# Pure and Mixed States

## J. C. A. Barata, M. Brum, V. Chabu & R. Correa da Silva

Springer

Springer

# Pure and Mixed States

J. C. A. Barata[1] · M. Brum[2] · V. Chabu[1] · R. Correa da Silva[1]

## Abstract

We present a review on the notion of pure states and mixtures as mathematical concepts that apply for both classical and quantum physical theories, as well as for any other theory depending on statistical description. Here, states will be presented as expectation values on suitable algebras of observables, in a manner intended for the non-specialist reader; accordingly, basic literature on the subject will be provided. Examples will be exposed together with a discussion on their meanings and implications. An example will be shown where a pure quantum state converges to a classical mixture of particles as Planck's constant tends to 0.

**Keywords** Quantum states · Classical states · Algebraic QFT · Semiclassical limit

## 1 Introduction

In many textbooks on quantum physics, there is some obscurity surrounding the notion of pure state, often giving rise to some misconceptions. In quantum mechanics, for instance, pure states are frequently associated to normalised vectors (or, more precisely, to rays) in adequate Hilbert spaces. This neither is an adequate definition nor is, strictly speaking, a correct one since, according to a mathematical construction known as GNS representation, any state (including mixed ones) in an adequately defined

algebra of observables can be represented by a vector state in some Hilbert space. Moreover, this pseudo-definition fails to capture the statistical quality of the notions of pure and of mixed states, which is actually quite simple and illuminating.

Our objective in this review is to present both the intuitive meaning of the concept of pure and of mixed states as well as to develop the mathematical (algebraic) formalism around these notions in order to clarify some of these issues.

Besides, we will explore this formalism in order to enhance our understanding about the physical nature hidden beneath these algebraic and statistical notions. One of the most notable results in this direction is the possibility of unifying the treatment of quantum and classical observables as elements in the same kind of algebra, known as C*-algebra, the only difference being that in the classical case these elements commute with respect to the algebra's operation (multiplication or composition), whereas in the quantum case they do not necessarily commute.

We also present in Section 6 a systematic procedure (known as Weyl quantisation) for transforming classical observables, namely, functions on a phase space, into quantum ones, i.e. self-adjoint operators acting on a Hilbert space. Then, by means of tools provided by semi-classical analysis we will introduce, we will exhibit an example of a family of *pure* quantum states that degenerates into classical *mixtures* in the limit when Planck's constant is taken small.

The physically important relation between purity of states and irreducibility of certain representations of the algebra of observables is discussed in Section 7. The notion of purification of states is presented in Section 8.

---

This paper is dedicated to the memory of Mahir S. Hussein, researcher, teacher and friend.

✉ J. C. A. Barata
jbarata@if.usp.br

M. Brum
marcos.brum@xerem.ufrj.br

V. Chabu
vbchabu@if.usp.br

R. Correa da Silva
ricardo.correa.silva@usp.br

1  Instituto de Física, Universidade de São Paulo, CP 66.318, 05314-970, São Paulo, SP, Brazil

2  Campus UFRJ-Duque de Caxias, Universidade Federal do Rio de Janeiro, Rodovia Washington Luiz, n. 19593, km 104,5, Duque de Caxias, CEP: 25.240-005, RJ, Brazil

Sections 9, 10 and 11 are more technical, mathematically. In Section 9, we introduce the important notion of normal state and present its relation to the so-called density matrix, a very important notion in quantum mechanics relevant in discussions concerning certain aspects of quantum information theory, as the notion of von Neumann entropy and entanglement entropy. In Section 10, we present some interesting remarks about the notion of purity and in Section 11 we present results, through Krein-Milman and Choquet's theorems, on the existence of pure states and on the decomposition of general states into pure ones.

## 1.1 Pure and Mixed Probability Distributions

We start our presentation considering the simpler and perhaps more familiar context of probability distributions, where the notions of purity and of mixture can be discussed in a quite elementary way.

Let us consider a probability space, which consists of a set $\Omega$, called *event space*, and a family $\mathscr{F}$ of subsets of $\Omega$, called *events*. For technical reasons, $\mathscr{F}$ has to be a $\sigma$-algebra of sets, but this point will not be relevant on what follows, except to point out that $\Omega$ itself and the empty set $\emptyset$ are possible events, that means, are elements of $\mathscr{F}$.

A probability measure $\mu$ in $(\Omega, \mathscr{F})$ is an assignment of each event $A \in \mathscr{F}$ to a real number in $[0, 1]$ such that the following conditions are fulfilled: $\mu(\emptyset) = 0$, $\mu(\Omega) = 1$ and, for any collection $\{A_n \in \mathscr{F}, n \in \mathbb{N}\}$ of disjoint subsets, $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$.

These postulates, widely known in the literature as *Kolmogorov axioms*, capture the essential ingredients of the intuitive notion of probability and many basic properties of probability measures can be directly derived from them. For instance, one of the easy consequences of the above postulates is that $\mu(A) \le \mu(B)$ for any $A$ and $B$ in $\mathscr{F}$ such that $A \subset B$.

As a simple example, let $\Omega = \mathbb{R}$ and let $\mu$ assign, to any measurable subset $A \subset \mathbb{R}$ (for instance, an open interval), the number:

$$\mu(A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-x^2/2} dx .$$

$\mu(A)$ is the probability of occurrence of event $A$ for the particular Gaussian distribution considered in the integral.

A probability measure $\mu$ is said to be a *mixture* if there are two other distinct probability measures $\mu_1$ and $\mu_2$, on the same probability space, and numbers $\lambda_1, \lambda_2 \in (0, 1)$ with $\lambda_1 + \lambda_2 = 1$ such that:

$$\mu(A) = \lambda_1\mu_1(A) + \lambda_2\mu_2(A) \tag{1}$$

holds for all events $A \in \mathscr{F}$. A probability measure is said to be *pure*, or extremal, if it is not a mixture. In the Bayesian parlance, the probabilities $\mu_1$ and $\mu_2$ are priors of $\mu$ and $\lambda_1$ and $\lambda_2$ are their respective likelihoods.

An expression like (1) is called a convex linear combination of $\mu_1$ and $\mu_2$. Notice that $\mu_1$ or $\mu_2$ may be mixtures themselves; and hence, we can say that a probability measure is a mixture if it can be written as a finite (or even infinite) convex sum of distinct probability measures: $\mu(A) = \sum_{k=1}^{n} \lambda_k\mu_k(A)$, for some $n \in \mathbb{N}$, with $\sum_{k=1}^{n} \lambda_k = 1$ and $\lambda_k \in (0, 1)$ for all $k$.

In order to explain the intuitive nature of a mixed probability distribution, let us consider a very simple situation where mixture occurs. Suppose we order a large amount of balls from two different factories, each factory having its own standard fabrication processes. The balls produced in each factory are not perfectly the same and will differ randomly from each other. If we consider one specific parameter for characterising the balls, say, their diameter, we can associate to each factory a probability distribution associated to the diameter: for $0 < d_1 < d_2$, the quantity $\mu_k((d_1, d_2))$ measures the probability for a ball produced in factory $k = 1, 2$, to have a diameter in the interval $(d_1, d_2)$.

Now, consider that we mix the balls produced in both factories, so that a fraction $\lambda_k \in (0, 1)$ comes from the production of factory $k = 1, 2$. Naturally, $\lambda_1 + \lambda_2 = 1$. If we measure the diameters of the balls in this mixed ensemble, it is intuitively clear that measurements of the diameters of the balls will be described by a probability measure $\mu$ given by $\mu((d_1, d_2)) = \lambda_1\mu_1((d_1, d_2)) + \lambda_2\mu_2((d_1, d_2))$, again with $0 < d_1 < d_2$.

The probability $\mu$ is therefore a mixture of the probabilities $\mu_1$ and $\mu_2$ with fractions $\lambda_1$ and $\lambda_2$, respectively, since the ensemble considered is a mixture (in the common sense of the word) of two ensembles described by the two probabilities $\mu_1$ and $\mu_2$.

Notice that the probabilities $\mu_1$ and $\mu_2$ can be themselves mixtures, as it can happen if, for instance, the balls are produced by different machines in each of the factories.

This example illustrates the intuitive idea behind the notion of a mixed probability distribution: it describes samples composed of objects of different origins which are placed together. In contrast, pure probability distributions describe systems that, in a sense, are not decomposable in simpler ones. As we will see, in the case of quantum systems, these notions are neatly reproduced in the algebraic formalism.

## 1.2 Mean Values and Variances

Given a probability distribution on a probability space, there is a series of statistical quantities that provide information on the distribution. They can also provide some insight on the nature of pure and mixed probability distributions.

Let $f : \Omega \to \mathbb{R}$ be a real function defined on the event space representing some observable quantity. (Technically, $f$ has to be a measurable function, but we will not stress such mathematical points by now.) We define its *expectation*, *average* or *mean value*, according to the probability measure $\mu$ by:

$$\mathbb{E}_\mu(f) \equiv \langle f \rangle_\mu := \int_\Omega f \, d\mu \,.$$

The *variance* of $f$ on $\mu$ is defined by:

$$\text{Var}_\mu(f) := \left\langle \left(f - \langle f \rangle_\mu\right)^2 \right\rangle_\mu = \langle f^2 \rangle_\mu - \langle f \rangle_\mu^2 \,.$$

As we see from the definition, $\text{Var}_\mu(f)$ measures how much $f$ typically deviates from its mean value $\langle f \rangle_\mu$. It is clear from the definition that $\text{Var}_\mu(f) \geq 0$. The quantity $\sigma_\mu(f) := \sqrt{\text{Var}_\mu(f)}$ is called the *standard deviation* of $f$ on $\mu$.

Consider a mixed probability measure $\mu = \lambda_1 \mu_1 + \lambda_2 \mu_2$, with $\mu_1$ and $\mu_2$ being two distinct probability measures in some probability space and $\lambda_1, \lambda_2 \in (0, 1)$ with $\lambda_1 + \lambda_2 = 1$. Then, one can easily see that

$$\langle f \rangle_\mu = \lambda_1 \langle f \rangle_{\mu_1} + \lambda_2 \langle f \rangle_{\mu_2} \,.$$

Moreover, one can also easily verify that

$$\text{Var}_\mu(f) = \lambda_1 \text{Var}_{\mu_1}(f) + \lambda_2 \text{Var}_{\mu_2}(f) + \lambda_1 \lambda_2 \left[\langle f \rangle_{\mu_1} - \langle f \rangle_{\mu_2}\right]^2 \,.$$

From this, we conclude that

$$\begin{aligned} \text{Var}_\mu(f) &\geq \lambda_1 \text{Var}_{\mu_1}(f) + \lambda_2 \text{Var}_{\mu_2}(f) \\ &\geq \min \left\{ \text{Var}_{\mu_1}(f), \text{Var}_{\mu_2}(f) \right\} \,. \end{aligned}$$

Hence, for the mixed probability measure $\mu$ the variance $\text{Var}_\mu(f)$ is always larger than or equal to the smallest of the numbers $\text{Var}_{\mu_1}(f)$ or $\text{Var}_{\mu_2}(f)$. Therefore, for a fixed function $f$, the smallest values of $\text{Var}_\mu(f)$ will be obtained for pure measures on this probability space. In this sense, pure probability measures are those for which the deviation of $f$ from its mean value is smallest.

## 2 The Notion of State

In physics, the word "state" is often used in a somewhat informal sense as a set $S$ of intrinsic characteristics of a system maximally specifying the possible outcomes of measurements of observable quantities. A given physical theory specifies which quantities are observable (i.e. measurable through experiments) and a state can be defined, with a little more precision, as a rule associating each observable $A$ and each set $S$ of a system's physical characteristics to a probability measure $\mu_{S, A}$ describing the statistical distribution of repeated measurement of $A$ on an ideally infinite ensemble of physical systems with the same set of characteristics $S$. Although this definition is still

vague, it is the base for the precise definition of state that we will present below, which is algebraic in its nature.

In classical mechanics, for instance, observables are (measurable) functions $A(q, p)$ defined in phase space and the state of a system is specified by a probability distribution $\rho(q, p)$ defined in phase space so that the mean values of repeated measurements of $A$ in the state $\rho$ are given by $\langle A \rangle_\rho = \int A(q, p) \rho(q, p) dq dp$.

A relevant case consists of states given by the probability distribution $\rho_0(q, p) = \delta(q - q_0) \delta(p - p_0)$, where $\delta$ represents the Dirac measure and where $(q_0, p_0)$ is a given point in phase space. In this case, we have $\langle A \rangle_{\rho_0} = A(q_0, p_0)$. Moreover, as one easily checks, $\text{Var}_{\rho_0}(A) = 0$, leading to the interpretation that all individual measurements of $A$ in the state $\rho_0$ will result in the same value $A(q_0, p_0)$. Hence, the state $\rho_0$ represents a deterministic state, fully characterised by $q_0$ and $p_0$, where measurements of observable quantities always lead to the same result.

On the other hand, in quantum mechanics, it is commonly thought that all the possible pure "states" of a physical system are described by normalised vectors of a Hilbert space, and the possible measurable observables by self-adjoint operators acting on them. Vectors in a Hilbert space $\mathcal{H}$ may indeed represent pure states (either in the intuitive notion explained above or in the formal one to be presented below). However, as we shall discuss, not every state can be represented as a vector in $\mathcal{H}$ and vector states are not necessarily pure.

For instance, let us consider the Hilbert space representing a two-level system (a *qubit*) described in the Hilbert space $\mathcal{H} = \mathbb{C}^2$. If we have several copies of this system in the same state $\psi \in \mathcal{H}$ (with $\|\psi\|^2 = \langle \psi, \psi \rangle = 1$)[1] and measure each one of them for an observable $A$, the mean value $\langle A \rangle_\psi$ of the measured results is given by the inner product $\langle A \rangle_\psi = \langle \psi, A\psi \rangle$. However, if the copies are composed by a few systems in a state $\phi_1 \in \mathcal{H}$ and a few other in a different state $\phi_2 \in \mathcal{H}$ (let us suppose a fraction $p_1$ of the total number of particles in $\phi_1$, and $p_2$ in $\phi_2$, so $p_1 + p_2 = 1$), then the mean value of the several measurements is expected to be $\langle A \rangle = p_1 \langle \phi_1, A\phi_1 \rangle + p_2 \langle \phi_2, A\phi_2 \rangle$. Calculating the average of the measures taken for different copies of a system is precisely what is meant by *average value* of an observable for a system in a defined state, so there *must* be a state in which the system can be that corresponds to this mean value $p_1 \langle \phi_1, A\phi_1 \rangle + p_2 \langle \phi_2, A\phi_2 \rangle$.

Is there any vector state that could represent such a mixture of states, i.e. a vector $\psi \in \mathbb{C}^2$ such that $\langle \psi, A\psi \rangle = p_1 \langle \phi_1, A\phi_1 \rangle + p_2 \langle \phi_2, A\phi_2 \rangle$? The answer is *no*, unless it is

---

[1] Scalar products (or inner products) in Hilbert spaces will be always denoted here by $\langle \phi, \psi \rangle$ rather than by $\langle \phi | \psi \rangle$. We follow the physicists' convention: they are antilinear in the first argument and linear in the second.

a trivial mixture where either $p_1$ or $p_2$ is 0, or $\phi_1 = \phi_2$ (this will follow from Theorem 3 in Section 4.2). An example of such impossibility is shown in Section 5.1, leading us to the conclusion that a more comprehensive way of representing states is needed in order to fully describe a quantum system.

Mixtures as those commented above are usually introduced in a quantum theory based on (separable) Hilbert spaces $\mathscr{H}$ as *density operators* $\rho$, which are positive trace class operators normalised so as $\mathrm{Tr}(\rho) = 1$. For a finite (possibly infinite) mixture of states $\phi_1, \ldots, \phi_n$ with weights $p_1, \ldots, p_n$ (and $\sum_{k=1}^{n} p_k = 1$), it is constructed as:

$$\rho = \sum_{k=1}^{n} p_k |\phi_k\rangle \langle\phi_k| \qquad (2)$$

(in the infinite case, the sum's convergence is uniform) and it is easy to see that the average value may be calculated by means of the formula $\langle A \rangle_\rho = \mathrm{Tr}(\rho A)$, since

$$\mathrm{Tr}(\rho A) = \sum_{k=1}^{n} p_k \langle\phi_k, \ A\phi_k\rangle.$$

We shall denote by $\mathscr{L}(\mathscr{H})$ the set of bounded (i.e. continuous) linear operators acting on a Hilbert space $\mathscr{H}$.

We remark that any operator such as that in Eq. (2) has the properties listed above for density operators. Conversely, if some $\rho \in \mathscr{L}(\mathscr{H})$ is positive and $\mathrm{Tr}(\rho) < \infty$, then it is a compact operator (see e.g. [1]) and, therefore, possesses discrete and finitely degenerate spectrum and a spectral decomposition $\rho = \sum_{k=1}^{\infty} \varrho_k |\phi_k\rangle \langle\phi_k|$, where the $\phi_k$ is normalised and mutually orthogonal, and $\varrho_k > 0$ with $\sum_{k=1}^{\infty} \varrho_k = 1$, allowing us to interpret $\rho$ as the density operator of an infinite mixture of states $\phi_k$ with statistical weights $\varrho_k$.

## 2.1 The Physically Motivated Topology on States

Any deeper analysis of quantum physics requires the introduction of topologies in the set of observables or in the set of states, that means, the introduction of the notion of *closeness* between different observables or between different states. This is particularly relevant if we intend to use the notion of convergence of observables or of states. The difficulty is that the space of observables (typically a C*-algebra) and the associated set of states are usually infinite dimensional and, therefore, are usually equipped with many non-equivalent topologies, i.e. with many non-equivalent notions of closeness between their elements and, therefore, equipped with many non-equivalent notions of convergence.

For the space of states, it is possible to point to a physically motivated topology, that we now describe. Other useful topologies will be mentioned later (as the norm topology in the set of states in C*-algebras; see Definition 2).

Let us take an observable $A$; for the sake of simplicity, imagine that it possesses a finite set of possible outcome values, and let $\lambda$ be one of them. Then, its relative frequency $\nu_A(\lambda, n)$ when $A$ is measured on $n$ copies of a system in the state $\omega$ should converge to a probability $p_A(\lambda)$ in the limit $n \to \infty$, and so must converge the measurements' mean value $\sum_k \lambda_k \nu(\lambda_k, n)$ to the expectation value $\omega(A)$. As we saw in the beginning of this section, saying that a system is in the state $\omega$ means precisely that whenever we take measurements on its copies for an observable $A$, we will obtain an outcome according to the probability distribution $\lambda \longmapsto p_A(\lambda) = \omega(E_\lambda^A)$, where $E_\lambda^A$ is the spectral projection of $A$ on the closed subspace of eigenvectors with eigenvalue $\lambda$. This is, in theory, how to connect the positive linear functional $\omega$ to the intrinsic property of a physical system that we called *state*; in practice, this connection is trickier.

It happens that it is not possible to take an infinite number of measurements resulting in a value $\lambda$, nor can we obtain all possible outcomes for an observable in the case where it has an infinite set of possible results, so we never know a system's exact probability distributions $\lambda \longmapsto p_A(\lambda)$. Worse, we may not measure a system for every possible observable $A$. As a consequence, we are bent not to know a system's state exactly; for the most we may hope to acquire from data are some approximative mean values for small sets of finite observables.

In order to treat this problem, consider first a particular bounded observable $A$. The first and the second issues are set by taking a number $n$ of measurements sufficiently large so as to have:

$$|v(A) - \omega(A)| < \varepsilon,$$

where $v(A) = \sum_k \lambda_k \nu_A(\lambda_k, n)$ is the measured mean value of $A$, and $\varepsilon > 0$ is an error that we may take arbitrarily small, by hypothesis. Unfortunately, it is possible that a sufficiently large $n$ that works for all observables does not exist. Yet, we may have arbitrarily precise information on $\omega$ for finite sets of observables: taking observables $A_1, \ldots, A_k$ and an error $\varepsilon$, we define a *neighbourhood* of $\omega$ with radius $\varepsilon$ in the set of all states $\Sigma$ as:

$$\Omega_\varepsilon(A_1, \ldots, A_k; \omega) = \{\alpha \in \Sigma : |\alpha(A_j) - \omega(A_j)| \\ < \varepsilon, \ \forall j \in \{1, \ldots, k\}\}.$$

Surely, $\omega \in \Omega_\varepsilon(A_1, \ldots, A_k; \omega)$. Also, this neighbourhood contains any state $\tilde{\omega}$ satisfying $\tilde{\omega}(A) = v(A)$ at least for $A = A_1, \ldots, A_k$, so under any measurement of observables $A_1, \ldots, A_k$, $\tilde{\omega}$ describes our actual physical system as well as $\omega$. In this sense, the states in $\Omega_\varepsilon(A_1, \ldots, A_k; \omega)$ are good approximations for $\omega$.

Besides, it is a mathematical fact that this kind of neighbourhoods induces in $\Sigma$ a topology, called in [2]

*physical* or *weak topology*, whose notion of convergence may be cast in the following way: one says that a sequence[2] of states $\omega_n$ converges *physically*, or *weakly*, to $\omega$ if, for any observable $A$, the mean values converge, i.e. if one has $\omega_n(A) \longrightarrow \omega(A)$ as $n \to \infty$. As a conclusion, $\omega$ may be completely determined by means of a process of taking limits in the weak sense.

It is important to clarify some differences in nomenclature: among mathematicians, the topology referred to as *weak topology* is called *weak-∗ topology*. The reason relies on the fact that sets defined like $\Omega_\varepsilon(A_1, \ldots, A_k; \omega)$, with $A_1, \ldots, A_k$ in a normed vector space $\mathcal{V}$, and $\omega$ any bounded linear functional on $\mathcal{V}$, form a basis of neighbourhoods for a locally convex topology on the dual of $\mathcal{V}$.

Topologies can be similarly introduced in the space of observables, providing notions of closeness between operators. We will make use of some of them in our final sections.

In the so-called *norm (or uniform) topology*, two operators $A$ and $A'$ acting on a Hilbert space $\mathcal{H}$ are considered close if for some prescribed $\varepsilon > 0$ one has $\|A - A'\| := \sup \|(A - A')\psi\| < \varepsilon$, where the supremum is taken over all vectors $\psi \in \mathcal{H}$ with $\|\psi\| = 1$. This means that $A\psi$ and $A'\psi$ differ in norm by an amount smaller than the prescribed error $\varepsilon$ regardless of the normalised vectors $\psi$. In this topology, we say that a sequence of operators $A'_n$ converges to $A$ if for any $\|A - A'_n\|$ goes to zero when $n$ goes to infinity.

In the so-called *strong operator topology* in a Hilbert space $\mathcal{H}$, two operators $A$ and $A'$ are considered close with respect to a distinct finite set of normalised vectors $\psi_j \in \mathcal{H}$, $j = 1, \ldots, N$, and for some prescribed $\varepsilon > 0$, if one has $\|(A - A')\psi_j\| < \varepsilon$, for all $j = 1, \ldots, N$. This means that the vectors $A\psi_j$ and $A'\psi_j$ differ in norm, for each $j = 1, \ldots, N$, by an amount smaller than the prescribed error $\varepsilon$. In this topology, we say that a sequence of operators $A'_n$ converges to $A$ if they eventually become close, when $n \to \infty$, with respect to all finite sets of normalised vectors $\psi_j \in \mathcal{H}$, $j = 1, \ldots, N$, and all $\varepsilon > 0$.

In the so-called *weak operator topology* in a Hilbert space $\mathcal{H}$, two operators $A$ and $A'$ are considered close with respect to a distinct set of normalised vectors $\psi_j \in \mathcal{H}$, $j = 1, \ldots, N$, and for some prescribed $\varepsilon > 0$, if $|\langle\psi_j, (A - A')\psi_j\rangle| < \varepsilon$, for all $j = 1, \ldots, N$. This means that $A$ and $A'$ provide the same expectation values for the vector states defined by the vectors $\psi_j$, $j = 1, \ldots, N$, up to an error smaller than the prescribed $\varepsilon$. In this topology, we say that a sequence of operators $A'_n$ converges to $A$ if they eventually become close when $n \to \infty$ with respect to all finite sets of normalised vectors $\psi_j \in \mathcal{H}$, $j = 1, \ldots, N$, and all $\varepsilon > 0$.

In the weak operator topology, the notion of closeness between operators is expressed in terms of their expectation values and, therefore, is directly linked to measurable quantities.

In infinite-dimensional Hilbert spaces all operator topologies defined above differ [6], leading to distinct notions of convergence between operators, a very significant fact for the mathematical analysis of quantum systems. For instance, convergence of sequences of operators in the uniform operator topology implies convergence in the strong operator topology. Analogously, convergence of sequences of operators in the strong operator topology implies convergence in the weak operator topology. In both cases, the opposite statements are not generally valid.

# 3 The Algebraic Approach to Quantum Systems

The familiar Hilbert space approach, however, has limitations when dealing with quantum systems with infinitely many degrees of freedom, as those considered in quantum field theory and quantum statistical mechanics, mainly due to important features commonly manifest in such systems, such as superselection sectors, phase transitions and the existence of some special states, for instance "thermal or finite temperature" states, that cannot be properly described in the Hilbert space formalism. A universal formalism that can be applied to general quantum systems was proposed by Haag, Kastler and many others (see e.g. [2, 3] or, for a more recent review, [4]). We will refer to this formalism as the algebraic approach to quantum systems. In general terms, it emphasises the dichotomy between observables (representing physically meaningful and measurable quantities) and states (dealing with the statistics of measurements of physically observable quantities).

In this formalism, observables are treated as abstract associative algebras of a certain kind (usually $C^*$ and/or von Neumann algebras are considered), while states are associated to positive normalised linear functionals on these algebras. Within the algebraic approach, one is no longer restricted to the use of the vectors or density operator on $\mathcal{H}$ in order to describe states. Moreover, this formalism allows the treatment of pure and mixed states in a very general and elegant fashion, a point that will be relevant for our purposes.

As we mentioned, it was due to the efforts of Haag and others that C*-algebras have been recognised as the relevant mathematical objects for the universal description

---

[2]Strictly speaking one should use *nets* to characterise convergence in those topologies, instead of sequences. For the sake of clarity, we will use the latter. See, e.g., [17].

of observables in quantum systems. Let us briefly describe such algebras.

Let us denote by $\mathscr{O}$ the set of observables of a physical system. $\mathscr{O}$ must have a real vector space structure; moreover, the composition of some observables must result in a new observable. This suggests that $\mathscr{O}$ must be contained in a larger set that is an associative algebra. Let us denote a minimal associative algebra satisfying these properties by $\mathfrak{A}$. Hence, $\mathfrak{A}$ is a (complex) vector space endowed with associative multiplication. We also require the existence of an operation $* : \mathfrak{A} \longrightarrow \mathfrak{A}$, called an *involution* in $\mathfrak{A}$, such that for all $A$, $B \in \mathfrak{A}$ and $z \in \mathbb{C}$ one has: $(A^*)^* = A$, $(A + B)^* = A^* + B^*$, $(zA)^* = \overline{z}A^*$ and[3] $(AB)^* = B^*A^*$. Here, $\overline{z}$ denotes the complex conjugate of $z \in \mathbb{C}$.

The algebra $\mathcal{L}(\mathscr{H})$, for example, possesses these properties and the involution is related the notion of the *adjoint* of a bounded operator acting on $\mathscr{H}$ with respect to the scalar product on $\mathscr{H}$.

The requirement of associativity in quantum systems deserves some physical clarification. Regarding the elements of $\mathfrak{A}$ as *operations* acting on a quantum system, the order of two successive operations matters and the algebra of observables is not supposed to be commutative. If we consider three successive operations, however, one has to guarantee that the last operation does not depend on the previous ones, which is achieved through the requirement of associativity.

Finally, since $\mathfrak{A}$ is an extension of the more physically relevant set $\mathscr{O}$, we need a way to distinguish $\mathscr{O}$ within $\mathfrak{A}$. The involution provides a method to identify the elements of $\mathscr{O}$ among all elements of $\mathfrak{A}$: let $A \in \mathfrak{A}$, if $A \in \mathscr{O}$, then $A = A^*$. One can wonder if $A^* = A$ implies $A \in \mathscr{O}$. In the first attempts to axiomatise quantum mechanics, it was required that any self-adjoint operator should represent an observable, but for physical reasons this requirement was discarded. There are many examples of self-adjoint operators not associated to observables. If $U$ is any unitary operator acting on a Hilbert space of physical states $\mathscr{H}$, then $U + U^*$ is self-adjoint, but it may not be associated to a measurable quantity. For example, when $U$ represents the shift operator on a separable Hilbert space, acting on an orthonormal base of vectors $\{\phi_n\}_{n \in \mathbb{Z}}$ as $U\phi_n = \phi_{n+1}$). Another example: consider a Fermionic field $\psi$ and take $\psi + \psi^*$, a self-adjoint operator not related to a measurable quantity. For a more detailed discussion on the axioms of quantum mechanics, see [5].

The last ingredient in our construction is a norm on $\mathfrak{A}$ which must be compatible with the multiplication and the involution operations. We now define:

**Definition 1** A C*-algebra is a set $\mathfrak{A}$ provided with a complex linear structure, an associative multiplication, an involution and a norm such that:

(i)   $\lambda(AB) = (\lambda A)B = A(\lambda B)$, for all $A$, $B \in \mathfrak{A}$ and $\lambda \in \mathbb{C}$;
(ii)  $A(B + C) = AB + AC$, for all $A$, $B$, $C \in \mathfrak{A}$;
(iii) $\|AB\| \leq \|A\|\|B\|$, for all $A$, $B \in \mathfrak{A}$;
(iv)  $\|A^*A\| = \|A\|^2$, for all $A \in \mathfrak{A}$;
(v)   $\mathfrak{A}$, as a vector space, is complete with the norm.

The algebra $\mathscr{L}(\mathscr{H})$, of all bounded (continuous) operators acting on a Hilbert space $\mathscr{H}$, is known to be a C*-algebra. One might ask whether the definition above leads to anything different from the usual description of quantum mechanics based in Hilbert spaces. The answer to this question is negative and is based on the following facts (respectively, Theorems 2.1.10 and 2.1.11A of [6]):

**Theorem 1** *Let $\mathfrak{A}$ be a C\*-algebra. There exists a Hilbert space $\mathscr{H}$ such that $\mathfrak{A}$ is isomorphic to some self-adjoint closed subalgebra of $\mathscr{L}(\mathscr{H})$.*

So, essentially, a C*-algebra is an abstract algebra of operators. Furthermore:

**Theorem 2** *If a C\*-algebra $\mathfrak{A}$ is commutative, then there exists a locally compact Hausdorff topological space $X$ such that $\mathfrak{A}$ is isomorphic to $C_0(X)$, i.e. the algebra of continuous complex functions on $X$ that vanish at infinity*[4].

In other words, (commutative) C*-algebras can be viewed as abstract algebras of functions. Hence, the notion of states as functionals over $\mathfrak{A}$ is suitable for both quantum and classical theories, as well as any other experimental theories.

Besides, on one hand, a C*-algebra can always be mapped to a closed $*$-subalgebra of $\mathscr{L}(\mathscr{H})$ for some suitable Hilbert space $\mathscr{H}$. On the other hand, such $\mathscr{H}$ may be obtained by means of the *GNS construction* that we will present in Section 7.

Finally, just like the operators in $\mathscr{L}(\mathscr{H})$, the elements of a C*-algebra possess adjoints, norms and even spectra: given $A \in \mathfrak{A}$, the spectrum of $A$ is the set:

$$\mathrm{spec}(A) = \{\lambda \in \mathbb{C} \ : \ A - \lambda \mathbb{1} \text{ is not invertible}\}.$$

Indeed, one may speak about the inverse of an element $A$ in a C*-algebra since either it has an identity $\mathbb{1}$, or we may map $\mathfrak{A}$ to a larger algebra $\tilde{\mathfrak{A}}$ containing an identity, a case where we consider the spectrum of $A$ in $\mathfrak{A}$ to be its spectrum

---

[3]These rules are not supposed to hold in the case of unbounded operators acting on Hilbert spaces. See e.g. [1].

[4]A function $f \in C_0(X)$ vanishes at infinite if, for any $\varepsilon > 0$, there is $K \subset X$ compact such that $|f(x)| < \varepsilon$ for any $x \in X \setminus K$.

as an element in $\tilde{\mathfrak{A}}$.[5] As a result, we can still have the usual interpretation of the spectrum of a self-adjoint operator as the possible outcomes of physical experiments. Moreover, some C*-algebras admit traces, i.e. positive functionals acting on them that generalise the usual notion of trace of an operator [6, 7].

A last comment worth mentioning is that if traditionally one considers unbounded operators such as momentum or the Hamiltonian as observables, in practice one never performs measurements capable of observing the entire set of possible values of momenta or energy when they are not bounded. Sensors and physical equipment have always a bounded range within which they are suitable for making measurements, so what is really done in a physical theory is to account for measurements of a real observable $A$ within a certain bounded interval $I \subset \mathbb{R}$, whose correspondent operator $A\chi_I(A)$ is bounded by $\mathrm{supp}\,|I|$. Here, $\chi_I$ is the so-called characteristic function of $I$: for real $x$, the function $\chi_I(x)$ equals 1 for $x \in I$ and 0 otherwise. The operator $\chi_I(A)$ is defined by the functional calculus for self-adjoint operators $A$. An unbounded observable $A$ shall thus be thought of as some kind of limit $A\chi_{I_n}(A)$ along an increasing sequence of intervals $I_n \nearrow \mathbb{R}$ (for more details, see the notion of affiliated operators in [6]).

## 4 States as Functionals on C*-Algebras

Since we have re-elaborated our concept of observables, that of states ought to be rediscussed too. Indeed, as we have argued in Section 2, a state could be thought of as some property of a physical system that associates to each observable $A$, a probability measure describing the statistical distribution of repeated measurements of $A$ on many copies of the same system.

This can be achieved if we associate to $A$ the number $\omega(A)$ corresponding to the mean value of all its outcomes over the several measurements, for if one knows the averages of any observables, including those like "the frequency of the outcome $a$ in an experiment measuring $A$", then it is possible to reconstruct the statistical distribution for $A$. For this reason, states will be defined as functions from $\mathfrak{A}$ into the theory's scalars, usually $\mathbb{C}$, with the special property that $\omega(A)$ must be a real number if $A$ is indeed a physical observable, i.e. if $A \in \mathcal{O} \subset \mathfrak{A}$.

Other reasonable properties for $\omega$ to have a physical meaning are:

– **Positivity** For a positive observable $A$, one must have $\omega(A) \geqslant 0$, for if any possible measurement of $A$ results in a positive value, so must be their average. This is equivalent to saying that $\omega(A^*A) \geqslant 0$ for any $A \in \mathfrak{A}$.

– **Boundedness** The average of a set of values cannot be greater than their supremum; the measurable values of an observable $A$ are the elements of its spectrum, which is bounded by $\|A\|$. It follows that $\omega(A) \leqslant \|A\|$, or, more shortly:

$$\|\omega\| = \sup_{\substack{A \neq 0 \\ \text{observable}}} \frac{|\omega(A)|}{\|A\|} \leqslant 1. \tag{3}$$

For a matter of convenience, we always take the supremum over the entire $\mathfrak{A} \setminus \{0\}$.

– **Normalisation** $\omega(A)$ is to be the average of measurements of $A$. If the algebra $\mathfrak{A}$ has an identity $\mathbb{1}$, whose only measurable value is 1, we must impose $\omega(\mathbb{1}) = 1$, consequently the supremum in (3) is achieved and one gets $\|\omega\| = 1$. If the algebra has no identity, there will be a bounded non-decreasing net[6] $(A_\nu)_{\nu \in I}$ of operators approximating the identity (see Theorem 2.2.18 of [6]), which causes $\omega(A_\nu)$ to approximate 1 and the supremum in (3) to be exactly 1. In any way, we end up with $\|\omega\| = 1$.

– **Linearity** In the usual description of quantum physics, the average of linear combinations of two observables $A$ and $B$, $\lambda \in \mathbb{R}$, is given by the combination of their averages:

$$\langle A + \lambda B \rangle = \langle A \rangle + \lambda \langle B \rangle,$$

irrespective to whether $A$ end $B$ are compatible observables or not, i.e. whether the corresponding operators $A$ and $B$ commute or not. Hence, we must have $\omega(A + \lambda B) = \omega(A) + \lambda\omega(B)$ at least for $A$ and $B$ being themselves observables and $\lambda \in \mathbb{R}$, and so being $A + \lambda B$. If one of these is not an observable (for instance if $A = i\mathbb{1}$, $B = -i\mathbb{1}$ and $\lambda = 1$), it is not physically clear what their linear combination should mean, nor even which interpretation one should give to $\omega(A)$. Therefore, imposing linearity on $\omega$ with respect to $A$ for any $A \in \mathfrak{A}$ is an arbitrary choice, so as to end up with a linear theory; linearity is apparently not a physical requirement unless for the restriction of $\omega$ to real linear combinations of elements of $\mathcal{O}$. Nonetheless, the theory we obtain doing so seems to be a good description of what is actually seen in the laboratory experiments.

---

[5] Concretely, $\tilde{\mathfrak{A}}$ is the algebra with elements $(\lambda, A) \in \mathbb{C} \times \mathfrak{A}$, involution $(\lambda, A)^* = (\bar{\lambda}, A^*)$ and multiplication $(\lambda_1, A_1)(\lambda_2, A_2) = (\lambda_1\lambda_2, \lambda_1 A_2 + \lambda_2 A_1 + A_1 A_2)$; one easily checks that $(1, 0)$ is an identity in $\tilde{\mathfrak{A}}$, and the convenient way to map $\mathfrak{A}$ into $\tilde{\mathfrak{A}}$ is through the application $A \longmapsto (0, A)$.

[6] An increasing net of operators is a net such that $A_\nu \geqslant A_\mu$ whenever $\nu \geqslant \mu$; the operators' order relation is defined in the following way: $A_\nu \geqslant A_\mu$ if $A_\nu - A_\mu$ is a positive operator, i.e. if its spectrum is included in $\mathbb{R}^+$.

We are thus led to:

**Definition 2** Given a C*-algebra $\mathfrak{A}$, $\omega$ is said to be a state over $\mathfrak{A}$ if it is a bounded positive linear functional with $\|\omega\| = \sup_{\substack{A \in \mathfrak{A} \\ A \neq 0}} \frac{|\omega(A)|}{\|A\|} = 1$.

*Remark 1* Supposing that a C*-algebra $\mathfrak{A}$ has an identity $\mathbb{1} \in \mathfrak{A}$, it is possible to show (see e.g. [6, 20, 21]) that a linear functional $\omega$ on $\mathfrak{A}$ is bounded with $\|\omega\| = \omega(\mathbb{1})$ if and only if it is positive. This close connection between positivity and boundedness allows us to require in Definition 2 that a state be merely a positive linear functional satisfying $\omega(\mathbb{1}) = 1$. If the algebra has no identity, then it has at least a non-decreasing net $(A_\nu)_{\nu \in I}$ approximating it (Theorem 2.2.18 of [6]), and this remark holds under the form $\|\omega\| = \omega\left(\sup_{\nu \in I} A_\nu\right)$.

The first evident virtue of this new notion of state is epistemological, as it applies not only to quantum physics, but also for any experimental science whose systems are in states about which we have information exclusively through series of measurements, so we can only determine them by the statistical profile of the data we gather. This holds regardless of any predefined deterministic notion of state, and regardless of having or having not complete knowledge about the system that one might obtain from a totally accurate measurement.

An obvious example of such an experimental theory is classical physics itself, and in Section 5.2 we show how this rigorous notion of state fits perfectly the classical situation from the laboratory's point of view (recall Theorem 2 above).

Another of this concept's advantages is that it does encompass more quantum states than previously; in Section 10.1, we exhibit a state that cannot be written neither as a vector nor as a density operator. In this case, however, the reader should pay attention to the point that this state will be defined through a process of taking limits, which is not fortuitous. In fact, any state on a C*-algebra $\mathfrak{A}$ may be approximated by a sequence of states that correspond to density operators when $\mathfrak{A}$ is realised as a concrete operator algebra $\mathcal{L}(\mathscr{H})$ (see Theorems 1, 7 and 8).

Finally. we notice that if $\omega_1$ and $\omega_2$ are two states in $\mathfrak{A}$ and $\lambda \in [0, 1]$, it can be easily verified from the definition that the convex combination $\lambda\omega_1 + (1 - \lambda)\omega_2$ is also a state in $\mathfrak{A}$. This remark will be essential for the notion of pure and mixed states below.

## 4.1 Pure and Mixed States

Now, we come to the central notions in this discussion. As informally described above, *pure states* are those that cannot be written as a convex combination of other states. This concept may be formalised for C*-algebras in the following way:

Let be $\mathfrak{A}$ a C*-algebra, and $\omega$ a state on it. $\omega$ is said to be pure if, given a scalar $\lambda \in [0, 1]$ and states $\omega_1 \neq \omega_2$ on $\mathfrak{A}$ such that:

$$\omega = \lambda\omega_1 + (1 - \lambda)\omega_2,$$

then necessarily $\lambda = 0$ or $\lambda = 1$. Otherwise, $\omega$ is said to be a mixture of $\omega_1$ and $\omega_2$.

In Section 7, we will present an important criterion for characterising a state $\omega$ on $\mathfrak{A}$ as pure. It will be done by means of evaluating the reducibility of the representation of $\mathfrak{A}$ into a specific algebra of operators on a Hilbert space especially constructed for $\mathfrak{A}$ and $\omega$, the so-called *GNS representation*. This characterisation is very important because it makes clear the point that a mixed state is a sort of composite system where each component can be analysed by an independent algebra of observables, namely, by an irreducible representation in the *GNS* Hilbert space of the abstract observable algebra.

## 4.2 Pure and Vector States

When the C*-algebra is a subalgebra of $\mathcal{L}(\mathscr{H})$ for some Hilbert space $\mathscr{H}$, there is a natural way of defining states on it, which is to take a $\Psi \in \mathscr{H}$ with norm $\|\Psi\| = 1$ and put

$$\omega_\Psi(A) = \langle \Psi, A\Psi \rangle$$

for any $A \in \mathfrak{A}$. Proving that this $\omega_\Psi$ is a true state in the sense of Definition 2 is straightforward: obviously, it is linear and positive (for $\omega_\Psi(A^*A) = \|A\Psi\|^2 \geqslant 0$), so by Remark 1, $\|\omega_\Psi\| = \|\Psi\|^2 = 1$.

A state $\omega$ on $\mathfrak{A} \subset \mathcal{L}(\mathscr{H})$ such that there exists a unit vector $\Psi$ in $\mathscr{H}$ satisfying:

$$\omega(A) = \langle \Psi, A\Psi \rangle,$$

for any $A \in \mathfrak{A}$ is said to be a vector state.

If the C*-algebra considered is the algebra of all continuous operators acting on a Hilbert space, one has the following important statement:

**Theorem 3** *Any vector state on the C*-algebra $\mathcal{L}(\mathscr{H})$ is pure.*

This theorem is a particular case of Theorem 4, whose proof is given below.

This theorem may explain why some quantum mechanics textbooks present the concepts of pure and of vector states as equivalent. It is important to emphasise, however, that these concepts are not equivalent, in general, and one may find situations, even physically relevant ones, where certain pure states are not vector states and certain vector states are not pure.

Indeed, in Section 10.1, we exhibit an example of a pure state that is not a vector state. Moreover, in the presence of superselection rules, one can see on physical grounds that not all vector states are pure.

As an example, consider the simple case where we have two superselection sectors corresponding to some conserved "charge" assuming two distinct values. The physical Hilbert space $\mathscr{H}$ is a direct sum of two mutually orthogonal subspaces $\mathscr{H} = \mathscr{H}_1 \oplus \mathscr{H}_2$, but the algebra of observables cannot be the whole $\mathscr{L}(\mathscr{H})$, since there are operators in $\mathscr{L}(\mathscr{H})$ that map $\mathscr{H}_1$ into $\mathscr{H}_2$ (and vice versa), violating the superselection rule. Thus, the algebra of observables has to be a subalgebra of $\mathscr{L}(\mathscr{H}_1) \oplus \mathscr{L}(\mathscr{H}_2)$. Let us assume for simplicity that the algebra of observables coincides with $\mathscr{L}(\mathscr{H}_1) \oplus \mathscr{L}(\mathscr{H}_2)$. Take a vector in $\mathscr{H}$ in the form $\Psi = (a_1 \psi_1) \oplus (a_2 \psi_2)$, where $\psi_1 \in \mathscr{H}_1$ and $\psi_2 \in \mathscr{H}_2$ are normalised vectors (i.e. $\|\psi_1\|_{\mathscr{H}_1} = \|\psi_2\|_{\mathscr{H}_2} = 1$) and where $a_1$ and $a_2$ are non-zero complex numbers with $|a_1|^2 + |a_2|^2 = 1$. Then, $\Psi$ is also normalised and, for any observable $A = A_1 \oplus A_2$, one has:

$$\omega_\Psi(A) = |a_1|^2 \omega_1(A) + |a_2|^2 \omega_2(A),$$

where $\omega_1(A) = \omega_1(A_1 \oplus A_2) := \omega_{\psi_1}(A_1)$ and, analogously, $\omega_2(A) = \omega_2(A_1 \oplus A_2) := \omega_{\psi_2}(A_2)$ are two states on $\mathscr{L}(\mathscr{H}_1) \oplus \mathscr{L}(\mathscr{H}_2)$. Thus, the vector state $\omega_\Psi$ is not a pure state on $\mathscr{L}(\mathscr{H}_1) \oplus \mathscr{L}(\mathscr{H}_2)$, but a mixture of $\omega_1$ and $\omega_2$. The relation between purity and indecomposability of the algebra will be further discussed in Section 7.

Although being a pure state does not imply being a vector state in the general picture, as stressed above, in some important cases this happens to be true; together with Theorem 3, this means that in these situations both concepts are indeed equivalent. In Section 10.2, for instance, we show that for a C*-algebra composed of all the compact operators on some Hilbert space, all pure states are vector ones.

We shall need a technical generalisation of Theorem 3:

**Theorem 4** *Let $\mathscr{H}$ be some Hilbert space and $\mathfrak{A} \subset \mathscr{L}(\mathscr{H})$ be a C*-subalgebra of $\mathscr{L}(\mathscr{H})$. Consider a normalised vector $\Phi \in \mathscr{H}$. If the orthogonal projection on the subspace generated by $\Phi$ is an element of $\mathfrak{A}$, then the vector state $\omega_\Phi$ on $\mathfrak{A}$ is pure.*

In the case when $\mathfrak{A} = \mathscr{L}(\mathscr{H})$, this implies Theorem 3, above, since in this case all orthogonal projectors on the unidimensional subspaces generated by the vectors of $\mathscr{H}$ belong to $\mathfrak{A}$.

*Proof* We follow closely the joint proof of Theorem 2.8 and of Lemma 2.9 in [2]. For simplicity, let us assume that $\mathfrak{A}$ contains a unit $\mathbb{1}$.

By contradiction, let us assume that $\omega_\Phi$ is a mixed state on $\mathfrak{A}$. Then, there are $\lambda \in (0, 1)$ and two distinct states $\omega_1$ and $\omega_2$ on $\mathfrak{A}$ such that:

$$\langle \Phi, A\Phi \rangle = \lambda \omega_1(A) + (1 - \lambda)\omega_2(A) \qquad (4)$$

for all $A \in \mathfrak{A}$. Let $E$ be the orthogonal projector on the one-dimensional subspace generated by $\Phi$. By assumption $E \in \mathfrak{A}$ and $E \in \mathfrak{A}$ and $\mathbb{1} - E \in \mathfrak{A}$, and we may write:

$$\begin{aligned} 0 &= \langle \Phi, (\mathbb{1} - E)\Phi \rangle \\ &= \lambda \omega_1(\mathbb{1} - E) + (1 - \lambda)\omega_2(\mathbb{1} - E). \end{aligned}$$

It follows from this that $\omega_1(\mathbb{1} - E) = \omega_2(\mathbb{1} - E) = 0$

$$\begin{aligned} |\omega_a((\mathbb{1} - E)B)|^2 &\leq \omega_a\left((\mathbb{1} - E)(\mathbb{1} - E)^*\right)\omega_a\left(B^*B\right) \\ &= \omega_a(\mathbb{1} - E)\omega_a\left(BB^*\right) = 0, \end{aligned}$$

which implies $\omega_a((\mathbb{1} - E)B) = 0$. Analogously, one has $\omega_a(B(\mathbb{1} - E)) = 0$, for both $a = 1, 2$. Since

$$\begin{aligned} B &= (E + (\mathbb{1} - E))B(E + (\mathbb{1} - E)) = EBE \\ &+ (\mathbb{1} - E)BE + EB(\mathbb{1} - E) + (\mathbb{1} - E)B(\mathbb{1} - E), \end{aligned}$$

one has $\omega_a(B) = \omega_a(EBE)$ for both $a = 1, 2$. Now, for any $\Psi \in \mathscr{H}$ one has $E\Psi = \langle \Phi, \Psi \rangle \Phi$ and, hence,

$$\begin{aligned} EBE\Psi &= \langle \Phi, \Psi \rangle EB\Phi = \langle \Phi, \Psi \rangle \langle \Phi, B\Phi \rangle \Phi \\ &= \langle \Phi, B\Phi \rangle (\langle \Phi, \Psi \rangle \Phi) = \langle \Phi, B\Phi \rangle E\Psi, \end{aligned}$$

which implies $EBE = \langle \Phi, B\Phi \rangle E$. It follows that $\omega_a(B) = \omega_a(EBE) = \langle \Phi, B\Phi \rangle \omega_a(E)$. Taking, in particular, $B = \mathbb{1}$, this says that $1 = \omega_a(E)$ for both $a = 1, 2$. Hence, $\omega_a(B) = \langle \Phi, B\Phi \rangle$ for each $a = 1, 2$ and for all $B \in \mathfrak{A}$, which contradicts (4) with $\omega_1$ and $\omega_2$ being distinct. This completes the proof.

The case when $\mathfrak{A}$ does not contain a unit can be treated similarly by using the so-called approximants of the identity (see e.g. [6]). □

## 5 Examples

### 5.1 Mixtures of Vector States in Quantum Mechanics

Let us consider the simple case where our C*-algebra is just $\mathscr{L}(\mathscr{H})$, for $\mathscr{H}$ a Hilbert space representing a non-interacting two-level system of particles, namely $\mathscr{H} = \mathbb{C}^2$, where level 1 is represented by $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and level 2 by $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. In this section, we are going to show that in general there is no vector state $\Psi \in \mathbb{C}^2$ that could represent a mixture of states $\Phi_1$ and $\Phi_2$ with respective statistical

weights $p_1$ and $p_2$, i.e. that it may be that no $\Psi$ is such that $\langle \Psi, A\Psi \rangle = p_1 \langle \Phi_1, A\Phi_1 \rangle + p_2 \langle \Phi_2, A\Phi_2 \rangle$ for all observables, that turn out to be self-adjoint operators $A \in \mathscr{L}(\mathscr{H})$ (in this case, Hermitian $2 \times 2$ complex matrices).

Take a mixture of $p_1 = \frac{1}{3}$ particles in state $\Phi_1 = \mathbf{e}_1$ and $p_2 = \frac{2}{3}$ in $\Phi_2 = \mathbf{e}_2$; put $\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$, with $\psi_1, \psi_2 \in \mathbb{C}$. In order that the mean values of an observable $A = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}$ coincide when measured for $\Psi$ and for the mixture, it is necessary that $|\psi_1|^2 = \frac{1}{3}$ and $|\psi_2|^2 = \frac{2}{3}$; we could thus choose $\psi_2 = \sqrt{\frac{2}{3}}$ and $\psi_1 = e^{i\gamma}\sqrt{\frac{1}{3}}$, with some phase $\gamma \in \mathbb{R}$ that can be fixed if we take an observable $B = \begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix}$ and impose again that the mean values of $B$ must coincide in both cases: we find out that $\gamma = \frac{\pi}{2}$. As a conclusion, one has $\Psi = \frac{1}{\sqrt{3}}\begin{pmatrix} i \\ \sqrt{2} \end{pmatrix}$.

Nonetheless, take the observable $C = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$. The mean value of $C$ for the mixture is 0, whereas $\langle \Psi, C\Psi \rangle = 2\sqrt{2}$. Therefore, no state $\Psi \in \mathbb{C}^2$ may represent the mixture consisting of $\frac{1}{3}$ of particles in the state $\mathbf{e}_1$ and $\frac{2}{3}$ of them in $\mathbf{e}_2$.

In fact, the algebra of observables of a two-level system coincides with the set of all complex $2 \times 2$ matrices. Moreover, we know (see Section 9) that a general state for this algebra is of the form $\omega_\rho(A) = \mathrm{Tr}(\rho A)$, where $\rho$, the so-called *density matrix*, is a self-adjoint and positive matrix such that $\mathrm{Tr}\rho = 1$. Since $\rho$ is a self-adjoint operator, it can be written in the form $\rho = \frac{1}{2}(a_0 \mathbb{1} + \vec{a} \cdot \vec{\sigma})$, where $\vec{a} \cdot \vec{\sigma}$ is a notation for $a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3$, with $a_k, k = 0, \ldots, 3$, being real numbers and $\sigma_l, l = 1, 2, 3$, being the Pauli matrices.[7] The condition $\mathrm{Tr}\rho = 1$ implies $a_0 = 1$. In this case, the eigenvalues of $\rho$ are $\varrho_1 = \frac{1+\|\vec{a}\|}{2}$ and $\varrho_2 = \frac{1-\|\vec{a}\|}{2}$; and hence, the condition of $\rho$ having strictly positive eigenvalues is $\|\vec{a}\| < 1$. For $\|\vec{a}\| = 1$, one has $\varrho_1 = 1$ and $\varrho_2 = 0$, implying that the matrix $\rho$ is an orthogonal projector. Hence, we can associate the space of states for a two-level system with a closed unit sphere centred at the origin in a three-dimensional space, the so-called Bloch sphere, with the pure states being those on the surface of the sphere (corresponding to $\|\vec{a}\| = 1$) and with the mixed states being inside the sphere (corresponding to $\|\vec{a}\| < 1$).

## 5.2 Experiment-Determined States in Classical Physics

Classically, the trajectory of a particle submitted to a smooth potential remains completely determined given its position

and momentum at a certain instant; besides, any other physical quantity may be expressed in terms of these data, like the energy and the angular momenta. Thus, the phase space becomes a natural environment for describing the states of classical systems, whose observables are scalar functions defined over this phase space.

Effectively, although in the classical theory the particle's states are understood as the points of the phase space, even the classical experiments alone cannot determine them with total accuracy: an experiment only furnishes a set of data permitting one to depict a (usually continuous) probability distribution with a mean value and a standard deviation, which allows us to calculate the probability of finding a particle within some region in the phase space, though not to assert that the particle will be precisely in one point or another. This does not imply that points are not states; they indeed are, since in classical mechanics, we admit the particle to have fully determined values of position and momentum, could we measure them or not. However, this rather indicates the need for more states with lesser localisation properties in order to take into account the fact that we may also have access to reality through experiments subjected to statistical errors.

Actually, it is possible to realise a commutative C*-algebra as an algebra of $C_0$ functions on some phase space (as reads Theorem 2), the states of which being linear positive and normalised functionals on such continuous functions. By the Riesz-Markov representation theorem (see e.g. [11]), a linear, bounded and positive functional $\omega$ acting on such a function $f$ can be written as an integral over a probability measure:

$$\omega(f) = \int f \, d\mu_\omega,$$

where $\mu_\omega$ is a positive measure over the phase space; the normalisation condition gives, $\int d\mu_\omega = 1$, so $\mu_\omega$ may be interpreted as a probability measure whose events are points and (open) regions of the phase space, just as one would need in order to represent classical physics as a theory seriously committed to experimental results and their intrinsic uncertainty.

Of course, this can only be done if $\mathfrak{A}$ is commutative; however, since commutativity basically means that the observables are compatible, i.e. can be measured at the same time, and in classical physics this is always the case (for it is supposed that a measurement does not modify a system's state), such a restriction on the C*-algebras is but very natural.

Notice, though, that Dirac delta measures centred at a single point in phase space of a classical mechanical system are genuine probability measures, so states of completely defined position and momentum are not excluded; on the contrary, these are precisely the classical theory's pure

---

[7]The matrices $\mathbb{1}, \sigma_1, \sigma_2$ and $\sigma_3$ are a basis in the real space of the $2 \times 2$ self-adjoint matrices.

states. Consequently, this framework does not exclude the possibility of a classical particle to have a definite position and momentum. The novelty here is that our inability of knowing them is not excluded as well, inasmuch as states are allowed that correspond to the actual results of our experiments.

In this context, it is also relevant to remark that, in classical mechanics, the variance of any observable (i.e. a measurable function in phase space) in a pure state vanishes identically, since pure states are expressed by Dirac delta measures centred at a point in phase space. Generally, this characteristic is not shared by pure states in quantum systems and this is one of the most relevant distinctions between these theories.

## 6 Pure Quantum States and Their Classical Limits

One of the physically interesting questions in this context is about what happens with the purity of a given state of a quantum system when the classical limit is taken. Naively, one could believe that the purity is preserved but, as we will now discuss, there are some interesting situations where a pure quantum state is transformed into a classical mixture.

As we have seen above, the pure states in classical physics are the one-particle well-defined position and momentum states, namely the measures $\delta_{x_0} \otimes \delta_{\xi_0}$ charging points $(x_0, \xi_0)$ of the phase space (for simplicity, we assume it to be $\mathbb{R}^2$). In the quantum picture for a system described by observables in $\mathfrak{A} = \mathscr{L}(\mathscr{H})$, with $\mathscr{H}$ a Hilbert space, given an adequate normalised vector state $\psi \in \mathscr{H}$, one can consider the following $\hbar$-dependent family of wave-packets:

$$\psi^{\hbar}(x) = \frac{1}{\hbar^{\frac{1}{4}}} \psi \left( \frac{x - x_0}{\sqrt{\hbar}} \right) e^{\frac{i}{\hbar} x \cdot \xi_0}, \quad \text{for } \hbar > 0. \tag{5}$$

These generally satisfy:

$$\lim_{\hbar \to 0} \left\langle \psi^{\hbar}, \hat{H}^{\hbar} \psi^{\hbar} \right\rangle_{\mathscr{H}} = E(x_0, \xi_0),$$

where $\hat{H}^{\hbar}$ is the usual Hamiltonian operator and $E$ is the corresponding classical energy:

$$\hat{H}^{\hbar} = -\frac{\hbar^2}{2} \Delta + V \quad \text{and} \quad E = \frac{1}{2} \xi^2 + V(x).$$

This calculation may be performed by noting that $\hat{H}^{\hbar}$ can be formally given by the integral:

$$\left( \hat{H}^{\hbar} \psi \right)(x) = \frac{1}{2\pi \hbar} \int_{\mathbb{R}_{\xi}} \int_{\mathbb{R}_y} e^{\frac{i}{\hbar} \xi \cdot (x-y)} E \left( \frac{x+y}{2}, \xi \right) \psi(y) \, dy \, d\xi,$$

whose convergence can be assured if we take $\psi$ with sufficient decay in frequency, or else if we put in place of $E$ a function $a \in C_0^{\infty}(\mathbb{R}^2)$, so as the operator $\mathrm{op}_{\hbar}(a)$ defined by

$$\left( \mathrm{op}_{\hbar}(a) \psi \right)(x) = \frac{1}{2\pi \hbar} \int_{\mathbb{R}_{\xi}} \int_{\mathbb{R}_y} e^{\frac{i}{\hbar} \xi \cdot (x-y)} a \left( \frac{x+y}{2}, \xi \right) \psi(y) \, dy \, d\xi$$

is bounded. Hence, $\mathrm{op}_h(a) \in \mathfrak{A}$ for all $\hbar > 0$ and one can show that:

$$\lim_{\hbar \to 0} \left\langle \psi^{\hbar}, \mathrm{op}_{\hbar}(a) \psi^{\hbar} \right\rangle_{\mathscr{H}} = a(x_0, \xi_0).$$

The operator $\mathrm{op}_{\hbar}(a)$ is known in the literature as the *Weyl quantisation of the symbol $a$* (see for instance [13, 14], or [9] and the references quoted therein), and the set $\Upsilon$ generated by the Weyl quantisation of symbols in $C_0^{\infty}(\mathbb{R}^2)$ is a $C^*$-subalgebra of $\mathfrak{A}$ when embedded with the Moyal or Weyl-Groenewold product, which is associative, but non-commutative; see Theorem 4.11 in [14].

With more generality, a probability measure $\mu$ on $\mathbb{R}^2$ is said to be the *Wigner* or *semiclassical measure*[9, 10, 12] associated to a normalised family of vectors $(\psi^{\hbar})_{\hbar>0} \subset \mathscr{H}$ if, for any $a \in C_0^{\infty}(\mathbb{R}^2)$, one has:

$$\lim_{\hbar \to 0} \left\langle \psi^{\hbar}, \mathrm{op}_{\hbar}(a) \psi^{\hbar} \right\rangle_{\mathscr{H}} = \int_{\mathbb{R}^2} a(x, \xi) \mu(dx, d\xi).$$

Within our framework of understanding states as linear functionals over a C*-algebra, we write down formally:

$$\mu = \underset{\hbar}{\mathrm{sclim}} \, \omega^{\hbar},$$

where $\omega^{\hbar}$ are states on the algebra $\Upsilon$ given by $\omega^{\hbar}(A) = \left\langle \psi^{\hbar}, A \psi^{\hbar} \right\rangle_{\mathscr{H}}$, for $A \in \Upsilon$. Since $\mu(\mathbb{R}^2) = 1$, one can understand $\mu$ as a classical state on a commutative C*-algebra describing the classical observables, which are $C_0^{\infty}(\mathbb{R}^2)$ functions, as in Section 5.2.

A natural question that arises at this point is whether the semiclassical limit of a family of pure states will always be pure or not. The answer was given in [8], where the author showed that a family of wave-packets not much different from (5) may split into the combination of two Dirac delta measures on the phase space. Even better: when regarding a time-dependent situation where the quantum states evolve according to a dynamical equation like Schrödinger's, and their associate Wigner measures evolve correspondingly, it is possible to have an initial state of quantum pure states concentrating to a classical pure state, and keeping like that for a while, but then degenerating into a classical mixture.

More precisely, it was shown in Theorem 1.12 of [8] that the pure states $\Psi_t^{\hbar}$ which are the propagation under $V(x) = -|x|$ of initial data:

$$\Psi_0^{\hbar}(x) = \frac{p_1}{\hbar^{\frac{1}{4}}} \Psi^1 \left( \frac{x}{\sqrt{\hbar}} \right) + \frac{p_2}{\hbar^{\frac{1}{4}}} \Psi^2 \left( \frac{x}{\sqrt{\hbar}} \right) e^{-i\hbar^{\beta-1} x},$$
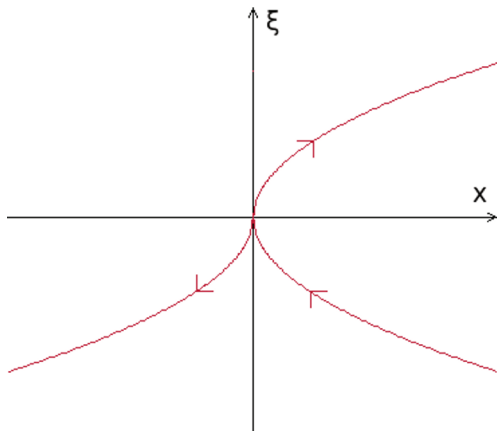
**Fig. 1** Trajectories followed by a pure state that splits into two at $t = 0$ on the phase-space origin, corresponding to the semiclassical limit of $\Psi_t^\hbar$; the arrows indicate the time's increasing direction. Source: [8]

with $p_1^2 + p_2^2 = 1$, will behave semiclassically, under appropriate choices of parameters,[8] as

$$\mu_t(x, \xi) = \delta\left(x - \frac{t^2}{2}\right) \otimes \delta\left(\xi + t\right)$$

for $t \leqslant 0$, so being a classical pure state, whereas for $t > 0$ we will have:

$$\mu_t(x, \xi) = p_1^2 \delta\left(x - \frac{t^2}{2}\right) \otimes \delta\left(\xi - t\right) + p_2^2 \delta\left(x + \frac{t^2}{2}\right) \otimes \delta\left(\xi + t\right),$$

which is a non-trivial convex combination of two pure classical states, thus a mixture. In Fig. 1, we can observe the trajectories followed by the semiclassical measures of the quantum states $\Psi_t^\hbar$, which are always, we recall, pure, as any vector state is.

## 7 Pure States and Irreducibility

We will now turn back to our general analysis and consider the important relation between purity and irreducibility of certain representations of the algebra of observables in quantum systems. This discussion is of particular relevance for the treatment of superselection sectors in quantum systems with infinitely many degrees of freedom, as in quantum field theory or quantum statistical mechanics.

It is possible to determine if a state $\omega$ on a C*-algebra $\mathfrak{A}$ is pure or mixed by analysing the reducibility of a certain representation of $\mathfrak{A}$ into the space of bounded operators on a Hilbert space suitably constructed from $\omega$ and $\mathfrak{A}$. This is achieved by means of the *GNS construction* (named after Gelfand, Naimark and Segal), which we sketch below.

[8]With $0 < \beta < \frac{1}{10}$, $\Psi^1, \Psi^2 \in C_0^\infty(\mathbb{R})$, $\Psi^1$ supported on $x > 0$ and $\operatorname{supp} \Psi^1 \cap \operatorname{supp} \Psi^2 = \emptyset$.

To begin with, define:

$$\mathcal{N}_\omega = \left\{ A \in \mathfrak{A}, \; \omega(A^*A) = 0 \right\}.$$

$\mathcal{N}_\omega$ is a vector subspace of $\mathfrak{A}$, as it may be verified by the Cauchy–Schwartz inequality (see e.g. [6]):

$$\left| \omega\left(A^*B\right) \right|^2 \leqslant \omega(A^*A)\omega(B^*B), \qquad (6)$$

valid for any $A, B \in \mathfrak{A}$ and positive linear functionals $\omega$. Thus, one obtains a well-defined inner product $\langle \cdot, \cdot \rangle_\omega$ on the quotient space $\mathfrak{A}/\mathcal{N}_\omega$ of equivalence classes:

$$\psi_A = \left\{ \hat{A} \; : \; \hat{A} = A + I, \; I \in \mathcal{N}_\omega \right\}$$

by posing $\langle \psi_A, \psi_B \rangle_\omega = \omega(A^*B)$ (its independence with respect to the classes' representatives can be verified once again by means of the Cauchy–Schwartz inequality).

The canonical completion of $\mathfrak{A}/\mathcal{N}_\omega$ with respect to the inner product $\langle \cdot, \cdot \rangle_\omega$, denoted $\mathscr{H}_\omega$, is called the representation space of $\mathfrak{A}$ for the state $\omega$.

Now, let us remark that $\mathcal{N}_\omega$ is also a left ideal of $\mathfrak{A}$,[9] so the linear operator $\pi_\omega(A)$ acting on the dense subspace $\mathfrak{A}/\mathcal{N}_\omega$ of $\mathscr{H}_\omega$ as $\pi_\omega(A)\psi_B = \psi_{AB}$ is well-defined; moreover:

$$\|\pi_\omega(A)\psi_B\|^2 = \omega(B^*A^*AB)$$
$$\leqslant \|A\|^2 \omega(B^*B) = \|A\|^2 \|\psi_B\|^2$$

(where we have used the same inequality as in footnote 9), showing that $\pi_\omega(A)$ is bounded and, therefore, may be continuously extended to the whole $\mathscr{H}_\omega$. Furthermore, it is easy to see that, for any $A_1, A_2 \in \mathfrak{A}$ and $\lambda \in \mathbb{C}$, we have $\pi_\omega(A_1)\pi_\omega(A_2) = \pi_\omega(A_1A_2)$, $\pi_\omega(A_1 + \lambda A_2) = \pi_{\omega A_1} + \lambda \pi_\omega(A_2)$ and $\pi_\omega\left(A_1^*\right) = \pi_\omega(A_1)^*$. It is sufficient to check these claims for vectors of the form $\psi_B$ with $B \in \mathfrak{A}$, since they are dense in $\mathscr{H}_\omega$; so, given any $B \in \mathfrak{A}$:

$$\pi_\omega(A_1)\pi_\omega(A_2)\psi_B = \psi_{A_1A_2B} = \pi_\omega(A_1A_2)\psi_B,$$

now, using the usual vector space operations for the quotient space:

$$\pi_\omega(A_1 + \lambda A_2)\psi_B = \psi_{(A_1+\lambda A_2)B} = \psi_{A_1B} + \lambda\psi_{A_2B}$$
$$= (\pi_\omega(A_1) + \lambda\pi_\omega(A_2))\psi_B,$$

and last

$$\left\langle \pi_\omega(B^*)\Psi_{A_1}, \Psi_{A_2} \right\rangle_\omega = \omega((B^*A_1)^*A_2) = \omega(A_1^*BA_2)$$
$$= \left\langle \pi_\omega(B)^*\Psi_{A_1}, \Psi_{A_2} \right\rangle_\omega,$$

leading us to the conclusion that $\pi_\omega : \mathfrak{A} \longrightarrow \mathscr{L}(\mathscr{H}_\omega)$ is actually a representation of the C*-algebra $\mathfrak{A}$ in the Hilbert space $\mathscr{H}_\omega$.

Finally, put $\Omega_\omega = \psi_{\mathbb{1}}$ if $\mathfrak{A}$ has an identity.[10]

[9] For any $A, B \in \mathfrak{A}$, one has $\omega((AB)^*(AB)) \leqslant \|A\|^2 \omega(B^*B)$ (see [6]), so clearly $AB \in \mathcal{N}_\omega$ whenever $B \in \mathcal{N}_\omega$.

[10]If $\mathfrak{A}$ does not have an identity, similar results can be obtained using nets converging to a unity. See e.g. [6].

Not only will we have that the set

$$\{\pi_\omega(A)\Omega_\omega \, : \, A \in \mathfrak{A}\}$$

is the dense in $\mathscr{H}_\omega$,[11] for $\pi_\omega(A)\Omega_\omega = \psi_A$, but also that

$$\langle\Omega_\omega, \pi_\omega(A)\Omega_\omega\rangle_\omega = \langle\psi_{\mathbb{1}}, \psi_A\rangle_\omega = \omega(A)$$

for any $A \in \mathfrak{A}$, i.e. within this particular representation crafted for $\omega$, this state appears as a vector state $\Omega_\omega$.

The triple $(\mathscr{H}_\omega, \pi_\omega, \Omega_\omega)$ is called the GNS representation of the C*-algebra $\mathfrak{A}$ for the state $\omega$.

Let be $\mathfrak{A}$ a C*-algebra and $\omega$ a state on it. A triple $(\mathscr{H}, \pi, \Omega)$ consisting of a Hilbert space $\mathscr{H}$, a representation $\pi$ of $\mathfrak{A}$ in $\mathscr{L}(\mathscr{H})$ and a vector $\Omega \in \mathscr{H}$ such that the set $\{\pi(A)\Omega : A \in \mathfrak{A}\}$ is dense in $\mathscr{H}$ and that $\langle\Omega, \pi(A)\Omega\rangle = \omega(A)$ is called a *cyclic representation* of $\mathfrak{A}$ for the state $\omega$.

As we see, for any state $\omega$ over the C*-algebra $\mathfrak{A}$, there exists at least one cyclic representation, the GNS one, $(\mathscr{H}_\omega, \pi_\omega, \Omega_\omega)$. Actually, it is not difficult to verify that the cyclic representations are unique up to unitary equivalences. See, for instance, Theorem 2.3.16 of [6] for details.

Yet, given that we ended up representing $\omega$ as a vector state despite of it being mixed or not, we may ask: is there a way to discover whether $\omega$, as a functional on $\mathfrak{A}$, was a pure state or a mixture by analysing its cyclic representations?

Now intervenes the notion of irreducibility. A representation $\pi$ of an algebra $\mathfrak{A}$ into a vector space $V$ is said to be irreducible if there is no closed subspace $U \subset V$ left invariant by the action of $\pi$, apart from the trivial spaces $U = \{0\}$ and $U = V$. Said otherwise, $\pi$ is irreducible if, should $\pi(A)U \subset U$ for $U \subset V$ closed and every $A \in \mathfrak{A}$, then either $U = \{0\}$ or $U = V$. If $\pi$ is not irreducible, it is said to be reducible.

The following theorem is of central importance for this discussion.

**Theorem 5** *Let $\mathfrak{A}$ be a C*-algebra and $\omega$ a state on it. Then, $\omega$ is pure if, and only if, given a cyclic representation $(\mathscr{H}, \pi, \Omega)$ of $\mathfrak{A}$ for $\omega$, $\pi$ is irreducible.*

A proof can be found in references like [6, 15]. For the convenience of the reader, we present it in Appendix 1.

As we see, purity manifests itself in irreducibility of the cyclic representation of the algebra. On the other hand, the fact that the cyclic representation of a mixed state is reducible (and, therefore, can be further decomposed into irreducible ones, as discussed in Section 11) means that such states can be interpreted as being built by elementary subsystems. This result is of major importance

---

for understanding the distinction between pure states and mixtures.

## 8 Purification of States

Given a quantum state defined in a separable Hilbert space $\mathscr{H}_I$ by a density matrix $\rho$, it is possible to find another (not uniquely defined) separable Hilbert space $\mathscr{H}_{II}$ such that the original state can be represented as a normalised *vector* state $\Psi_\rho$ in the tensor product space $\mathscr{H}_I \otimes \mathscr{H}_{II}$. Therefore, the vector state defined by $\Psi_\rho$ is a pure state in the enlarged algebra $\mathscr{L}(\mathscr{H}_I \otimes \mathscr{H}_{II})$ but, except when $\rho$ is a one-dimensional projection, it is not a pure state for the original algebra $\mathscr{L}(\mathscr{H}_I) \simeq \mathscr{L}(\mathscr{H}_I) \otimes \mathbb{1}_{\mathscr{H}_{II}}$. The vector $\Psi_\rho \in \mathscr{H}_I \otimes \mathscr{H}_{II}$ is called a *purification* of the state defined by $\rho$.

These facts, which we are going to establish below, have an interesting physical interpretation, since they say that a given mixed state of a quantum system can always be thought of as a pure state of a larger quantum system.

The vector $\Psi_\rho$ is not to be confused with the GNS vector presented above, since the vector $\Psi_\rho$ is not cyclic and separating for the original algebra $\mathscr{L}(\mathscr{H}_I) \simeq \mathscr{L}(\mathscr{H}_I) \otimes \mathbb{1}_{II}$ and since the algebra $\mathscr{L}(\mathscr{H}_I \otimes \mathscr{H}_{II})$ is indeed "larger" than the "original" algebra $\mathscr{L}(\mathscr{H}_I) \otimes \mathbb{1}_{II}$.

Let $\mathscr{H}_I$ be a separable Hilbert space and let $\rho$ be a density matrix acting on $\mathscr{H}_I$, i.e. a bounded, trace class, self-adjoint and positive operator acting on $\mathscr{H}_I$ with $\mathrm{Tr}_{\mathscr{H}_I}(\rho) = 1$. The expression $\mathscr{L}(\mathscr{H}_I) \ni D \mapsto \mathrm{Tr}_{\mathscr{H}_I}(\rho D) \in \mathbb{C}$ represents a state on the C*-algebra $\mathscr{L}(\mathscr{H}_I)$.

Consider any other separable Hilbert space $\mathscr{H}_{II}$ with $\dim(\mathscr{H}_{II}) \geq \dim(\mathscr{H}_I)$. We can find a normalised vector $\Psi_\rho \in \mathscr{H} := \mathscr{H}_I \otimes \mathscr{H}_{II}$ such that:

$$\begin{aligned}\mathrm{Tr}_{\mathscr{H}_I}(\rho D) &= \langle\Psi_\rho, (D \otimes \mathbb{1}_{\mathscr{H}_{II}})\Psi_\rho\rangle_{\mathscr{H}} \\ &= \mathrm{Tr}_{\mathscr{H}}\left(P_{\Psi_\rho}(D \otimes \mathbb{1}_{\mathscr{H}_{II}})\right),\end{aligned} \tag{7}$$

where $P_{\Psi_\rho}$ is the projector on the subspace generated by $\Psi_\rho$.

Let $\varrho_k$, $k \in \mathbb{N}$ be the eigenvalues of $\rho$ (including multiplicity), with $\varrho_k \geq 0$ for all $k$, and let $\mathbf{v}_k$, $k \in \mathbb{N}$ be the corresponding normalised eigenvectors, building a complete orthonormal basis in $\mathscr{H}_I$: $\langle\mathbf{v}_i, \mathbf{v}_j\rangle_{\mathscr{H}_I} = \delta_{i,j}$. The spectral decomposition of $\rho$ is $\rho = \sum_{i=1}^\infty \varrho_i P_{\mathbf{v}_i}$, where $P_{\mathbf{v}_i}$ is the orthogonal projector on the one-dimensional subspace generated by $\mathbf{v}_i$.

Let $\{\mathbf{w}_l, \, l \in \mathbb{N}\}$ be an arbitrary orthonormal set of vectors (not necessarily complete) in $\mathscr{H}_{II}$ (i.e. with $\langle\mathbf{w}_i, \mathbf{w}_j\rangle_{\mathscr{H}_{II}} = \delta_{i,j}$). Define $\Psi_\rho \in \mathscr{H}$ by

$$\Psi_\rho := \sum_{i=1}^\infty \sqrt{\varrho_i}\,(\mathbf{v}_i \otimes \mathbf{w}_i). \tag{8}$$

The sequence in the r.h.s. converges in $\mathscr{H}$, because the vectors $\mathbf{v}_i \otimes \mathbf{w}_i$, $i \in \mathbb{N}$, are orthonormal and because $\{\sqrt{\varrho_i},\ i \in \mathbb{N}\}$ is a square summable numerical sequence, since $\sum_{i=1}^{\infty} \varrho_i = \mathrm{Tr}_{\mathscr{H}_{\mathrm{I}}}(\rho) = 1$. We have,

$$
\begin{aligned}
\langle \Psi_\rho,\ (D \otimes \mathbb{1}_{\mathscr{H}_{\mathrm{II}}}) \Psi_\rho \rangle_{\mathscr{H}} &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sqrt{\varrho_i} \sqrt{\varrho_j} \langle \mathbf{v}_i \otimes \mathbf{w}_i,\ (D\mathbf{v}_j) \otimes \mathbf{w}_j \rangle_{\mathscr{H}} \\
&= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sqrt{\varrho_i} \sqrt{\varrho_j} \langle \mathbf{v}_i,\ D\mathbf{v}_j \rangle_{\mathscr{H}_{\mathrm{I}}} \underbrace{\langle \mathbf{w}_i,\ \mathbf{w}_j \rangle_{\mathscr{H}_{\mathrm{II}}}}_{\delta_{i,j}} \\
&= \sum_{i=1}^{\infty} \varrho_i \langle \mathbf{v}_i,\ D\mathbf{v}_i \rangle_{\mathscr{H}_{\mathrm{I}}} = \mathrm{Tr}_{\mathscr{H}_{\mathrm{I}}}(\rho D),
\end{aligned}
$$

establishing (7). For $D = \mathbb{1}_{\mathscr{H}_{\mathrm{I}}}$, in particular, this relation shows that $\|\psi_\rho\|_{\mathscr{H}} = 1$.

The vector $\Psi_\rho$ defined in (8) depends on $\rho$ (through the eigenvalues $\varrho_i$ and the eigenvectors $\mathbf{v}_i$) and on the arbitrary choice of the orthonormal vectors $\mathbf{w}_j$ of $\mathscr{H}_{\mathrm{II}}$. The l.h.s. of (7), however, does not depend on the choice of the $\mathbf{w}_j$'s.

It is relevant to notice that, except when $\rho$ is a one-dimensional projection, the vector state defined by $\Psi_\rho$ is not a pure state for the algebra $\mathscr{L}(\mathscr{H}_{\mathrm{I}}) \otimes \mathbb{1}_{\mathrm{II}} \simeq \mathscr{L}(\mathscr{H}_{\mathrm{I}})$. It is clear from (7) that, for two density matrices $\rho$ e $\rho'$ and for $\lambda \in [0,\ 1]$, one has:

$$
\begin{aligned}
&\langle \Psi_{\lambda\rho+(1-\lambda)\rho'},\ (D \otimes \mathbb{1}_{\mathscr{H}_{\mathrm{II}}}) \Psi_{\lambda\rho+(1-\lambda)\rho'} \rangle_{\mathscr{H}} \\
&= \lambda \langle \Psi_\rho,\ (D \otimes \mathbb{1}_{\mathscr{H}_{\mathrm{II}}}) \Psi_\rho \rangle_{\mathscr{H}} + (1-\lambda) \langle \Psi_{\rho'},\ (D \otimes \mathbb{1}_{\mathscr{H}_{\mathrm{II}}}) \Psi_{\rho'} \rangle_{\mathscr{H}}.
\end{aligned}
$$

From this, it also follows that

$$
\langle \Psi_\rho,\ (D \otimes \mathbb{1}_{\mathscr{H}_{\mathrm{II}}}) \Psi_\rho \rangle_{\mathscr{H}} = \sum_{i=1}^{\infty} \varrho_i \left\langle \Psi_{P_{\mathbf{v}_i}},\ (D \otimes \mathbb{1}_{\mathscr{H}_{\mathrm{II}}}) \Psi_{P_{\mathbf{v}_i}} \right\rangle_{\mathscr{H}},
$$

since $\Psi_{P_{\mathbf{v}_i}} = \mathbf{v}_i \otimes \mathbf{w}_i$. Hence, except when $\rho$ is a one-dimensional projection, the vector state defined by $\Psi_\rho$ is not a pure state for the algebra $\mathscr{L}(\mathscr{H}_{\mathrm{I}}) \simeq \mathscr{L}(\mathscr{H}_{\mathrm{I}}) \otimes \mathbb{1}_{\mathscr{H}_{\mathrm{II}}}$.

One can easily show that $\rho = \mathrm{Ptr}_{\mathscr{H}_{\mathrm{II}}}(P_{\Psi_\rho})$, where $\mathrm{Ptr}_{\mathscr{H}_{\mathrm{II}}}$ denotes the partial trace with respect to the Hilbert space $\mathscr{H}_{\mathrm{II}}$.

A relevant question is whether a physical process (through a completely positive map) can be identified leading to one of the purifications associated to a given state. Such a process is known as *physical purification* and we refer the reader to [22] for further discussions on this issue.

## 9 Normal States and Density Matrices

We now come to the important relation between purity and normality of states. In spite of being mathematically a more technical discussion, it is of central importance to physics due to its relation to the notion of density matrices and to other issues.

The underlying question relevant to physics is as follows: under which circumstances can a state $\omega$ be defined by a density matrix, i.e. can be written in the form $\omega(A) = \mathrm{Tr}(\rho A)$?

Definition 2 gives rise to far more exotic states than those we are used to, represented by density operators when $\mathfrak{A} \subset \mathscr{L}(\mathscr{H})$ for some separable Hilbert space $\mathscr{H}$. In this case, these density operators amount to an important part of the set of all states: as a corollary of Theorem 8, the density operators are dense in the whole set of states with respect to the weak operator topology, also known as physical topology (details ahead). As a result, we are led to looking for further characterisations regarding states in order to better understand them.

In particular, we will see that density operators are the realisation of *normal states* in the case where $\mathfrak{A} \subset \mathscr{L}(\mathscr{H})$.

Recall that an increasing net of operators is a net $(A_\nu)_{\nu \in I} \subset \mathfrak{A}^{12}$ such that $A_\nu \geqslant A_\mu$ whenever $\nu \geqslant \mu$; the operators' order relation is defined in the following way: $A_\nu \geqslant A_\mu$ if $A_\nu - A_\mu$ is a positive operator, i.e. if its spectrum is included in $\mathbb{R}^+$.

**Definition 3** A state (or any positive functional) $\omega$ on a C*-algebra $\mathfrak{A}$ is said to be normal if

$$
\omega\left(\sup_{\nu \in I} A_\nu\right) = \sup_{\nu \in I} \omega(A_\nu)
$$

for any bounded increasing net of positive operators $(A_\nu)_{\nu \in I} \subset \mathfrak{A}$.

As one can see, the definition of normal states heavily depends on the notion of operator order and seems to express some kind of "order continuity" property. Indeed, one may wonder whether there exists some topology in $\mathfrak{A}$ for which a state being normal would merely mean that it is continuous. A topology as such does exist, actually more than one. Let us define them.

**Definition 4** Let $\mathscr{H}$ be a Hilbert space. We define the following families of seminorms on $\mathscr{L}(\mathscr{H})$:

– **Ultra-strong seminorms** For a sequence $(\psi_n)_{n \in \mathbb{N}} \subset \ell_2(\mathscr{H})$:

$$
\|A\|_\psi := \left(\sum_{n=1}^{\infty} \|A\psi_n\|^2\right)^{\frac{1}{2}}.
$$

---

[12]$I$ is the *index set*, the set where the indices take their values from.

– **Ultra-weak seminorms** For a pair of sequences $(\psi_n)_{n\in\mathbb{N}}$, $(\phi_n)_{\mathbb{N}} \subset \ell_2(\mathscr{H})$:

$$\|A\|_{\psi,\phi} := \left(\sum_{n=1}^{\infty} |\langle \psi_n, A\phi_n \rangle|^2\right)^{\frac{1}{2}}.$$

The topologies induced on $\mathscr{L}(\mathscr{H})$ by these families are known, respectively, as the ultra-strong operator topology and the ultra-weak operator topology.

Although these two topologies are distinct, being ultra-weakly or ultra-strongly continuous is equivalent for linear functionals.

**Theorem 6** *Let $\mathfrak{A} \subset \mathscr{L}(\mathscr{H})$ be a C\*-algebra, and $\omega$ a positive functional on it. The following conditions are equivalent:*

*(i)   $\omega$ is normal;*
*(ii)   $\omega$ is ultra-strongly continuous;*
*(iii)   $\omega$ is ultra-weakly continuous;*
*(iv)   there exists $(\phi_n)_{n\in\mathbb{N}} \subset \mathscr{H}$ with $\sum_{n=1}^{\infty} \|\phi_n\|^2 < \infty$ such that*

$$\omega = \sum_{n=1}^{\infty} \omega_{\phi_n}, \quad \text{(convergence in norm),} \qquad (9)$$

*where $\omega_{\phi_n}$ is given, for any $A \in \mathfrak{A}$, by $\omega_{\phi_n}(A) = \langle \phi_n, A\phi_n \rangle$.*

A proof of this theorem can be found in [6, 7].

If $\mathscr{H}$ is a separable Hilbert space, it is now easy to conclude from (9) that a normal state on $\mathfrak{A} \subset \mathscr{L}(\mathscr{H})$ can be represented by:

$$\omega(A) = \text{Tr}(\rho A), \qquad \forall A \in \mathfrak{A}, \qquad (10)$$

where

$$\rho = \sum_{n=1}^{\infty} \|\phi_n\|^2 |\Phi_n\rangle \langle \Phi_n|,$$

$(\phi_n)_{n\in\mathbb{N}} \subset \mathscr{H}$ being the sequence obtained in the above theorem, and $\Phi_n = \frac{\phi_n}{\|\phi_n\|}$.

Besides, since a state is by hypothesis normalised, it follows that we will also have:

$$\text{Tr}(\rho) = \sum_{n=1}^{\infty} \|\phi_n\|^2 = 1,$$

so $\rho$ is a genuine density operator.

**Theorem 7** *Let $\mathscr{H}$ be separable. Normal states are realised by density operators when $\mathfrak{A}$ is mapped into $\mathscr{L}(\mathscr{H})$. Conversely, any states realised by density operator is a normal state.*

*Proof* From Theorem 1, it is known that $\mathfrak{A}$ may be isomorphically mapped onto a closed subalgebra $\tilde{\mathfrak{A}} \subset \mathscr{L}(\mathscr{H})$, so the states on $\mathfrak{A}$ may be mapped onto the states on $\tilde{\mathfrak{A}}$. By Theorem 6 and the subsequent discussion, the proof is complete.                                                    $\square$

Now, this section's central result:

**Theorem 8** *Normal states are dense in the set of all states according to the the weak (physical) topology, i.e. given any state $\omega$ on a C\*-algebra $\mathfrak{A}$, there exists a sequence of normal states $(\omega_n)_{n\in\mathbb{N}} \subset \mathfrak{A}$ such that, for any $A \in \mathfrak{A}$, $\omega(A) = \lim_n \omega_n(A)$.*

*Proof* Theorem 1.1 in [16].                                                    $\square$

*Remark 2* As we saw in Section 2.1, an actual state $\omega$ of a physical system may be determined by a process of taking limits within the physical topology. In view of the previous theorem, we may only consider normal states to partially describe $\omega$ within weak neighbourhood $\Omega_\varepsilon(A_1, \ldots, A_k; \omega)$; however, it is important to bear in mind that, even in this case, the limit, i.e. $\omega$, may end up being non-normal.

If a state is normal, it is easier to identify it as pure or mixed:

**Theorem 9** *Let $\mathfrak{A}$ be a C\*-algebra isomorphic to $\tilde{\mathfrak{A}} \subset \mathscr{L}(\mathscr{H})$ for some separable Hilbert space $\mathscr{H}$. Let $\omega$ be a normal state on $\mathfrak{A}$ and $\rho \in \tilde{\mathfrak{A}}$ the density operator associated to $\omega$ (see Theorem 7 above). Then $\rho$ has Hilbert-Schmidt norm*

$$\|\rho\|_{HS} = \sqrt{Tr(\rho^*\rho)} = 1$$

*if and only if $\omega$ is pure. If $\omega$ is a mixed state, $\|\rho\|_{HS} < 1$.*

*Proof* Since $\rho$ is a density operator, it is positive and compact, so it is possible to choose a Hilbertian basis $(e_n)_{n\in\mathbb{N}} \subset \mathscr{H}$ and a sequence of scalars $\varrho_n \geqslant 0$ such that $\rho = \sum_{n=1}^{\infty} \varrho_n |e_n\rangle \langle e_n|$. Given that $\text{Tr}(\rho) = \sum_{n=1}^{\infty} \varrho_n = 1$, we must have $\varrho_n \leqslant 1$ for every $n \in \mathbb{N}$, which implies $\varrho_n^2 \leqslant \varrho_n$, the equality holding only in the case $\varrho_n = 0$ or $\varrho_n = 1$; as a consequence, calculating its Hilbert-Schmidt norm:

$$\|\rho\|_{HS}^2 = \text{Tr}(\rho^*\rho) = \sum_{n=1}^{\infty} \varrho_n^2,$$

it becomes clear that $\|\rho\|_{HS}$ equals 1 if and only if there is exactly one $\varrho_n$ non-null, say for $n = N$. Thus, we have $\rho = |e_N\rangle \langle e_N| \in \mathfrak{A}$, which means that $\omega$ is mapped to a state

that acts on $A \in \tilde{\mathfrak{A}}$ as $\mathrm{Tr}\,(\rho A) = \langle e_N,\, A e_N \rangle$, a vector state. By Theorem 4, states of this form are always pure, implying the same for $\omega$ itself. Conversely, if $\|\rho\|_{HS} < 1$ strictly, then we have at least two non-zero $\varrho_n$'s; calling one of them $\varrho_N$, one has:

$$\rho = \lambda\,|e_N\rangle\,\langle e_N| + (1-\lambda)\rho',$$

with $0 < \lambda = \varrho_N < 1$ and $\rho' = \frac{1}{1-\varrho_N}\sum_{\substack{n=1 \\ n\neq N}}^{\infty}\varrho_n\,|e_n\rangle\,\langle e_n|$.

Since $\rho'$ also represents a state, for it is a density operator, we have that $\rho$ is a non-trivial mixture, so it is not pure.

The converse affirmations, i.e. that being pure implies $\|\rho\|_{HS} = 1$, and being a mixture $\|\rho\|_{HS} < 1$, are also true, as any density operator has Hilbert-Schmidt norm within $(0, 1]$. $\qquad\square$

As a side remark, we would like to highlight that it is not trivial to see that $|e_N\rangle\,\langle e_N|$ and $\rho'$ define states on $\tilde{\mathfrak{A}}$, since the spectral projections of $\rho$ have no reason to belong to $\tilde{\mathfrak{A}}$. Fortunately, the spectral projections are elements of the closure of $\tilde{\mathfrak{A}} \subset \mathscr{L}(\mathscr{H})$ in the weak operator topology (wot), hence $|e_N\rangle\,\langle e_N|$ and $\rho'$ define normal states in the von Neumann algebra $\overline{\tilde{\mathfrak{A}}}^{wot}$, the closure of $\tilde{\mathfrak{A}}$ in the weak operator topology. Finally, as seen in Theorem 6, normal operators are (ultra-)weakly continuous and the conclusion holds.

We should emphasise that normal states may be pure or mixed. From their characterisation given in Theorem 9, normal pure states acting on $\mathfrak{M} \subset \mathscr{L}(\mathscr{H})$ are density operators on $\mathscr{H}$ with unitary Hilbert-Schmidt norm, and from the very proof of this result one sees that:

**Theorem 10** *Any normal pure state on the* C*-algebra $\mathscr{L}(\mathscr{H})$ *is a vector state.*

This last result justifies a very common statement, found in many quantum mechanics textbooks, that pure states are those whose Hilbert-Schmidt norm equals 1, and mixtures those with $\|\rho\|_{HS} < 1$. This is only correct for normal states, which tells us not the whole picture, as we will just see in the next section.

# 10 More Issues About Purity

In this section, we discuss some important issues and examples concerning the relation between pure and vector states.

## 10.1 A Pure State That Is Not a Vector State

Here, we will extend the study in Section 4.2 by exhibiting an example of a pure state that is not a vector state. For

simplicity, let us suppose that $\mathscr{H}$ is a separable infinite-dimensional Hilbert space with orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$. Let $(a_n)_{n\in\mathbb{N}} \subset [0, 1)$ be a sequence such that $a_n \underset{n\to\infty}{\longrightarrow} 1$, for instance: $a_n = 2^{-\frac{1}{n}}$.

Now, define an operator $A \in \mathscr{L}(\mathscr{H})$ acting on a vector $\Psi = \sum_{n=1}^{\infty}\psi_n e_n \in \mathscr{H}$ as:

$$A\Psi = \sum_{n\in\mathbb{N}}a_n\,\psi_n e_n.$$

Clearly, $\|A\Psi\| < \|\Psi\|$ for any $\Psi \in \mathscr{H}$, and together with $\|Ae_n\| \longrightarrow 1$, we obtain $\|A\| = 1$. Finally, Theorem 5.1.11 in [15] states that there exists a pure state $\omega$ on $\mathscr{L}(\mathscr{H})$ such that $\omega(A) = \|A\| = 1$; we claim that this pure state cannot be vector. In fact, if it were the case, we would have, for some $\phi \in \mathscr{H}$ with $\|\phi\| = 1$:

$$1 = \omega(A) = \langle \phi,\, A\phi \rangle \leqslant \|A\phi\| < 1,$$

which is absurd.

## 10.2 All Pure States on $\mathscr{K}(\mathscr{H})$ Are Vector States

The present example (inspired in Section 5.1.1 of [15]) contrasts with the previous one. Above, we have shown that there may be in general pure states which are not vector; here we will see the opposite, i.e. a special case where all pure states are also vector, stressing the importance of the particular algebra that we take for observables.

To begin with, noting by $\mathscr{K}(\mathscr{H})$ the algebra of compact operators on a separable Hilbert space $\mathscr{H}$, it is known that its dual is composed by the set of trace class operators acting on $\mathscr{H}$, in symbols: $\mathscr{K}(\mathscr{H})^* = \mathscr{L}_1(\mathscr{H})$. For the reader's convenience, let us quickly proof this fact by remarking that the linear function $\mathscr{L}_1(\mathscr{H}) \ni A \longmapsto \mathrm{Tr}_A \in \mathscr{K}(\mathscr{H})^*$, where $\mathrm{Tr}_A(K) = \mathrm{Tr}\,(AK)$ for any $K \in \mathscr{K}(\mathscr{H})$, is an isometric isomorphism.

Indeed, taking an element $\omega \in \mathscr{K}(\mathscr{H})^*$, Riesz's representation theorem implies that there is a bounded operator $A_\omega \in \mathscr{L}(\mathscr{H})$ such that the sesquilinear form:

$$\mathscr{H} \times \mathscr{H} \ni (x, y) \longmapsto \omega\,(|x\rangle\,\langle y|) \in \mathbb{C}$$

can be written as $\omega\,(|x\rangle\,\langle y|) = \langle x,\, A_\omega y\rangle$. $A_\omega$ is trace-class, for picking up a Hilbertian basis $\{e_n\}_{n\in\mathbb{N}}$ of $\mathscr{H}$:

$$\mathrm{Tr}\,(A_\omega) = \sum_{n=1}^{\infty}\langle e_n,\, A_\omega e_n\rangle = \sum_{n=1}^{\infty}\omega\,(|e_n\rangle\,\langle e_n|)$$
$$= \sum_{n=1}^{\infty}\omega\,(\mathbb{1}) \leqslant \|\omega\|.$$

Using continuity and linearity of $\omega$, denseness of the finite-rank operators in $K(\mathscr{H})$ and further remarks about the injectivity $A \mapsto \mathrm{Tr}_A$, we obtain the desired duality.

Now, concerning a state $\omega$ on $\mathscr{K}(\mathscr{H})$, it is easy to see that the corresponding $A_\omega$ will be positive; since it is also

compact (as trace-class implies compact), there exists a Hilbertian basis $\{e_n\}_{n\in\mathbb{N}}$ of $\mathscr{H}$ for which $A_\omega$ is diagonal, i.e. $A_\omega e_n = \lambda_n e_n$, with $\lambda_n \geqslant 0$, and $\sum_{n=1}^{\infty} \lambda_n = 1$ (this sum comes from the normalisation of $\omega$). For $K \in \mathscr{K}(\mathscr{H})$, we have:

$$\omega(K) = \mathrm{Tr}\,(A_\omega K) = \sum_{n=1}^{\infty} \lambda_n \langle e_n, K e_n \rangle. \tag{11}$$

Hence, any state $\omega$ on $\mathscr{K}(\mathscr{H})$ is a convex combination of states like $\langle e_n, K e_n \rangle$; if $\omega$ is pure, then $\lambda_N = 1$ for some $N \in \mathbb{N}$ and $\lambda_n = 0$ for $n \neq N$, so we conclude that it is also a vector state.

## 11 Krein-Milman's and Choquet's Theorems

From the very beginning, we have been talking about pure states, but until now we have not answered a crucial question: do they exist?

Notice that pure states constitute some kind of "fundamental brick" in the construction of states, that is, if we have a state $\omega$ we can wonder if it is a mixed state. Then, if it is a mixed state, we have $\omega = \lambda_1 \omega_1 + \lambda_2 \omega_2$ for two distinct states $\omega_1$ and $\omega_2$ and $\lambda_1, \lambda_2 \in (0, 1)$, and we can wonder now if $\omega_1$ and $\omega_2$ are themselves mixed states. Proceeding this way, after some steps, we write the original state as a convex combination $\omega = \sum_{i=1}^{n} \lambda_i \omega_i$, with $\lambda_i \in (0, 1)$ and $\sum_{i=1}^{n} \lambda_i = 1$. This procedure is very similar to the one used to prove that a positive integer number has a prime decomposition, but there is a very important difference because you cannot divide positive integers forever by divisors bigger than 1. This difference creates the possibility that the process we suggested for decomposing states never stops, leaving unanswered the question about the very existence of pure states (apart from some concrete examples, like the cases where $\mathfrak{A} = \mathcal{L}(\mathscr{H})$, where it is known that vector states are pure).

Is there a way to circumvent the problem of our infinite process appealing for topology, that means can we assure that at least the sequence obtained by our steps is convergent, in which case we could write $\omega = \sum_{i=1}^{\infty} \lambda_i \omega_i$, with $\lambda_i \in (0, 1)$ and $\sum_{i=1}^{\infty} \lambda_i = 1$?

Fortunately, it is possible to prove that there exist pure states and, as our previous discussion suggests, that they exist in such a number that all states can be written as limits (in a suitable topology) of convex combination of them. This is Krein-Milman's theorem.

In order to state Krein-Milman's theorem, we need to define what is a face and what is an extremal point.

**Definition 5** Let $V$ be a topological vector space and $C \subset V$ be a non-empty convex subset. A non-empty closed and convex subset $F \subset C$ is said to be a face (or extremal) set of $C$ if, given $x, y \in C$ and $\lambda \in (0, 1)$, the fact that $\lambda x + (1 - \lambda)y \in F$ imply $x, y \in F$.

Notice that a face is a set such that, if it contains any internal point of a line segment of $C$, then it contains the whole segment. A good intuition on this definition comes from polyhedra, which are in fact, the origin of the name "face". If we think of a cube, its squared faces are, indeed, six faces in the sense above.

Notice now that a face is again a non-empty convex set; hence, we can ask about the faces of a face. It is not difficult to notice that a face of a face is also a face of the original set (see Lemma 2.10.5 of [17]). Back to our example, we can ask about the faces of the six squares and, it is easy to verify, they are the squares' edges (and the cube's vertices), and the edges' faces are the ending points of the edges, i.e. faces without subfaces that are unitary sets containing each of the cube's vertices. Let us give a special name to these points.

**Definition 6** Let $V$ be a Hausdorff topological vector space and let $C \subset V$ be a non-empty convex set. An extremal point of $C$ is a element $x \in C$ such that $\{x\}$ is a face of $C$. We denote $\mathrm{Ext}(C) = \{x \in C \ : \ x \text{ is an extremal point of C}\}$.

Of course the cube is a very simple instance of the general question, but this example gives us a general idea on what is going on: the extremal points of the cube, namely, its vertices, can be used to obtain any other of the cube's points by taking convex combinations: first obtaining the edges, after the faces and finally the interior of the cube. That is, the cube is the smallest convex set containing its vertices. We call this smallest convex set the convex hull, that is, the convex hull of a set $A$ is the intersection of all convex subsets of the vector space containing $A$. The convex hull of a set $A$ is denoted by $co\,(A)$. An analogous definition can be done by taking the closed convex subsets, that is the closed convex hull, which is denoted by $\overline{co}\,(A1)$.
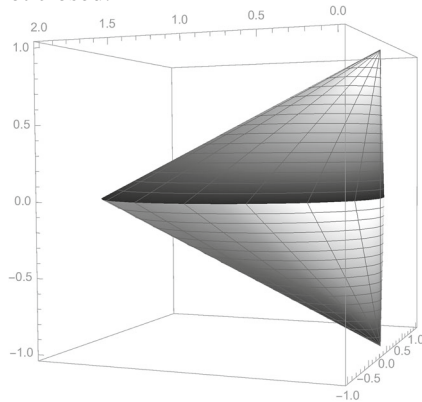
Another interesting fact is that, even in finite dimension, $\mathrm{Ext}(K)$ is not necessarily closed. Consider for example the $y$-displaced double cone in $\mathbb{R}^3$:

$$C = \overline{co}\left(\{(x, y, 0) \in \mathbb{R}^3 \ : \ x^2 + (y - 1)^2 = 1\} \cup \{(0, 0, 1), (0, 0, -1)\}\right).$$

Notice that $(0, 0, 0)$ cannot be an extremal point of $C$, since $(0, 0, 0) = \frac{1}{2}(0, 0, 1) + \frac{1}{2}(0, 0, -1)$. In fact,

$$\mathrm{Ext}(C) = \{(x, y, 0) \in \mathbb{R}^3 \ : \ x^2 + (y-1)^2 = 1, \ x \neq 0\} \cup \{(0, 0, 1), (0, 0, -1)\},$$

which is not closed.



The double cone

Finally, we are fit for stating the general result.

**Theorem 11** (Krein-Milman) *Let $V$ be a Hausdorff locally convex topological vector space and let $K \subset V$ be compact and convex. Then, one has $K = \overline{co}\,(Ext(K))$.*

It is important to reinforce that compactness plays a central role in the proof of Krein-Milman's theorem. In fact, we need compactness to ensure that this process of taking faces of faces does end and does not end up in an empty set, which is similar to the issue described in this section's second paragraph.

Now that we have Krein-Milman's theorem in hand, we can clarify our previous claim on the existence of pure states. First of all, the closed unit ball of any infinity dimensional Banach space is not compact in the norm topology, but the closed unit ball in the dual of a normed vector space is compact in the weak-$*$ topology, thanks to the Banach-Alaoglu's theorem (see Theorem 2.6.18 of [17]). The conclusion is that the states of any C$^*$-algebra $\mathfrak{A} \subset \mathcal{L}(\mathscr{H})$ constitute a weak-$*$ closed subset $\mathscr{S}$ of the weak-$*$ compact set of all bounded and normalised linear functionals on $\mathfrak{A}$. Hence, $\mathscr{S}$ is weak-$*$ compact itself and, by Krein-Milman's theorem, $\mathscr{S} = \overline{co}\,(Ext\,(\mathscr{S}))$. Since the extremal points of $\mathscr{S}$ are the states that do not lay in the interior of any segment line of $\mathscr{S}$, what is a way to say that they are not convex combinations of other states, $Ext\,(\mathscr{S})$ is the set of all pure states.

There are several interesting consequences of Krein-Milman's theorem, among them, it can be used to prove that given $A \in \mathfrak{A}$, there is a pure state $\omega_A$ such that $\omega_A(A) = \|A\|$. This result is used in the construction presented in Section 10.1.

Let us return to Krein-Milman's theorem. Suppose $V$ is a Hausdorff locally convex topological vector space and $K \subset V$ is compact and convex. Notice now that, for every $x \in co\,(Ext(K))$, there exists $n \in \mathbb{N}$, $\{\lambda_i\}_{i=1}^n \subset [0, 1]$, with $\sum_{i=1}^n \lambda_i = 1$, and $\{x_i\}_{i=1}^n \subset Ext(K)$ such that $x = \sum_{i=1}^n \lambda_i x_i$.

Then, for every continuous linear functional $f$ on $V$, $f(x) = f(\sum_{i=1}^n \lambda_i x_i) = \sum_{i=1}^n \lambda_i f(x_i)$. When $Ext(K)$ is finite, we could also write $f(x) = \sum_{e \in Ext(K)} \lambda_e f(e)$, for every $f \in V^*$, with $\lambda_e \in [0, 1]$. Notice that the different $\lambda_e$'s work as weights in the sum, with $\sum_{e \in Ext(K)} \lambda_e = 1$. We can define the probability measures (Dirac measures) $\delta_e : K \to \mathbb{R}$ by

$$\delta_e(X) = \begin{cases} 1, & \text{if } e \in X, \\ 0, & \text{if } e \notin X, \end{cases}$$

for each $e \in Ext(K)$ and $\mu = \sum_{e \in Ex(K)} \lambda_e \delta_e$. It is quite easy to check that $f(x) = \int_K f\,d\mu$.

Notice that the measure $\mu$ is a regular Borel measure satisfying $\mu(K) = 1$, in other words, $\mu$ is a probability measure. Furthermore, $\mu\,(K \setminus Ext(K)) = 0$. A measure satisfying this is said to be supported in $Ext(K)$. In addition, we have that $f(x) = \int_K f\,d\mu$ for all continuous linear functional on $V$. A measure satisfying this property is said to represent $x \in K$.

This is not an isolated case, and its generalisation is given by:

**Theorem 12** (Choquet) *Let $V$ be a Hausdorff locally convex space and $K \in V$ a metrizable convex compact set. Then, for every $x \in K$, there exists a probability measure $\mu$ supported in $Ext(K)$ representing $x$.*

Krein-Milman's and Choquet's theorems are equivalent when $Ext(K)$ is closed see [18].

Choquet's theorem has a very interesting consequence in von Neumann algebras. Let $\mathcal{M} \in \mathcal{L}(H)$ be a von Neumann algebra. Since a von Neumann algebra is a dual space of some Banach space $\mathcal{M}_*$, its closed unit ball is weak-$*$ compact. In addition, if $\mathscr{H}$ is separable, the closed unit ball $B_1 \subset \mathcal{L}(\mathscr{H})$ with the weak-$*$ topology is metrizable. Hence, for $K = B_1$, we are in the conditions of Choquet's theorem, so for any $A \in B_1$ there exists a probability measure $\mu_A$ supported in $Ext(B_1)$ such that, for all continuous linear functionals on $\mathcal{L}(\mathscr{H})$, $f(A) = \int_{B_1} f\,d\mu_A$.

## Appendix 1: Proof of Theorem 5

Here, we follow [15] closely. First, take a non-null positive linear functional $\omega'$ such that, $\forall A \in \mathfrak{A}$, $\omega'(A^*A) \leqslant$

$\omega(A^*A)$. By the Cauchy-Schwartz inequality (6),

$$|\omega'(A^*B)|^2 \leqslant \omega'(A^*A)\omega'(B^*B) \leqslant \omega(A^*A)\omega(B^*B)$$
$$= \|\pi(A)\Omega\|^2\|\pi(B)\Omega\|^2,$$

which implies, by Riesz representation theorem (see e.g. [1]), the existence of a positive operator $T \in \mathscr{L}(\mathscr{H})$ such that $\langle\pi(A)\Omega, T\pi(B)\Omega\rangle = \omega'(A^*B)$ (remark that the sesquilinear form $(\pi(A)\Omega, \pi(B)\Omega) \longmapsto \omega'(A^*B)$ may be extended to the whole $\mathscr{H}$, and thus the domain of $T$, only because $\Omega$ is cyclic). Taking $A$, $B$, $C \in \mathfrak{A}$ arbitrary:

$$\langle\pi(A)\Omega, T\pi(B)\pi(C)\Omega\rangle = \omega'(A^*BC) = \omega'\left((B^*A)^*C\right)$$
$$= \langle\pi(A)\Omega, \pi(B)T\pi(C)\Omega\rangle,$$

which implies that $[T, \pi(B)] = 0$ for any $B$.

As known in representation theory (as a consequence of Schur's Lemma, see e.g. [19]), if $\pi$ is an irreducible representation, any self-adjoint operator commuting with it must be a multiple of the identity and vice versa. It happens that, if $T = \lambda\mathbb{1}$ for some $\lambda \in \mathbb{C}$, then $\omega' = \lambda\omega$. Let us show an implication of this fact, that $\omega$ is a mixture if and only if $\pi$ is reducible, which is enough for the theorem's statement.

Indeed, if $\omega$ is a mixture, one may find $\omega'$ such that $\omega'(A^*A) \leqslant \omega(A^*A)$ which is not a multiple of $\omega$: in this case, there is a scalar $\sigma \in (0, 1)$ and states $\omega_1$ and $\omega_2$ (none of them multiples of $\omega$) such that $\omega = \sigma\omega_1 + (1 - \sigma)\omega_2$, so just take $\omega' = \sigma\omega_1$. The corresponding $T$ will not be a multiple of the identity, hence the reducibility of $\pi$.

Conversely, supposing that $\pi$ is reducible, one may find a self-adjoint $S \in \mathfrak{A}$ commuting with every $\pi(B)$ and not being a multiple of $\mathbb{1}$. As a consequence, any non-trivial spectral projector $P$ of $S$ will also commute with $\pi$, not be a multiple of the identity, and further satisfy $0 < P < \mathbb{1}$. Define the functional $\omega'(A) = \langle\Omega, P\pi(A)\Omega\rangle$ and remark that it is positive, not a multiple of $\omega$, and that

$$\omega(A^*A) - \omega'(A^*A) = \langle\pi(A)\Omega, (\mathbb{1} - P)\pi(A)\Omega\rangle \geqslant 0.$$

This implies that $\omega$ is a non-trivial mixture $\omega = \lambda\omega_1 + (1 - \lambda)\omega_2$, with $\lambda = \|\omega'\| \in (0, 1)$ and states $\omega_1 = \frac{1}{\|\omega'\|}\omega'$ and $\omega_2 = \frac{1}{1-\|\omega'\|}(\omega - \omega')$ (see Remark 1 for a quick justification that $\|\omega_2\| = 1$).

# References

1. M. Reed, B. Simon, *Methods of Modern Mathematical Physics I – Functional Analysis* (Academic, Cambridge, 1980)

2. H. Araki, Vol. 101, *Mathematical theory of quantum fields, International series of monographs in Physics* (Oxford University Press, London, 2009). ISBN 0 19 851773 4

3. R. Haag, *Local Quantum Physics*, 2nd edn. (Springer, Berlin, 1996). ISSN 0172-5998

4. C.J. Fewster, K. Rejzner, Algebraic Quantum Field Theory - an introduction. arXiv:1904.04051 [hep-th] (2019)

5. G.érard.G. Emch, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory* (Wiley-Interscience, New York, 1972). ISBN-13 978-0-486-47209-6

6. O. Bratteli, D.W. Robinson, *Operator algebras and Quantum Statistical Mechanics I – C\*- and W\*-algebras, symmetry groups, decomposition of states*, 2nd edn. (Springer, Berlin, 1987). ISBN 3-540-09187-4

7. J. Dixmier, *Von Neumann algebras* (North Holland, Amsterdam, 2011). ISBN-13: 978-0444557407

8. V. Chabu, Semiclassical analysis of the schrödinger equation with conical singularities. Asympt. Anal. **103**(4), 165–220 (2017)

9. V. Chabu, Semiclassical analysis of the schrödinger equation with singular potentials thèse de l'université Paris-Est (2016)

10. P. Gérard, P.A. Markowich, N.J. Mauser, F. Poupaud, Homogenization limits and Wigner transforms. Pure Appl. Math. **50**(4), 323–379 (1997)

11. W. Rudin, *Real And Complex Analysis. Higher Mathematics series*, 3rd edn. (McCraw-Hill Education, New York, 1986). ISBN-13: 978-0070542341

12. P. Lions, T. Paul, Sur les mesures de Wigner. Revista Matemá,tica Iberoamericana **9**, 553–618 (1993)

13. M. Dimassi, S. Sjostrand, Spectral asymptotics in the semiclassical limit, London Math. Society lecture note series 268: Cambridge University Press, Cambridge (1999)

14. M. Zworski, Semiclassical Analysis. Graduate studies in Mathematics, Vol. 138: American Mathematical Society, Providence (2012)

15. G.J. Murphy, C\*-algebras and Operator Theory (Academic, Cambridge, 1990). ISBN 0-12-511360-9

16. J.M.G. Fell, The dual spaces of C\*-algebras. Transactions of the American Mathematical Society. https://doi.org/10.2307/1993431 (1960)

17. R.E. Megginson, *An Introduction to Banach Space Theory* (Springer, Berlin, 1998). ISBN-13: 978-1461268352

18. R.R. Phelps, *Lectures on Choquet's Theorem* (Springer, Berlin, 2001). ISBN-13: 978-3540418344

19. K. Erdmann, T. Holm, *Algebras and Representation Theory* (Springer, Berlin, 2018). ISBN-13: 978-3319919973

20. K.R. Davidson, C\*-Algebras by Example. Fields Institute Monographs (American Mathematical Society, Providence, 1996). ISBN-13: 978-8185931111

21. W.B. Arveson, *An Anvitation to C\*-Algebras* (Springer, Berlin, 1998). ISBN-13: 978-1461263739

22. M. Kleinmann, H. Kampermann, T. Meyer, D. Bruss, Physical purification of quantum states. Phys. Rev. A **73**, 062309 (2006)