetries and conservation laws in the quantum measurement of a physical system. In particular, we will see that the conservation laws put constraints on the possible measurements that can be performed on a physical system [8] and similarly, by requiring that a POVM or a quantum instrument is covariant under the action of a symmetry group, in a way that we will specify, permits to restrict the class of possible POVMs/quantum

instruments that can be considered for the given physical system [36, 38, 37].
In what follows, we briefly discuss, in a first subsection, the restrictions imposed by a conserved quantity on a quantum measurement process. In particular, we will state and prove under simplified hypotheses a theorem known as the **Wigner-Araki-Yanase** theorem [8]. Then, we left a second subsection for a concise discussion on the concepts of covariant POVMs and covariant quantum instruments.

Wigner-Araki-Yanase theorem

The Wigner-Araki-Yanase theorem is based on the observation due to Wigner that a conservation law puts constraints on the possible measurements that can be performed on a physical system. This observation was then transformed into a theorem due to Araki and Yanase. We will discuss this theorem, following [8], in the particular case in which the measurement model considered is normal and repeatable. However, it is also possible to extend the validity of this theorem under more general hypotheses [9].
We begin by defining the concept of a normal measurement model [8]:

Definition 4.5.1. A measurement model $\mathcal{M} = \langle \mathcal{H}_P, \sigma, \mathscr{C}, Z \rangle$ is said to be normal if the following conditions are satisfied:

1) The pointer observable is a sharp nondegenerate observable associated with an orthonormal basis $\{\psi_i\}_i$ of the Hilbert space $\mathcal{H}_P$,

2) the initial state of the probe system σ is a pure state $\sigma = |\psi_0\rangle\langle\psi_0|$,

3) the quantum channel $\mathscr{C}$ is a unitary quantum channel, i.e. its action is described by a unitary operator U on $\mathcal{H} \otimes \mathcal{H}_P$.

Clearly, an example of a normal measurement model is the Von Neumann measurement model discussed in the previous section.
We pass then to briefly introduce the concept of repeatability for quantum instruments. The main idea is to consider an instrument such that it produces always the same outcome state when performed repeatedly on the system. In other terms, this means that a repeatable instrument makes the measurement outcome in further measurements completely predictable [8]. This, in turn implies that if A is a sharp observable, and $\mathscr{I}$ is a compatible instrument for it, it is repeatable if $tr[A(x)\tilde{\rho}_x] = 1$, where $\tilde{\rho}_x$ is the conditioned output state corresponding to the outcome x. Summing up what we have said so far, we can then give the following [8]

Definition 4.5.2. An instrument $\mathscr{I}$ is said to be repeatable if $tr[\tilde{\rho}_x A(x)] = 1$ holds, where $\mathscr{I}$ is a A-compatible instrument, A is a sharp observable and $\tilde{\rho}_x$ is the conditional output state of the A measurement.

Usually, one says that a measurement model $\mathcal{M}$ is repeatable if so is the induced instrument $\mathscr{J}^{\mathcal{M}}$. We do not go further in the discussion of repeatable instruments, and refer to the literature [8, 10] for further details on this argument.
We pass now to define the concept of a conserved quantity for a physical system in a state ρ:

Definition 4.5.3. A bounded selfadjoint operator R acting on $\mathcal{H} \otimes \mathcal{H}_P$ is called a conserved quantity if

$$tr[\mathscr{C}(\rho)R] = tr[\rho R] \tag{4.5.1}$$

for all states ρ, and where $\rho \mapsto \mathscr{C}(\rho)$ is the channel describing the interaction between the probe and the system during the measurement process.
Moreover, if the operator R can be written as

$$R = R_1 \otimes I + I \otimes R_2,$$

then R is said to be an additive conserved quantity [8].

In order to prove the Wigner-Araki-Yanase theorem, we need another technical tool that is discussed in the following proposition, whose proof can be found in [8] (see proposition 5.24 of chapter 5):

Proposition 4.5.1. *Let $\mathcal{M}$ be a normal measurement model for an observable A. Then there exists a set of vectors $\{\tilde{\phi}_j\}_j$ of $\mathcal{H}$ such that*

$$U(\phi_j \otimes \psi_0) = \tilde{\phi}_j \otimes \psi_j. \tag{4.5.2}$$

If the measurement model $\mathcal{M}$ is repeatable, then the vectors $\tilde{\phi}_j$ are orthogonal.

We are now ready to state the Wigner-Araki-Yanase theorem:

Theorem 4.5.1. *(Wigner-Araki-Yanase)*
Let $\mathcal{M}$ be a normal measurement model for a sharp observable B, and let R be an additive conserved quantity. If $\mathcal{M}$ is repeatable than it must be true that $[R_1, B(i)] = 0$ for every i in the sample space of the sharp observable.

Proof. We assume that B is a sharp observable such that $B(i) = |\phi_i\rangle\langle\phi_i|$ for $\{\phi_i\}_i$ an orthonormal basis of $\mathcal{H}$. Further we consider a state $\rho \in \mathcal{S}(\mathcal{H})$ of the system written as $\rho = \sum_{i,j} c_i c_j^* |\phi_i\rangle\langle\phi_j|$, while we consider the initial state of the probe to be $|\psi_0\rangle\langle\psi_0|$. Then, since R is a conserved quantity it must be true that

$$tr[(\rho \otimes |\psi_0\rangle\langle\psi_0|)R] = tr[U(\rho \otimes |\psi_0\rangle\langle\psi_0|)U^* R]. \tag{4.5.3}$$

Expanding the trace on the left hand side of eq.(4.5.3) we find, for $i \neq j$, the term

$$\langle \phi_i \otimes \psi_0 | R(\phi_j \otimes \psi_0) \rangle = \langle \phi_i \otimes \psi_0 | (R_1 \otimes I + I \otimes R_2)(\phi_j \otimes \psi_0) \rangle = \\ \langle \phi_i | R_1 \phi_j \rangle + \langle \phi_i | \phi_j \rangle \langle \psi_0 | R_2 \psi_0 \rangle = \langle \phi_i | R_1 \phi_j \rangle.$$

The term on the right hand side of eq.(4.5.3) gives instead

$$\langle U(\phi_i \otimes \psi_0) | RU(\phi_j \otimes \psi_0) \rangle = \langle \tilde{\phi}_i \otimes \psi_i | R(\tilde{\phi}_j \otimes \psi_j) \rangle \\ \langle \tilde{\phi}_i | \tilde{\phi}_j \rangle \langle \psi_i | R_2 \psi_j \rangle + \langle \psi_i | \psi_j \rangle \langle \tilde{\phi}_i | R_1 \tilde{\phi}_j \rangle = 0.$$

We can then conclude that $\langle \phi_i | R_1 \phi_j \rangle = \delta_{ij} \langle \phi_i | R_1 \phi_j \rangle$. If we now consider

$$\langle \phi_k | [R_1, B(i)] \phi_l \rangle = \langle \phi_k | (R_1 | \phi_i \rangle \langle \phi_i | - | \phi_i \rangle \langle \phi_i | R_1) \phi_l \rangle = \\ \langle \phi_k | R_1 \phi_i \rangle \delta_{il} - \langle \phi_i | R_1 \phi_l \rangle \delta_{ki} = \langle \phi_k | R_1 \phi_i \rangle \delta_{il} \delta_{ki} - \langle \phi_i | R_1 \phi_l \rangle \delta_{il} \delta_{ki} = 0,$$

from which we conclude that $[R_1, B(i)] = 0$ for all i □

So we observe that the Wigner-Araki-Yanase theorem imposes that only sharp observables that commute with R can be measured on the system, if R is a conserved quantity.

Covariant Instruments and covariant POVMs

In the last subsection, we have seen the limitations that a conserved quantity imposes on the measurement process. In this subsection, we will see that the requirement to have a POVM or a quantum instrument that is covariant under the action of a given symmetry group, in a sense that will be soon clarified, selects a specific class of POVMs or quantum instruments that satisfy this condition. In what follows, we start by defining the concept of covariance for a POVM stating in particular a theorem (see theorem 4.5.3) that characterizes, under suitable hypotheses, the mathematical form of covariant POVMs. Then, we concentrate on covariant instruments repeating a similar analysis and giving at the end an analogous characterization theorem (theorem 4.5.5).

The analysis of covariant POVMs can be carried out under very general hypotheses. One can start, in fact, by considering a locally compact second countable Hausdorff topological group G (briefly a l.c.s.c. group), assumed to be in general non-unimodular [7], that acts both on the sample space of the POVM and on the Hilbert space of the physical system, by means of a continuous (left)-action and a strongly continuous unitary representation respectively. Then, the necessary and sufficient condition to have a covariant POVM is that the representation is square

integrable [36, 38]. However, in our treatment, we will restrict to the case in which the group G is assumed to be a l.c.s.c. unimodular group. Here, for the moment, we start by defining the condition of covariance [36] for a POVM:

Definition 4.5.4. Let G be a group that acts on $\mathcal{H}$ by means of a strongly continuous unitary representation π, and on a topological space M through a (left)-action α_g. A POVM Z with sample space $(M, \mathcal{B}(M))$ is said to be covariant with respect to G if for all $g \in G$ the following condition holds:

$$\pi(g)Z(X)\pi(g^{-1}) = Z(\alpha_g(X)), \tag{4.5.4}$$

for all $X \in \mathcal{B}(M)$.

In the discussion that follows we will consider POVMs whose outcome space is defined on the group itself. Thus, the definition of covariant POVMs can be rephrased in the following terms:

Definition 4.5.5. Let G be a locally compact topological group. Let $\mathcal{B}(G)$ denotes the Borel σ-algebra defined over G. Then, a positive operator valued measure Z defined over G that satisfies:

i) $Z(G) = I$,

ii) for all $X \in \mathcal{B}(G)$

$$\pi(g)Z(X)\pi(g^{-1}) = Z(g[X]) \qquad \forall g \in G, \tag{4.5.5}$$

where π is a strongly continuous unitary representation of G on $\mathcal{H}$ and $g[X]$ denotes the (left)-action of $g \in G$ on $X \in \mathcal{B}(G)$, is called a π-covariant POVM.

Remark 4.5.1. The definition of a covariant POVM can be seen as a generalization of the concept of imprimitivity system [7]. Indeed, given a locally compact group G, the triple $\Sigma = (\pi, S, P)$ where

i) π is a unitary representation of G on a Hilbert space $\mathcal{H}$,

ii) S is a G-space, i.e. a locally compact Hausdorff space equipped with a continuous left G-action,

iii) P is a projection valued measure defined on S and satisfying the condition

$$\pi(g)P(X)\pi(g)^{-1} = P(gX) \tag{4.5.6}$$

for all $g \in G$ and $X \in \mathcal{B}(S)$,

is called a system of imprimitivity on G, and eq.(4.5.6) is known as the imprimitivity condition.
There are important physical examples of imprimitivity systems [3]. Indeed, the triple formed by $\mathbb{R}^3$, the position observable P_S^X and the projective unitary representation of the group of isometries of $\mathbb{R}^3$, i.e. the group formed by the semidirect product between the abelian group of translations and the group $O(3)$, constitutes an imprimitivity system ([3], chapter 12 example 12.17). So we see that eq.(4.5.4) can be seen as a generalization of the imprimitivity condition to POVMs, i.e. to more general measurements that can be performed on a system, as discussed in chapter 2.

It is now natural to ask what is the form that a POVM must have to be covariant, i.e. so that eq.(4.5.4) is satisfied. First of all, we restrict our discussion to the case in which the group considered is a l.c.s.c. unimodular group. The locally compact hypothesis assures the existence of a left Haar measure λ_G (uniquely defined up to a multiplicative positive constant) for the group G (see [7] theorem 2.10), while the condition to be unimodular means that the so called modular function Δ is identically equal to 1, i.e. $\Delta \equiv 1$. Moreover, this also implies that every Haar measure is both left and right invariant. Further, we will assume that G admits a strongly continuous irreducible unitary representation π over a separable complex Hilbert space $\mathcal{H}$.
In order to clarify some of the concepts that we will soon introduce, we prefer to remember here the notion of a square integrable representation [39].
Given a locally compact group G, and a strongly continuous irreducible unitary representation $\pi : G \to \mathcal{U}(\mathcal{H})$, we can define the so called matrix elements of the representation:

$$c^{\pi}_{u,v} : G \ni g \mapsto \langle \pi(g)u, v\rangle \in \mathbb{C} \tag{4.5.7}$$

for $u, v \in \mathcal{H}$ and where $\langle\ ,\ \rangle$ is the scalar product on $\mathcal{H}$ linear in the second argument. The matrix coefficients provide a map sending elements of the group G in $\mathbb{C}$. We are interested in the case in which such maps are elements of $L^2(G, \lambda_G, \mathbb{C})$, i.e. they are square integrable with respect to the left Haar measure λ_G. For this reason, we introduce the set $\mathcal{A}(\pi)$, known as the set of the admissible vectors for the representation π, defined as

$$\mathcal{A}(\pi) \coloneqq \{u \in \mathcal{H} |\ \exists v \in \mathcal{H}, v \neq 0\ :\ c^{\pi}_{u,v} \in L^2(G, \lambda_G, \mathbb{C})\} \tag{4.5.8}$$

In particular, this set is always non-empty because $0 \in \mathcal{A}(\pi)$. We can then make the following

Definition 4.5.6. A unitary irreducible representation π is said to be square integrable if $\mathcal{A}(\pi) \neq \{0\}$.

The main results concerning the theory of square integrable representations are summarized in the following theorem due to Duflo and Moore [40]:

Theorem 4.5.2. *Let G be l.c.s.c. group and let the strongly continuous irreducible unitary representation π of the group G in the Hilbert space $\mathcal{H}$ be square integrable. Then, the set $\mathcal{A}(\pi)$ is a dense linear subspace in $\mathcal{H}$. For every vector $u \in \mathcal{A}(\pi)$ and $v \in \mathcal{H}$ the coefficient of the representation $c^{\pi}_{u,v}$ is square integrable with respect to the left Haar measure λ_G on G.*
Further, there exists a unique positive selfadjoint injective linear operator K_π in $\mathcal{H}$ such that

$$Dom(K_\pi) = \mathcal{A}(\pi) \tag{4.5.9}$$

and such that

$$\int_G d\lambda_G(g)\langle v_1, \pi(g)u_1\rangle\langle \pi(g)u_2, v_2\rangle = \int_G d\lambda_G(g) c^{\pi}_{u_1,v_1}(g)^* c^{\pi}_{u_2,v_2}(g) = \langle v_1, v_2\rangle\langle K_\pi u_2, K_\pi u_1\rangle, \tag{4.5.10}$$

for every $v_1, v_2 \in \mathcal{H}$ and $u_1, u_2 \in \mathcal{A}(\pi) = Dom(K_\pi)$.

We can give some comments on this theorem which will be useful in the following:

1) The operator K_π, usually known as the Duflo Moore operator, is in general unbounded. However, it can be shown that it is bounded if and only if the group G is unimodular. Further, in this case, K_π is a multiple of the identity $I_{\mathcal{H}}$, i.e. $K_\pi = k_\pi I_{\mathcal{H}}$, for $k_\pi > 0$.

2) For a compact group G, it is known that every strongly continuous irreducible unitary representation is square integrable. Further, compact group are unimodular. From this, one can conclude that the Duflo Moore operator is a multiple of the identity, in particular that $K_\pi = d_\pi^{-1/2} I_{\mathcal{H}}$ [39], where d_π is the dimension of the Hilbert space of the representation π (always finite dimensional for an irreducible representation of a compact group). Lastly, the relation in eq.(4.5.10) reduces to the Schur orthogonality relations [39].

3) For every $u \in \mathcal{A}(\pi)$ $u \neq 0$, we can define the so called generalized Wavelet transform with analyzing vector u:

$$\mathcal{W}^{\pi}_u : \mathcal{H} \ni v \mapsto \|K_\pi u\|^{-1}_{L^2} c^{\pi}_{u,v} \in L^2(G, \lambda_G). \tag{4.5.11}$$

More precisely, it can be shown that $\mathcal{W}^{\pi}_u$ is a linear operator which is in particular an isometry.

Having described the main results concerning square integrable representations, we are now ready to state the characterization theorem for the so called π-covariant POVMs:

Theorem 4.5.3. *Let G be a l.c.s.c. unimodular group, and let π be a square integrable representation for it. Let T be a trace one positive trace class operator in $\mathcal{T}(\mathcal{H})$. Then the map*

$$\mathcal{B}(G) \ni X \mapsto Z_T(X) \coloneqq \int_X \pi(g) T \pi(g^{-1}) d\lambda_G(g), \tag{4.5.12}$$

for a suitable normalization of λ_G, defines a POVM Z_T on G that is covariant with respect to π. Conversely, if Z is a POVM on G covariant with respect to π, then π is square integrable and there exists a positive trace class trace one operator T such that $Z = Z_T$.

Remark 4.5.2. If the group G is compact, it suffices to consider a strongly continuous irreducible unitary representation to construct a covariant POVM.
In this case, we also observe that if the trace one positive operator T is a pure state, i.e. $T = |\xi\rangle\langle\xi|$ for $\xi \in \mathcal{H}$ and $\|\xi\| = 1$, eq.(4.5.12) can be rewritten in terms of the wavelet transform. Indeed we have

$$\begin{aligned}\langle Z_T(X)u, v\rangle = d_\pi \int_X \langle \pi(g^{-1})u, \xi\rangle\langle \xi, \pi(g^{-1})v\rangle d\lambda_G(g) = \\ d_\pi \int_X \langle u, \pi(g)\xi\rangle\langle \pi(g)\xi, v\rangle d\lambda_G(g) = \\ \int_X \overline{(W_\xi u)(g)}(W_\xi v)(g) d\lambda_G(g),\end{aligned}$$

for all $u, v \in \mathcal{H}$.

Theorem 4.5.3 gives the characterization of covariant POVMs under the hypotheses that the the group G is locally compact and unimodular and its representation π is square integrable. It is possible to extend this theorem also under more general hypotheses, considering groups that are non unimodular, or by requiring the square integrability of the representation only on a suitable quotient space instead of on the whole group G [38]. In these cases, what one observes is that covariant POVMs have a form similar to that of eq.(4.5.12), but with the explicit presence of the Duflo Moore operator.

An important example: the Galilei group

We discuss the important example of the POVMs covariant under the action of the Galilei group, obtaining in this way the so called **phase space measurements** [36]. The importance of this example lies on the fact that the Galilei group, as it is known [41], is the symmetry group of classical mechanics and of non relativistic quantum mechanics. We start by remembering some basic facts about this group, focusing in particular on the isochronous Galilei group.
The isochronous Galilei group is the group $\mathcal{G} = (\mathbb{R}^3 \times \mathbb{V}^3) \times' SO(3)$, with $\mathbb{R}^3$ the group of translations, $\mathbb{V}^3$ the three dimensional group of Galilean boosts and $SO(3)$ the group of rotations, whose elements are the triple $(\vec{c}, \vec{v}, \mathbf{R})$ subjected to the product law:

$$(\vec{c}_2, \vec{v}_2, \mathbf{R}_2)(\vec{c}_1, \vec{v}_1, \mathbf{R}_1) = (\vec{c}_2 + \mathbf{R}_2\vec{c}_1, \vec{v}_2 + \mathbf{R}_2\vec{v}_1, \mathbf{R}_2\mathbf{R}_1).$$

Further, we can consider the action of the group on an element $(\vec{q}, \vec{p})$ of the phase space $\mathcal{F} = \mathbb{R}^3 \times \mathbb{P}^3$ given by:

$$(\vec{c}, \vec{v}, \mathbf{R})(\vec{q}, \vec{p}) = (\vec{c} + \mathbf{R}\vec{q}, m\vec{v} + \mathbf{R}\vec{p}),$$

where m denotes the mass of a non relativistic particle. The action of $\mathcal{G}$ on $\mathcal{F}$ is transitive, and by considering also the fact that the stability subgroup at the origin $(\vec{0}, \vec{0})$ is the compact group $SO(3)$, one can conclude [36] that $\mathcal{F} \cong \mathcal{G}/SO(3)$.
The Hilbert space of a spinless non relativistic particles of mass m is the space of functions $L^2(\mathbb{R}^3, d\vec{x})$, and $\mathcal{G}$ acts on it by means of a irreducible projective unitary representation.
We remember that given a l.c.s.c. group G, a projective representation [41] in a separable complex Hilbert space $\mathcal{H}$ is a map $U : G \to \mathcal{U}(\mathcal{H})$ such that:

i) U is a weakly Borel map, i.e. $g \mapsto \langle v, U(g)w\rangle \in \mathbb{C}$ is a Borel map for every $v, w \in \mathcal{H}$,

ii) $U(e) = I$ for e the unit element in G,

iii) $U(g_1)U(g_2) = \mu(g_1, g_2)U(g_1g_2)$,

where the map $\mu : G \times G \to \mathbb{T}$ is a Borel map called the multiplier of the representation, for $\mathbb{T} = \{z \in \mathbb{C} |\ |z| = 1\}$ the circle group (a multiplier is more in general defined as a map $\mu : G \times G \to \mathbb{A}$, for $\mathbb{A}$ an abelian group. Then μ is called an $\mathbb{A}$-multiplier). The multiplier μ associated with the representation U satisfies the following conditions:

$$\mu(g_1, g_2g_3)\mu(g_2, g_3) = \mu(g_1g_2, g_3)\mu(g_1, g_2), \qquad \forall g_1, g_2, g_3 \in G, \tag{4.5.13}$$

and

$$\mu(g,e) = \mu(e,g) = 1, \ \ \forall g \in G. \tag{4.5.14}$$

Further, two multipliers μ_1, μ_2 for a group G are said to be similar, and we write $\mu_1 \sim \mu_2$, if there exists a Borel function $\beta : G \to \mathbb{T}$ such that

$$\mu_1(g_1,g_2) = \beta(g_1,g_2)\beta(g_1)^{-1}\beta(g_2)^{-1}\mu_2(g_1,g_2) \ \ \forall g_1,g_2 \in G. \tag{4.5.15}$$

We say that μ is exact if $\mu \sim 1$, i.e. if there exists a map $\beta : G \to \mathbb{T}$ such that

$$\mu(g_1,g_2) = \beta(g_1,g_2)\beta(g_1)^{-1}\beta(g_2)^{-1}. \tag{4.5.16}$$

Returning now to the case of our interest, the isochronous Galilei group $\mathcal{G}$ acts on the Hilbert space $L^2(\mathbb{R}^3, d\vec{x})$ by means of an irreducible projective unitary representation $U(\vec{c},\vec{v},\mathbf{R})$ as

$$(U(\vec{c},\vec{v},\mathbf{R})\psi)(x) = e^{im\vec{v}\cdot(\vec{x}-\vec{c})}\psi(\mathbf{R}^{-1}(\vec{x}-\vec{c})),$$

for every $\psi \in L^2(\mathbb{R}^3, d\vec{x})$.

Having described some features of the isochronous Galilei group $\mathcal{G}$, we turn now our attention to the construction of the phase space measurements. Since we are interested in POVMs defined on $\mathcal{F}$ and covariant under the action of $\mathcal{G}$, we can consider POVMs with outcome space defined on $\mathcal{G}/SO(3)$. Clearly, the definition 4.5.4 can be extended to the case in which the POVM is defined on G/H for H a subgroup of G, with the only difference that now the action of $g \in G$ on $\tilde{a} \in G/H$ is given by $g[\tilde{a}] = \tilde{g}\tilde{a}$. Also theorem 4.5.3 can be extended, and one can show [38] that the map

$$\mathcal{B}(G/H) \ni X \mapsto Z_T(X) \coloneqq \int_X U_g T U_g^* d\lambda_{G/H}(\tilde{g}), \tag{4.5.17}$$

for a suitable positive trace one trace class operator T, is a U-covariant POVM if and only if U is a square integrable representation.

Then, returning to the case of our interest we can construct covariant POVMs under the action of the Galilei group using eq.(4.5.17). Clearly in this case, $G/H = \mathcal{G}/SO(3) \cong \mathcal{F}$, while we can consider the invariant measure $d\lambda_{\mathcal{G}/SO(3)} = \frac{m}{(2\pi)^3} d\vec{c}\, d\vec{v}$ on $\mathcal{G}/SO(3)$. What we must observe is if the representation $U(\vec{c},\vec{v},\mathbf{R})$ is square integrable, and what are the additional conditions that must be imposed on the trace class trace one operator T in order to satisfy the covariance condition.

About the second question, we observe that for a l.c.s.c. group G and every compact subgroup H the following measure decomposition holds:

$$\int_G f(g) d\lambda_G(g) = \int_{G/H} d\lambda_{G/H}(\tilde{g}) \int_H f(gh) d\lambda_H(h), \tag{4.5.18}$$

for $f \in L^1(G, \lambda_G)$, and where $\lambda_{G/H}$ and λ_H are invariant measures on G/H and H respectively. By using this measure decomposition, and eq.(4.5.17) it is easily shown [36] that, in the case G is the Galilei group and H is the compact subgroup $SO(3)$, the trace one trace class operator appearing in eq.(4.5.17) must be rotation invariant, i.e. it must be true that $TU(\vec{0}, \vec{0}, \mathbf{R}) = U(\vec{0}, \vec{0}, \mathbf{R})T$. So we arrive at the conclusion that the trace class trace one operator that can be used in the construction of a covariant POVM under the action of the Galilei group must be rotation invariant.

We pass now to the first question. We must verify that the representation $U(\vec{c}, \vec{v}, \mathbf{R})$ is square integrable. As we have said, $\mathcal{G}$ acts on $L^2(\mathbb{R}^3, d\vec{x})$ as an irreducible projective representation. In particular, one can find from a direct calculation that the multiplier of the representation is given by

$$\mu((\vec{c}_2, \vec{v}_2, \mathbf{R}_2)(\vec{c}_1, \vec{v}_1, \mathbf{R}_1)) = e^{im\vec{v}_2 \cdot \mathbf{R}_2 \vec{c}_1}. \tag{4.5.19}$$

that, moreover, has a direct physical significance since it depends on the mass m of the particle. However, one sees that the multiplier is not exact, i.e. does not exist a function $\beta : G \to \mathbb{T}$ such that $\mu \sim 1$.

Nevertheless, it is possible to extend $U(\vec{c}, \vec{v}, \mathbf{R})$ to a unitary representation on the central extension $\mathcal{G}_\mu$ of the Galilei group. If fact, it is known that given a l.c.s.c. group G one can define the so called central extension, that is the group $G_\mu = \mathbb{T} \times_\mu G$ endowed with the product law

$$(z_1, g_1)(z_2, g_2) = (z_1 z_2 \mu(g_1, g_2), g_1 g_2). \tag{4.5.20}$$

In particular, a theorem due to Mackey [41] shows that there exists a unique topology, called the Weil topology, respect to which G_μ is a l.c.s.c. group. Further, this is the only topology on G_μ that generates a Borel structure coinciding with the product Borel structure of $\mathbb{T} \times_\mu G$.

If G admits a projective representation, it is possible to extend it to a unitary representation on the extended group G_μ, i.e. if U is a projective representation then $\hat{U}$ defined as

$$\hat{U}(z, g) \coloneqq z^{-1} U(g) \tag{4.5.21}$$

is a unitary representation, that is irreducible if U is. Returning now to our case, we can consider the central extension $\mathcal{G}_\mu = \mathbb{T} \times_\mu \mathcal{G}$, with composition law

$$(z_1, \vec{c}_1, \vec{v}_1, \mathbf{R}_1)(z_2, \vec{c}_2, \vec{v}_2, \mathbf{R}_2) = (z_1 z_2 e^{im\vec{v}_2 \cdot \mathbf{R}_2 \vec{c}_1}, \vec{c}_2 + \mathbf{R}_2 \vec{c}_1, \vec{v}_2 + \mathbf{R}_2 \vec{v}_1, \mathbf{R}_2 \mathbf{R}_1).$$

Moreover, the irreducible projective representation U lifts to an irreducible unitary representation $\hat{U}$ of $\mathcal{G}_\mu$ given by

$$(\hat{U}(z, \vec{c}, \vec{v}, \mathbf{R})\psi)(x) = z^{-1} e^{im\vec{v} \cdot (\vec{x} - \vec{c})} \psi(\mathbf{R}^{-1}(\vec{x} - \vec{c})), \tag{4.5.22}$$

for $\psi \in L^2(\mathbb{R}^3, d\vec{x})$. Eventually, it can be shown that $\hat{U}$ is square integrable [36]. Now, by considering that U is square integrable if and only if $\hat{U}$ is, and that the phase space $\mathcal{F} \cong \mathcal{G}_\mu/\mathbb{T} \times SO(3)$, we arrive at the following characterization [36] of POVMs covariant under the action of $\mathcal{G}$, that sums up all we have said until now:

Theorem 4.5.4. *Let T be a positive trace class trace one operator such that*

$$TU(\vec{0}, \vec{0}, \mathbf{R}) = U(\vec{0}, \vec{0}, \mathbf{R})T \tag{4.5.23}$$

$\forall \mathbf{R} \in SO(3)$, and where $U(\vec{c}, \vec{v}, \mathbf{R})$ is the projective unitary representation of the Galilei group $\mathcal{G}$ defined by eq.(4.5.22). Then, the map

$$\mathcal{B}(\mathcal{F}) \ni X \mapsto \frac{1}{(2\pi)^3} \int_X U(\vec{c}, \frac{\vec{p}}{m}, I)TU(\vec{c}, \frac{\vec{p}}{m}, I)^* d\vec{c}\, d\vec{p}, \tag{4.5.24}$$

is a POVM on $\mathcal{F}$ covariant with respect to U. Conversely, every POVM on $\mathcal{F}$ covariant with respect to U is of the form in eq.(4.5.24), for a trace class trace one operator T invariant under rotations.

Covariant Quantum Instruments

Having described some peculiarities of the covariant POVMs, we turn now our attention to covariant quantum instruments. As in the case of a POVM, we will consider a unimodular locally compact second countable topological group G. We also assume that this group admits an irreducible unitary representation and that it acts continuously on the sample space $\mathcal{B}(\Omega)$ on which the instrument is defined. This requirement means that there exists a continuous mapping $G \times \Omega \ni (g, x) \mapsto gx \in \Omega$ such that for fixed g it represents a homeomorphism of Ω. Eventually we will use the alternative definition of quantum instruments given in remark 4.3.2.
At this point, we can give the following

Definition 4.5.7. Let G be a group and let $U : G \to \mathcal{U}(\mathcal{H})$ be a unitary representation. An instrument $\mathscr{I}$ is said to be covariant with respect the action of U if

$$U(g)^* \mathscr{I}_{gX}(T)U(g) = \mathscr{I}_X(U(g)^* TU(g)), \tag{4.5.25}$$

for every $X \in \mathcal{B}(\Omega)$, $g \in G$ and $T \in \mathcal{T}(\mathcal{H})$.

Remark 4.5.3. We observe that a U-covariant instrument induces always a U-covariant POVM. We know, in fact, that every instrument $\mathscr{I}$ induces an observable

$A^{\mathscr{I}}$, that is defined as $A^{\mathscr{I}}(X) = \mathscr{I}_X^*(I)$. Thus if $\mathscr{I}$ is a U-covariant instrument, the following relations hold:

$$tr[\rho A^{\mathscr{I}}(gX)] = tr[\mathscr{I}_{gX}(\rho)] = tr[U(g)\mathscr{I}_X(U(g)^*\rho U(g))U(g)^*] =$$
$$tr[U(g)^*\rho U(g)\mathscr{I}_X^*(I)] = tr[U(g)^*\rho U(g)A^{\mathscr{I}}(X)] = tr[\rho U(g)A^{\mathscr{I}}(X)U(g)^*],$$

for all $g \in G$, $X \in \mathcal{B}(\Omega)$ and $\rho \in \mathcal{S}(\mathcal{H})$. We can then conclude that

$$A^{\mathscr{I}}(gX) = U(g)A^{\mathscr{I}}(X)U(g)^*, \tag{4.5.26}$$

i.e. that the observable $A^{\mathscr{I}}$ is U-covariant.

As we have said, every instrument admits a measurement model $\mathcal{M}$ that induces it, as proved in Ozawa's theorem 4.3.1. We can then consider an analogous theorem for covariant instruments

Corollary 4.5.1. *(of theorem 4.3.1)*
Let $\mathscr{I}$ be a U-covariant instrument. Then, there exist a measurement model $\mathcal{M} = \langle \mathcal{H}_P, \sigma, E, \mathscr{C}\rangle$ and a unitary representation $U : G \to \mathcal{U}(\mathcal{H}_P)$ such that $\mathscr{I} = \mathscr{I}^{\mathcal{M}}$, and the pointer observable E satisfies the condition

$$U(g)E(X)U(g)^* = E(gX) \qquad \forall X \in \mathcal{B}(\Omega),\ \ g \in G \tag{4.5.27}$$

i.e. it is a U-covariant observable.

Similarly to what we have done for POVMs, we can state a theorem that characterizes the aspect of a covariant quantum instrument under the hypotheses of having a l.c.s.c. unimodular group [37]:

Theorem 4.5.5. *If G is a l.c.s.c. unimodular group and $\pi : G \to \mathcal{U}(\mathcal{H})$ is a strongly continuous irreducible unitary representation for it, there exists a π-covariant quantum instrument based on G if and only if π is square integrable. In this case, if $C \in \mathcal{H} \otimes \mathcal{H}^* \cong \mathcal{B}_2(\mathcal{H})$ has norm 1, where $\mathcal{B}_2(\mathcal{H})$ denotes the space of Hilbert-Schmidt operators, there exists a unique instrument $\mathscr{I}^C : \mathcal{T}(\mathcal{H}) \to \mathcal{M}(G, \mathcal{T}(\mathcal{H}))$, such that for every $T = |u\rangle\langle v| \in \mathcal{T}(\mathcal{H})$ for $u, v \in \mathcal{H}$*

$$(\mathscr{I}^C T)(X) = \int_X |\pi(g)C\pi(g)^*u\rangle\langle\pi(g)C\pi(g)^*v| d\lambda_G(g) \qquad \forall X \in \mathcal{B}(G), \tag{4.5.28}$$

and it is π-covariant.

Remark 4.5.4. Likewise to what we have said in remark 4.5.2, we can conclude that if the group G is compact, it suffices that the representation is unitary and irreducible for eq. 4.5.28 to give a covariant instrument.

We conclude this subsection, discussing an interesting generalization of theorem 4.5.5 to the case of projective representations, that, as it is known [3, 41], are of fundamental importance in physics.

Corollary 4.5.2. *(of theorem 4.5.5)*
Let G be a l.c.s.c group and let U be an irreducible projective representation. Then, there exists a U-covariant quantum instrument based on G if and only if U is square integrable. In this case, if $C \in \mathcal{H} \otimes \mathcal{H}^* \cong \mathcal{B}_2(\mathcal{H})$ has norm 1, there exists a unique instrument $\mathscr{I}^C : \mathcal{T}(\mathcal{H}) \to \mathcal{M}(G, \mathcal{T}(\mathcal{H}))$ such that for $T = |u\rangle\langle v| \in \mathcal{T}(\mathcal{H})$, with $u, v \in \mathcal{H}$

$$(\mathscr{I}^C)(X) = \int_X |U(g)CU(g)^* u\rangle\langle U(g)CU(g)^* v| d\lambda_G(g) \qquad \forall X \in \mathcal{B}(G), \quad (4.5.29)$$

and it is U-covariant.

Conclusions

The main aim of this thesis work has been to introduce the concepts of measurement model and quantum instrument that play a key role in the modern theory of quantum measurement. To introduce these notions, we have preliminary discussed the concepts of positive operator-valued measure (chapter 2), and of quantum channel and operation (chapter 3). Each of these topics plays an important role in information theory and has several applications. However, we tried to select those aspects that seemed most relevant and useful in view of their application to the concepts of quantum instrument and measurement model developed in the last chapter. This has inevitably led in some cases to the omission of a series of important results, and in others to a partial discussion. Interesting concepts such as the decoherence or the conceptual aspects of the quantum measurement problem have not been treated, while others such as symmetry transformations or the reconstruction of quantum states have only been partially discussed.

The principle that we have tried to follow in the choice of the topics covered is to describe a picture, as precise and detailed as possible, of the different aspects that come into play in the description of a measurement process in quantum mechanics. In doing this, we started by the assumption that every experiment can be divided into a preparation and a measurement part. For what concerns the preparation part, we have introduced the concept of quantum state, whose mathematical character we have derived from the Gleason theorem. This choice, on the one hand, allowed us to directly characterize quantum states as positive trace class trace one operators, and on the other shows how states behave as positive functionals on the set of projection operators $\mathcal{P}(\mathcal{H})$. For the measurement part, we have gradually introduced the concepts of quantum effect, POVM, operation and quantum channel. Each of these objects describes a certain aspect of the measurement process. More

precisely, effects and POVMs provide us with information about the probabilities of the measurement outcomes, while quantum operations give information on the output states. Quantum channels concern, eventually, the deterministic transmission of an input state. These different aspects of the measurement process are described by means of a unique object represented by the notion of quantum instrument. A quantum instrument in fact, on the one hand summarizes in itself the concepts of operation and quantum channel, while on the other, it uniquely induces a POVM. Eventually, it shows how a measurement process can be used to prepare a quantum state suitably, thus providing a conditional state preparation process.
Further, a slightly more general concept, namely that of the quantum measurement model, which incorporates instruments, represents the most complete description that can be given of a measurement process since, in fact, it takes into consideration the interaction between the measured system and the measuring apparatus. More precisely, it is described by a quadruple $\mathcal{M} = \langle \mathcal{H}_P, \sigma, \mathscr{C}, Z \rangle$, where $\mathcal{H}_P$ is the Hilbert space of the probe system, σ is its initial state, $\mathscr{C}$ the channels describing the interaction between the open system and the probe, and Z is the pointer observable, i.e. the observable implemented on the probe and assumed to reproduce the measurement outcome probabilities of the initial observable A that we intend to measure. To conclude, the idea that led us from the concept of quantum effect to that of measurement model, was to describe processes that gave an increasingly richer and detailed description of the quantum measurement process. Of each notion introduced, an attempt has been made to highlight the physical idea behind it, but without neglecting the mathematical structure with which it is described. In this spirit, we have discussed the convex structure of the set of states and effects, finding their extremal elements. Further, a similar analysis was repeated also in the context of POVMs where we have shown the hypotheses that make the set of POVMs a convex one, finding then its extremal elements using the method of perturbations [13]. Where necessary, we have deepened the mathematical notions used in the various definitions, such as the discussion relating to the Neumark theorem, which highlights the link between PVMs and POVMs and at the same time offers the possibility of extending the concept of spectral decomposition also to symmetric operators. On completely positive functions, we have studied the general case of linear maps from a generic *-algebra in $\mathcal{L}(H)$, thus arriving at the Stinespring theorem treated in its most general form, and then particularizing it to the specific case of channels and quantum operations. Of each argument, we have tried to show the links with the concepts of which they represent a generalization, such as the relations between projections and effects, between the Gleason and Busch theorem, between self-adjoint operators and POVMs. Eventually, each chapter ends with a concrete application of the ideas developed, such as the qubit tomography, treated in

the last part of the second chapter, in which the idea of informational completeness of a set of POVMs finds an application or the discussion made in the conclusion of the third chapter on the generator of the dynamical semigroup for open quantum systems, that represents an important application of quantum channels.
An exception is represented by the last chapter, which ends with a discussion on a recent research topic represented by covariant POVMs and covariant quantum instruments. Its purpose is to show the role played by symmetries in a measurement process. In particular, theorems 4.5.3 and 4.5.5, show that the necessary and sufficient condition that must be satisfied in order to have a covariant POVM or a covariant quantum instrument is the square-integrability of the representation of the group. The mathematical setting that we have fixed for these theorems is that of a locally compact and unimodular group. Such a choice, on the one hand, was made in order not to overload the chapter with mathematical details, and on the other hand, because in these hypotheses we have discussed the important example of the POVMs covariant under the action of the Galilei group, finding the so called **phase space measurements** [36].
What we have tried to highlight with this thesis work is that although measurement models do not solve completely the quantum measurement problem or, more specifically, they do not answer the so-called *objectification problem* [4], they give a scheme as detailed and precise as possible of the different aspects that come into play in its description. This, eventually, allows us to precisely investigate the measurement problem (by changing, for example, a specific hypothesis or by making a suitable assumption) thus favoring an ever deeper understanding. We conclude with an observation on the possible continuations of this work. On the one hand, in fact, we have seen that one of the basic hypotheses in the definition of a measurement model is the unitary character of the interaction between the state of the object system and the probe. Recently, however, efforts are being made to describe measurement processes in which this assumption is released, thus allowing more general descriptions of the interaction.
On the other hand, another possible continuation of this work is offered by the investigation of the measurement problem in relativistic quantum mechanics. Indeed, in the relativistic domain, the quantum measurement theory shows new aspects that are not present in non-relativistic quantum mechanics [12]. In particular, the non-Lorentz covariant character of the standard collapse postulate, and the possibility of measuring and preparing non-local observables, i.e. to perform non-local measurements, give rise to a series of paradoxes concerning the measurement in relativistic quantum mechanics. Although some of these paradoxes have been solved [42], many problems concerning measurement in a relativistic quantum theory still remain open, especially in the characterization of observables that can be measured in the rela-

tivistic domain. In this context, the characterization of the observables, and more generally of the POVMs, covariant under the action of the Poincarè group, plays a central role. In fact, it is known that the Poincarè group does not possess square integrable representations [43]. This, therefore, implies that a characterization of the covariant POVMs under its action cannot be carried out similarly to what has been shown, in this work, for the Galilei group in the non-relativistic regime, thus making it necessary to introduce a new and different approach to this problem.

Bibliography

[1] J. Von Neumann, *Mathematical Foundations of Quantum Mechanics: New Edition*. Princeton university press, 2018.

[2] P. A. M. Dirac, *The principles of quantum mechanics*. No. 27, Oxford university press, 1981.

[3] V. Moretti, *Spectral theory and quantum mechanics*, vol. 64. Springer, 2013.

[4] P. Busch, P. Lahti, J.-P. Pellonpää, and K. Ylinen, *Quantum measurement*, vol. 22. Springer, 2016.

[5] R. Haag and D. Kastler, "An algebraic approach to quantum field theory," *Journal of Mathematical Physics*, vol. 5, no. 7, pp. 848–861, 1964.

[6] F. Strocchi, *An introduction to the mathematical structure of quantum mechanics: a short course for mathematicians*, vol. 28. World Scientific, 2008.

[7] G. B. Folland, *A course in abstract harmonic analysis*. Chapman and Hall/CRC, 2016.

[8] T. Heinosaari and M. Ziman, *The mathematical language of quantum theory: from uncertainty to entanglement*. Cambridge University Press, 2011.

[9] M. Ozawa, "Quantum measuring processes of continuous observables," *Journal of Mathematical Physics*, vol. 25, no. 1, pp. 79–87, 1984.

[10] P. Busch, M. Grabowski, and P. J. Lahti, *Operational quantum physics*, vol. 31. Springer Science & Business Media, 1997.

[11] M. M. Wilde, *Quantum information theory*. Cambridge University Press, 2013.

[12] H.-P. Breuer, F. Petruccione, *et al.*, *The theory of open quantum systems*. Oxford University Press on Demand, 2002.

[13] G. M. D'Ariano, P. L. Presti, and P. Perinotti, "Classical randomness in quantum measurements," *Journal of Physics A: Mathematical and General*, vol. 38, no. 26, p. 5979, 2005.

[14] P. Busch, "Quantum states and generalized observables: a simple proof of gleason's theorem," *Physical Review Letters*, vol. 91, no. 12, p. 120403, 2003.

[15] C. M. Caves, C. A. Fuchs, K. K. Manne, and J. M. Renes, "Gleason-type derivations of the quantum probability rule for generalized measurements," *Foundations of Physics*, vol. 34, no. 2, pp. 193–209, 2004.

[16] G. Cassinelli, E. De Vito, and A. Levrero, "On the decompositions of a quantum state," *Journal of Mathematical Analysis and Applications*, vol. 210, no. 2, pp. 472–483, 1997.

[17] M. Reed, *Methods of modern mathematical physics: Functional analysis*. Elsevier, 2012.

[18] R. Cooke, M. Keane, and W. Moran, "An elementary proof of gleason's theorem," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 98, pp. 117–128, Cambridge University Press, 1985.

[19] W. M. De Muynck, *Foundations of quantum mechanics, an empiricist approach*, vol. 127. Springer Science & Business Media, 2006.

[20] N. I. Akhiezer and I. M. Glazman, *Theory of linear operators in Hilbert space*. Courier Corporation, 2013.

[21] F. Riesz and S. Nagy, "B.(1990). functional analysis," *Dover Publications, Inc., New York. First published in*, vol. 3, no. 6, p. 35, 1955.

[22] G. B. Folland, *Real analysis: modern techniques and their applications*. John Wiley & Sons, 2013.

[23] G. M. D'Ariano, G. Chiribella, and P. Perinotti, *Quantum theory from first principles: an informational approach*. Cambridge University Press, 2017.

[24] A. Messiah, "Quantum mechanics, two volumes," 1999.

[25] E. Prugovečki, "Information-theoretical aspects of quantum measurement," *International Journal of Theoretical Physics*, vol. 16, no. 5, pp. 321–331, 1977.

[26] U. Fano, "Description of states in quantum mechanics by density matrix and operator techniques," *Reviews of Modern Physics*, vol. 29, no. 1, p. 74, 1957.

[27] M. Paris and J. Rehacek, *Quantum state estimation*, vol. 649. Springer Science & Business Media, 2004.

[28] G. Cassinelli, G. D'ariano, E. De Vito, and A. Levrero, "Group theoretical quantum tomography," *Journal of Mathematical Physics*, vol. 41, no. 12, pp. 7940–7951, 2000.

[29] W. Rudin, *Real and complex analysis*. Tata McGraw-hill education, 2006.

[30] E. B. Davies, *Quantum theory of open systems*. Academic Press, 1976.

[31] G. Lindblad, "On the generators of quantum dynamical semigroups," *Communications in Mathematical Physics*, vol. 48, no. 2, pp. 119–130, 1976.

[32] M. A. Nielsen and I. Chuang, "Quantum computation and quantum information," 2002.

[33] J. B. Conway, *A course in functional analysis*, vol. 96. Springer Science & Business Media, 2013.

[34] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, "Completely positive dynamical semigroups of n-level systems," *Journal of Mathematical Physics*, vol. 17, no. 5, pp. 821–825, 1976.

[35] C. Gardiner and P. Zoller, *Quantum noise: a handbook of Markovian and non-Markovian quantum stochastic methods with applications to quantum optics*, vol. 56. Springer Science & Business Media, 2004.

[36] C. Carmeli, G. Cassinelli, E. DeVito, A. Toigo, and B. Vacchini, "A complete characterization of phase space measurements," *arXiv preprint quant-ph/0405026*, 2004.

[37] C. Carmeli, T. Heinosaari, and A. Toigo, "Covariant quantum instruments," *Journal of Functional Analysis*, vol. 257, no. 11, pp. 3353–3374, 2009.

[38] G. Cassinelli, E. De Vito, and A. Toigo, "Positive operator valued measures covariant with respect to an irreducible representation," *Journal of mathematical physics*, vol. 44, no. 10, pp. 4768–4775, 2003.

[39] S. A. Gaal, *Linear analysis and representation theory*, vol. 198. Springer Science & Business Media, 2012.

[40] M. Duflo and C. C. Moore, "On the regular representation of a nonunimodular locally compact group," *Journal of functional analysis*, vol. 21, no. 2, pp. 209–243, 1976.

[41] V. Varadarajan, *Geometry of Quantum Theory*. Springer Science & Business Media, 2007.

[42] H.-P. Breuer and F. Petruccione, "State vector reduction in relativistic quantum mechanics: An introduction," in *Open Systems and Measurement in Relativistic Quantum Theory*, pp. 1–30, Springer, 1999.

[43] P. ANIELLO, "Extended wavelet transforms," *International Journal of Geometric Methods in Modern Physics*, vol. 3, no. 03, pp. 341–373, 2006.