

Mathematical Foundations of Quantum Mechanics: An Advanced Short Course

Valter Moretti

*Department of Mathematics of the University of Trento and INFN-TIFPA,
via Sommarive 14, I-38122 Povo (Trento), Italy
valter.moretti@unitn.it*

Abstract. This paper collects and extends the lectures I gave at the “XXIV International Fall Workshop on Geometry and Physics” held in Zaragoza (Spain) during September 2015. Within these lectures I review the formulation of Quantum Mechanics, and quantum theories in general, from a mathematically advanced viewpoint, essentially based on the orthomodular lattice of elementary propositions, discussing some fundamental ideas, mathematical tools and theorems also related to the representation of physical symmetries. The final step consists of an elementary introduction to the so-called (C*-) algebraic formulation of quantum theories.

Keywords: Mathematical Formulation of Quantum Mechanics; Spectral Theory; von Neumann algebra, C*-algebra.

Contents

1	Introduction: Summary of elementary facts of QM	2
1.1	Physical facts about Quantum Mechanics	3
1.1.1	When a physical system is quantum	3
1.1.2	General properties of quantum systems	4
1.2	Elementary formalism for the finite dimensional case	6
1.3	Time evolution	10
1.4	Composite systems	11
1.5	A first look to the infinite dimensional case, CCR and quantization procedures	13
2	Observables in infinite dimensional Hilbert spaces: Spectral Theory	18
2.1	Hilbert spaces	18
2.2	Classes of (especially unbounded) operators in Hilbert spaces	21
2.3	Criteria for (essential) selfadjointness	31
2.4	Spectrum of an operator	35
2.5	Spectral measures	38
2.6	Spectral Decomposition and Representation Theorems	43
2.7	Mesurable functional calculus	49
2.8	Elementary formalism for the infinite dimensional case	50

2.9	Technical Intermezzo: Three Operator Topologies	53
3	More Fundamental Quantum Structures	55
3.1	The Boolean logic of CM	55
3.2	The non-Boolean Logic of QM, the reason why observables are self-adjoint operators.	58
3.3	Recovering the Hilbert space structure	62
3.4	States as measures on $\mathcal{L}(\mathcal{H})$: Gleason's Theorem	64
3.4.1	Trace class operators	65
3.4.2	The notion of quantum state and the crucial theorem by Gleason	67
3.5	von Neumann algebra of observables, superselection rules	73
3.5.1	von Neumann algebras	73
3.5.2	Lattices of von Neumann algebras	75
3.5.3	General algebra of observables and its centre	75
3.5.4	Superselection charges and coherent sectors	78
3.5.5	States in the presence of superselection rules	80
3.6	Quantum Symmetries: unitary projective representations	83
3.6.1	Wigner and Kadison theorems, groups of symmetries	84
3.6.2	Groups of quantum symmetries	86
3.6.3	One-parameter strongly continuous unitary groups: von Neumann and Stone theorems	90
3.6.4	Time evolution, Heisenberg picture and quantum Noether theorem	91
3.6.5	Strongly continuous unitary representations of Lie groups, Nelson theorem	94
3.6.6	Selfadjoint version of Stone - von Neumann - Mackey Theorem	101
4	Just few words about the Algebraic Approach	103
4.1	Algebraic formulation	104
4.1.1	The GNS reconstruction theorem	106
4.1.2	Pure states and irreducible representations	107

1. Introduction: Summary of elementary facts of QM

This paper collects and technically extends the lectures given by the author at the “XXIV International Fall Workshop on Geometry and Physics” held in Zaragoza, August 31 - September 4, 2015. These lecture notes contain much more written material than the lectures themselves.

A concise account of the basic structure of quantum mechanics and *quantization procedures* has already been presented in [1] with several crucial examples. In the rest of Section 1, we quickly review again some elementary facts and properties, either of physical or mathematical nature, related to Quantum Mechanics, without fully entering into the mathematical details.

Section 2 is instead devoted to present some technical definitions and results of spectral analysis in complex Hilbert spaces, especially the basic elements of spectral theory, including the classic theorem about spectral decomposition of (generally unbounded) selfadjoint operators and the so called measurable functional calculus. A brief presentation of the three most important operator topologies for applications in QM closes Section 2.

Within Section 3, the *corpus* of the lectures, we pass to analyse the mathematical structure of QM from a finer and advanced viewpoint, adopting the framework based on orthomodular lattices' theory. This approach permits one to justify some basic assumptions of QM, like the mathematical nature of the observables represented by selfadjoint operators and the quantum states viewed as trace class operators. QM is essentially a probability measure on the non-Boolean lattice $\mathcal{L}(\mathcal{H})$ of elementary observables. A key tool of that analysis is the theorem by Gleason characterising the notion probability measure on $\mathcal{L}(\mathcal{H})$ in terms of certain trace class operators. We also discuss the structure of the algebra of observables in the presence of superselection rules after having introduced the mathematical notion of von Neumann algebra. The subsequent part of the third section is devoted to present the idea of quantum symmetry, illustrated in terms of Wigner and Kadison theorems. Some basic mathematical facts about groups of quantum symmetries are introduced and discussed, especially in relation with the problem of their unitarisation. Bargmann's condition is stated. The particular case of a strongly continuous one-parameter unitary group will be analysed in some more detail, mentioning von Neumann's theorem and the celebrated Stone theorem, remarking its use to describe the time evolution of quantum systems. A quantum formulation of Noether theorem ends this part. The last part of Section 3 aims to introduce some elementary results about continuous unitary representations of Lie groups, discussing in particular a theorem by Nelson which proposes sufficient conditions for lifting a (anti)selfadjoint representation of a Lie algebra to a unitary representation of the unique simply connected Lie group associated to that Lie algebra.

The last section closes the paper focussing on elementary ideas and results of the so called algebraic formulation of quantum theories. Many examples and exercises (with solutions) accompany the theoretical text at every step.

1.1. *Physical facts about Quantum Mechanics*

Let us quickly review the most relevant and common features of quantum systems. Next we will present a first elementary mathematical formulation which will be improved in the rest of the lectures, introducing a suitable mathematical technology.

1.1.1. *When a physical system is quantum*

Loosely speaking, Quantum Mechanics is the physics of microscopic world (elementary particles, atoms, molecules). That realm is characterized by a universal physical constant denoted by h and called **Planck constant**. A related constant

– nowadays much more used – is the **reduced Planck constant**, pronounced “h-bar”,

$$\hbar := \frac{h}{2\pi} = 1.054571726 \times 10^{-34} J \cdot s .$$

The physical dimensions of h (or \hbar) are those of an *action*, i.e. *energy* \times *time*. A rough check on the appropriateness of a quantum physical description for a given physical system is obtained by comparing the value of some characteristic action of the system with \hbar . For a macroscopic pendulum (say, length $\sim 1m$, mass $\sim 1kg$ maximal speed $\sim 1ms^{-1}$), multiplying the period of oscillations and the maximal kinetic energy, we obtain a typical action of $\sim 2Js \gg h$. In this case quantum physics is expected to be largely inappropriate, exactly as we actually know from our experience of every days. Conversely, referring to a hydrogen electron orbiting around its proton, the first *ionization energy* multiplied with the orbital period of the electron (computed using the classical formula with a value of the radius of the order of 1 \AA) produces a typical action of the order of h . Here quantum mechanics is necessary.

1.1.2. General properties of quantum systems

Quantum Mechanics (QM) enjoys a triple of features which seem to be very far from properties of Classical Mechanics (CM). These remarkable general properties concern the physical quantities of physical systems. In QM physical quantities are called *observables*.

(1) Randomness. When we perform a measurement of an observable of a quantum system, the outcomes turn out to be *stochastic*: Performing measurements of the same observable A on completely identical systems prepared in the *same* physical state, one generally obtains different results $a, a', a'' \dots$

Referring to the standard interpretation of the formalism of QM (see [3] for a nice up-to-date account on the various interpretations), the randomness of measurement outcomes should not be considered as due to an incomplete knowledge of the state of the system as it happens, for instance, in Classical Statistical Mechanics. Randomness is not *epistemic*, but it is *ontological*. It is a fundamental property of quantum systems.

On the other hand, *QM permits one to compute the probability distribution of all the outcomes of a given observable, once the state of the system is known.*

Moreover, it is always possible to prepare a state ψ_a where a certain observable A is *defined* and takes its value a . That is, repeated measurements of A give rise to the same value a with probability 1. (Notice that we can perform simultaneous measurements on identical systems all prepared in the state ψ_a , or we can perform different subsequent measurements on the same system in the state ψ_a . In the second case, these measurements have to be performed very close to each other in time to prevent the state of the system from evolving in view of Schrödinger evolution

as said in (3) below.) Such states, where observable take definite values, cannot be prepared for *all* observables simultaneously as discussed in (2) below.

(2) Compatible and Incompatible Observables. The second noticeable feature of QM is the existence of *incompatible observables*. Differently from CM, there are physical quantities which cannot be measured simultaneously. There is no physical instrument capable to do it. If an observable A is *defined* in a given state ψ – i.e. it attains a precise value a with probability 1 in case of a measurement – an observable B *incompatible* with A turns out to be *not defined* in the state ψ – i.e., it generally attains several different values $b, b', b'' \dots$, *none with probability 1*, in case of measurement. So, if we perform a measurement of B , we generally obtain a spectrum of values described by a probabilistic distribution as preannounced in (1) above.

Incompatibility is a *symmetric* property: A is incompatible with B if and only if B is incompatible with A . However it is not *transitive*.

There are also *compatible observables* which, by definition, can be measured simultaneously. An example is the component x of the position of a particle and the component y of the momentum of that particle, referring to a given inertial reference frame. A popular case of incompatible observables is a pair of *canonically conjugated observables* [1] like the position X and the momentum P of a particle both along the same fixed axis of a reference frame. In this case there is a lower bound for the product of the standard deviations, resp. $\Delta X_\psi, \Delta P_\psi$, of the outcomes of the measurements of these observables in a given state ψ (these measurement has to be performed on different identical systems all prepared in the same state ψ). This lower bound does not depend on the state and is encoded in the celebrated mathematical formula of the *Heisenberg principle* (a theorem in the modern formulations):

$$\Delta X_\psi \Delta P_\psi \geq \hbar/2, \quad (1)$$

where Planck constant shows up.

(3) Post measurement Collapse of the State. In QM, measurements *generally change the state of the system* and produce a post-measurement state from the state on which the measurement is performed. (We are here referring to idealized measurement procedures, since measurement procedures are very often destructive.) If the measured state is ψ , immediately after the measurement of an observable A obtaining the value a among a plethora of possible values a, a', a'', \dots , the state changes to ψ' generally different from ψ . In the new state ψ' , the distribution of probabilities of the outcomes of A changes to 1 for the outcome a and 0 for all other possible outcomes. A is therefore *defined* in ψ' .

When we perform repeated and alternated measurements of a pair of incompatible observables, A, B , the outcomes disturb each other: If the first outcome of A is a , after a measurement of B , a subsequent measurement of A produces $a' \neq a$ in general. Conversely, if A and B are compatible, the outcomes of their subsequent measurements do not disturb each other.

In CM there are measurements that, in practice, disturb and change the state of the system. It is however possible to decrease the disturbance arbitrarily, and nullify it in ideal measurements. In QM it is not always possible as for instance witnessed by (1).

In QM, there are two types of time evolution of the state of a system. One is the usual one due to the dynamics and encoded in the famous *Schrödinger equation* we shall see shortly. It is nothing but a quantum version of classical *Hamiltonian evolution* [1]. The other is the sudden change of the state due to measurement procedure of an observable, outlined in (3): The *collapse of the state* (or *wavefunction*) of the system.

The nature of the second type of evolution is still source of an animated debate in the scientific community of physicists and philosophers of Science. There are many attempts to reduce the collapse of the state to the standard time evolution referring to the quantum evolution of the whole physical system, also including the measurement apparatus and the environment (*de-coherence processes*) [3,12]. None of these approaches seem to be completely satisfactory up to now.

Remark 1. *Unless explicitly stated, we henceforth adopt a physical unit system such that $\hbar = 1$.*

1.2. Elementary formalism for the finite dimensional case

To go on with this introduction, let us add some further technical details to the presented picture to show how practically (1)-(3) have to be mathematically interpreted (reversing the order of (2) and (3) for our convenience). The rest of the paper is devoted to make technically precise, justify and widely develop these ideas from a mathematically more advanced viewpoint than the one of [1].

To mathematically simplify this introductory discussion, throughout this section, except for Sect 1.5, we assume that \mathcal{H} denotes a *finite dimensional* complex vector space equipped with a Hermitian scalar product, denoted by $\langle \cdot | \cdot \rangle$, where the linear entry is the second one. With \mathcal{H} as above, $L(\mathcal{H})$ will denote the complex algebra of operators $A : \mathcal{H} \rightarrow \mathcal{H}$. We remind the reader that, if $A \in L(\mathcal{H})$ with \mathcal{H} finite dimensional, the *adjoint operator*, $A^* \in L(\mathcal{H})$, is the unique linear operator such that

$$\langle A^* x | y \rangle = \langle x | Ay \rangle \quad \text{for all } x, y \in \mathcal{H}. \quad (2)$$

A is said to be *selfadjoint* if $A = A^*$, so that, in particular

$$\langle Ax | y \rangle = \langle x | Ay \rangle \quad \text{for all } x, y \in \mathcal{H}. \quad (3)$$

Since $\langle \cdot | \cdot \rangle$ is linear in the second entry and antilinear in the first entry, we immediately have that all eigenvalues of a selfadjoint operator A are real.

Our assumptions on the mathematical description of quantum systems are the following ones.

- (1) A quantum mechanical system S is always associated to a complex vector space \mathcal{H} (here finite dimensional) equipped with a Hermitian scalar product $\langle \cdot | \cdot \rangle$;
- (2) observables are pictured in terms of *selfadjoint* operators A on \mathcal{H} ;
- (3) states are equivalence classes of *unit* vectors $\psi \in \mathcal{H}$, where $\psi \sim \psi'$ iff $\psi = e^{ia}\psi'$ for some $a \in \mathbb{R}$.

Remark 2.

(a) It is clear that states are therefore one-to-one represented by all of the elements of the complex projective space $P\mathcal{H}$. The states we are considering within this introductory section are called *pure* states. A more general notion of state, already introduced in [1], will be discussed later.

(b) \mathcal{H} is an elementary version of complex Hilbert space since it is automatically complete it being finite dimensional.

(c) Since $\dim(\mathcal{H}) < +\infty$, every self-adjoint operator $A \in L(\mathcal{H})$ admits a spectral decomposition

$$A = \sum_{a \in \sigma(A)} a P_a^{(A)}, \quad (4)$$

where $\sigma(A)$ is the *finite* set of eigenvalues – which must be *real* as A is self-adjoint – and $P_a^{(A)}$ is the orthogonal projector onto the eigenspace associated to a . Notice that $P_a P_{a'} = 0$ if $a \neq a'$ as eigenvectors with different eigenvalue are orthogonal. ■

Let us show how the mathematical assumptions (1)-(3) permit us to set the physical properties of quantum systems (1)-(3) into a mathematically nice form.

(1) Randomness: The eigenvalues of an observable A are physically interpreted as the possible values of the outcomes of a measurement of A .

Given a state, represented by the unit vector $\psi \in \mathcal{H}$, the probability to obtain $a \in \sigma(A)$ as an outcome when measuring A is

$$\mu_\psi^{(A)}(a) := \|P_a^{(A)}\psi\|^2.$$

Going along with this interpretation, the expectation value of A , when the state is represented by ψ , turns out to be

$$\langle A \rangle_\psi := \sum_{a \in \sigma(A)} a \mu_\psi^{(A)}(a) = \langle \psi | A \psi \rangle.$$

So that the identity holds

$$\langle A \rangle_\psi = \langle \psi | A \psi \rangle. \quad (5)$$

Finally, the standard deviation ΔA_ψ results to be

$$\Delta A_\psi^2 := \sum_{a \in \sigma(A)} (a - \langle A \rangle_\psi)^2 \mu_\psi^{(A)}(a) = \langle \psi | A^2 \psi \rangle - \langle \psi | A \psi \rangle^2. \quad (6)$$

Remark 3.

(a) Notice that the arbitrary phase affecting the unit vector $\psi \in \mathcal{H}$ ($e^{ia}\psi$ and ψ represent the same quantum state for every $a \in \mathbb{R}$) is armless here.

(b) If A is an observable and $f : \mathbb{R} \rightarrow \mathbb{R}$ is given, $f(A)$ is interpreted as an observable whose values are $f(a)$ if $a \in \sigma(A)$: Taking (4) into account,

$$f(A) := \sum_{a \in \sigma(A)} f(a)P_a^{(A)}. \quad (7)$$

For polynomials $f(x) = \sum_{k=0}^n a_k x^k$, it results $f(A) = \sum_{k=0}^n a_k A^k$ as expected. The selfadjoint operator A^2 can naturally be interpreted this way as the natural observable whose values are a^2 when $a \in \sigma(A)$. For this reason, looking at the last term in (6) and taking (5) into account,

$$\Delta A_\psi^2 = \langle A^2 \rangle_\psi - \langle A \rangle_\psi^2 = \langle (A - \langle A \rangle_\psi I)^2 \rangle_\psi = \langle \psi | (A - \langle A \rangle_\psi I)^2 \psi \rangle. \quad (8)$$

■

(3) Collapse of the state: If a is the outcome of the (idealized) measurement of A when the state is represented by ψ , the new state immediately after the measurement is represented by the unit vector

$$\psi' := \frac{P_a^{(A)}\psi}{\|P_a^{(A)}\psi\|}. \quad (9)$$

Remark 4. Obviously this formula does not make sense if $\mu_\psi^{(A)}(a) = 0$ as expected. Moreover the arbitrary phase affecting ψ does not lead to troubles, due to the linearity of $P_a^{(A)}$.

(2) Compatible and Incompatible Observables: Two observables are compatible – i.e. they can be simultaneously measured – if and only if the associated operators *commute*, that is

$$AB - BA = 0.$$

Using the fact that \mathcal{H} has finite dimension, one easily proves that the observables A and B are compatible if and only if the associated spectral projectors commute as well

$$P_a^{(A)}P_b^{(B)} = P_b^{(B)}P_a^{(A)} \quad a \in \sigma(A), b \in \sigma(B).$$

In this case

$$\|P_a^{(A)}P_b^{(B)}\psi\|^2 = \|P_b^{(B)}P_a^{(A)}\psi\|^2$$

has the natural interpretation of the probability to obtain the outcomes a and b for a simultaneous measurement of A and B . If instead A and B are incompatible, it may happen that

$$\|P_a^{(A)}P_b^{(B)}\psi\|^2 \neq \|P_b^{(B)}P_a^{(A)}\psi\|^2.$$

Sticking to the case of A and B incompatible, exploiting (9),

$$\|P_a^{(A)}P_b^{(B)}\psi\|^2 = \left\| P_a^{(A)} \frac{P_b^{(B)}\psi}{\|P_b^{(B)}\psi\|} \right\|^2 \|P_b^{(B)}\psi\|^2 \quad (10)$$

has the natural meaning of *the probability of obtaining first b and next a in a subsequent measurement of B and A .*

Remark 5.

(a) Notice that, in general, we cannot interchange the rôle of A and B in (10) because, in general, $P_a^{(A)}P_b^{(B)} \neq P_b^{(B)}P_a^{(A)}$ if A and B are incompatible. The measurement procedures “disturb each other” as already said.

(b) The interpretation of (10) as probability of subsequent measurements can be given also if A and B are compatible. In this case, the probability of obtaining first b and next a in a subsequent measurement of B and A is identical to the probability of measuring a and b simultaneously and, in turn, it coincides with the probability of obtaining first a and next b in a subsequent measurement of A and B .

(c) A is always compatible with itself. Moreover $P_a^{(A)}P_a^{(A)} = P_a^{(A)}$ just due to the definition of projector. This fact has the immediate consequence that if we obtain a measuring A so that the state immediately after the measurement is represented by $\psi_a = \|P_a^{(A)}\psi\|^{-1}\psi$, it will remain ψ_a even after other subsequent measurements of A and the outcome will result to be always a . Versions of this phenomenon, especially regarding the decay of unstable particles, are experimentally confirmed and it is called the *quantum Zeno effect*. ■

Example 6. An electron admits a triple of observables, S_x , S_y , S_z , known as the components of the *spin*. Very roughly speaking, the spin can be viewed as the angular momentum of the particle referred to a reference frame always at rest with the centre of the particle and carrying its axes parallelly to the ones of the reference frame of the laboratory, where the electron moves. In view of its peculiar properties, the spin cannot actually have a complete classical corresponding and thus that interpretation is untenable. For instance, one cannot “stop” the spin of a particle or change the constant value of $S^2 = S_x^2 + S_y^2 + S_z^2$: It is a given property of the particle like the mass. The electron spin is described within an *internal* Hilbert space \mathcal{H}_s , which has dimension 2. Identifying \mathcal{H}_s with \mathbb{C}^2 , the three spin observables are defined in terms of the three Hermitian matrices (occasionally re-introducing the constant \hbar)

$$S_x = \frac{\hbar}{2}\sigma_x, \quad S_y = \frac{\hbar}{2}\sigma_y, \quad S_z = \frac{\hbar}{2}\sigma_z, \quad (11)$$

where we have introduced the well known *Pauli matrices*,

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (12)$$

Notice that $[S_a, S_b] \neq 0$ if $a \neq b$ so that the components of the spin are incompatible observables. In fact one has

$$[S_x, S_y] = i\hbar S_z$$

and this identity holds also cyclically permuting the three indices. These commutation relations are the same as for the observables L_x, L_y, L_z describing the angular momentum referred to the laboratory system which have classical corresponding (we shall return on these observables in example 166). So, differently from CM, the observables describing the components of the angular momentum are incompatible, they cannot be measured simultaneously. However the failure of the compatibility is related to the appearance of \hbar on the right-hand side of

$$[L_x, L_y] = i\hbar L_z .$$

That number is extremely small if compared with macroscopic scales. This is the ultimate reason why the incompatibility of L_x and L_z is negligible for macroscopic systems.

Direct inspection proves that $\sigma(S_a) = \{\pm\hbar/2\}$. Similarly $\sigma(L_a) = \{n\hbar \mid n \in \mathbb{Z}\}$. Therefore, differently from CM, the values of angular momentum components form a discrete set of reals in QM. Again notice that the difference of two closest values is extremely small if compared with typical values of the angular momentum of macroscopic systems. This is the practical reason why this discreteness disappears at macroscopic level. ■

Just a few words about the time evolution and composite systems [1] are necessary now, a wider discussion on the time evolution will take place later in this paper.

1.3. Time evolution

Among the class of observables of a quantum system described in a given inertial reference frame, an observable H called the (quantum) *Hamiltonian* plays a fundamental rôle. We are assuming here that the system interacts with a stationary environment. The one-parameter group of unitary operators associated to H (exploiting (7) to explain the notation)

$$U_t := e^{-itH} := \sum_{h \in \sigma(H)} e^{-ith} P_h^{(H)}, \quad t \in \mathbb{R} \quad (13)$$

describes the *time evolution of quantum states* as follows. If the state at time $t = 0$ is represented by the unit vector $\psi \in \mathcal{H}$, the state at the generic time t is represented by the vector

$$\psi_t = U_t \psi .$$

Remark 7. Notice that ψ_t has norm 1 as necessary to describe states, since U_t is norm preserving it being unitary. ■

Taking (13) into account, this identity is equivalent to

$$i \frac{d\psi_t}{dt} = H\psi_t. \quad (14)$$

Equation (14) is nothing but a form of the celebrated *Schrödinger equation*. If the environment is not stationary, a more complicated description can be given where H is replaced by a class of Hamiltonian (selfadjoint) operators parametrized in time, $H(t)$, with $t \in \mathbb{R}$. This time dependence accounts for the time evolution of the external system interacting with our quantum system. In that case, it is simply assumed that the time evolution of states is again described by the equation above where H is replaced by $H(t)$:

$$i \frac{d\psi_t}{dt} = H(t)\psi_t.$$

Again, this equation permits one to define a two-parameter *groupoid* of unitary operators $U(t_2, t_1)$, where $t_2, t_1 \in \mathbb{R}$, such that

$$\psi_{t_2} = U(t_2, t_1)\psi_{t_1}, \quad t_2, t_1 \in \mathbb{R}.$$

The groupoid structure arises from the following identities: $U(t, t) = I$ and $U(t_3, t_2)U(t_2, t_1) = U(t_3, t_1)$ and $U(t_2, t_1)^{-1} = U(t_2, t_1)^* = U(t_1, t_2)$.

Remark 8. In our elementary case where \mathcal{H} is finite dimensional, *Dyson's formula* holds with the simple hypothesis that the map $\mathbb{R} \ni t \mapsto H_t \in L(\mathcal{H})$ is continuous (adopting any topology compatible with the vector space structure of $L(\mathcal{H})$) [6]

$$U(t_2, t_1) = \sum_{n=0}^{+\infty} \frac{(-i)^n}{n!} \int_{t_1}^{t_2} \cdots \int_{t_1}^{t_2} T[H(\tau_1) \cdots H(\tau_n)] d\tau_1 \cdots d\tau_n.$$

Above, we define $T[H(\tau_1) \cdots H(\tau_n)] = H(\tau_{\pi(1)}) \cdots H(\tau_{\pi(n)})$, where the bijective function $\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ is any permutation with $\tau_{\pi(1)} \geq \cdots \geq \tau_{\pi(n)}$. ■

1.4. Composite systems

If a quantum system S is made of two parts, S_1 and S_2 , respectively described in the Hilbert spaces \mathcal{H}_1 and \mathcal{H}_2 , it is assumed that the whole system is described in the space $\mathcal{H}_1 \otimes \mathcal{H}_2$ equipped with the unique Hermitian scalar product $\langle \cdot | \cdot \rangle$ such that $\langle \psi_1 \otimes \psi_2 | \phi_1 \otimes \phi_2 \rangle = \langle \psi_1 | \phi_1 \rangle_1 \langle \psi_2 | \phi_2 \rangle_2$ (in the infinite dimensional case $\mathcal{H}_1 \otimes \mathcal{H}_2$ is the Hilbert completion of the afore-mentioned algebraic tensor product).

If $\mathcal{H}_1 \otimes \mathcal{H}_2$ is the space of a composite system S as before and A_1 represents an observable for the part S_1 , it is naturally identified with the selfadjoint operator $A_1 \otimes I_2$ defined in $\mathcal{H}_1 \otimes \mathcal{H}_2$. A similar statement holds swapping 1 and 2. Notice that $\sigma(A_1 \otimes I_2) = \sigma(A_1)$ as one easily proves. (The result survives the extension to the infinite dimensional case.)

Remark 9.

(a) Composite systems are in particular systems made of many (either identical or not) particles. If we have a pair of particles respectively described in the Hilbert space \mathcal{H}_1 and \mathcal{H}_2 , the full system is described in $\mathcal{H}_1 \otimes \mathcal{H}_2$. Notice that the dimension of the final space is the *product* of the dimension of the component spaces. In CM the system would instead be described in a space of phases which is the Cartesian product of the two spaces of phases. In that case the dimension would be the *sum*, rather than the product, of the dimensions of the component spaces.

(b) $\mathcal{H}_1 \otimes \mathcal{H}_2$ contains the so-called *entangled states*. They are states represented by vectors *not* factorized as $\psi_1 \otimes \psi_2$, but they are *linear combinations* of such vectors. Suppose the whole state is represented by the entangled state

$$\Psi = \frac{1}{\sqrt{2}} (\psi_a \otimes \phi + \psi_{a'} \otimes \phi') ,$$

where $A_1 \psi_a = a \psi_a$ and $A_1 \psi_{a'} = a' \psi_{a'}$ for a certain observable A_1 of the part S_1 of the total system. Performing a measurement of A_1 on S_1 , due to the collapse of state phenomenon, we automatically act on the whole state and on the part describing S_2 . As a matter of fact, up to normalization, the state of the full system after the measurement of A_1 will be $\psi_a \otimes \phi$ if the outcome of A_1 is a , or it will be $\psi_{a'} \otimes \phi'$ if the outcome of A_1 is a' . It could happen that the two measurement apparatuses, respectively measuring S_1 and S_2 , are localized very far in the physical space. Therefore acting on S_1 by measuring A_1 , we “instantaneously” produce a change of S_2 which can be seen performing measurements on it, even if the measurement apparatus of S_2 is very far from the one of S_1 . This seems to contradict the fundamental relativistic postulate, the *locality* postulate, that there is a maximal speed, the one of light, for propagating physical information. After the famous analysis of Bell, improving the original one by Einstein, Podolsky and Rosen, the phenomenon has been experimentally observed. Locality is truly violated, but in a such subtle way which does not allow superluminal propagation of physical information. Non-locality of QM is nowadays widely accepted as a real and fundamental feature of Nature [2,3]. ■

Example 10. An electron also possesses an *electric charge*. That is another *internal* quantum observable, Q , with two values $\pm e$, where $e = 1.602176565 \times 10^{-19} C$ is the value elementary electrical charge. So there are two types of electrons. *Proper electrons*, whose internal state of charge is an eigenvector of Q with eigenvalue $-e$ and *positrons*, whose internal state of charge is a eigenvector of Q with eigenvalue e . The simplest version of the internal Hilbert space of the electrical charge is therefore \mathcal{H}_c which^a, again, is isomorphic to \mathbb{C}^2 . With this representation $Q = e\sigma_3$. The full Hilbert space of an electron must contain a factor $\mathcal{H}_s \otimes \mathcal{H}_c$. Obviously this is by no means sufficient to describe an electron, since we must introduce at least the

^aAs we shall say later, in view of a *superselection rule* not all normalized vectors of \mathcal{H}_c represent (pure) states.

observables describing the position of the particle in the physical space at rest with a reference (inertial) frame. ■

1.5. A first look to the infinite dimensional case, CCR and quantization procedures

All the described formalism, barring technicalities we shall examine in the rest of the paper, holds also for quantum systems whose complex vector space of the states is *infinite* dimensional.

To extend the ideas treated in Sect. 1.2 to the general case, dropping the hypothesis that \mathcal{H} is finite dimensional, it seems to be natural to assume that \mathcal{H} is complete with respect to the norm associated to $\langle \cdot | \cdot \rangle$. In particular, completeness assures the existence of spectral decompositions, generalizing (4) for instance when referring to *compact* selfadjoint operators (e.g., see [6]). In other words, \mathcal{H} is a *complex Hilbert space*.

The most elementary example of a quantum system described in an infinite dimensional Hilbert space is a quantum particle whose position is along the axis \mathbb{R} . In this case [1], the Hilbert space is $\mathcal{H} := L^2(\mathbb{R}, dx)$, dx denoting the standard Lebesgue measure on \mathbb{R} . States are still represented by elements of $P\mathcal{H}$, namely equivalence classes $[\psi]$ of measurable functions $\psi : \mathbb{R} \rightarrow \mathbb{C}$ with unit norm, $\|[\psi]\| = \int_{\mathbb{R}} |\psi(x)|^2 dx = 1$.

Remark 11. We therefore have here *two* quotient procedures. ψ and ψ' define the same element $[\psi]$ of $L^2(\mathbb{R}, dx)$ iff $\psi(x) - \psi'(x) \neq 0$ on a zero Lebesgue measure set. Two unit vectors $[\psi]$ and $[\phi]$ define the same state if $[\psi] = e^{ia}[\phi]$ for some $a \in \mathbb{R}$. ■

Notation 12. In the rest of the paper we adopt the standard convention of many textbooks on functional analysis denoting by ψ , instead of $[\psi]$, the elements of spaces L^2 and tacitly identifying pair of functions which are different on a zero measure set. ■

The functions ψ defining (up to zero-measure set and phases) states, are called *wavefunctions*. There is a pair of fundamental observables describing our quantum particle moving in \mathbb{R} . One is the *position observable*. The corresponding selfadjoint operator, denoted by X , is defined as follows

$$(X\psi)(x) := x\psi(x), \quad x \in \mathbb{R}, \quad \psi \in L^2(\mathbb{R}, dx).$$

The other observable is the one associated to the momentum and indicated by P . Restoring \hbar for the occasion, the *momentum operator* is

$$(P\psi)(x) := -i\hbar \frac{d\psi(x)}{dx}, \quad x \in \mathbb{R}, \quad \psi \in L^2(\mathbb{R}, dx).$$

We immediately face several mathematical problems with these, actually quite naive, definitions. Let us first focus on X . First of all, generally $X\psi \notin L^2(\mathbb{R}, dx)$

even if $\psi \in L^2(\mathbb{R}, dx)$. To fix the problem, we can simply restrict the domain of X to the linear subspace of $L^2(\mathbb{R}, dx)$

$$D(X) := \left\{ \psi \in L^2(\mathbb{R}, dx) \mid \int_{\mathbb{R}} |x\psi(x)|^2 dx < +\infty \right\}. \quad (15)$$

Though it holds

$$\langle X\psi | \phi \rangle = \langle \psi | X\phi \rangle \quad \text{for all } \psi, \phi \in D(X), \quad (16)$$

we cannot say that X is selfadjoint simply because we have not yet given the definition of adjoint operator of an operator defined in a non-maximal domain in an infinite dimensional Hilbert space. In this general case, the identity (2) does not define a (unique) operator X^* without further technical requirements. We just say here, to comfort the reader, that X is truly selfadjoint with respect to a general definition we shall give in the next section, when its domain is (15).

Like (3) in the finite dimensional case, the identity (16) implies that all eigenvalues of X must be real if any. Unfortunately, for every fixed $x_0 \in \mathbb{R}$ there is no $\psi \in L^2(\mathbb{R}, dx)$ with $X\psi = x_0\psi$ and $\psi \neq 0$. (A function ψ satisfying $X\psi = x_0\psi$ must also satisfy $\psi(x) = 0$ if $x \neq x_0$, due to the definition of X . Hence $\psi = 0$, as an element of $L^2(\mathbb{R}, dx)$ just because $\{x_0\}$ has zero Lebesgue measure!) All that seems to prevent the existence of a spectral decomposition of X like the one in (4), since X does not admit eigenvectors in $L^2(\mathbb{R}, dx)$ (and *a fortiori* in $D(X)$).

The definition of P suffers from similar troubles. The domain of P cannot be the whole $L^2(\mathbb{R}, dx)$ but should be restricted to a subset of (weakly) differentiable functions with derivative in $L^2(\mathbb{R}, dx)$. The simplest definition is

$$D(P) := \left\{ \psi \in L^2(\mathbb{R}, dx) \mid \exists \text{w-} \frac{d\psi(x)}{dx}, \int_{\mathbb{R}} \left| \text{w-} \frac{d\psi(x)}{dx} \right|^2 dx < +\infty \right\}. \quad (17)$$

Above $\text{w-} \frac{d\psi(x)}{dx}$ denotes the *weak derivative* of ψ^b . As a matter of fact $D(P)$ coincides with the *Sobolev space* $H^1(\mathbb{R})$.

Again, without a precise definition of adjoint operator in an infinite dimensional Hilbert space (with non-maximal domain) we cannot say anything more precise about the selfadjointness of P with that domain. We say however that P turns out to be selfadjoint with respect to the general definition we shall give in the next section provided its domain is (17).

From the definition of the domain of P and passing to the Fourier-Plancherel transform, one finds again (it is not so easy to see it)

$$\langle P\psi | \phi \rangle = \langle \psi | P\phi \rangle \quad \text{for all } \psi, \phi \in D(P), \quad (18)$$

so that, eigenvalues are real if exist. However P does not admit eigenvectors. The naive eigenvectors with eigenvalue $p \in \mathbb{R}$ are functions proportional to the map

^b $f : \mathbb{R} \rightarrow \mathbb{C}$, defined up to zero-measure set, is the weak derivative of $g \in L^2(\mathbb{R}, dx)$ if it holds $\int_{\mathbb{R}} g \frac{dh}{dx} dx = - \int_{\mathbb{R}} fh dx$ for every $h \in C_0^\infty(\mathbb{R})$. If g is differentiable, its standard derivative coincide with the weak one.

$\mathbb{R} \ni x \mapsto e^{ipx/\hbar}$, which does not belong to $L^2(\mathbb{R}, dx)$ nor $D(P)$. We will tackle all these issues in the next section in a very general fashion.

We observe that the space of *Schwartz functions*, $\mathcal{S}(\mathbb{R})$ ^c satisfies

$$\mathcal{S}(\mathbb{R}) \subset D(X) \cap D(P)$$

and furthermore $\mathcal{S}(\mathbb{R})$ is dense in $L^2(\mathbb{R}, dx)$ and *invariant* under X and P : $X(\mathcal{S}(\mathbb{R})) \subset \mathcal{S}(\mathbb{R})$ and $P(\mathcal{S}(\mathbb{R})) \subset \mathcal{S}(\mathbb{R})$.

Remark 13. Though we shall not pursue this approach within these notes, we stress that X admits a set of eigenvectors if we extend the domain of X to the space $\mathcal{S}'(\mathbb{R})$ of *Schwartz distributions* in a standard way: If $T \in \mathcal{S}'(\mathbb{R})$,

$$\langle X(T), f \rangle := \langle T, X(f) \rangle \quad \text{for every } f \in \mathcal{S}(\mathbb{R}).$$

With this extension, the eigenvectors in $\mathcal{S}'(\mathbb{R})$ of X with eigenvalues $x_0 \in \mathbb{R}$ are the distributions $c\delta(x - x_0)$ [1]. This class of eigenvectors can be exploited to build a spectral decomposition of X similar to that in (4).

Similarly, P admits eigenvectors in $\mathcal{S}'(\mathbb{R})$ with the same procedure. They are just the above exponential functions. Again, this class of eigenvectors can be used to construct a spectral decomposition of P like the one in (4). The idea of this procedure can be traced back to Dirac [4] and, in fact, something like ten years later it gave rise to the rigorous *theory of distributions* by L. Schwartz. The modern formulation of this approach to construct spectral decompositions of selfadjoint operators was developed by Gelfand in terms of the so called *rigged Hilbert spaces* [5]. ■

Referring to a quantum particle moving in \mathbb{R}^n , whose Hilbert space is $L^2(\mathbb{R}^n, dx^n)$, one can introduce observables X_k and P_k representing position and momentum with respect to the k -th axis, $k = 1, 2, \dots, n$. These operators, which are defined analogously to the case $n = 1$, have domains smaller than the full Hilbert space. We do not write the form of these domain (where the operators turn out to be properly selfadjoint referring to the general definition we shall state in the next section). We just mention the fact that all these operators admit $\mathcal{S}(\mathbb{R}^n)$ as common invariant subspace included in their domains. Thereon

$$(X_k\psi)(x) = x_k\psi(x), \quad (P_k\psi)(x) = -i\hbar \frac{\partial\psi(x)}{\partial x_k}, \quad \psi \in \mathcal{S}(\mathbb{R}^n) \quad (19)$$

and so

$$\langle X_k\psi|\phi \rangle = \langle \psi|X_k\phi \rangle, \quad \langle P_k\psi|\phi \rangle = \langle \psi|P_k\phi \rangle \quad \text{for all } \psi, \phi \in \mathcal{S}(\mathbb{R}^n), \quad (20)$$

By direct inspection one easily proves that the *canonical commutation relations* (CCR) hold when all the operators in the subsequent formulas are supposed to be

^c $\mathcal{S}(\mathbb{R}^n)$ is the vector space of the C^∞ complex valued functions on \mathbb{R}^n which, together with their derivatives of all orders in every set of coordinate, decay faster than every negative integer power of $|x|$ for $|x| \rightarrow +\infty$.

restricted to $\mathcal{S}(\mathbb{R}^n)$

$$[X_h, P_k] = i\hbar\delta_{hk}I, \quad [X_h, X_k] = 0, \quad [P_h, P_k] = 0. \quad (21)$$

We have introduced the *commutator* $[A, B] := AB - BA$ of the operators A and B generally with different domains, defined on a subspace where both compositions AB and BA makes sense, $\mathcal{S}(\mathbb{R}^n)$ in the considered case. Assuming that (5) and (8) are still valid for X_k and P_k referring to $\psi \in \mathcal{S}(\mathbb{R}^n)$, (21) easily leads to the *Heisenberg uncertainty relations*,

$$\Delta X_{k\psi}\Delta P_{k\psi} \geq \frac{\hbar}{2}, \quad \text{for } \psi \in \mathcal{S}(\mathbb{R}^n), \quad \|\psi\| = 1. \quad (22)$$

Exercise 14. Prove inequality (22) assuming (5) and (8).

Solution. Using (5), (8) and the Cauchy-Schwarz inequality, one easily finds (we omit the index k for simplicity),

$$\Delta X_\psi\Delta P_\psi = \|X'\psi\|\|P'\psi\| \geq |\langle X'\psi|P'\psi\rangle|.$$

where $X' := X - \langle X \rangle_\psi I$ and $P' := P - \langle P \rangle_\psi I$. Next notice that

$$|\langle X'\psi|P'\psi\rangle| \geq |\text{Im}\langle X'\psi|P'\psi\rangle| = \frac{1}{2}|\langle X'\psi|P'\psi\rangle - \langle P'\psi|X'\psi\rangle|$$

Taking advantage from (20) and the definitions of X' and P' and exploiting (21), $|\langle X'\psi|P'\psi\rangle - \langle P'\psi|X'\psi\rangle| = |\langle \psi|(X'P' - P'X')\psi\rangle| = |\langle \psi|(XP - PX)\psi\rangle| = \hbar|\langle \psi|\psi\rangle|$. Since $\langle \psi|\psi\rangle = \|\psi\|^2 = 1$ by hypotheses, (22) is proved. Obviously the open problem is to justify the validity of (5) and (8) also in the infinite dimensional case. \square

Another philosophically important consequence of the CCR (21) is that they resemble the *classical canonical commutation relations* of the Hamiltonian variables q^h and p_k , referring to the standard *Poisson brackets* $\{\cdot, \cdot\}_P$,

$$\{q^h, p_k\}_P = \delta_k^h, \quad \{q^h, q^k\}_P = 0, \quad \{p_h, p_k\}_P = 0. \quad (23)$$

as soon as one identifies $(i\hbar)^{-1}[\cdot, \cdot]$ with $\{\cdot, \cdot\}_P$. This fact, initially noticed by Dirac [4], leads to the idea of “quantization” of a classical Hamiltonian theory [1].

One starts from a classical system described on a symplectic manifold (Γ, ω) , for instance $\Gamma = \mathbb{R}^{2n}$ equipped with the standard symplectic form as ω and considers the (real) Lie algebra $(C^\infty(\Gamma, \mathbb{R}), \{\cdot, \cdot\}_P)$. To “quantize” the system one looks for a map associating classical observables $f \in C^\infty(\Gamma, \mathbb{R})$ to quantum observables O_f , i.e. selfadjoint operators restricted^d to a common invariant domain \mathcal{S} of a certain Hilbert space \mathcal{H} . (In case $\Gamma = T^*Q$, \mathcal{H} can be chosen as $L^2(Q, d\mu)$ where μ is some natural measure.) The map $f \mapsto O_f$ is expected to satisfy a set of constraints. The most important are listed here

^dThe restriction should be such that it admits a unique selfadjoint extension. A sufficient requirement on \mathcal{S} is that every O_f is *essentially selfadjoint* thereon, notion we shall discuss in the next section.

- (1) \mathbb{R} -linearity;
- (2) $O_{id} = I|_{\mathcal{S}}$;
- (3) $O_{\{f,g\}_P} = -i\hbar[O_f, O_g]$
- (4) If (Γ, ω) is \mathbb{R}^{2n} equipped with the standard symplectic form, they must hold $O_{x_k} = X_k|_{\mathcal{S}}$ and $O_{p_k} = P_k|_{\mathcal{S}}$, $k = 1, 2, \dots, n$.

The penultimate requirement says that the map $f \mapsto O_f$ transforms the real Lie algebra $(C^\infty(\Gamma, \mathbb{R}), \{\cdot, \cdot\}_P)$ into a real Lie algebra of operators whose Lie bracket is $i\hbar[O_f, O_g]$. A map fulfilling these constraints, in particular the third one, is possible if f, g are both functions of only the q or the p coordinates separately or if they are linear in them. But it is false already if we consider elementary physical systems [1]. The ultimate reason of this obstructions due to the fact that the operators P_k, X_k do not commute, contrary to the functions p_k, q^k which do. The problem can be solved, in the paradigm of the so-called *Geometric Quantization*[1], replacing $(C^\infty(\Gamma, \mathbb{R}), \{\cdot, \cdot\}_P)$ with a sub-Lie algebra (as large as possible). There are other remarkable procedures of “quantization” in the literature, we shall not insist on them any further here [1].

Example 15.

(a) The full Hilbert space of an electron is therefore given by the tensor product $L^2(\mathbb{R}^3, d^3x) \otimes \mathcal{H}_s \otimes \mathcal{H}_c$.

(b) Consider a particle in $3D$ with mass m , whose potential energy is a bounded-below real function $U \in C^\infty(\mathbb{R}^3)$ with polynomial growth. Classically, its Hamiltonian function reads

$$h := \sum_{k=1}^3 \frac{p_k^2}{2m} + U(x).$$

A brute force quantization procedure in $L^2(\mathbb{R}^3, d^3x)$ consists of replacing every classical object with corresponding operators. It may make sense at most when there are no ordering ambiguities in translating functions like p^2x , since classically $p^2x = pxp = xp^2$, but these identities are false at quantum level. In our case these problems do not arise so that

$$H := \sum_{k=1}^3 \frac{P_k^2}{2m} + U, \tag{24}$$

where $(U\psi)(x) := U(x)\psi(x)$, could be accepted as first quantum model of the Hamiltonian function of our system. The written operator is at least defined on $\mathcal{S}(\mathbb{R}^3)$, where it satisfies $\langle H\psi|\phi \rangle = \langle \psi|H\phi \rangle$. The existence of selfadjoint extensions is a delicate issue [6] we shall not address here. Taking (19) into account, always on $\mathcal{S}(\mathbb{R}^3)$, one immediately finds

$$H := -\frac{\hbar^2}{2m}\Delta + U,$$

where Δ is the standard Laplace operator in \mathbb{R}^3 . If we assume that the equation describing the evolution of the quantum system is again^e (14), in our case we find the known form of the Schrödinger equation,

$$i\hbar \frac{d\psi_t}{dt} = -\frac{\hbar^2}{2m} \Delta \psi_t + U \psi_t,$$

when $\psi_\tau \in \mathcal{S}(\mathbb{R}^3)$ for τ varying in a neighborhood of t (this requirement may be relaxed). Actually the meaning of the derivative on the left-hand side should be specified. We only say here that it is computed with respect to the natural topology of $L^2(\mathbb{R}^3, d^3x)$. ■

2. Observables in infinite dimensional Hilbert spaces: Spectral Theory

The main goal of this section is to present a suitable mathematical theory, sufficient to extend to the infinite dimensional case the mathematical formalism of QM introduced in the previous section. As seen in Sect. 1.5, the main issue concerns the fact that, in the infinite dimensional case, there are operators representing observables which do not have proper eigenvalues and eigenvectors, like X and P . So, naive expansions as (4) cannot be literally extended to the general case. These expansions, together with the interpretation of the eigenvalues as values attained by the observable associated with a selfadjoint operator, play a crucial rôle in the mathematical interpretation of the quantum phenomenology introduced in Sect. 1.1 and mathematically discussed in Sect. 1.2. In particular we need a precise definition of selfadjoint operator and a result regarding a spectral decomposition in the infinite dimensional case. These tools are basic elements of the so called *spectral theory in Hilbert spaces*, literally invented by von Neumann in his famous book [7] to give a rigorous form to Quantum Mechanics and successively developed by various authors towards many different directions of pure and applied mathematics. The same notion of abstract Hilbert space, as nowadays known, was born in the second chapter of that book, joining and elaborating previous mathematical constructions by Hilbert and Riesz. The remaining part of this section is devoted to introduce the reader to some basic elements of that formalism. Reference books are, e.g., [8,6,9,10]

2.1. Hilbert spaces

A **Hermitian scalar product** over the complex Hilbert space \mathcal{H} is a map

$$\langle \cdot | \cdot \rangle : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$$

such that, for $a, b \in \mathbb{C}$ and $x, y, z \in \mathcal{H}$,

^eA factor \hbar has to be added in front of the left-hand side of (14) if we deal with a unit system where $\hbar \neq 1$.

- (i) $\langle x|y \rangle = \overline{\langle y|x \rangle}$,
- (ii) $\langle x|ay + bz \rangle = a\langle x|y \rangle + b\langle x|z \rangle$,
- (iii) $\langle x|x \rangle \geq 0$ with $x = 0$ if $\langle x|x \rangle = 0$.

The space \mathcal{H} is said to be a (complex) **Hilbert space** if is complete with respect to the natural norm $\|x\| := \sqrt{\langle x|x \rangle}$, $x \in \mathcal{H}$.

In particular, just in view of positivity of the scalar product and regardless the completeness property, the **Cauchy-Schwarz inequality** holds

$$|\langle x|y \rangle| \leq \|x\| \|y\|, \quad x, y \in \mathcal{H}.$$

Another elementary purely algebraic fact is the **polarization identity** concerning the Hermitian scalar product (here, \mathcal{H} is not necessarily complete)

$$4\langle x|y \rangle = \|x + y\|^2 - \|x - y\|^2 - i\|x + iy\|^2 + i\|x - iy\|^2 \quad \text{for of } x, y \in \mathcal{H}, \quad (25)$$

which immediately implies the following elementary result.

Proposition 16. *If \mathcal{H} is a complex vector space with Hermitian scalar product $\langle \cdot | \cdot \rangle$, a linear map $L : \mathcal{H} \rightarrow \mathcal{H}$ which is an isometry - $\|Lx\| = \|x\|$ if $x \in \mathcal{H}$ - also preserves the scalar product - $\langle Lx|Ly \rangle = \langle x|y \rangle$ for $x, y \in \mathcal{H}$.*

The converse proposition is obviously true.

Similarly to the above identity we have another useful identity, for every operator $A : \mathcal{H} \rightarrow \mathcal{H}$:

$$\begin{aligned} 4\langle x|Ay \rangle &= \langle x + y|A(x + y) \rangle - \langle x - y|A(x - y) \rangle - i\langle x + iy|A(x + iy) \rangle \\ &\quad + i\langle x - iy|A(x - iy) \rangle \quad \text{for of } x, y \in \mathcal{H}. \end{aligned} \quad (26)$$

That in particular proves that *if $\langle x|Ax \rangle = 0$ for all $x \in \mathcal{H}$, then $A = 0$.*

(Observe that this fact is not generally true if dealing with real vector spaces equipped with a real symmetric scalar product.)

We henceforth assume that the reader be familiar with the basic theory of normed, Banach and Hilbert spaces and notions like *orthogonal of a set*, *Hilbertian basis* (also called *complete orthonormal systems*) and that their properties and use be well known [8,6]. We summarize here some basic results concerning *orthogonal sets* and *Hilbertian bases*.

Notation 17. If $M \subset \mathcal{H}$, $M^\perp := \{y \in \mathcal{H} \mid \langle y|x \rangle = 0 \quad \forall x \in M\}$ denotes the **orthogonal** of M . ■

Evidently M^\perp is a closed subspace of \mathcal{H} . $^\perp$ enjoys several nice properties quite easy to prove (e.g., see [8,6]), in particular,

$$\overline{\text{span}(M)} = (M^\perp)^\perp \quad \text{and} \quad \mathcal{H} = \overline{\text{span}(M)} \oplus M^\perp \quad (27)$$

where the bar denotes the topological closure and \oplus the direct sum of (orthogonal) subspaces.

Definition 18. A Hilbertian basis N of a Hilbert space \mathcal{H} is a set of **orthonormal vectors**, i.e., $\|u\| = 1$ and $\langle u|v \rangle = 0$ for $u, v \in N$ with $u \neq v$, such that if $s \in \mathcal{H}$ satisfies $\langle s|u \rangle = 0$ for every $u \in N$ then $s = 0$.

Hilbertian bases always exist as a consequence of Zorn's lemma. Notice that, as consequence of (27),

Proposition 19. A set of orthonormal vectors $N \subset \mathcal{H}$ is a Hilbertian basis of the Hilbert space \mathcal{H} if and only if $\overline{\text{span}(N)} = \mathcal{H}$.

Furthermore a generalized version of *Pythagorean theorem* holds true.

Proposition 20. $N \subset \mathcal{H}$ is a Hilbertian basis of the Hilbert space \mathcal{H} if and only if

$$\|x\|^2 = \sum_{u \in N} |\langle u|x \rangle|^2 \quad \text{for every } x \in \mathcal{H}.$$

The sum appearing in the above decomposition of x is understood as the *supremum* of the sums $\sum_{u \in F} |\langle u|x \rangle|^2$ for every finite set $F \subset N$. As a consequence (e.g., see [8,6]), only a set at most countable of elements $|\langle u|x \rangle|^2$ do not vanish and the sum is therefore interpreted as a standard series and it can be re-ordered preserving the sum because it absolutely converges. It turns out that all Hilbertian bases of \mathcal{H} have the same cardinality and \mathcal{H} is separable (it admits a dense countable subset) if and only if \mathcal{H} has an either finite or countable Hilbertian basis. If $M \subset \mathcal{H}$ is a set of orthonormal vectors, the weaker so-called **Bessel inequality** holds

$$\|x\|^2 \geq \sum_{u \in M} |\langle u|x \rangle|^2 \quad \text{for every } x \in \mathcal{H},$$

so that Hilbertian bases are exactly orthonormal sets saturating the inequality.

Finally, if $N \subset \mathcal{H}$ is an Hilbertian basis, the decompositions hold for every $x, y \in \mathcal{H}$

$$x = \sum_{u \in N} \langle u|x \rangle u, \quad \langle x|y \rangle = \sum_{u \in N} \langle x|u \rangle \langle u|y \rangle.$$

In view of the already mentioned fact that only a finite or countable set of elements $\langle u_n|x \rangle$ do not vanish, the first sum is actually a finite sum or at most series $\lim_{m \rightarrow +\infty} \sum_{n=0}^m \langle u_n|x \rangle u_n$ computed with respect to the norm of \mathcal{H} , where the order used to label the element u_n does not matter as the series absolutely converges. The second sum is similarly absolutely convergent so that, again, it can be re-ordered arbitrarily.

We remind the reader the validity of an elementary though fundamental technical result (e.g., see [8,6]):

Theorem 21 (Riesz' lemma). Let \mathcal{H} be a complex Hilbert space. $\phi : \mathcal{H} \rightarrow \mathbb{C}$ is linear and continuous if and only if has the form $\phi = \langle x| \cdot \rangle$ for some $x \in \mathcal{H}$. The vector x is uniquely determined by ϕ .

2.2. Classes of (especially unbounded) operators in Hilbert spaces

Our goal is to present some basic results of *spectral analysis*, useful in QM. From now on, an **operator** A in \mathcal{H} always means a *linear* map $A : D(A) \rightarrow \mathcal{H}$, whose **domain**, $D(A) \subset \mathcal{H}$, is a *subspace* of \mathcal{H} . In particular, I always denotes the **identity operator** defined on the *whole* space ($D(I) = \mathcal{H}$)

$$I : \mathcal{H} \ni x \mapsto x \in \mathcal{H} .$$

If A is an operator in \mathcal{H} , $Ran(A) := \{Ax \mid x \in D(A)\}$ is the **image** or **range** of A .

Notation 22. If A and B are operators in \mathcal{H}

$$A \subset B \text{ means that } D(A) \subset D(B) \text{ and } B|_{D(A)} = A,$$

where $|_S$ is the standard “restriction to S ” symbol. We also adopt usual conventions regarding **standard domains** of combinations of operators A, B :

- (i) $D(AB) := \{x \in D(B) \mid Bx \in D(A)\}$ is the domain of AB ,
- (ii) $D(A + B) := D(A) \cap D(B)$ is the domain of $A + B$,
- (ii) $D(\alpha A) = D(A)$ for $\alpha \neq 0$ is the domain of αA . ■

With these definitions it is easy to prove that

- (1) $(A + B) + C = A + (B + C)$,
- (2) $A(BC) = (AB)C$,
- (3) $A(B + C) = AB + AC$,
- (4) $(B + C)A \supset BA + CA$,
- (5) $A \subset B$ and $B \subset C$ imply $A \subset C$,
- (6) $A \subset B$ and $B \subset A$ imply $A = B$,
- (7) $AB \subset BA$ implies $A(D(B)) \subset D(B)$ if $D(A) = \mathcal{H}$,
- (8) $AB = BA$ implies $D(B) = A^{-1}(D(B))$ if $D(A) = \mathcal{H}$ (so, $A(D(B)) = D(B)$ is A is surjective).

To go on, we define some abstract algebraic structures naturally arising in the space of operators on a Hilbert space.

Definition 23. Let \mathfrak{A} be an associative complex algebra \mathfrak{A} .

(1) \mathfrak{A} is a **Banach algebra** if it is a Banach space such that $\|ab\| \leq \|a\| \|b\|$ for $a, b \in \mathfrak{A}$. An **unital** Banach algebra is a Banach algebra with unit multiplicative element $\mathbb{1}$, satisfying $\|\mathbb{1}\| = 1$.

(2) \mathfrak{A} is an (unital) ***-algebra** if it is an (unital) algebra equipped with an anti linear map $\mathfrak{A} \ni a \mapsto a^* \in \mathfrak{A}$, called **involution**, such that $(a^*)^* = a$ and $(ab)^* = b^*a^*$ for $a, b \in \mathfrak{A}$.

(3) \mathfrak{A} is a (unital) **C^* -algebra** if it is a (unital) Banach algebra \mathfrak{A} which is also a *-algebra and $\|a^*a\| = \|a\|^2$ for $a \in \mathfrak{A}$.

A ***-homomorphism** from the *-algebra \mathcal{A} to the *-algebra \mathcal{B} is an algebra homomorphism preserving the involutions (and the unities if both present). A bijective *-homomorphism is called ***-isomorphism**. ■

Exercise 24. Prove that $\mathbb{1}^* = \mathbb{1}$ in a unital $*$ -algebra and that $\|a^*\| = \|a\|$ if $a \in \mathfrak{A}$ when \mathfrak{A} is a C^* -algebra.

Solution. From $\mathbb{1}a = a\mathbb{1} = a$ and the definition of $*$, we immediately have $a^*\mathbb{1}^* = \mathbb{1}^*a^* = a^*$. Since $(b^*)^* = b$, we have found that $b\mathbb{1}^* = \mathbb{1}^*b = b$ for every $b \in \mathfrak{A}$. Uniqueness of the unit implies $\mathbb{1}^* = \mathbb{1}$. Regarding the second property, $\|a\|^2 = \|a^*a\| \leq \|a^*\| \|a\|$ so that $\|a\| \leq \|a^*\|$. Everywhere replacing a for a^* and using $(a^*)^*$, we also obtain $\|a^*\| \leq \|a\|$, so that $\|a^*\| = \|a\|$. \square

We remind the reader that a linear map $A : X \rightarrow Y$, where X and Y are normed complex vector spaces with resp. norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, is said to be **bounded** if

$$\|Ax\|_Y \leq b\|x\|_X \quad \text{for some } b \in [0, +\infty) \text{ and all } x \in X. \quad (28)$$

As is well known [8,6], it turns out that

Proposition 25. An operator $A : X \rightarrow Y$, with X, Y normed spaces, is continuous if and only if it is bounded.

Proof. It is evident that A bounded is continuous because, for $x, x' \in X$, $\|Ax - Ax'\|_Y \leq b\|x - x'\|_X$. Conversely, if A is continuous then it is continuous for $x = 0$, so $\|Ax\|_Y \leq \epsilon$ for $\epsilon > 0$ if $\|x\|_X < \delta$ for $\delta > 0$ sufficiently small. If $\|x\|_X = \delta/2$ we therefore have $\|Ax\|_Y < \epsilon$ and thus, dividing by $\delta/2$, we also find $\|Ax'\|_Y < 2\epsilon/\delta$, where $\|x'\|_X = 1$. Multiplying for $\lambda > 0$, $\|A\lambda x'\|_Y < 2\lambda\epsilon/\delta$ which can be re-written $\|Ax\|_Y < 2\frac{\epsilon}{\delta}\|x\|_X$ for every $x \in X$, proving that A is bounded. \square

For bounded operators it is possible to define the **operator norm**,

$$\|A\| := \sup_{0 \neq x \in X} \frac{\|Ax\|_Y}{\|x\|_X} \quad \left(= \sup_{x \in X, \|x\|_X=1} \|Ax\|_Y \right).$$

It is easy to prove that this is a true norm on the complex vector spaces $\mathfrak{B}(X, Y)$ of bounded operators $T : X \rightarrow Y$, with X, Y complex normed spaces when defining the linear combination of operators $\alpha A + \beta B \in \mathfrak{B}(X, Y)$ for $\alpha, \beta \in \mathbb{C}$ and $A, B \in \mathfrak{B}(X, Y)$ by $(\alpha A + \beta B)x := \alpha Ax + \beta Bx$ for every $x \in X$.

An important elementary technical result is stated in the following proposition.

Proposition 26. Let $A : S \rightarrow Y$ be a bounded operator defined on the subspace $S \subset X$, where X, Y are normed spaces with Y complete. If S is dense in X , then A can be continuously extended to a unique bounded operator $A_1 : X \rightarrow Y$. It holds $\|A_1\| = \|A\|$.

Proof. Uniqueness is obvious from continuity: if $S \ni x_n \rightarrow x \in X$ and A_1, A'_1 are continuous extensions, $A_1x - A'_1x = \lim_{n \rightarrow +\infty} A_1x_n - A'_1x_n = \lim_{n \rightarrow +\infty} 0 = 0$. Let us construct a linear continuous extension. If $x \in X$ there is a sequence $S \ni x_n \rightarrow x \in X$ since S is dense. $\{Ax_n\}_{n \in \mathbb{N}}$ is Cauchy because $\{x_n\}_{n \in \mathbb{N}}$ is and $\|Ax_n - Ax_m\|_Y \leq \|A\| \|x_n - x_m\|_X$. So the limit $A_1x := \lim_{n \rightarrow +\infty} Ax_n$ exists because Y is complete.

The limit does not depend on the sequence: if $S \ni x'_n \rightarrow x$, then $\|Ax_n - Ax'_n\| \leq \|A\| \|x_n - x'_n\| \rightarrow 0$, so A_1 is well defined. It is immediately proved that A_1 is linear form linearity of A and therefore A_1 is an operator extending A on the whole X . By construction $\|A_1x\|_Y = \lim_{n \rightarrow +\infty} \|Ax_n\|_Y \leq \lim_{n \rightarrow +\infty} \|A\| \|x_n\|_X \leq \|A\| \|x\|_X$, so $\|A_1\| \leq \|A\|$, in particular A_1 is bounded. On the other hand

$$\begin{aligned} \|A_1\| &= \sup\{\|A_1x\| \|x\|^{-1} \mid x \in X \setminus \{0\}\} \geq \sup\{\|A_1x\| \|x\|^{-1} \mid x \in S \setminus \{0\}\} \\ &= \sup\{\|Ax\| \|x\|^{-1} \mid x \in S \setminus \{0\}\} = \|A\|, \end{aligned}$$

so that $\|A_1\| \geq \|A\|$ and therefore $\|A_1\| = \|A\|$. \square

From now on, $\mathfrak{B}(\mathcal{H}) := \mathcal{B}(\mathcal{H}, \mathcal{H})$ denotes the vecotor space of bounded operators $A : \mathcal{H} \rightarrow \mathcal{H}$ over the complex Hilbert space \mathcal{H} .

$\mathfrak{B}(\mathcal{H})$ acquires the structure of a *unital Banach algebra*: The complex vector space structure is the standard one of operators, the associative algebra product is the composition of operators with unit given by I , and the norm being the above defined operator norm,

$$\|A\| := \sup_{0 \neq x \in \mathcal{H}} \frac{\|Ax\|}{\|x\|}.$$

This definition of $\|A\|$ can be given also for an operator $A : D(A) \rightarrow \mathcal{H}$, if A is bounded and $D(A) \subset \mathcal{H}$ but $D(A) \neq \mathcal{H}$. It immediately arises that

$$\|Ax\| \leq \|A\| \|x\| \quad \text{if } x \in D(A).$$

As we already know, $\|\cdot\|$ is a norm over $\mathfrak{B}(\mathcal{H})$. Furthermore it satisfies

$$\|AB\| \leq \|A\| \|B\| \quad A, B \in \mathfrak{B}(\mathcal{H}).$$

It is also evident that $\|I\| = 1$. Actually $\mathfrak{B}(\mathcal{H})$ is a Banach space so that: $\mathfrak{B}(\mathcal{H})$ is a *unital Banach algebra*. In fact, a fundamental result is the following theorem.

Theorem 27. *If \mathcal{H} is a Hilbert space, $\mathfrak{B}(\mathcal{H})$ is a Banach space with respect to the norm of operators.*

Proof. The only non-trivial property id completeness of $\mathfrak{B}(\mathcal{H})$. Let us prove it. Consider a Cauchy sequence $\{T_n\}_{n \in \mathbb{N}} \subset \mathfrak{B}(\mathcal{H})$. We want to prove that there exists $T \in \mathfrak{B}(\mathcal{H})$ with $\|T - T_n\| \rightarrow 0$ as $n \rightarrow +\infty$. Let us define $Tx = \lim_{n \rightarrow +\infty} T_n x$ for every $x \in \mathcal{H}$. The limit exists because $\{T_n x\}_{n \in \mathbb{N}}$ is Cauchy from $\|T_n x - T_m x\| \leq \|T_n - T_m\| \|x\|$. Linearity of T is easy to prove from linearity of every T_n . Next observe that $\|Tx - T_m x\| = \|\lim_n T_n x - T_m x\| = \lim_n \|T_n x - T_m x\| \leq \epsilon \|x\|$ is m is sufficiently large. Assuming that $T \in \mathfrak{B}(\mathcal{H})$, the found inequality, dividing by $\|x\|$ and taking the sup over x with $\|x\| \neq 0$ proves that $\|T - T_m\| \leq \epsilon$ and thus $\|T - T_m\| \rightarrow 0$ for $m \rightarrow +\infty$ as wanted. This concludes the proof because $T \in \mathfrak{B}(\mathcal{H})$: $\|Tx\| \leq \|Tx - T_m x\| + \|T_m x\| \leq \epsilon \|x\| + \|T_m\| \|x\|$ and thus $\|T\| \leq (\epsilon + \|T_m\|) < +\infty$. \square

Remark 28. The result, with the same proof is valid for the above defined complex vectorspace $\mathfrak{B}(X, Y)$, provided the normed space Y is $\|\cdot\|_Y$ -complete. In particular the **topological dual of X** $X' = \mathfrak{B}(X, \mathbb{C})$ with X complex normed space is always complete since \mathbb{C} is complete.

$\mathfrak{B}(\mathcal{H})$ is more strongly a unital C^* -algebra, if we introduce the notion of adjoint of an operator. To this end, we have the following general definition concerning also unbounded operators defined on non-maximal domains.

Definition 29. Let A be a densely defined operator in the complex Hilbert space \mathcal{H} . Define the subspace of \mathcal{H} ,

$$D(A^*) := \{y \in \mathcal{H} \mid \exists z_y \in \mathcal{H} \text{ s.t. } \langle y|Ax \rangle = \langle z_y|x \rangle \forall x \in D(A)\} .$$

The linear map $A^* : D(A^*) \ni y \mapsto z_y$ is called the **adjoint** operator of A . ■

Remark 30.

(a) Above, z_y is uniquely determined by y , since $D(A)$ is dense. If both z_y, z'_y satisfy $\langle y|Ax \rangle = \langle z_y|x \rangle$ and $\langle y|Ax \rangle = \langle z'_y|x \rangle$, then $\langle z_y - z'_y|x \rangle = 0$ for every $x \in D(A)$. Taking a sequence $D(A) \ni x_n \rightarrow z_y - z'_y$, we conclude that $\|z_y - z'_y\| = 0$. Thus $z_y = z'_y$. The fact that $y \mapsto z_y$ is linear can immediately be checked.

(b) By construction, we immediately have that

$$\langle A^*y|x \rangle = \langle y|Ax \rangle \quad \text{for } x \in D(A) \text{ and } y \in D(A^*)$$

and also

$$\langle x|A^*y \rangle = \langle Ax|y \rangle \quad \text{for } x \in D(A) \text{ and } y \in D(A^*) ,$$

if taking the complex conjugation of the former identity. ■

Exercise 31. Prove that $D(A^*)$ can equivalently be defined as the set (subspace) of $y \in \mathcal{H}$ such that the linear functional $D(A) \ni x \mapsto \langle y|Ax \rangle$ is continuous.

Solution. It is a simple application of Riesz' lemma, after having uniquely extended $D(A) \ni x \mapsto \langle y|Ax \rangle$ to a continuous linear functional defined on $\overline{D(A)} = \mathcal{H}$ by continuity. □

Remark 32.

(a) If A is densely defined and A^* is also densely defined then $A \subset (A^*)^*$. The proof immediately follows from the definition of adjoint.

(b) If A is densely defined and $A \subset B$ then $B^* \subset A^*$. The proof immediately follows from the definition of adjoint.

(c) If $A \in \mathfrak{B}(\mathcal{H})$ then $A^* \in \mathfrak{B}(\mathcal{H})$ and $(A^*)^* = A$. Moreover

$$\|A^*\|^2 = \|A\|^2 = \|A^*A\| = \|AA^*\| .$$

(d) Directly from given definition of adjoint one has, for densely defined operators A, B on \mathcal{H} ,

$$A^* + B^* \subset (A + B)^* \quad \text{and} \quad A^*B^* \subset (BA)^* .$$

Furthermore

$$A^* + B^* = (A + B)^* \quad \text{and} \quad A^*B^* = (BA)^*, \quad (29)$$

whenever $B \in \mathfrak{B}(\mathcal{H})$ and A is densely defined.

(e) From (c) and the last statement in (d) in particular, it is clear that $\mathfrak{B}(\mathcal{H})$ is a unital C^* -algebra with involution $\mathfrak{B}(\mathcal{H}) \ni A \mapsto A^* \in \mathfrak{B}(\mathcal{H})$. ■

Definition 33 (*-representation). If \mathfrak{A} is a (unital) $*$ -algebra and \mathcal{H} a Hilbert space, a ***-representation** on \mathcal{H} is a $*$ -homomorphism $\pi : \mathfrak{A} \rightarrow \mathfrak{B}(\mathcal{H})$ referring to the natural (unital) $*$ -algebra structure of $\mathfrak{B}(\mathcal{H})$. ■

Exercise 34. Prove that $A^* \in \mathfrak{B}(\mathcal{H})$ if $A \in \mathfrak{B}(\mathcal{H})$ and that, in this case $(A^*)^* = A$, $\|A\| = \|A^*\|$ and $\|A^*A\| = \|AA^*\| = \|A\|^2$.

Solution. If $A \in \mathfrak{B}(\mathcal{H})$, for every $y \in \mathcal{H}$, the linear map $\mathcal{H} \ni x \mapsto \langle y|Ax \rangle$ is continuous ($|\langle y|Ax \rangle| \leq \|y\| \|Ax\| \leq \|y\| \|A\| \|x\|$) therefore Theorem 21 proves that there exists a unique $z_{y,A} \in \mathcal{H}$ with $\langle y|Ax \rangle = \langle z_{y,A}|x \rangle$ for all $x, y \in \mathcal{H}$. The map $\mathcal{H} \ni y \mapsto z_{y,A}$ is linear as consequence of the said uniqueness and the antilinearity of the left entry of scalar product. The map $\mathcal{H} \ni y \mapsto z_{y,A}$ fits the definition of A^* , so it coincides with A^* and $D(A^*) = \mathcal{H}$. Since $\langle A^*x|y \rangle = \langle x|Ay \rangle$ for $x, y \in \mathcal{H}$ implies (taking the complex conjugation) $\langle y|A^*x \rangle = \langle Ay|x \rangle$ for $x, y \in \mathcal{H}$, we have $(A^*)^* = A$. To prove that A^* is bounded observe that $\|A^*x\|^2 = \langle A^*x|A^*x \rangle = \langle x|AA^*x \rangle \leq \|x\| \|A\| \|A^*x\|$ so that $\|A^*x\| \leq \|A\| \|x\|$ and $\|A^*\| \leq \|A\|$. Using $(A^*)^* = A$ we have $\|A^*\| = \|A\|$. Regarding the last identity, it is evidently enough to prove that $\|A^*A\| = \|A\|^2$. First of all $\|A^*A\| \leq \|A^*\| \|A\| = \|A\|^2$, so that $\|A^*A\| \leq \|A\|^2$. On the other hand $\|A\|^2 = (\sup_{\|x\|=1} \|Ax\|)^2 = \sup_{\|x\|=1} \|Ax\|^2 = \sup_{\|x\|=1} \langle Ax|Ax \rangle = \sup_{\|x\|=1} \langle x|A^*Ax \rangle \leq \sup_{\|x\|=1} \|x\| \|A^*Ax\| = \sup_{\|x\|=1} \|A^*Ax\| = \|A^*A\|$. We have found that $\|A^*A\| \leq \|A\|^2 \leq \|A^*A\|$ so that $\|A^*A\| = \|A\|^2$. □

Definition 35. Let A be an operator in the complex Hilbert space \mathcal{H} .

(1) A is said to be **closed** if the **graph** of A , that is the set

$$G(A) := \{(x, Ax) \in \mathcal{H} \times \mathcal{H} \mid x \in D(A)\},$$

is closed in the product topology of $\mathcal{H} \times \mathcal{H}$.

(2) A is **closable** if it admits extensions in terms of closed operators. This is equivalent to saying that the closure of the graph of A is the graph of an operator, denoted by \overline{A} , and called the **closure** of A .

(3) If A is closable, a subspace $S \subset D(A)$ is called **core** for A if $\overline{A|_S} = \overline{A}$. ■

Referring to (2), we observe that, given an operator A , we can always define the closure of the graph $\overline{G(A)}$ in $\mathcal{H} \times \mathcal{H}$. In general this closure is not the graph of an operator, because there could exist sequences $D(A) \ni x_n \rightarrow x$ and $D(A) \ni x'_n \rightarrow x$ such that $Tx_n \rightarrow y$ and $Tx'_n \rightarrow y'$ with $y \neq y'$. However, both pairs (x, y) and (x, y') belong to $\overline{G(A)}$. If this case does not take place – and this is equivalent to condition (a) below when making use of linearity – $\overline{G(A)}$ is the graph of an

operator, denoted by \overline{A} , that is closed by definition. Therefore A admits closed operatorial extensions: at least \overline{A} . If, conversely, A admits extensions in terms of closed operators, the intersection of the (closed) graphs of all these extensions $\overline{G(A)}$ is still closed and it is still the graph of an operator again, which must coincide with \overline{A} by definition.

Remark 36.

- (a) Directly from the definition and using linearity, A is closable if and only if there are no sequences of elements $x_n \in D(A)$ such that $x_n \rightarrow 0$ and $Ax_n \rightarrow y$ with $y \neq 0$ as $n \rightarrow +\infty$. Since $\overline{G(A)}$ is the union of $G(A)$ and its accumulation points in $\mathcal{H} \times \mathcal{H}$ and, if A is closable, it is also the graph of the operator \overline{A} , we conclude that
- (i) $D(\overline{A})$ is made of the elements $x \in \mathcal{H}$ such that $x_n \rightarrow x$ and $Ax_n \rightarrow y_x$ for some sequences $\{x_n\}_{n \in \mathbb{N}} \subset D(A)$ and some $y_x \in D(A)$ and
 - (ii) $\overline{A}x = y_x$.

(b) As a consequence of (a) one has that, if A is closable, then $aA + bI$ is closable and $\overline{aA + bI} = a\overline{A} + bI$ for every $a, b \in \mathbb{C}$.

N.B. *This result generally fails if replacing I for some closable operator B .*

(c) Directly from the definition, A is closed if and only if $D(A) \ni x_n \rightarrow x \in \mathcal{H}$ and $Ax_n \rightarrow y \in \mathcal{H}$ imply both $x \in D(A)$ and $y = Ax$. ■

An useful proposition is the following.

Proposition 37. *Consider an operator $A : D(A) \rightarrow \mathcal{H}$ with $D(A)$ dense in the Hilbert space \mathcal{H} . The following facts hold.*

- (a) A^* is closed.
- (b) A is closable if and only if $D(A^*)$ is dense and, in this case, $\overline{A} = (A^*)^*$.

Proof. Define on the Hermitian scalar product $((x, y)|(x'y')) := \langle x|x' \rangle + \langle y|y' \rangle$ on the vector space defined as the standard direct sum of vector spaces, we henceforth denote by $\mathcal{H} \oplus \mathcal{H}$. It is immediately proved that $\mathcal{H} \oplus \mathcal{H}$ becomes a Hilbert space when equipped with that scalar product. Next define the operator

$$\tau : \mathcal{H} \oplus \mathcal{H} \ni (x, y) \mapsto (-y, x) \in \mathcal{H} \oplus \mathcal{H} .$$

It is easy to check that $\tau \in \mathfrak{B}(\mathcal{H} \oplus \mathcal{H})$ and also (referring the adjoint to the space $\mathcal{H} \oplus \mathcal{H}$),

$$\tau^* = \tau^{-1} = -\tau . \tag{30}$$

Finally, by direct computation one sees that τ and $^\perp$ (referred to $\mathcal{H} \oplus \mathcal{H}$ with the said scala product) commute

$$\tau(F^\perp) = (\tau(F))^\perp \quad \text{if } F \subset \mathcal{H} \oplus \mathcal{H} . \tag{31}$$

Let us pass to prove (a). The following noticeable identity holds true, for an operator $A : D(A) \rightarrow \mathcal{H}$ with $D(A)$ dense in \mathcal{H} (so that A^* is defined)

$$G(A^*) = \tau(G(A))^\perp . \tag{32}$$

Since the right hand side is closed (it being the orthogonal space of a set), the graph of A^* is closed and A^* is therefore closed by definition. To prove (32) observe that, in view of the definition of τ ,

$$\tau(G(A))^\perp = \{(y, z) \in \mathcal{H} \oplus \mathcal{H} \mid \langle (y, z) | (-Ax, x) \rangle = 0, \forall x \in D(A)\}$$

that is

$$\tau(G(A))^\perp = \{(y, z) \in \mathcal{H} \oplus \mathcal{H} \mid \langle y | Ax \rangle = \langle z | x \rangle, \forall x \in D(A)\}$$

Since A^* exists, the pairs $(y, z) \in \tau(G(A))^\perp$ can be written (y, A^*y) according to the definition of A^* . Therefore $\tau(G(A))^\perp = G(A^*)$ proving (a).

(b) From the properties of $^\perp$ we immediately have $\overline{G(D(A))} = (G(A)^\perp)^\perp$. Since τ and $^\perp$ commute by (31), and $\tau\tau = -I$ (30),

$$\overline{G(D(A))} = -\tau \circ \tau((G(A)^\perp)^\perp) = -\tau(\tau(G(A))^\perp)^\perp = \tau(\tau(G(A))^\perp)^\perp = \tau(G(A^*))^\perp,$$

where we have omitted the minus sign since it appears in front of a subspace that, by definition, is closed with respect multiplication by scalars and we make use of (32). Now suppose that $D(A^*)$ is dense so that $(A^*)^*$ exists. In this case, taking advantage of (32) again, we have $\overline{G(A)} = G((A^*)^*)$. Notice that the right-hand side is the graph of an operator so we have obtained that, if $D(A^*)$ is dense, then A is closable. In this case, by definition of closure of an operator, we also have $\overline{A} = (A^*)^*$.

Vice versa, suppose that A is closable, so the operator \overline{A} exists and $G(\overline{A}) = \overline{G(A)}$. In this case $\tau(G(A^*))^\perp = \overline{G(A)}$ is the graph of an operator and thus its graph cannot include pairs $(0, y)$ with $y \neq 0$ by linearity. In other words, if $(0, y) \in \tau(G(A^*))^\perp$ then $y = 0$. This is the same as saying that $\langle (0, y) | (-A^*x, x) \rangle = 0$ for all $x \in D(A^*)$ implies $y = 0$. Summing up, $\langle y | x \rangle = 0$ for all $x \in D(A^*)$ implies $y = 0$. Since $\mathcal{H} = D(A^*)^\perp \oplus (D(A^*)^\perp)^\perp = D(A^*)^\perp \oplus \overline{D(A^*)}$, we conclude that $\overline{D(A^*)} = \mathcal{H}$, that is $D(A^*)$ is dense. \square

An immediate corollary follows.

Corollary 38. *Let $A : D(A) \rightarrow \mathcal{H}$ an operator in the Hilbert space \mathcal{H} . If both $D(A)$ and $D(A^*)$ are densely defined then*

$$A^* = \overline{A^*} = \overline{A^*} = (((A^*)^*)^*).$$

The Hilbert space version of the **closed graph theorem** holds (e.g., see [6]).

Theorem 39 (Closed graph Theorem). *Let $A : \mathcal{H} \rightarrow \mathcal{H}$ be an operator, \mathcal{H} being a complex Hilbert space. A is closed if and only if $A \in \mathfrak{B}(\mathcal{H})$.*

Exercise 40. *Prove that, if $B \in \mathfrak{B}(\mathcal{H})$ and A is a closed operator in \mathcal{H} such that $\text{Ran}(B) \subset D(A)$, then $AB \in \mathfrak{B}(\mathcal{H})$.*

Solution. AB is well defined by hypothesis and $D(AB) = \mathcal{H}$. Exploiting (c) in remark 36 and continuity of B , one easily finds that AB is closed as well. Theorem 39 finally proves that $AB \in \mathfrak{B}(\mathcal{H})$. \square

Definition 41. An operator A in the complex Hilbert space \mathcal{H} is said to be

- (0) **Hermitian** if $\langle Ax|y\rangle = \langle x|Ay\rangle$ for $x, y \in D(A)$,
- (1) **symmetric** if it is densely defined and Hermitian, which is equivalent to say that $A \subset A^*$.
- (2) **selfadjoint** if it is symmetric and $A = A^*$,
- (3) **essentially self-adjoint** if it is symmetric and $(A^*)^* = A^*$.
- (4) **unitary** if $A^*A = AA^* = I$,
- (5) **normal** if it is closed, densely defined and $AA^* = A^*A$. ■

Remark 42.

(a) If A is unitary then $A, A^* \in \mathfrak{B}(\mathcal{H})$. Furthermore $A : \mathcal{H} \rightarrow \mathcal{H}$ is unitary if and only if it is surjective and norm preserving. (See the exercises 46 below). These operators are nothing but the **automorphisms** of the given Hilbert space. Considering two Hilbert spaces $\mathcal{H}, \mathcal{H}'$ **isomorphisms** are linear maps $T : \mathcal{H} \rightarrow \mathcal{H}'$ which are isometric and surjective. Notice that T also preserve the scalar products in view of Proposition 17.

(b) A selfadjoint operator A does not admit proper symmetric extensions and essentially selfadjoint operators admit only one self-adjoint extension. (See Proposition 43 below).

(c) A symmetric operator A is always closable because $A \subset A^*$ and A^* is closed (Proposition 37), moreover for that operator the following conditions are equivalent in view of Proposition 37 and Corollary 38, as the reader immediately proves:

- (i) $(A^*)^* = A^*$ (A is essentially self adjoint),
- (ii) $\overline{A} = A^*$,
- (iii) $\overline{A} = (\overline{A})^*$.

(d) Unitary and selfadjoint operators are cases of normal operators. ■

The pair of elementary results in on (essentially) selfadjoint operators stated in (b) are worth to be proved.

Proposition 43. *Let $A : D(A) \rightarrow \mathcal{H}$ be a densely defined operator in the Hilbert space \mathcal{H} . The following facts are true.*

- (a) *If A is selfadjoint, then A does not admit proper symmetric extensions.*
- (b) *If A is essentially selfadjoint, then A admits a unique selfadjoint extension, and that this extension is A^* .*

Proof. (a) Let B be a symmetric extension of A . $A \subset B$ then $B^* \subset A^*$ for (b) in remark 32. As $A = A^*$ we have $B^* \subset A \subset B$. Since $B \subset B^*$, we conclude that $A = B$.

(b) Let B be a selfadjoint extension of the essentially selfadjoint operator A , so that $A \subset B$. Therefore $A^* \supset B^* = B$ and $(A^*)^* \subset B^* = B$. Since A is essentially selfadjoint, we have found $A^* \subset B$. Here A^* is selfadjoint and B is symmetric because selfadjoint. (b) implies $A^* = B$. That is, every selfadjoint extension of A coincides with A^* □

Another elementary though important result, helping understand why in QM observables are very often described by selfadjoint operators which are unbounded and defined in proper subspaces, is the following proposition (see (c) in remark 75).

Theorem 44 (Hellinger-Toeplitz theorem). *Let A be a self-adjoint operator in the complex Hilbert space \mathcal{H} . A is bounded if and only if $D(A) = \mathcal{H}$ (thus $A \in \mathfrak{B}(\mathcal{H})$).*

Proof. Assume $D(A) = \mathcal{H}$. As $A = A^*$ we have $D(A^*) = \mathcal{H}$. Since A^* is closed, Theorem 39 implies the $A^*(= A)$ is bounded. Conversely, if $A = A^*$ is bounded, since $D(A)$ is dense, we can continuously extend it to a bounded operator $A_1 : \mathcal{H} \rightarrow \mathcal{H}$. That extension, by continuity, trivially satisfies $\langle A_1 x | y \rangle = \langle x | A_1 y \rangle$ for all $x, y \in \mathcal{H}$ thus A_1 is symmetric. Since $A^* = A \subset A_1 \subset A_1^*$, (a) in Proposition 43 implies $A = A_1$. \square

Let us pass to focus on unitary operators. The relevance of unitary operators is evident from the following proposition where it is proved that they preserve the nature of operators with respect to the Hermitian conjugation.

Proposition 45. *Let $U : \mathcal{H} \rightarrow \mathcal{H}$ be a unitary operator in the complex Hilbert space \mathcal{H} and A another operator in \mathcal{H} . Prove that UAU^* with domain $U(D(A))$ (resp. U^*AU with domain $U^*(D(A))$) is symmetric, selfadjoint, essentially selfadjoint, unitary, normal if A is respectively symmetric, selfadjoint, essentially selfadjoint, unitary, normal.*

Proof. Since U^* is unitary when U is and $(U^*)^* = U$, it is enough to establish the thesis for UAU^* . First of all notice that $D(UAU^*) = U(D(A))$ is dense if $D(A)$ is dense since U is bijective and isometric and $U(D(A)) = \mathcal{H}$ if $D(A) = \mathcal{H}$ because U is bijective. By direct inspection, applying the definition of adjoint operator, one sees that $(UAU^*)^* = UA^*U^*$ and $D((UAU^*)^*) = U(D(A^*))$. Now, if A is symmetric $A \subset A^*$ which implies $UAU^* \subset UA^*U^* = (UAU^*)^*$ so that UAU^* is symmetric as well. If A is selfadjoint $A = A^*$ which implies $UAU^* = UA^*U^* = (UAU^*)^*$ so that UAU^* is self adjoint as well. If A is essentially self adjoint it is symmetric and $(A^*)^* = A^*$, so that UAU^* is symmetric and $U(A^*)^*U^* = UA^*U^*$ that is $(UA^*U^*)^* = UA^*U^*$ which means $((UAU^*)^*)^* = (UAU^*)^*$ so that UA^*U^* is essentially selfadjoint. If A is unitary, we have $A^*A = AA^* = I$ so that $UA^*AU^* = UAA^*U^* = UU^*$ which, since $U^*U = I = UU^*$, is equivalent to $UA^*U^*UAU^* = UAU^*UA^*U^* = U^*U = I$, that is $(UA^*U^*)UAU^* = (UAU^*)UA^*U^* = I$ and thus UAU^* is unitary as well. If A is normal UAU^* is normal too, with the same reasoning as in the unitary case. \square

Exercise 46.

(1) *Prove that if A is unitary then $A, A^* \in \mathfrak{B}(\mathcal{H})$.*

Solution. It holds $D(A) = D(A^*) = D(I) = \mathcal{H}$ and $\|Ax\|^2 = \langle Ax | Ax \rangle = \langle x | A^*Ax \rangle = \|x\|^2$ if $x \in \mathcal{H}$, so that $\|A\| = 1$. Due to (c) in remark 32, $A^* \in \mathfrak{B}(\mathcal{H})$.

□

(2) Prove that $A : \mathcal{H} \rightarrow \mathcal{H}$ is unitary if and only if it is surjective and norm preserving.

Solution. If A is unitary ((3) Def 41), it is evidently bijective, moreover as $D(A^*) = \mathcal{H}$, $\|Ax\|^2 = \langle Ax|Ax \rangle = \langle x|A^*Ax \rangle = \langle x|x \rangle = \|x\|^2$, so A is isometric. If $A : \mathcal{H} \rightarrow \mathcal{H}$ is isometric its norm is 1 and thus $A \in \mathfrak{B}(\mathcal{H})$. Therefore $A^* \in \mathfrak{B}(\mathcal{H})$. The condition $\|Ax\|^2 = \|x\|^2$ can be re-written $\langle Ax|Ax \rangle = \langle x|A^*Ax \rangle = \langle x|x \rangle$ and thus $\langle x|(A^*A - I)x \rangle = 0$ for $x \in \mathcal{H}$. Using $x = y \pm z$ and $x = y \pm iz$, the found identity implies $\langle z|(A^*A - I)y \rangle = 0$ for all $y, z \in \mathcal{H}$. Taking $z = (A^*A - I)y$, we finally have $\|(A^*A - I)y\| = 0$ for all $y \in \mathcal{H}$ and thus $A^*A = I$. In particular A is injective as it admits the left inverse A^* . Since A is also surjective, it is bijective and thus its left inverse (A^*) is also a right inverse, that is $AA^* = I$.

(3) Prove that, if $A : \mathcal{H} \rightarrow \mathcal{H}$ satisfies $\langle x|Ax \rangle \in \mathbb{R}$ for all $x \in \mathcal{H}$ (and in particular if $A \geq 0$, which means $\langle x|Ax \rangle \geq 0$ for all $x \in \mathcal{H}$), then $A^* = A$ and $A \in \mathfrak{B}(\mathcal{H})$.

Solution. We have $\langle x|Ax \rangle = \overline{\langle x|Ax \rangle} = \langle Ax|x \rangle = \langle x|A^*x \rangle$ where, as $D(A) = \mathcal{H}$, the adjoint A^* is well defined everywhere on \mathcal{H} . Thus $\langle x|(A - A^*)x \rangle = 0$ for every $x \in \mathcal{H}$. Using there $x = y \pm z$ and $x = y \pm iz$ we obtain $\langle y|(A - A^*)z \rangle = 0$ for all $y, z \in \mathcal{H}$. Choosing $y = (A - A^*)z$, we conclude that $A = A^*$. Theorem 44 concludes the proof. □

Example 47. The **Fourier transform**, $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$, defined as^f

$$(\mathcal{F}f)(k) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ik \cdot x} f(x) d^n x \quad (33)$$

($k \cdot x$ being the standard \mathbb{R}^n scalar product of k and x) is a bijective linear map with inverse $\mathcal{F}_- : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$,

$$(\mathcal{F}_-g)(x) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ik \cdot x} g(k) d^n k. \quad (34)$$

Both \mathcal{F} and \mathcal{F}_- preserve the scalar product

$$\langle \mathcal{F}f | \mathcal{F}g \rangle = \langle f | g \rangle, \quad \langle \mathcal{F}_-f | \mathcal{F}_-g \rangle = \langle f | g \rangle \quad \forall f, g \in \mathcal{S}(\mathbb{R}^n)$$

and thus they also preserve the norm of $L^2(\mathbb{R}^n, d^n x)$, in particular, $\|\mathcal{F}\| = \|\mathcal{F}_-\| = 1$. As a consequence of Proposition 26, using the fact that $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n, d^n x)$, one easily proves that \mathcal{F} and \mathcal{F}_- uniquely continuously extend to unitary operators, respectively, $\hat{\mathcal{F}} : L^2(\mathbb{R}^n, d^n x) \rightarrow L^2(\mathbb{R}^n, d^n k)$ and $\hat{\mathcal{F}}_- : L^2(\mathbb{R}^n, d^n k) \rightarrow L^2(\mathbb{R}^n, d^n x)$ such that $\hat{\mathcal{F}}^* = \hat{\mathcal{F}}^{-1} = \hat{\mathcal{F}}_-$. The map $\hat{\mathcal{F}}$ is the **Fourier-Plancherel (unitary) operator**. ■

^fIn QM, adopting units with $\hbar \neq 1$, $k \cdot x$ has to be replaced for $\frac{k \cdot x}{\hbar}$ and $(2\pi)^{n/2}$ for $(2\pi\hbar)^{n/2}$.

2.3. Criteria for (essential) selfadjointness

Let us briefly introduce, without proofs (see [6]), some commonly used mathematical technology to study (essential) selfadjointness of symmetric operators. If A is a densely defined symmetric operator in the complex Hilbert space \mathcal{H} , define the **deficiency indices**, $n_{\pm} := \dim \mathcal{H}_{\pm}$ (cardinal numbers in general) where \mathcal{H}_{\pm} are the (closed) subspaces of the solutions of $(A^* \pm iI)x_{\pm} = 0$ [8,6,9].

Proposition 48. *If A is a densely defined symmetric operator in the complex Hilbert space \mathcal{H} the following holds.*

(a) *A is essentially selfadjoint (thus it admits an unique selfadjoint extension) if $n_{\pm} = 0$, that is $\mathcal{H}_{\pm} = \{0\}$.*

(b) *A admits selfadjoint extensions if and only if $n_+ = n_-$ and these extension are labelled by means of n_+ parameters.*

Remark 49. If A is symmetric, an easy sufficient condition, due to von Neumann, for $n_+ = n_-$ is that $CA \subset AC$ where $C : \mathcal{H} \rightarrow \mathcal{H}$ is a **conjugation** that is an isometric surjective *antilinear*[§] map with $CC = I$.

Indeed, using the definition of A^* and $D(A^*)$ and observing that (from the polarization identity (25)) $\langle Cy|Cx \rangle = \overline{\langle y|x \rangle}$, the condition

$$AC \subset CA \quad \text{implies the condition} \quad CA^* \subset A^*C.$$

Therefore, remembering that $CC = I$, we have that $A^*x = \pm ix$ if and only if $A^*Cx = C(\pm ix) = \mp iCx$. Since C preserves orthogonality and norm of vectors, it transform a Hilbertian basis of \mathcal{H}_+ into a Hilbertian basis of \mathcal{H}_- and *vice versa*. We conclude that $n_+ = n_-$.

Taking C as the standard conjugation of functions in $L^2(\mathbb{R}^n, d^n x)$, this result proves in particular that all operators in QM of the *Schödinger* form as (24) admit self-adjoint extensions when defined on dense domains. ■

Exercise 50. *Prove that a symmetric operator that admits a unique self-adjoint extension is necessarily essentially selfadjoint.*

Solution. By (b) of Proposition 48, $n_+ = n_-$ if the operator admits self-adjoint extensions. Furthermore, if $n_{\pm} \neq 0$ there are many selfadjoint extension again for (b). The only possibility for the uniqueness of the selfadjoint extension is $n_{\pm} = 0$. (a) of Proposition 48 implies that A is essentially selfadjoint. □

A very useful criterion to establish the essentially selfadjointness of a symmetric operator is due to Nelson. It relies upon an important definition.

Definition 51. Let A be an operator in the complex Hilbert space \mathcal{H} .

[§]In other words $C(\alpha x + \beta y) = \bar{\alpha}Cx + \bar{\beta}Cy$ if $\alpha, \beta \in \mathbb{C}$ and $x, y \in \mathcal{H}$.

If $\psi \in \bigcap_{n \in \mathbb{N}} D(A^n)$ satisfies

$$\sum_{n=0}^{+\infty} \frac{t^n}{n!} \|A^n \psi\| < +\infty \quad \text{for some } t > 0,$$

then ψ is said to be an **analytic vector** of A . ■

We can state Nelson's criterion here [6].

Theorem 52 (Nelson's essentially selfadjointness criterium). *Let A be a symmetric operator in the complex Hilbert space \mathcal{H} , A is essentially selfadjoint if $D(A)$ contains a dense set D of analytic vectors (or – which is equivalent – a set D of analytic vectors whose finite span is dense in \mathcal{H}).*

The above equivalence is due to the fact that a finite linear combination of analytic vector is an analytic vector as well, the proof being elementary. We have the following evident corollary due to the fact that the (finite) span of a Hilbert basis is dense and that, if $A\psi = a\psi$, then

$$\sum_{n=0}^{+\infty} \frac{t^n}{n!} \|A^n \psi\| = \sum_{n=0}^{+\infty} \frac{a^n t^n}{n!} \|\psi\| = e^{at} \|\psi\| < +\infty \quad \text{for some } t \in \mathbb{R}$$

Corollary 53. *If A is a symmetric operator admitting a Hilbertian basis of eigenvectors in $D(A)$, then A is essentially selfadjoint.*

Example 54.

(1) For $m \in \{1, 2, \dots, n\}$, consider the operators X'_m and X''_m in $L^2(\mathbb{R}^n, d^n x)$ with dense domains $D(X'_m) = C_0^\infty(\mathbb{R}^n; \mathbb{C})$, $D(X''_m) = \mathcal{S}(\mathbb{R}^n)$ for $x \in \mathbb{R}^n$ and, for ψ, ϕ in the respective domains,

$$(X'_m \psi)(x) := x_m \psi(x), \quad (X''_m \phi)(x) := x_m \phi(x),$$

where x_m is the m -th component of $x \in \mathbb{R}^n$. Both operators are symmetric but not selfadjoint. They admit selfadjoint extensions because they commute with the standard complex conjugation of functions (see remark 49). It is furthermore possible to prove that both operators are essentially selfadjoint as follows. First define the **k -axis position operator** X_m in $L^2(\mathbb{R}^n, d^n x)$ with domain

$$D(X_m) := \left\{ \psi \in L^2(\mathbb{R}^n, d^n x) \mid \int_{\mathbb{R}^n} |x_m \psi(x)|^2 d^n x < +\infty \right\}$$

and

$$(X_m \psi)(x) := x_m \psi(x), \quad x \in \mathbb{R}^n. \tag{35}$$

Just by applying the definition of adjoint one sees that $X_m^* = X_m$ so that X_m is selfadjoint [6]. Again applying the definition of adjoint, one sees (see below) that $X_m'^* = X_m''^* = X_m$ where we know that the last one is selfadjoint. By definition, X_m' and X_m'' are therefore essentially selfadjoint. By (b) in proposition 43 X_m' and X_m'' admit a unique selfadjoint extension which must coincide with X_m itself. We

conclude that $C_0^\infty(\mathbb{R}^n; \mathbb{C})$ and $\mathcal{S}(\mathbb{R}^n)$ are *cores* (Def. 35) for the m -axis position operator.

Let us prove that $X_m'^* = X_m$, the proof for $X_m''^*$ is identical. By direct inspection one easily sees that $X_m'^* \subset X_m$. Let us prove the converse inclusion. We have that $\phi \in D(X_m'^*)$ if and only if there exists $\eta_\phi \in L^2(\mathbb{R}^n, d^n x)$ such that $\int \overline{\phi(x)} x_m \psi(x) dx = \int \overline{\eta_\phi(x)} \psi(x) dx$, that is $\int (\overline{\phi(x)} x_m - \overline{\eta_\phi(x)}) \psi(x) dx = 0$, for every $\psi \in C_0^\infty(\mathbb{R}^n; \mathbb{C})$. Fix a compact $K \subset \mathbb{R}^n$ of the form $[a, b]^n$, obviously the function $K \ni x \mapsto \phi(x) x_m - \eta_\phi(x)$ is $L^2(K, dx)$ (we could not conclude the same result if K were replaced by \mathbb{R}^n). Since we can $L^2(K)$ -approximate that function with a sequence of $\psi_n \in C_0^\infty(\mathbb{R}^n; \mathbb{C})$ such that $\text{supp}(\psi_n) \subset K$, we conclude that $\int_K |\phi(x) x_m - \eta_\phi(x)|^2 dx = 0$, so that $K \ni x \mapsto \phi(x) x_m - \eta_\phi(x)$ is zero a.e.. Since $K = [a, b]^n$ was arbitrary, we conclude that $\mathbb{R}^n \ni x \mapsto \phi(x) x_m = \eta_\phi(x)$ a.e.. In particular, both ϕ and $\mathbb{R}^n \ni x \mapsto x_m \phi(x)$ are $L^2(\mathbb{R}^n, dx)$ (the latter because it is a.e. identical to $\eta_\phi \in L^2(\mathbb{R}^n, dx)$), namely, $D(X_m'^*) \ni \phi$ implies $\phi \in D(X_m)$. This proves that $D(X_m'^*) \subset D(X_m)$ and consequently $X_m'^* \subset X_m$ as wanted.

(2) For $m \in \{1, 2, \dots, n\}$, the k -axis **momentum operator**, P_m , is obtained from the position operator using the Fourier-Plancherel unitary operator $\hat{\mathcal{F}}$ introduced in example 47.

$$D(P_m) := \left\{ \psi \in L^2(\mathbb{R}^n, d^n x) \mid \int_{\mathbb{R}^n} |k_m (\hat{\mathcal{F}}\psi)(k)|^2 d^n k \right\}$$

and

$$(P_m \psi)(x) := (\hat{\mathcal{F}}^{-1} K_m \hat{\mathcal{F}} \psi)(x), \quad x \in \mathbb{R}^n. \quad (36)$$

Above K_m is the m -axis *position operator* just written for functions (in $L^2(\mathbb{R}^n, d^n k)$) whose variable, for pure convenience, is denoted by k instead of x . Since K_m is self-adjoint, P_m is selfadjoint as well, as established in Proposition 45 as a consequence of the fact that $\hat{\mathcal{F}}$ is unitary.

It is possible to give a more explicit form to P_m if restricting its domain. Taking $\psi \in C_0^\infty(\mathbb{R}^n; \mathbb{C}) \subset \mathcal{S}(\mathbb{R}^n)$ or directly $\psi \in \mathcal{S}(\mathbb{R}^n)$, $\hat{\mathcal{F}}$ reduces to the standard integral Fourier transform (33) with inverse (34). Using these integral expressions we easily obtain

$$(P_m \psi)(x) = (\hat{\mathcal{F}}^{-1} K_m \hat{\mathcal{F}} \psi)(x) = -i \frac{\partial}{\partial x_m} \psi(x) \quad (37)$$

because in $\mathcal{S}(\mathbb{R}^n)$, which is invariant under the Fourier (and inverse Fourier) integral transformation,

$$\int_{\mathbb{R}^n} e^{ik \cdot x} k_m (\mathcal{F}\psi)(k) d^n k = -i \frac{\partial}{\partial x_m} \int_{\mathbb{R}^n} e^{-ik \cdot x} (\mathcal{F}\psi)(k) d^n k.$$

This way leads us to consider the operators P_m' and P_m'' in $L^2(\mathbb{R}^n, d^n x)$ with

$$D(P_m') = C_0^\infty(\mathbb{R}^n; \mathbb{C}), \quad D(P_m'') = \mathcal{S}(\mathbb{R}^n)$$

and, for $x \in \mathbb{R}^n$ and ψ, ϕ in the respective domains,

$$(P'_m \psi)(x) := -i \frac{\partial}{\partial x_m} \psi(x), \quad (P''_m \phi)(x) := -i \frac{\partial}{\partial x_m} \phi(x).$$

Both operators are symmetric as one can easily prove by integrating by parts, but not selfadjoint. They admit selfadjoint extensions because they commute with the conjugation $(C\psi)(x) = \overline{\psi(-x)}$ (see remark 49). It is furthermore possible to prove that both operators are essentially self-adjoint by direct use of Proposition 48 [6]. However we already know that P''_m is essentially selfadjoint as it coincides with the essentially selfadjoint operator $\hat{\mathcal{F}}^* K''_m \hat{\mathcal{F}}$ because $\mathcal{S}(\mathbb{R}^n)$ is invariant under $\hat{\mathcal{F}}$.

The unique selfadjoint extension of both operators turns out to be P_m . We conclude that $C_0^\infty(\mathbb{R}^n; \mathbb{C})$ and $\mathcal{S}(\mathbb{R}^n)$ are *cores* for the m -axis momentum operator.

With the given definitions of selfadjoint operators X_k and P_k , $\mathcal{S}(\mathbb{R}^n)$ turns out to be an invariant domain and thereon the CCR (21) hold rigorously.

As a final remark to conclude, we say that, if $n = 1$, $D(P)$ coincides to the already introduced domain (17). In that domain P is nothing but the weak derivative times the factor $-i$.

(3) The most elementary example of application of Nelson's criterion is in $L^2([0, 1], dx)$. Consider $A = -\frac{d^2}{dx^2}$ with dense domain $D(A)$ given by the functions in $C^\infty([0, 1]; \mathbb{C})$ such that $\psi(0) = \psi(1)$ and $\frac{d\psi}{dx}(0) = \frac{d\psi}{dx}(1)$. A is symmetric thereon as it arises immediately by integration by parts, in particular its domain is dense since it includes the Hilbert basis of exponentials $e^{i2\pi n x}$, $n \in \mathbb{Z}$, which are eigenvectors of A . Thus A is also essentially selfadjoint on the above domain.

(4) A more interesting case is the **Hamiltonian operator of the harmonic oscillator**, H [1] obtained as follows. One starts by

$$H_0 = -\frac{1}{2m} \frac{d^2}{dx^2} + \frac{m\omega^2}{2} x^2$$

with $D(H_0) := \mathcal{S}(\mathbb{R})$. Above, x^2 is the multiplicative operator and $m, \omega > 0$ are constants.

To go on, we start by defining a triple of operators, $A, A^\dagger, \mathcal{N} : \mathcal{S}(\mathbb{R}) \rightarrow L^2(\mathbb{R}, dx)$ as

$$A^\dagger := \sqrt{\frac{m\omega}{2\hbar}} \left(x - \frac{\hbar}{m\omega} \frac{d}{dx} \right), \quad A := \sqrt{\frac{m\omega}{2\hbar}} \left(x + \frac{\hbar}{m\omega} \frac{d}{dx} \right), \quad \mathcal{N} := A^\dagger A.$$

These operators have the same domain which is also invariant: $A(\mathcal{S}(\mathbb{R})) \subset \mathcal{S}(\mathbb{R})$, $A^\dagger(\mathcal{S}(\mathbb{R})) \subset \mathcal{S}(\mathbb{R})$, $\mathcal{N}(\mathcal{S}(\mathbb{R})) \subset \mathcal{S}(\mathbb{R})$. It is also easy to see, using integration by parts, that $A^\dagger \subset A^*$, $A \subset (A^\dagger)^*$, and that \mathcal{N} is Hermitian and also symmetric because $\mathcal{S}(\mathbb{R})$ is dense in $L^2(\mathbb{R}, dx)$. By direct computation, exploiting the given definitions one immediately sees that

$$H_0 = \hbar \left(A^\dagger A + \frac{1}{2} I \right) = \hbar \left(\mathcal{N} + \frac{1}{2} I \right).$$

Finally we have the commutation relations *where both sides are evaluated on $\mathcal{S}(\mathbb{R})$* ,

$$[A, A^\dagger] = I. \tag{38}$$

Supposing that there exists $\psi_0 \in \mathcal{S}(\mathbb{R})$ such that

$$\|\psi_0\| = 1, \quad A\psi_0 = 0. \quad (39)$$

Starting from (38) and using an inductive procedure on the vectors

$$\psi_n := \frac{1}{\sqrt{n!}} (A^\dagger)^n \psi_0 \quad (40)$$

it is quite easy to prove that (e.g., see [6] for details), for $n, m = 0, 1, 2, \dots$, the relations hold

$$A\psi_n = \sqrt{n}\psi_{n-1}, \quad A^\dagger\psi_n = \sqrt{n+1}\psi_{n+1}, \quad \langle \psi_n | \psi_m \rangle = \delta_{nm}. \quad (41)$$

Finally, the ψ_n are eigenvectors of H_0 (and \mathcal{N}) since

$$\begin{aligned} H_0\psi_n &= \hbar\omega \left(A^\dagger A\psi_n + \frac{1}{2}\psi_n \right) = \hbar\omega \left(A^\dagger \sqrt{n}\psi_{n-1} + \frac{1}{2}\psi_n \right) = \hbar\omega \left(\sqrt{n^2}\psi_n + \frac{1}{2}\psi_n \right) \\ &= \hbar\omega \left(n + \frac{1}{2} \right) \psi_n. \end{aligned} \quad (42)$$

As a consequence $\{\psi_n\}_{n \in \mathbb{N}}$ is an orthonormal set of vectors. This set is actually a Hilbertian basis because (1), a solution in $\mathcal{S}(\mathbb{R})$ of (39) exist (and is unique):

$$\psi_0(x) = \frac{1}{\pi^{1/4} \sqrt{s}} e^{-\frac{x^2}{2s^2}}, \quad s := \sqrt{\frac{\hbar}{m\omega}},$$

(2) the first identity in (41) is a well-known recurrence relation of the Hilbert basis of $L^2(\mathbb{R}, dx)$ made of Schwartz' functions known as *Hermite functions* $\{H_n\}_{n \in \mathbb{N}}$, and $\psi_0(x) = H_0(x)$.

Exploiting Nelson's criterium, we conclude that the symmetric operator H_0 is essentially selfadjoint in $D(H_0) = \mathcal{S}(\mathbb{R})$ and $H := \overline{H_0} = H_0^*$, because H_0 admits a Hilbert basis of eigenvectors with corresponding eigenvalues $\hbar\omega(n + \frac{1}{2})$. ■

2.4. Spectrum of an operator

Our goal is to extend (4) to a formula valid in the infinite dimensional case. As we shall see shortly, passing to the infinite dimensional case, the sum is replaced by an integral and $\sigma(A)$ must be enlarged with respect to the pure set of eigenvalues of A . This is because, as already noticed in the first section, there are operators which should be decomposed with the prescription (4) but they do not have eigenvalues, though they play a crucial rôle in QM.

Notation 55. If $A : D(A) \rightarrow \mathcal{H}$ is injective, A^{-1} indicates its inverse when the co-domain of A is restricted to $Ran(A)$. In other words, $A^{-1} : Ran(A) \rightarrow D(A)$. ■

The definition of *spectrum* of the operator $A : D(A) \rightarrow \mathcal{H}$ extends the notion of set of eigenvalues. The eigenvalues of A are the numbers $\lambda \in \mathbb{C}$ such that $(A - \lambda I)^{-1}$ does not exist. When passing to infinite dimensions, topological issues take place.

As a matter of fact, even if $(A - \lambda I)^{-1}$ exists, it may be bounded or unbounded and its domain $\text{Ran}(A - \lambda I)$ may or may not be dense. These features permit us to define a suitable extension of the notion of set of eigenvalues.

Definition 56. Let A be an operator in the complex Hilbert space \mathcal{H} . The **resolvent set** of A is the subset of \mathbb{C} ,

$$\rho(A) := \{\lambda \in \mathbb{C} \mid (A - \lambda I) \text{ is injective, } \overline{\text{Ran}(A - \lambda I)} = \mathcal{H}, (A - \lambda I)^{-1} \text{ is bounded}\}$$

The **spectrum** of A is the complement $\sigma(A) := \mathbb{C} \setminus \rho(A)$ and it is given by the union of the following pairwise disjoint three parts:

- (i) the **point-spectrum**, $\sigma_p(A)$, where $A - \lambda I$ not injective ($\sigma_p(A)$ is the set of *eigenvalues* of A),
- (ii) the **continuous spectrum**, $\sigma_c(A)$, where $A - \lambda I$ injective, $\overline{\text{Ran}(A - \lambda I)} = \mathcal{H}$ and $(A - \lambda I)^{-1}$ not bounded,
- (iii) the **residual spectrum**, $\sigma_r(A)$, where $A - \lambda I$ injective and $\overline{\text{Ran}(A - \lambda I)} \neq \mathcal{H}$. ■

Remark 57.

- (a) It turns out that $\rho(A)$ is always *open*, so that $\sigma(A)$ is always *closed* [8,6,9].
- (b) If A is closed and normal, in particular, if A is either selfadjoint or unitary), $\sigma_r(A) = \emptyset$ (e.g., see [6]). Furthermore, if A is closed (if $A \in \mathfrak{B}(\mathcal{H})$ in particular), $\lambda \in \rho(A)$ if and only if $A - \lambda I$ admits inverse in $\mathfrak{B}(\mathcal{H})$ (see (2) in exercise 58).
- (c) If A is selfadjoint, one finds $\sigma(A) \subset \mathbb{R}$ (see (1) in exercise 58).
- (d) If A is unitary one finds $\sigma(A) \subset \mathbb{T} := \{e^{ia} \mid a \in \mathbb{R}\}$ (e.g., see [6]).
- (e) If $U : \mathcal{H} \rightarrow \mathcal{H}$ is unitary and A is any operator in the complex Hilbert space \mathcal{H} , just by applying the definition one finds $\sigma(UAU^*) = \sigma(A)$ and in particular,

$$\sigma_p(UAU^*) = \sigma_p(A), \quad \sigma_c(UAU^*) = \sigma_c(A), \quad \sigma_r(UAU^*) = \sigma_r(A). \quad (43)$$

The same result holds replacing $U : \mathcal{H} \rightarrow \mathcal{H}$ for $U : \mathcal{H} \rightarrow \mathcal{H}'$ and U^* for U^{-1} , where U is now a Hilbert space isomorphism (an isometric surjective linear map) and \mathcal{H}' another complex Hilbert space. ■

To conclude let us mention two useful technical facts which will turn out to be useful several times in the rest of this paper. From the definition of adjoint, one easily has for $A : D(A) \rightarrow \mathcal{H}$ densely defined and $\lambda \in \mathbb{C}$,

$$\text{Ker}(A^* - \bar{\lambda}I) = [\text{Ran}(A - \lambda I)]^\perp, \quad \text{Ker}(A - \lambda I) \subset [\text{Ran}(A^* - \bar{\lambda}I)]^\perp \quad (44)$$

where the inclusion becomes an identity if $A \in \mathfrak{B}(\mathcal{H})$.

Exercise 58.

- (1) Prove that if A is a selfadjoint operator in the complex Hilbert space \mathcal{H} then
 - (i) $\sigma(A) \subset \mathbb{R}$,
 - (ii) $\sigma_r(A) = \emptyset$,
 - (iii) *eigenvectors with different eigenvalues are orthogonal.*

Solution. Let us begin with (i). Suppose $\lambda = \mu + i\nu$, $\nu \neq 0$ and let us prove $\lambda \in \rho(A)$. If $x \in D(A)$,

$$\langle (A - \lambda I)x | (A - \lambda I)x \rangle = \langle (A - \mu I)x | (A - \mu I)x \rangle + \nu^2 \langle x | x \rangle + i\nu [\langle Ax | x \rangle - \langle x | Ax \rangle].$$

The last summand vanishes for A is selfadjoint. Hence

$$\|(A - \lambda I)x\| \geq |\nu| \|x\|.$$

With a similar argument we obtain

$$\|(A - \bar{\lambda}I)x\| \geq |\nu| \|x\|.$$

The operators $A - \lambda I$ and $A - \bar{\lambda}I$ are injective, and $\|(A - \lambda I)^{-1}\| \leq |\nu|^{-1}$, where $(A - \lambda I)^{-1} : \text{Ran}(A - \lambda I) \rightarrow D(A)$. Notice that, from (44),

$$\overline{\text{Ran}(A - \lambda I)}^\perp = [\text{Ran}(A - \lambda I)]^\perp = \text{Ker}(A^* - \bar{\lambda}I) = \text{Ker}(A - \bar{\lambda}I) = \{0\},$$

where the last equality makes use of the injectivity of $A - \bar{\lambda}I$. Summarising: $A - \lambda I$ is injective, $(A - \lambda I)^{-1}$ bounded and $\overline{\text{Ran}(A - \lambda I)}^\perp = \{0\}$, i.e. $\text{Ran}(A - \lambda I)$ is dense in \mathcal{H} ; therefore $\lambda \in \rho(A)$, by definition of resolvent set. Let us pass to (ii). Suppose $\lambda \in \sigma(A)$, but $\lambda \notin \sigma_p(A)$. Then $A - \lambda I$ must be one-to-one and $\text{Ker}(A - \lambda I) = \{0\}$. Since $A = A^*$ and $\lambda \in \mathbb{R}$ by (i), we have $\overline{\text{Ker}(A^* - \bar{\lambda}I)} = \{0\}$, so $[\text{Ran}(A - \lambda I)]^\perp = \text{Ker}(A^* - \bar{\lambda}I) = \{0\}$ and $\overline{\text{Ran}(A - \lambda I)} = \mathcal{H}$. Consequently $\lambda \in \sigma_c(A)$. Proving (iii) is easy: if $\lambda \neq \mu$ and $Au = \lambda u$, $Av = \mu v$, then

$$(\lambda - \mu)\langle u | v \rangle = \langle Au | v \rangle - \langle u | Av \rangle = \langle u | Av \rangle - \langle u | Av \rangle = 0;$$

from $\lambda, \mu \in \mathbb{R}$ and $A = A^*$. But $\lambda - \mu \neq 0$, so $\langle u | v \rangle = 0$. \square

(2) Let $A : D(A) \rightarrow \mathcal{H}$ be a closed operator in \mathcal{H} (in particular $A \in \mathfrak{B}(\mathcal{H})$). Prove that $\lambda \in \rho(A)$ if and only if $A - \lambda I$ admits an inverse which belongs to $\mathfrak{B}(\mathcal{H})$.

Solution. If $(A - \lambda I)^{-1} \in \mathfrak{B}(\mathcal{H})$, it must be $\overline{\text{Ran}(A - \lambda I)} = \text{Ran}(A - \lambda I) = \mathcal{H}$ and $(A - \lambda I)^{-1}$ is bounded, so that $\lambda \in \rho(A)$ by definition. Let us prove the converse. Suppose that $\lambda \in \rho(A)$. We know that $(A - \lambda I)^{-1}$ is defined on the dense domain $\text{Ran}(A - \lambda I)$ and is bounded. To conclude, it is therefore enough proving that $y \in \mathcal{H}$ implies $y \in \text{Ran}(A - \lambda I)$. To this end, notice that if $y \in \mathcal{H} = \overline{\text{Ran}(A - \lambda I)}$, then $y = \lim_{n \rightarrow +\infty} (A - \lambda I)x_n$ for some $x_n \in D(A - \lambda I)$. The sequence of x_n converges. Indeed \mathcal{H} is complete and $\{x_n\}_{n \in \mathbb{N}}$ is Cauchy as (1) $x_n = (A - \lambda I)^{-1}y_n$, (2) $\|x_n - x_m\| \leq \|(A - \lambda I)^{-1}\| \|y_n - y_m\|$, and (3) $y_n \rightarrow y$. To end the proof, we observe that, $A - \lambda I$ is closed since A is such ((b) in remark 36). It must consequently be ((c) in remark 36) $x = \lim_{n \rightarrow +\infty} x_n \in D(A - \lambda I)$ and $y = (A - \lambda I)x \in \text{Ran}(A - \lambda I)$. \square

Example 59. The m -axis position operator X_m in $L^2(\mathbb{R}^n, d^n x)$ introduced in (1) of example 54 satisfies

$$\sigma(X_m) = \sigma_c(X_m) = \mathbb{R}. \quad (45)$$

The proof can be obtained as follows. First observe that $\sigma(X_m) \subset \mathbb{R}$ since the operator is selfadjoint. However $\sigma_p(X_m) = \emptyset$ as observed in the first section and $\sigma_r(X_m) = \emptyset$ because X_m is self-adjoint ((1) in exercise 58). Suppose that, for some $r \in \mathbb{R}$, $(X_m - rI)^{-1}$ is bounded. If $\psi \in D(X_m - rI) = D(X_m)$ with $\|\psi\| = 1$ we have $\|\psi\| = \|(X_m - rI)^{-1}(X_m - rI)\psi\|$ and thus $\|\psi\| \leq \|(X_m - rI)^{-1}\| \|(X_m - rI)\psi\|$. Therefore

$$\|(X_m - rI)^{-1}\| \geq \frac{1}{\|(X_m - rI)\psi\|}$$

For every fixed $\epsilon > 0$, it is simply constructed $\psi \in D(X_m)$ with $\|\psi\| = 1$ and $\|(X_m - rI)\psi\| < \epsilon$. Therefore $(X_m - rI)^{-1}$ cannot be bounded and thus $r \in \sigma_c(X_m)$. In view of (e) in remark 57, we also conclude that

$$\sigma(P_m) = \sigma_c(P_m) = \mathbb{R}, \quad (46)$$

just because the momentum operator P_m is related to the position one by means of a unitary operator given by the Fourier-Plancherel operator $\hat{\mathcal{F}}$ as discussed in (2) of example 54. ■

2.5. Spectral measures

Definition 60. Let \mathcal{H} be a complex Hilbert space. $P \in \mathfrak{B}(\mathcal{H})$ is called **orthogonal projector** when $PP = P$ and $P^* = P$. $\mathcal{L}(\mathcal{H})$ denotes the set of orthogonal projectors of \mathcal{H} . ■

We have the well known relation between orthogonal projectors and closed subspaces [8,6]

Proposition 61. *If $P \in \mathcal{L}(\mathcal{H})$, then $P(\mathcal{H})$ is a closed subspace. If $\mathcal{H}_0 \subset \mathcal{H}$ is a closed subspace, there exists exactly one $P \in \mathcal{L}(\mathcal{H})$ such that $P(\mathcal{H}) = \mathcal{H}_0$. Finally, $I - P \in \mathcal{L}(\mathcal{H})$ and it projects onto \mathcal{H}_0^\perp (e.g., see [6]).*

We can now state one of the most important definitions in spectral theory.

Definition 62. Let \mathcal{H} be a complex Hilbert space and $\Sigma(X)$ a σ -algebra over X . A **projector-valued measure (PVM)** on X , P , is a map $\Sigma(X) \ni E \mapsto P_E \in \mathcal{L}(\mathcal{H})$ such that

- (i) $P_X = I$,
- (ii) $P_E P_F = P_{E \cap F}$,
- (iii) If $N \subset \mathbb{N}$ and $\{E_k\}_{k \in N} \subset \Sigma(X)$ satisfies $E_j \cap E_k = \emptyset$ for $k \neq j$, then

$$\sum_{j \in N} P_{E_j} x = P_{\cup_{j \in N} E_j} x \quad \text{for every } x \in \mathcal{H}.$$

(If N is infinite, the sum on the left hand side of (iii) is computed referring to the topology of \mathcal{H}) ■

Remark 63.

(a) (i) and (iii) with $N = \{1, 2\}$ imply that $P_\emptyset = 0$ using $E_1 = X$ and $E_2 = \emptyset$. Next (ii) entails that $P_E P_F = 0$ if $E \cap F = \emptyset$. An important consequence is that for N infinite, the vector given by the sum on the left hand side of (iii) is independent from the chosen order because that vector is a sum of pairwise orthogonal vectors $P_{E_j} x$.

(b) If $x, y \in \mathcal{H}$, $\Sigma(X) \ni E \mapsto \langle x | P_E y \rangle =: \mu_{xy}^{(P)}(E)$ is a *complex measure* whose (finite) *total variation* [8] will be denoted by $|\mu_{xy}^{(P)}|$. From the definition of μ_{xy} , we immediately have:

- (i) $\mu_{xy}^{(P)}(X) = \langle x | y \rangle$,
- (ii) $\mu_{xx}^{(P)}$ is always positive and finite and $\mu_{xx}^{(P)}(X) = \|x\|^2$;
- (iii) if $s = \sum_{k=1}^n s_k \chi_{E_k}$ is a *simple function* [8], $\int_X s d\mu_{xy} = \langle x | \sum_{k=1}^n s_k P_{E_k} y \rangle$.

■

Example 64.

(1) The simplest example of PVM is related to a countable Hilbertian basis N in a separable Hilbert space \mathcal{H} . We can define $\Sigma(N)$ as the class of all subsets of N itself. Next, for $E \in \Sigma(N)$ and $z \in \mathcal{H}$ we define

$$P_E z := \sum_{x \in E} \langle x | z \rangle x$$

and $P_\emptyset := 0$. It is easy to prove that the class of all P_E defined this way form a PVM on N . (This definition can be also given if \mathcal{H} is non-separable and N is uncountable, since for every $y \in \mathcal{H}$ only an at most countable subset of elements $x \in E$ satisfy $\langle x | y \rangle \neq 0$). In particular $\mu_{xy}(E) = \langle x | P_E y \rangle = \sum_{z \in E} \langle x | z \rangle \langle z | y \rangle$ and $\mu_{xx}(E) = \sum_{z \in E} |\langle x | z \rangle|^2$.

(2) A more complicated version of (1) consists of a PVM constructed out of a orthogonal Hilbertian decomposition of a separable Hilbert space, $\mathcal{H} = \bigoplus_{n \in \mathbb{N}} \mathcal{H}_n$, where $\mathcal{H}_n \subset \mathcal{H}$ is a closed subspace and $\mathcal{H}_n \perp \mathcal{H}_m$ if $n \neq m$. Again defining $\Sigma(\mathbb{N})$ as the set of subsets of \mathbb{N} , for $E \in \Sigma(N)$ and $z \in \mathcal{H}$ we define

$$P_E z := \sum_{x \in E} Q_n z$$

where Q_n is the orthogonal projector onto \mathcal{H}_n (the reader can easily check that the sum always converges using Bessel's inequality). It is easy to prove that the class of P_E s defined this way form a PVM on \mathbb{N} . In particular $\mu_{xy}(E) = \langle x | P_E y \rangle = \sum_{n \in E} \langle x | Q_n y \rangle$ and $\mu_{xx}(E) = \sum_{n \in E} \|Q_n x\|^2$.

(3) In $L^2(\mathbb{R}, dx)$ a simple PVM, not related with a Hilbertian basis, is made as follows. To every $E \in \mathcal{B}(\mathbb{R})$, the Borel σ -algebra, associate the orthonormal projector P_E such that, if χ_E is the **characteristic function of E** - $\chi_E(x) = 0$ if $x \notin E$ and $\chi_E(x) = 1$ if $x \in E$ -

$$(P_E \psi)(x) := \chi_E(x) \psi(x) \quad \forall \psi \in L^2(\mathbb{R}, dx) .$$

Moreover $P_\emptyset := 0$. It is easy to prove that the collection of the P_E is a PVM. In particular $\mu_{fg}(E) = \langle f | P_E g \rangle = \int_E \overline{f(x)} g(x) dx$ and $\mu_{ff}(E) = \int_E |f(x)|^2 dx$. ■

We have the following fundamental result [8,6,9].

Proposition 65. *Let \mathcal{H} be a complex Hilbert space and $P : \Sigma(X) \rightarrow \mathcal{L}(\mathcal{H})$ a PVM. If $f : X \rightarrow \mathbb{C}$ is measurable, define*

$$\Delta_f := \left\{ x \in \mathcal{H} \mid \int_X |f(\lambda)|^2 \mu_{xx}^{(P)}(\lambda) < +\infty \right\} .$$

Δ_f is a dense subspace of \mathcal{H} and there is a unique operator

$$\int_X f(\lambda) dP(\lambda) : \Delta_f \rightarrow \mathcal{H} \quad (47)$$

such that

$$\left\langle x \mid \int_X f(\lambda) dP(\lambda) y \right\rangle = \int_X f(\lambda) \mu_{xy}^{(P)}(\lambda) \quad \forall x \in \mathcal{H}, \forall y \in \Delta_f \quad (48)$$

The operator in (47) turns out to be closed and normal. It finally satisfies

$$\left(\int_X f(\lambda) dP(\lambda) \right)^* = \int_X \overline{f(\lambda)} dP(\lambda) \quad (49)$$

and

$$\left\| \int_X f(\lambda) dP(\lambda) x \right\|^2 = \int_X |f(\lambda)|^2 d\mu_{xx}^{(P)}(\lambda) \quad \forall x \in \Delta_f . \quad (50)$$

Idea of the existence part of the Proof. The idea of the proof of existence of the operator in (47) relies upon the validity of the inequality ((1) in exercises 67 below)

$$\int_X |f(\lambda)| d|\mu_{xy}^{(P)}|(\lambda) \leq \|x\| \sqrt{\int_X |f(\lambda)|^2 d\mu_{yy}^{(P)}(\lambda)} \quad \forall y \in \Delta_f, \forall x \in \mathcal{H} . \quad (51)$$

This inequality also proves that $f \in L^2(X, d\mu_{yy}^{(P)})$ implies $f \in L^1(X, d|\mu_{xy}^{(P)}|)$ for $x \in \mathcal{H}$, so that (48) makes sense. Since from the general measure theory

$$\left| \int_X f(\lambda) d\mu_{xy}^{(P)}(\lambda) \right| \leq \int_X |f(\lambda)| d|\mu_{xy}^{(P)}|(\lambda) ,$$

(51) implies that $\mathcal{H} \ni x \mapsto \int_X f(\lambda) d\mu_{xy}^{(P)}(\lambda)$ is continuous at $x = 0$. This map is also anti-linear as follows from the definition of $\mu_{x,y}$. An elementary use of Riesz' lemma proves that there exists a vector, indicated by $\int_X f(\lambda) dP(\lambda) y$, satisfying (48). That is the action of an operator on a vector $y \in \Delta_f$ because $\Delta_f \ni y \mapsto \int_X f(\lambda) d\mu_{xy}^{(P)}(\lambda)$ is linear. \square

Remark 66. Identity (50) gives Δ_f a direct meaning in terms of boundedness of $\int_X f(\lambda) dP(\lambda)$. Since $\mu_{xx}(X) = \|x\|^2 < +\infty$, (50) together with the definition of Δ_f

immediately implies that: if f is bounded or, more weakly P -essentially bounded^h on X , then

$$\int_X f(\lambda) dP(\lambda) \in \mathfrak{B}(\mathcal{H})$$

and

$$\left\| \int_X f(\lambda) dP(\lambda) \right\| \leq \|f\|_\infty^{(P)} \leq \|f\|_\infty.$$

The P -essentially boundedness is also a *necessary* (not only sufficient) condition for $\int_X f(\lambda) dP(\lambda) \in \mathfrak{B}(\mathcal{H})$ [8,6]. ■

Exercise 67.

(1) *Prove inequality (51).*

Solution. Let $x \in \mathcal{H}$ and $y \in \Delta_f$. If $s : X \rightarrow \mathbb{C}$ is a *simple function* and $h : X \rightarrow \mathbb{C}$ is the *Radon-Nikodym derivative* of μ_{xy} with respect to $|\mu_{xy}|$ so that $|h(x)| = 1$ and $\mu_{xy}(E) = \int_E h d|\mu_{xy}|$ (see, e.g., [6]), we have for an increasing sequence of simple functions $z_n \rightarrow h$ pointwise, with $|z_n| \leq |h^{-1}| = 1$, due to the dominate convergence theorem,

$$\int_X |s| d|\mu_{xy}| = \int_X |s| h^{-1} d\mu_{xy} = \lim_{n \rightarrow +\infty} \int_X |s| z_n d\mu_{xy} = \lim_{n \rightarrow +\infty} \left\langle x \left| \sum_{k=1}^{N_n} z_{n,k} P_{E_{n,k}} y \right. \right\rangle.$$

In the last step we have made use of (iii)(b) in remark 63 for the simple function $|s| z_n = \sum_{k=1}^{N_n} z_{n,k} \chi_{E_{n,k}}$. Cauchy Schwartz inequality immediately yields

$$\int_X |s| d|\mu_{xy}| \leq \|x\| \lim_{n \rightarrow +\infty} \left\| \sum_{k=1}^{N_n} z_{n,k} P_{E_{n,k}} y \right\| = \|x\| \lim_{n \rightarrow +\infty} \sqrt{\int_X |s z_n|^2 d\mu_{yy}},$$

where we have used $P_{E_{n,k}}^* P_{E_{n,k'}} = P_{E_{n,k}} P_{E_{n,k'}} = \delta_{kk'} P_{E_{n,k}}$ since $E_{n,k} \cap E_{n,k'} = \emptyset$ for $k \neq k'$. Next observe that, as $|s z_n|^2 \rightarrow |s h^{-1}|^2 = |s|^2$, dominate convergence theorem leads to

$$\int_X |s| d|\mu_{xy}| \leq \|x\| \sqrt{\int_X |s|^2 d\mu_{yy}}.$$

Finally, replace s above for a sequence of simple functions $|s_n| \rightarrow f \in L^2(X, d\mu_{yy})$ pointwise, with $s_n \leq |s_{n+1}| \leq |f|$. Monotone convergence theorem and dominate convergence theorem, respectively applied to the left and right-hand side of the found inequality, produce inequality (51).

(2) *Prove that, with the hypotheses of Proposition 65, it holds*

$$\int_X \chi_E(\lambda) dP(\lambda) = P_E, \quad \text{if } E \in \Sigma(X) \quad (52)$$

^hAs usual, $\|f\|_\infty^{(P)}$ is the infimum of positive reals r such that $P(\{x \in X \mid |f(x)| > r\}) = 0$.

and in particular

$$\int_X 1 dP(\lambda) = I. \quad (53)$$

Solution. It is sufficient to prove (52) since we know that $P_X = I$. To this end, notice that, by direct inspection

$$\langle x | P_E y \rangle = \int_X \chi_E(\lambda) \mu_{xy}^{(P)}(\lambda) \quad \forall x \in \mathcal{H}, \forall y \in \Delta_{\chi_E} = \mathcal{H}.$$

By the uniqueness property stated in Proposition 65 (52) holds. \square

(3) Prove that if P a PVM on \mathcal{H} and T is an operator in \mathcal{H} with $D(T) = \Delta_f$ such that

$$\langle x | Tx \rangle = \int_X f(\lambda) \mu_{xx}^{(P)}(\lambda) \quad \forall x \in \Delta_f \quad (54)$$

then

$$T = \int_X f(\lambda) dP(\lambda).$$

Solution. From the definition of μ_{xy} we easily have (everywhere omitting (P) for simplicity)

$$4\mu_{xy}(E) = \mu_{x+y, x+y}(E) - \mu_{x-y, x-y}(E) - i\mu_{x+iy, x+iy}(E) + i\mu_{x-iy, x-iy}(E)$$

This identity implies that, if $x, y \in \Delta_f$,

$$4 \int_X f d\mu_{xy} = \int_X f d\mu_{x+y, x+y} - \int_X f d\mu_{x-y, x-y} - i \int_X f d\mu_{x+iy, x+iy} + i \int_X f d\mu_{x-iy, x-iy}$$

Similarly, from the elementary properties of the scalar product, when $x, y \in D(T)$

$$4\langle x | Ty \rangle = \langle x+y | T(x+y) \rangle - \langle x-y | T(x-y) \rangle - i\langle x+iy | T(x+iy) \rangle + i\langle x-iy | T(x-iy) \rangle.$$

It is then obvious that (54) implies

$$\langle x | Ty \rangle = \int_X f(\lambda) \mu_{xy}^{(P)}(\lambda) \quad \forall x, y \in \Delta_f,$$

so that

$$\left\langle x \left| \left(T - \int_X f(\lambda) dP(\lambda) \right) y \right. \right\rangle = 0 \quad \forall x, y \in \Delta_f$$

Since x varies in a dense set Δ_f , $Ty - \int_X f(\lambda) dP(\lambda)y = 0$ for every $y \in \Delta_f$ which is the thesis. \square

Example 68.

(1) Referring to the PVM in (2) of example 64, directly from the definition of $\int_X f(\lambda) dP(\lambda)$ or exploiting (3) in exercises 67 we have that

$$\int_{\mathbb{N}} f(\lambda) dP(\lambda) z = \sum_{n \in \mathbb{N}} f(n) Q_n z$$

for every $f : \mathbb{N} \rightarrow \mathbb{C}$ (which is necessarily measurable with our definition of $\Sigma(\mathbb{N})$). Correspondingly, the domain of $\int_{\mathbb{N}} f(\lambda) dP(\lambda)$ results to be

$$\Delta_f := \left\{ z \in \mathcal{H} \mid \sum_{n \in \mathbb{N}} |f(n)|^2 \|Q_n z\|^2 < +\infty \right\}$$

We stress that we have found a direct generalization of the expansion (4) if the operator A is now hopefully written as

$$Az = \sum_{n \in \mathbb{N}} n Q_n z.$$

We shall see below that it is the case.

(3) Referring to the PVM in (3) of example 64, directly from the definition of $\int_X f(\lambda) dP(\lambda)$ or exploiting (3) in exercises 67 we have that

$$\left(\int_{\mathbb{R}} f(\lambda) dP(\lambda) \psi \right) (x) = f(x) \psi(x), \quad x \in \mathbb{R}$$

Correspondingly, the domain of $\int_{\mathbb{R}} f(\lambda) dP(\lambda)$ results to be

$$\Delta_f := \left\{ \psi \in L^2(\mathbb{R}, dx) \mid \int_{\mathbb{R}} |f(x)|^2 |\psi(x)|^2 dx < +\infty \right\}$$

■

2.6. Spectral Decomposition and Representation Theorems

We are in a position to state the fundamental result of the spectral theory of selfadjoint operators, which extend the expansion (4) to an integral formula valid also in the infinite dimensional case, and where the set of eigenvalues is replaced by the full spectrum of the selfadjoint operator.

To state the theorem, we preventively notice that (49) implies that $\int f(\lambda) dP(\lambda)$ is selfadjoint if f is real: The idea of the theorem is to prove that every selfadjoint operator can be written this way for a specific f and with respect to a PVM on \mathbb{R} associated with the operator itself.

Notation 69. From now on $\mathcal{B}(T)$ denotes the Borel σ -algebra on the topological space T . ■

Theorem 70 (Spectral Decomposition Theorem for Selfadjoint Operators). *Let A be a selfadjoint operator in the complex Hilbert space \mathcal{H} .*

(a) *There is a unique PVM, $P^{(A)} : \mathcal{B}(\mathbb{R}) \rightarrow \mathcal{L}(\mathcal{H})$, such that*

$$A = \int_{\mathbb{R}} \lambda dP^{(A)}(\lambda).$$

In particular $D(A) = \Delta_{id}$, where $id : \mathbb{R} \ni \lambda \mapsto \lambda$.

(b) *Defining the **support** of $P^{(A)}$, $\text{supp}(P^{(A)})$, as the complement in \mathbb{R} of the union of all open sets $O \subset \mathbb{C}$ with $P_O^{(A)} = 0$ it results*

$$\text{supp}(P^{(A)}) = \sigma(A)$$

so that

$$P^{(A)}(E) = P^{(A)}(E \cap \sigma(A)), \quad \forall E \in \mathcal{B}(\mathbb{R}). \quad (55)$$

(c) $\lambda \in \sigma_p(A)$ if and only if $P^{(A)}(\{\lambda\}) \neq 0$, this happens in particular if λ is an isolated point of $\sigma(A)$.

(d) $\lambda \in \sigma_c(A)$ if and only if $P^{(A)}(\{\lambda\}) = 0$ but $P^{(A)}(E) \neq 0$ if $E \ni \lambda$ is an open set of \mathbb{R} .

The proof can be found, e.g., in [8,6,9].

Remark 71. Theorem 70 is a particular case of a more general theorem (see [8,6] and especially [9]) valid when A is a (densely defined closed) normal operator. The general statement is identical, it is sufficient to replace everywhere \mathbb{R} for \mathbb{C} . A particular case is the one of A unitary. In this case the statement can be rephrased replacing everywhere \mathbb{R} for \mathbb{T} since it includes the spectrum of A in this case ((d) remark 57). ■

Notation 72. In view of the said theorem, and (b) in particular, if $f : \sigma(A) \rightarrow \mathbb{C}$ is measurable (with respect to the σ -algebra obtained by restricting $\mathcal{B}(\mathbb{R})$ to $\sigma(A)$), we use the notation

$$f(A) := \int_{\sigma(A)} f(\lambda) dP^{(A)}(\lambda) := \int_{\mathbb{R}} g(\lambda) dP^{(A)}(\lambda) =: g(A). \quad (56)$$

where $g : \mathbb{R} \rightarrow \mathbb{C}$ is the extension of f to the zero function outside $\sigma(A)$ or any other measurable function which coincides with f on $\text{supp}(P^{(A)}) = \sigma(A)$. Obviously $g(A) = g'(A)$ if $g, g' : \mathbb{R} \rightarrow \mathbb{C}$ coincide in $\text{supp}(P^{(A)}) = \sigma(A)$. ■

Exercise 73. Prove that if A is a selfadjoint operator in the complex Hilbert space \mathcal{H} , it holds $A \geq 0$ - that is $\langle x|Ax \rangle \geq 0$ for every $x \in D(A)$ - if and only if $\sigma(A) \subset [0, +\infty)$.

Solution. Suppose that $\sigma(A) \subset [0, +\infty)$. If $x \in D(A)$ we have $\langle x|Ax \rangle = \int_{\sigma(A)} \lambda d\mu_{x,x} \geq 0$ in view of (48), the spectral decomposition theorem, since $\mu_{x,x}$ is a positive measure ad $\sigma(A) \in [0, +\infty)$. To conclude, we prove that $A \geq 0$ is false if $\sigma(A)$ includes negative elements. To this end assume that, conversely, $\sigma(A) \ni \lambda_0 < 0$. Using (c) and (d) of Theorem 70, one finds an interval $[a, b] \subset \sigma(A)$ with $[a, b] \subset (-\infty, 0)$ and $P_{[a,b]}^{(A)} \neq 0$ (possibly $a = b = \lambda_0$). If $x \in P_{[a,b]}^{(A)}(\mathcal{H})$ with $x \neq 0$, it holds $\mu_{xx}(E) = \langle x|P_E x \rangle = \langle x|P_{[a,b]}^* P_E x P_{[a,b]} \rangle = \langle x|P_{[a,b]} P_E P_{[a,b]} x \rangle = \langle x|P_{[a,b] \cap E} x \rangle = 0$ if $[a, b] \cap E = \emptyset$. Therefore, $\langle x|Ax \rangle = \int_{\sigma(A)} \lambda d\mu_{x,x} = \int_{[a,b]} \lambda d\mu_{x,x} \leq \int_{[a,b]} b d\mu_{x,x} < b \|x\|^2 < 0$. ■

Example 74.

(1) Let us focus on the m -axis position operator X_m in $L^2(\mathbb{R}^n, d^n x)$ introduced in (1) of example 54. We know that $\sigma(X_m) = \sigma_c(X_m) = \mathbb{R}$ from example 59. We are interested in the PVM $P^{(X_m)}$ of X_m defined on $\mathbb{R} = \sigma(X_m)$. Let us fix $m = 1$ the other cases are analogous. The PVM associated to X_1 is

$$(P_E^{(X_1)} \psi)(x) = \chi_{E \times \mathbb{R}^{n-1}}(x) \psi(x) \quad \psi \in L^2(\mathbb{R}^n, d^n x), \quad (57)$$

where $E \in \mathcal{B}(\mathbb{R})$ is here identified with a subset of the first factor of $\mathbb{R} \times \mathbb{R}^{n-1} = \mathbb{R}^n$. Indeed, indicating by P the right-hand side of (57), one easily verifies that $\Delta_{x_1} = D(X_1)$ and¹

$$\langle \psi | X_1 \psi \rangle = \int_{\mathbb{R}} \lambda \mu_{\psi, \psi}^{(P)}(\lambda) \quad \forall \psi \in D(X_1) = \Delta_{x_1}$$

where $\mu_{\psi, \psi}^{(P)}(E) = \langle \psi | P_E \psi \rangle = \int_{E \times \mathbb{R}^{n-1}} |\psi(x)|^2 d^n x$. (2) in exercise 64 proves that $X_1 = \int_{\mathbb{R}} \lambda dP(\lambda)$ and thus (57) holds true.

(2) Considering the m -axis momentum operator P_m in $L^2(\mathbb{R}^n, d^n x)$ introduced in (2) of example 54, taking (36) into account where $\hat{\mathcal{F}}$ (and thus $\hat{\mathcal{F}}^*$) is unitary, in view of (i) in Proposition 82 we immediately have that the PVM of P_m is

$$Q_E^{(P_m)} := \hat{\mathcal{F}}^* P_E^{(K_m)} \hat{\mathcal{F}}.$$

Above K_m is the operator X_m represented in $L^2(\mathbb{R}^n, d^n x)$ as in (1) of example 54.

(3) More complicated cases exist. Considering an operator of the form

$$H := \frac{1}{2m} P^2 + U$$

where P is the momentum operator in $L^2(\mathbb{R}, dx)$, $m > 0$ is a constant and U is a real valued function on \mathbb{R} used as multiplicative operator. If $U = U_1 + U_2$ with $U_1 \in L^2(\mathbb{R}, dx)$ and $U_2 \in L^\infty(\mathbb{R}, dx)$ real valued, and $D(H) = C^\infty(\mathbb{R}; \mathbb{C})$, H turns out to be (trivially) symmetric but also essentially selfadjoint [6] as a consequence of a well known result (*Kato-Rellich's theorem*). The unique selfadjoint extension $\overline{H} = (H^*)^*$ of H physically represent the Hamiltonian operator of a quantum particle living along \mathbb{R} with a potential energy described by U . In this case, generally speaking, $\sigma(\overline{H})$ has both point and continuous part. $\int_{\sigma_p(\overline{H})} \lambda dP^{(\overline{H})}(\lambda)$ has a form like this

$$\int_{\sigma_p(\overline{H})} \lambda dP^{(\overline{H})}(\lambda) = \sum_{\lambda \in \sigma_p(\overline{H})} \lambda P_\lambda$$

where P_λ is the orthogonal projector onto the eigenspace of \overline{H} with eigenvalue λ . Conversely, $\int_{\sigma_c(\overline{H})} \lambda dP^{(\overline{H})}(\lambda)$ has an expression much more complicated and, under a unitary transform is similar to the integral decomposition of X . ■

Remark 75.

(a) It is worth stressing that the notion (56) of a function of a selfadjoint operator is just an extension of in the analogous notion introduced for the finite dimensional case (7) and thus may be used in QM applications.

It is possible to prove that if $f : \sigma(A) \rightarrow \mathbb{R}$ is continuous, then

$$\sigma(f(A)) = \overline{f(\sigma(A))} \tag{58}$$

¹More generally $\int_{\mathbb{R}} \int_{\mathbb{R}^{n-1}} g(x_1) |\psi(x)|^2 dx d^{n-1} x = \int_{\mathbb{R}} g(x_1) d\mu_{\psi, \psi}^{(P)}(x_1)$ is evidently valid for simple functions and then it extends to generic measurable functions when both sides make sense in view of, for instance, Lebesgue's dominate convergence theorem for positive measures.

where the bar denotes the closure and, if $f : \sigma(A) \rightarrow \mathbb{R}$ is measurable,

$$\sigma_p(f(A)) \supset f(\sigma_p(A)). \quad (59)$$

More precise statements based on the notion of *essential range* can be found in [6]. It turns out that, for A selfadjoint and $f : \sigma(A) \rightarrow \mathbb{C}$ measurable, $z \in \sigma(f(A))$ if and only if $P^{(A)}(E_z) \neq 0$ for some open set $E_z \ni z$. Now $z \in \sigma(f(A))$ is in $\sigma_p(f(A))$ iff $P^{(A)}(f^{-1}(z)) \neq 0$ or it is in $\sigma_c(f(A))$ iff $P^{(A)}(f^{-1}(z)) = 0$.

(b) It is fundamental to stress that in, QM, (58) permits us to adopt the standard operational approach on observables $f(A)$ as the observable whose set of possible values is (the closure of) the set of reals $f(a)$ where a is a possible value of A .

(c) The following important fact holds

Proposition 76. *A selfadjoint operator is bounded (and its domain coincide to the whole \mathcal{H}) if and only if $\sigma(A)$ is bounded.*

Proof. It essentially follows from (50) restricting the integration space to $X = \sigma(A)$. In fact, if $\sigma(A)$ is bounded and thus compact it being closed, the continuous function $id : \sigma(A) \ni \lambda \rightarrow \lambda$ is bounded and (50) implies that $A = \int_{\sigma(A)} id dP^{(A)}$ is bounded and the inequality holds

$$\|A\| \leq \sup\{|\lambda| \mid \lambda \in \sigma(A)\}. \quad (60)$$

In this case it also hold $D(A) = \Delta_{id} = \mathcal{H}$.

If, conversely, $\sigma(A)$ is not bounded, we can find a sequence $\lambda_n \in \sigma(A)$ with $|\lambda_n| \rightarrow \infty$ as $n \rightarrow +\infty$. With the help of (c) and (d) in Theorem 70, it is easy to construct vectors x_n with $\|x_n\| \neq 0$ and $x_n \in P_{B(\lambda_n)}^{(A)}(\mathcal{H})$ where $B(\lambda_n) := [\lambda_n - 1, \lambda_n + 1]$. (50) implies

$$\|Ax_n\|^2 \geq \|x_n\|^2 \inf_{z \in B(\lambda_n)} |id(z)|^2$$

Since $\inf_{z \in B(\lambda_n)} |id(z)|^2 \rightarrow +\infty$, we have that $\|Ax_n\|/\|x_n\|$ is not bounded and A , in turn, cannot be bounded. In this case, since $A = A^*$, Theorem 44 entails that $D(A)$ is *strictly* included in \mathcal{H} . \square

It is possible to prove [6] that (60) can be turned into an identity when $A \in \mathfrak{B}(\mathcal{H})$ also if A is not selfadjoint but only normal

$$\|A\| = \sup\{|\lambda| \mid \lambda \in \sigma(A)\}, \quad (61)$$

This is the well known *spectral radius formula*, the **spectral radius** of $A \in \mathfrak{B}(\mathcal{H})$ being, by definition, the number in the right hand side.

(d) The result stated in (c) explains the reason why observables A in QM are very often represented by unbounded selfadjoint operators. $\sigma(A)$ is the set of values of the observable A . When, as it happens very often, that observable is allowed to take arbitrarily large values (think of X or P), it cannot be represented by a bounded selfadjoint operator just because its spectrum is not bounded.

(e) If P is a PVM on \mathbb{R} and $f : \mathbb{R} \rightarrow \mathbb{C}$ is measurable, we can always write

$$\int_{\mathbb{R}} f(\lambda) dP(\lambda) = f(A)$$

where we have introduced the selfadjoint operator A obtained as

$$A = \int_{\mathbb{R}} id(\lambda) dP(\lambda), \quad (62)$$

due to (49) and where $id : \mathbb{R} \ni \lambda \rightarrow \lambda$. Evidently $P^{(A)} = P$ due to the uniqueness part of the spectral theorem. This fact leads to the conclusion that, *in a complex Hilbert space \mathcal{H} , all the PVM over \mathbb{R} are one-to-one associated to all selfadjoint operators in \mathcal{H} .*

(f) An element $\lambda \in \sigma_c(A)$ is not an eigenvalue of A . However there is the following known result arising from (d) in Theorem 70 [6] which proves that we can have approximated eigenvalues with arbitrary precision: With the said hypotheses, for every $\epsilon > 0$ there is $x_\epsilon \in D(A)$ such that

$$\|Ax_\epsilon - \lambda x_\epsilon\| < \epsilon, \quad \text{but } \|x_\epsilon\| = 1.$$

(g) If A is selfadjoint and U unitary, UAU^* , with $D(UAU^*) = U(D(A))$, is selfadjoint as well (Proposition 45). It is very simple to prove that the PVM of UAU^* is nothing but $UP^{(A)}U^*$. ■

The next theorem we state here concerns a general explicit form of the integral decomposition $f(A) = \int_{\sigma(A)} f(\lambda) dP^{(A)}(\lambda)$. As a matter of facts, up to multiplicity, one can always reduce to a multiplicative operator in a L^2 space, as it happens for the position operator X . Again, this theorem can be restated for generally normal operators.

Theorem 77 (Spectral Representation Theorem for Selfadjoint Operators). *Let A be a selfadjoint operator in the complex Hilbert space \mathcal{H} . The following facts hold.*

(a) \mathcal{H} may be decomposed a Hilbert sum^j $\mathcal{H} = \bigoplus_{a \in S} \mathcal{H}_a$, whose summands \mathcal{H}_a are closed and orthogonal. Moreover:

(i) for any $a \in S$,

$$A(\mathcal{H}_a \cap D(A)) \subset \mathcal{H}_a$$

and, more generally, for any measurable $f : \sigma(A) \rightarrow \mathbb{C}$,

$$f(A)(\mathcal{H}_a \cap D(f(A))) \subset \mathcal{H}_a$$

(ii) for any $a \in S$ there exist a unique finite positive Borel measure μ_a on $\sigma(A) \subset \mathbb{R}$, and a surjective isometric operator $U_a : \mathcal{H}_a \rightarrow L^2(\sigma(A), \mu_a)$, such that:

$$U_a f(A)|_{\mathcal{H}_a} U_a^{-1} = f.$$

^j S is countable, at most, if \mathcal{H} is separable.

for any measurable $f : \sigma(A) \rightarrow \mathbb{C}$, where $f \cdot$ is the point-wise multiplication by f on $L^2(\sigma(A), \mu_a)$.

(b) If $\text{supp}\{\mu_a\}_{a \in S}$ is the complementary set to the numbers $\lambda \in \mathbb{R}$ for which there exists an open set $O_\lambda \subset \mathbb{R}$ with $O_\lambda \ni \lambda$, $\mu_a(O_\lambda) = 0$ for any $a \in S$, then

$$\sigma(A) = \text{supp}\{\mu_a\}_{a \in S}.$$

Remark 78. Notice that the theorem encompasses the case of an operator A in \mathcal{H} with $\sigma(A) = \sigma_p(A)$. Suppose in particular that every eigenspace is one-dimensional and the whole Hilbert space is separable. Let $\sigma(A) = \sigma_p(A) = \{\lambda_n \mid n \in \mathbb{N}\}$. In this case

$$A = \sum_{n \in \mathbb{N}} \lambda_n \langle x_n | \cdot \rangle x_n,$$

where x_λ is a unit eigenvector with eigenvalue λ_n . Consider the σ -algebra on $\sigma(A)$ made of all subsets and define $\mu(E) :=$ number of elements of $E \subset \sigma(A)$. In this case \mathcal{H} is isomorphic to $L^2(\sigma(A), \mu)$ and the isomorphism is $U : \mathcal{H} \ni x \mapsto \psi_x \in L^2(\sigma(A), \mu)$ with $\psi_x(n) := \langle x_n | x \rangle$ if $n \in \mathbb{N}$. With this surjective isometry, trivially

$$U f(A) U^{-1} = U \int_{\sigma(A)} f(\lambda) dP^{(A)}(\lambda) U^{-1} = U \sum_{n \in \mathbb{N}} f(\lambda_n) \langle x_n | \cdot \rangle x_n U^{-1} = f \cdot.$$

If all eigenspaces have dimension 2, exactly two copies of $L^2(\sigma(A), \mu)$ are sufficient to improve the construction. If the dimension depends on the eigenspace, the construction can be rebuilt exploiting many copies of different $L^2(S_k, \mu_k)$, where the S_k are suitable (not necessarily disjoint) subsets of $\sigma(A)$ and μ_k the measure which counts the elements of S_k . ■

The last tool we introduce is the notion of *joint spectral measure*. Everything is stated in the following theorem [6].

Theorem 79 (Joint spectral measure). Consider selfadjoint operators A_1, A_2, \dots, A_n in the complex separable Hilbert space \mathcal{H} . Suppose that the spectral measures of those operators pairwise commute:

$$P_{E_k}^{(A_k)} P_{E_h}^{(A_h)} = P_{E_h}^{(A_h)} P_{E_k}^{(A_k)} \quad \forall k, h \in \{1, \dots, n\}, \forall E_k, E_h \in \mathcal{B}(\mathbb{R}).$$

There is a unique PVM, $P^{(A_1 \times \dots \times A_n)}$, on \mathbb{R}^n such that

$$P^{(A_1 \times \dots \times A_n)}(E_1 \times \dots \times E_n) = P_{E_1}^{(A_1)} \dots P_{E_n}^{(A_n)}, \quad \forall E_1, \dots, E_n \in \mathcal{B}(\mathbb{R}).$$

For every $f : \mathbb{R} \rightarrow \mathbb{C}$ measurable, it holds

$$\int_{\mathbb{R}^n} f(x_k) dP^{(A_1 \times \dots \times A_n)}(x) = f(A_k), \quad k = 1, \dots, n \quad (63)$$

where $x = (x_1, \dots, x_k, \dots, x_n)$.

Definition 80. Referring to Theorem 79, the PVM $P^{(A_1 \times \dots \times A_n)}$ is called the **joint spectral measure** of A_1, A_2, \dots, A_n and its support $\text{supp}(P^{(A_1 \times \dots \times A_n)})$, i.e. the

complement in \mathbb{R}^n to the largest open set A with $P_A = 0$, is called the **joint spectrum** of A_1, A_2, \dots, A_n . ■

Example 81. The simplest example is provided by considering the n position operators X_m in $L^2(\mathbb{R}^n, d^n x)$. It should be clear that the n spectral measures commute because $P_E^{(X_k)}$, for $E \in \mathcal{B}(\mathbb{R})$, is the multiplicative operator for $\chi_{\mathbb{R} \times \dots \times \mathbb{R} \times E \times \mathbb{R} \times \dots \times \mathbb{R}}$ the factor E staying in the k -th position among the n Cartesian factors. In this case the joint spectrum of the n operators X_m coincides with \mathbb{R}^n itself.

A completely analogous discussion holds for the n momentum operators P_k , since they are related to the position ones by means of the unitary Fourier-Plancherel operator as already seen several times. Again the joint spectrum of the n operators P_m coincides with \mathbb{R}^n itself. ■

2.7. Mesurable functional calculus

The following proposition states some useful properties of $f(A)$, where A is self-adjoint and $f : \mathbb{R} \rightarrow \mathbb{C}$ is Borel measurable. These properties define the so called *measurable functional calculus*. We suppose here that $A = A^*$, but the statements can be reformulated for normal operators [6].

Proposition 82. *Let A be a selfadjoint operator in the complex Hilbert space \mathcal{H} , $f, g : \sigma(A) \rightarrow \mathbb{C}$ measurable functions, $f \cdot g$ and $f + g$ respectively denote the point-wise product and the point-wise sum of functions. The following facts hold.*

(a) $f(A) = \sum_{k=0}^n a_k A^k$ where the right-hand side is defined in its standard domain $D(A^n)$ when $f(\lambda) = \sum_{k=0}^n a_k \lambda^k$ with $a_n \neq 0$.

(b) $f(A) = P^{(A)}(E)$ if $f = \chi_E$ the characteristic function of $E \in \mathcal{B}(\sigma(A))$;

(c) $f(A)^* = \bar{f}(A)$ where the bar denotes the complex conjugation;

(d) $f(A) + g(A) \subset (f + g)(A)$ and $D(f(A) + g(A)) \subset \Delta_f \cap \Delta_g$
(the symbol “ \subset ” can be replaced by “ $=$ ” if and only if $\Delta_{f+g} = \Delta_f \cap \Delta_g$),

(e) $f(A)f(B) \subset (f \cdot g)(A)$ and $D(f(A)f(B)) = \Delta_{f \cdot g} \cap \Delta_g$
(the symbol “ \subset ” can be replaced by “ $=$ ” if and only if $\Delta_{f \cdot g} \subset \Delta_g$),

(f) $f(A)^* f(A) = |f|^2(A)$ so that $D(f(A)^* f(A)) = \Delta_{|f|^2}$,

(g) $\langle x | f(A)x \rangle \geq 0$ for $x \in \Delta_f$ if $f \geq 0$.

(h) $\|f(A)x\|^2 = \int_{\sigma(A)} |f(\lambda)|^2 d\mu_{xx}(\lambda)$, if $x \in \Delta_f$.

In particular, if f is bounded or $P^{(A)}$ -essentially bounded^k on $\sigma(A)$, $f(A) \in \mathfrak{B}(\mathcal{H})$ and

$$\|f(A)\| \leq \|f\|_{\infty}^{P^{(A)}} \leq \|f\|_{\infty}.$$

(i) If $U : \mathcal{H} \rightarrow \mathcal{H}$ is unitary, $Uf(A)U^* = f(UAU^*)$ and, in particular, $D(f(UAU^*)) = UD(f(A)) = U(\Delta_f)$.

(j) If $\phi : \mathbb{R} \rightarrow \mathbb{R}$ is measurable, then $\mathcal{B}(\mathbb{R}) \ni E \mapsto P'(E) := P^{(A)}(\phi^{-1}(E))$ is a PVM on \mathbb{R} . Introducing the selfadjoint operator

$$A' = \int_{\mathbb{R}} \lambda' dP'(\lambda')$$

such that $P^{(A')} = P'$, we have

$$A' = \phi(A).$$

Moreover, if $f : \mathbb{R} \rightarrow \mathbb{C}$ is measurable,

$$f(A') = (f \circ \phi)(A) \quad \text{and} \quad \Delta'_{f'} = \Delta_{f \circ \phi}.$$

2.8. Elementary formalism for the infinite dimensional case

To complete the discussion in the introduction, let us show how practically the physical hypotheses on quantum systems (1)-(3) have to be mathematically interpreted (again reversing the order of (2) and (3) for our convenience) in the general case of infinite dimensional Hilbert spaces. Our general assumptions on the mathematical description of quantum systems are the following ones.

- (1) A quantum mechanical system S is always associated to complex Hilbert space \mathcal{H} , finite or infinite dimensional;
- (2) observables are pictured in terms of (generally unbounded) *self-adjoint* operators A in \mathcal{H} ,
- (3) states are of equivalence classes of *unit* vectors $\psi \in \mathcal{H}$, where $\psi \sim \psi'$ iff $\psi = e^{ia}\psi'$ for some $a \in \mathbb{R}$.

Let us show how the mathematical assumptions (1)-(3) permit us to set the physical properties of quantum systems (1)-(3) of Section 1.1.2 into mathematically nice form in the general case of an infinite dimensional Hilbert space \mathcal{H} .

(1) Randomness: The Borel subset $E \subset \sigma(A)$, represents the outcomes of measurement procedures of the observable associated with the selfadjoint operator A . (In case of continuous spectrum the outcome of a measurement is at least an interval in view of the experimental errors.) Given a state represented by the unit

^kRemark 66.

vector $\psi \in \mathcal{H}$, the probability to obtain $E \subset \sigma(A)$ as an outcome when measuring A is

$$\mu_{\psi, \psi}^{(P(A))}(E) := \|P_E^{(A)}\psi\|^2,$$

where we have used the PVM $P^{(A)}$ of the operator A .

Going along with this interpretation, the **expectation value**, $\langle A \rangle_\psi$, of A when the state is represented by the unit vector $\psi \in \mathcal{H}$, turns out to be

$$\langle A \rangle_\psi := \int_{\sigma(A)} \lambda d\mu_{\psi, \psi}^{(P(A))}(\lambda). \quad (64)$$

This identity makes sense provided $id : \sigma(A) \ni \lambda \rightarrow \lambda$ belongs to $L^1(\sigma(A), \mu_{\psi, \psi}^{(P(A))})$ (which is equivalent to say that $\psi \in \Delta_{|id|^{1/2}}$ and, in turn, that $\psi \in D(|A|^{1/2})$), otherwise the expectation value is not defined.

Since

$$L^2(\sigma(A), \mu_{\psi, \psi}^{(P(A))}) \subset L^1(\sigma(A), \mu_{\psi, \psi}^{(P(A))})$$

because $\mu_{\psi, \psi}^{(P(A))}$ is finite, we have the popular identity arising from (48),

$$\langle A \rangle_\psi = \langle \psi | A \psi \rangle \quad \text{if } \psi \in D(A). \quad (65)$$

The associated **standard deviation**, ΔA_ψ , results to be

$$\Delta A_\psi^2 := \int_{\sigma(A)} (\lambda - \langle A \rangle_\psi)^2 d\mu_{\psi, \psi}^{(P(A))}(\lambda). \quad (66)$$

This definition makes sense provided $id \in L^2(\sigma(A), \mu_{\psi, \psi}^{(P(A))})$ (which is equivalent to say that $\psi \in \Delta_{id}$ and, in turn, that $\psi \in D(A)$).

As before, the functional calculus permits us to write the other popular identity

$$\Delta A_\psi^2 = \langle \psi | A^2 \psi \rangle - \langle \psi | A \psi \rangle^2 \quad \text{if } \psi \in D(A^2) \subset D(A). \quad (67)$$

We stress that now, Heisenberg inequalities, as established in exercise 14, are now completely justified as the reader can easily check.

(3) Collapse of the state: If the Borel set $E \subset \sigma(A)$ is the outcome of the (idealized) measurement of A , when the state is represented by the unit vector $\psi \in \mathcal{H}$, the new state immediately after the measurement is represented by the unit vector

$$\psi' := \frac{P_E^{(A)}\psi}{\|P_E^{(A)}\psi\|}. \quad (68)$$

Remark 83. Obviously this formula does not make sense if $\mu_{\psi, \psi}^{(P(A))}(E) = 0$ as expected. Moreover the arbitrary phase affecting ψ does not give rise to troubles due to the linearity of $P_E^{(A)}$. ■

(2) Compatible and Incompatible Observables: Two observables A, B are compatible – i.e. they can be simultaneously measured – if and only if their **spectral measures commute** which means

$$P_E^{(A)} P_F^{(B)} = P_F^{(B)} P_E^{(A)}, \quad E \in \mathcal{B}(\sigma(A)), \quad F \in \mathcal{B}(\sigma(B)). \quad (69)$$

In this case

$$\|P_E^{(A)} P_F^{(B)} \psi\|^2 = \|P_F^{(B)} P_E^{(A)} \psi\|^2 = \|P_{E \times F}^{(A,B)} \psi\|^2$$

where $P^{(A,B)}$ is the joint spectral measure of A and B , has the natural interpretation of the probability to obtain the outcomes E and F for a simultaneous measurement of A and B . If instead A and B are incompatible it may happen that

$$\|P_E^{(A)} P_F^{(B)} \psi\|^2 \neq \|P_F^{(B)} P_E^{(A)} \psi\|^2.$$

Sticking to the case of A and B incompatible, exploiting (68),

$$\|P_E^{(A)} P_F^{(B)} \psi\|^2 = \left\| P_E^{(A)} \frac{P_F^{(B)} \psi}{\|P_F^{(B)} \psi\|} \right\|^2 \|P_F^{(B)} \psi\|^2 \quad (70)$$

has the natural meaning of *the probability of obtaining first F and next E in a subsequent measurement of B and A .*

Remark 84.

(a) It is worth stressing that the notion of probability we are using here cannot be a classical notion because of the presence of incompatible observables. The theory of conditional probability cannot follow the standard rules. The probability $\mu_\psi(E_A|F_B)$, that (in a state defined by a unit vector ψ) a certain observable A takes the value E_A when the observable B has the value F_B , cannot be computed by the standard procedure

$$\mu_\psi(E_A|F_B) = \frac{\mu_\psi(E_A \text{ AND } F_B)}{\mu_\psi(F_B)}$$

if A and B are incompatible, just because, in general, nothing exists which can be interpreted as the event “ E_A AND F_B ” if $P_E^{(A)}$ and $P_F^{(B)}$ do not commute! The correct formula is

$$\mu_\psi(E_A|F_B) = \frac{\langle \psi | P_F^{(B)} P_E^{(A)} P_F^{(B)} \psi \rangle}{\|P_F^{(B)} \psi\|^2}$$

which leads to well known different properties with respect to the classical theory, the so called combination of “probability amplitudes” in particular. As a matter of fact, up to now we do not have a clear notion of (quantum) probability. This issue will be clarified in the next section.

(b) The reason to pass from operators to their spectral measures in defining compatible observables is that, if A and B are selfadjoint and defined on different domains, $AB = BA$ does not make sense in general. Moreover it is possible to find counterexamples (due to Nelson) where commutativity of A and B on common

dense invariant subspaces does not implies that their spectral measures commute. However, from general results again due to Nelson, one has the following nice result (see exercise 165).

Proposition 85. *If selfadjoint operators, A and B , in a complex Hilbert space \mathcal{H} commute on a common dense invariant domain D where $A^2 + B^2$ is essentially selfadjoint, then the spectral measures of A and B commute.*

The following result, much easier to prove, is also true [6].

Proposition 86. *Let A, B be selfadjoint operators in the complex Hilbert space \mathcal{H} . If $B \in \mathfrak{B}(\mathcal{H})$ the following facts are equivalent,*

- (i) *the spectral measures of A and B commute (i.e. (69) holds),*
- (ii) *$BA \subset AB$,*
- (iii) *$Bf(A) \subset f(A)B$, if $f : \sigma(A) \rightarrow \mathbb{R}$ is Borel measurable,*
- (iv) *$P_E^{(A)}B = BP_E^{(A)}$ if $E \in \mathfrak{B}(\sigma(A))$,*

Another useful result toward the converse direction [6] is the following.

Proposition 87. *Let A, B be selfadjoint operators in the complex Hilbert space \mathcal{H} such that their spectral measures commute. The following facts hold.*

- (a) *$ABx = BAx$ if $x \in D(AB) \cap D(BA)$.*
- (b) *$\langle Ax|By \rangle = \langle Bx|Ay \rangle$ if $x, y \in D(A) \cap D(B)$.*

■

2.9. Technical Intermezzo: Three Operator Topologies

In QM there are at least 7 relevant topologies [13] which enter the game discussing sequences of operators, here we limit ourselves to quickly illustrate the three most important ones. We assume that \mathcal{H} is a complex Hilbert space though the illustrated examples may be extended to more general context with some re-adaptation.

(a) The strongest topology is the **uniform operator topology** in $\mathfrak{B}(\mathcal{H})$: It is the topology induced by the operator norm $\| \cdot \|$ defined in (29).

As a consequence of the definition of this topology, a sequence of elements $A_n \in \mathfrak{B}(\mathcal{H})$ is said to **uniformly** converge to $A \in \mathfrak{B}(\mathcal{H})$ when $\|A_n - A\| \rightarrow 0$ for $n \rightarrow +\infty$. We already know that $\mathfrak{B}(\mathcal{H})$ is a Banach algebra with respect to that norm and also a C^* algebra.

(b) If $\mathfrak{L}(D; \mathcal{H})$ with $D \subset \mathcal{H}$ a subspace, denotes the complex vector space of the operators $A : D \rightarrow \mathcal{H}$, the **strong operator topology** on $\mathfrak{L}(D; \mathcal{H})$ is the topology induced by the seminorms p_x with $x \in D$ and $p_x(A) := \|Ax\|$ if $A \in \mathfrak{L}(D; \mathcal{H})$.

As a consequence of the definition of this topology, a sequence of elements $A_n \in$

$\mathfrak{L}(D; \mathcal{H})$ is said to **strongly** converge to $A \in \mathfrak{L}(D; \mathcal{H})$ when $\|(A_n - A)x\| \rightarrow 0$ for $n \rightarrow +\infty$ for every $x \in D$.

It should be evident that, if we restrict ourselves to work in $\mathfrak{B}(\mathcal{H})$, the uniform operator topology is stronger than the strong operator topology.

(c) The **weak operator topology** on $\mathfrak{L}(D; \mathcal{H})$ is the topology induced by the seminorms $p_{x,y}$ with $x \in \mathcal{H}$, $y \in D$ and $p_{x,y}(A) := |\langle x|Ay \rangle|$ if $A \in \mathfrak{L}(D; \mathcal{H})$.

As a consequence of the definition of this topology, a sequence of elements $A_n \in \mathfrak{L}(D; \mathcal{H})$ is said to **weakly** converge to $A \in \mathfrak{L}(D; \mathcal{H})$ when $|\langle x|(A_n - A)y \rangle| \rightarrow 0$ for $n \rightarrow +\infty$ for every $x \in \mathcal{H}$ and $y \in D$.

It should be evident that, the strong operator topology is stronger than the weak operator topology.

Example 88.

(1) If $f : \mathbb{R} \rightarrow \mathbb{C}$ is Borel measurable, and A a selfadjoint operator in \mathcal{H} , consider the sets

$$R_n := \{r \in \mathbb{R} \mid |f(r)| < n\} \quad \text{for } n \in \mathbb{N}.$$

It is clear that $\chi_{R_n} f \rightarrow f$ pointwise as $n \rightarrow +\infty$ and that $|\chi_{R_n} f|^2 \leq |f|^2$. As a consequence restricting the operators on the left hand side to Δ_f ,

$$\int_{\sigma(A)} \chi_{R_n} f dP^{(A)} \Big|_{\Delta_f} \rightarrow f(A) \quad \text{strongly, for } n \rightarrow +\infty,$$

as an immediate consequence of Lebesgue's dominate convergence theorem and the first part of (h) in Proposition 82.

(2) If in the previous example f is bounded on $\sigma(A)$, and $f_n \rightarrow f$ uniformly on $\sigma(A)$, (or P -essentially uniformly $\|f - f_n\|_{\infty}^{(P^{(A)})} \rightarrow 0$ for $n \rightarrow +\infty$) then

$$f_n(A) \rightarrow f(A) \quad \text{uniformly, as } n \rightarrow +\infty,$$

again for the (second part of (h) in Proposition 82. ■

Exercise 89. Prove that a selfadjoint operator A in the complex Hilbert \mathcal{H} admits a dense set of analytic vectors in its domain.

Solution. Consider the class of functions $f_n = \chi_{[-n,n]}$ where $n \in \mathbb{N}$. As in (1) of example 88, we have $\psi_n := f_n(A)\psi = \int_{[-n,n]} 1 dP^{(A)}\psi \rightarrow \int_{\mathbb{R}} 1 dP^{(A)}\psi = P_{\mathbb{R}}^{(A)}\psi = \psi$ for $n \rightarrow +\infty$. Therefore the set $D := \{\psi_n \mid \psi \in \mathcal{H}, n \in \mathbb{N}\}$ is dense in \mathcal{H} . The elements of D are analytic vectors for A as we go to prove. Clearly $\psi_n \in D(A^k)$ since $\mu_{\psi_n, \psi_n}^{(P^{(A)})}(E) = \mu_{\psi, \psi}^{(P^{(A)})}(E \cap [-n, n])$ as immediate consequence of the definition of the measure $\mu_{x,y}^{(P)}$, therefore $\int_{\mathbb{R}} |\lambda^k|^2 d\mu_{\psi_n, \psi_n}^{(P^{(A)})}(\lambda) = \int_{[-n,n]} |\lambda|^{2k} d\mu_{\psi, \psi}^{(P^{(A)})}(\lambda) \leq \int_{[-n,n]} |n|^{2k} d\mu_{\psi, \psi}^{(P^{(A)})}(\lambda) \leq |n|^{2k} \int_{\mathbb{R}} d\mu_{\psi, \psi}^{(P^{(A)})}(\lambda) = |n|^{2k} \|\psi\|^2 < +\infty$. Similarly $\|A^k \psi_n\|^2 = \langle A^k \psi_n | A^k \psi_n \rangle = \langle \psi_n | A^{2k} \psi_n \rangle = \int_{\mathbb{R}} \lambda^{2k} d\mu_{\psi_n, \psi_n}^{(P^{(A)})}(\lambda) \leq |n|^{2k} \|\psi\|^2$. We conclude that $\sum_{k=0}^{+\infty} \frac{(it)^k}{k!} \|A^k \psi_n\|$ converges for every $t \in \mathbb{C}$ as it is dominated by the series $\sum_{k=0}^{+\infty} \frac{|t|^k}{k!} |n|^{2k} \|\psi\|^2 = e^{|t||n|^2} \|\psi\|^2$. ■

3. More Fundamental Quantum Structures

The question we want to answer now is the following:

Is there anything more fundamental behind the phenomenological facts (1), (2), and (3) discussed in the first section and their formalization presented in Sect. 2.8?

An appealing attempt to answer that question and justify the formalism based on the spectral theory is due to von Neumann [7] (and subsequently extended by Birkhoff and von Neumann). This section is devoted to quickly review an elementary part of those ideas, adding however several more modern results (see also [11] for a similar approach).

3.1. The Boolean logic of CM

Consider a classical Hamiltonian system described in symplectic manifold (Γ, ω) , where $\omega = \sum_{k=1}^n dq^k \wedge dp_k$ in any system of local symplectic coordinates $q^1, \dots, q^n, p_1, \dots, p_n$. The state of the system at time t is a point $s \in \Gamma$, in local coordinates $s \equiv (q^1, \dots, q^n, p_1, \dots, p_n)$, whose evolution $\mathbb{R} \ni t \mapsto s(t)$ is a solution of the *Hamiltonian equation* of motion. Always in local symplectic coordinates, they read

$$\frac{dq^k}{dt} = \frac{\partial H(t, q, p)}{\partial p_k}, \quad k = 1, \dots, n \quad (71)$$

$$\frac{dp_k}{dt} = -\frac{\partial H(t, q, p)}{\partial q^k} \quad ;, \quad k = 1, \dots, n, \quad (72)$$

H being the Hamiltonian function of the system, depending on the (inertial) reference frame. Every physical *elementary property*, E , that the system may possess at a certain time t , i.e., which can be true or false at that time, can be identified with a subset $E \subset \Gamma$. The property is true if $s \in E$ and it is not if $s \notin E$. From this point of view, the standard set theory operations $\cap, \cup, \subset, \neg$ (where $\neg E := \Gamma \setminus E$ from now on is the **complement operation**) have a logical interpretation:

- (i) $E \cap F$ corresponds to the property “ E AND F ”,
- (ii) $E \cup F$ corresponds to the property “ E OR F ”,
- (iii) $\neg E$ corresponds to the property “NOT F ”,
- (iv) $E \subset F$ means “ E IMPLIES F ”.

In this context:

- (v) Γ is the property which is always true
- (vi) \emptyset is the property which is always false.

This identification is possible because, as is well known, the logical operations have the same algebraic structure of the set theory operations.

As soon as we admit the possibility to construct statements including *countably*

infinite number of disjunctions or conjunctions, we can enlarge our interpretation towards the abstract measure theory, interpreting the states as *probability Dirac measures* supported on a single point. To this end we first restrict the class of possible elementary properties to the Borel σ -algebra of Γ , $\mathcal{B}(\Gamma)$. For various reasons this class of sets seems to be sufficiently large to describe physics (in particular $\mathcal{B}(\Gamma)$ includes the preimages of measurable sets under continuous functions). A state at time t , $s \in \Gamma$, can be viewed as a Dirac measure, δ_s , supported on s itself. If $E \in \mathcal{B}(\Gamma)$, $\delta_s(E) = 0$ if $s \notin E$ or $\delta_s(E) = 1$ if $s \in E$.

If we do not have a perfect knowledge of the system, as for instance it happens in *statistical mechanics*, the state at time t , μ , is a proper probability measure on $\mathcal{B}(\Gamma)$ which now, is allowed to attain all values of $[0, 1]$. If $E \in \mathcal{B}(\Gamma)$ is an elementary property of the physical system, $\mu(E)$ denotes the probability that the property E is true for the system at time t .

Remark 90. The evolution equation of μ , in statistical mechanics is given by the well-known *Liouville's equation* associate with the Hamiltonian flow. In that case μ is proportional to the natural symplectic volume measure of Γ , $\Omega = \omega \wedge \dots \wedge \omega$ (n -times, where $2n = \dim(\Gamma)$). In fact we have $\mu = \rho\Omega$, where the non-negative function ρ is the so-called **Liouville density** satisfying the famous *Liouville's equation*. In symplectic local coordinates that equation reads

$$\frac{\partial \rho(t, q, p)}{\partial t} + \sum_{k=1}^n \left(\frac{\partial \rho}{\partial q^k} \frac{\partial H}{\partial p_k} - \frac{\partial \rho}{\partial p_k} \frac{\partial H}{\partial q^k} \right) = 0.$$

We shall not deal any further with this equation in this paper. ■

More complicated classical quantities of the system can be described by *Borel measurable* functions $f : \Gamma \rightarrow \mathbb{R}$. Measurability is a good requirement as it permits one to perform physical operations like computing, for instance, the *expectation value* (at a given time) when the state is μ :

$$\langle f \rangle_\mu = \int_\Gamma f \mu.$$

Also elementary properties can be pictured by measurable functions, in fact they are one-to-one identified with all the Borel measurable functions $g : \Gamma \rightarrow \{0, 1\}$. The Borel set E_g associated to g is $g^{-1}(\{1\})$ and in fact $g = \chi_{E_g}$.

A generic physical quantity, a measurable function $f : \Gamma \rightarrow \mathbb{R}$, is completely determined by the class of Borel sets (elementary properties) $E_B^{(f)} := f^{-1}(B)$ where $B \in \mathcal{B}(\mathbb{R})$. The meaning of $E_B^{(f)}$ is

$$E_B^{(f)} = \text{“the value of } f \text{ belongs to } B\text{”} \tag{73}$$

It is possible to prove [6] that the map $\mathcal{B}(\mathbb{R}) \ni B \mapsto E_B^{(f)}$ permits one to reconstruct the function f . The sets $E_B^{(f)} := f^{-1}(B)$ form a σ -algebra as well and the class of sets $E_B^{(f)}$ satisfies the following elementary properties when B ranges in $\mathcal{B}(\mathbb{R})$.

(Fi) $E_{\mathbb{R}}^{(f)} = \Gamma$,

$$\text{(Fii)} \quad E_B^{(f)} \cap E_C^{(f)} = E_{B \cap C}^{(f)},$$

(Fiii) If $N \subset \mathbb{N}$ and $\{B_k\}_{k \in N} \subset \mathcal{B}(\mathbb{R})$ satisfies $B_j \cap B_k = \emptyset$ if $k \neq j$, then

$$\cup_{j \in N} E_{B_j}^{(f)} = E_{\cup_{j \in N} B_j}^{(f)}.$$

These conditions just say that $\mathcal{B}(\mathbb{R}) \ni B \mapsto E_B^{(f)}$ is a **homomorphism of σ -algebras**.

For future convenience we observe that our model of *classical* elementary properties can be also viewed as another mathematical structure, when referring to the notion of *lattice*.

Definition 91. A partially ordered set (X, \geq) is a **lattice** when, for any $a, b \in X$,

(a) $\sup\{a, b\}$ exists, denoted $a \vee b$ (sometimes called ‘join’);

(b) $\inf\{a, b\}$ exists, written $a \wedge b$ (sometimes ‘meet’).

(The partially ordered set is not required to be totally ordered.) ■

Remark 92.

(a) In our considered concrete case $X = \mathcal{B}(\mathbb{R})$ and \geq is nothing but \supset and thus \vee means \cup and \wedge has the meaning of \cap .

(b) In the general case \vee and \wedge turn out to be separately *associative*, therefore it make sense to write $a_1 \vee \cdots \vee a_n$ and $a_1 \wedge \cdots \wedge a_n$ in a lattice. Moreover they are also separately *commutative* so

$$a_1 \vee \cdots \vee a_n = a_{\pi(1)} \vee \cdots \vee a_{\pi(n)} \quad \text{and} \quad a_1 \wedge \cdots \wedge a_n = a_{\pi(1)} \wedge \cdots \wedge a_{\pi(n)}$$

for every permutation $\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$. ■

Definition 93. A lattice (X, \geq) is said to be:

(a) **distributive** if \vee and \wedge distribute over one another: for any $a, b, c \in X$,

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c), \quad a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c);$$

(b) **bounded** if it admits a minimum $\mathbf{0}$ and a maximum $\mathbf{1}$ (sometimes called ‘bottom’ and ‘top’);

(c) **orthocomplemented** if bounded and equipped with a mapping $X \ni a \mapsto \neg a$, where $\neg a$ is the **orthogonal complement** of a , such that:

(i) $a \vee \neg a = \mathbf{1}$ for any $a \in X$,

(ii) $a \wedge \neg a = \mathbf{0}$ for any $a \in X$,

(iii) $\neg(\neg a) = a$ for any $a \in X$,

(iv) $a \geq b$ implies $\neg b \geq \neg a$ for any $a, b \in X$;

(d) **σ -complete**, if every countable set $\{a_n\}_{n \in \mathbb{N}} \subset X$ admits least upper bound $\vee_{n \in \mathbb{N}} a_n$.

A lattice with properties (a), (b) and (c) is called a **Boolean algebra**. A Boolean algebra satisfying (d) is a **Boolean σ -algebra**. ■

Definition 94. If X, Y are lattices, a map $h : X \rightarrow Y$ is a **(lattice) homomorphism** when

$$h(a \vee_X b) = h(a) \vee_Y h(b), \quad h(a \wedge_X b) = h(a) \wedge_Y h(b), \quad a, b \in X$$

(with the obvious notations.) If X and Y are bounded, a homomorphism h is further required to satisfy

$$h(\mathbf{0}_X) = \mathbf{0}_Y, \quad h(\mathbf{1}_X) = \mathbf{1}_Y.$$

If X and Y are orthocomplemented, a homomorphism h also satisfies

$$h(\neg_X a) = \neg_Y h(a).$$

If X, Y are σ -complete, h further fulfills

$$h(\bigvee_{n \in \mathbb{N}} a_n) = \bigvee_{n \in \mathbb{N}} h(a_n), \text{ if } \{a_n\}_{n \in \mathbb{N}} \subset X.$$

In all cases (bounded, orthocomplemented, σ -complete lattices, Boolean (σ -)algebras) if h is bijective it is called **isomorphism** of the relative structures.

It is clear that, just because it is a concrete σ -algebra, the lattice of the elementary properties of a classical system is a lattice which is *distributive*, *bounded* (here $0 = \emptyset$ and $1 = \Gamma$), *orthocomplemented* (the orthocomplement being the complement with respect to Γ) and *σ -complete*. Moreover, as the reader can easily prove, the above map, $\mathcal{B}(\mathbb{R}) \ni B \mapsto E_B^{(f)}$, is also a homomorphism of Boolean σ -algebras.

Remark 95. Given an abstract Boolean σ -algebra X , does there exist a concrete σ -algebra of sets that is isomorphic to the previous one? In this respect the following general result holds, known as *Loomis-Sikorski theorem*.¹ This guarantees that every Boolean σ -algebra is isomorphic to a quotient Boolean σ -algebra Σ/\mathcal{N} , where Σ is a concrete σ -algebra of sets over a measurable space and $\mathcal{N} \subset \Sigma$ is closed under countable unions; moreover, $\emptyset \in \mathcal{N}$ and for any $A \in \Sigma$ with $A \subset N \in \mathcal{N}$, then $A \in \mathcal{N}$. The equivalence relation is $A \sim B$ iff $A \cup B \setminus (A \cap B) \in \mathcal{N}$, for any $A, B \in \Sigma$. It is easy to see the coset space Σ/\mathcal{N} inherits the structure of Boolean σ -algebra from Σ with respect to the (well-defined) partial order relation $[A] \geq [B]$ if $A \supset B$, $A, B \in \Sigma$. ■

3.2. The non-Boolean Logic of QM, the reason why observables are selfadjoint operators.

It is evident that the classical like picture illustrated in Sect. 3.1 is untenable if referring to quantum systems. The deep reason is that there are pair of elementary properties E, F of quantum systems which are incompatible. Here an elementary property is an observable which, if measured by means of a corresponding experimental apparatus, can only attain two values: 0 if it is false or 1 if it is true. For

¹Sikorski S.: *On the representation of Boolean algebras as field of sets*. Fund. Math. **35**, 247-256 (1948).

instance, $E =$ “the component S_x of the electron is $\hbar/2$ ” and $F =$ “the component S_y of the electron is $\hbar/2$ ”. There is no physical instrument capable to establish if E AND F is true or false. We conclude that some of elementary observables of quantum systems cannot be logically combined by the standard operation of the logic. The model of Borel σ -algebra seems not to be appropriate for quantum systems. However one could try to use some form of lattice structure different from the classical one. The fundamental ideas by von Neumann were the following pair.

(vN1) Given a quantum system, there is a complex separable Hilbert space \mathcal{H} such that the **elementary observables** – the ones which only assume values in $\{0, 1\}$ – are one-to-one represented by all the elements of $\mathcal{L}(\mathcal{H})$, the orthogonal projectors in $\mathfrak{B}(\mathcal{H})$.

(vN2) Two elementary observables P, Q are compatible if and only if they commute as projectors.

Remark 96.

(a) As we shall see later (vN1) has to be changed for those quantum systems which admit *superselection rules*. For the moment we stick to the above version of (vN1).

(b) The technical requirement of separability will play a crucial role in several places. ■

Let us analyse the reasons for von Neumann’s postulates. First of all we observe that $\mathcal{L}(\mathcal{H})$ is in fact a lattice if one remembers the relation between orthogonal projectors and closed subspaces stated in Proposition 61.

Notation 97. Referring to Proposition 61, if $P, Q \in \mathcal{L}(\mathcal{H})$, we write $P \geq Q$ if and only if $P(\mathcal{H}) \supset Q(\mathcal{H})$. ■

$P(\mathcal{H}) \supset Q(\mathcal{H})$ is equivalent to $PQ = Q$. Indeed, if $P(\mathcal{H}) \supset Q(\mathcal{H})$ then there is a Hilbert basis of $P(\mathcal{H})$ $N_P = N_Q \cup N'_Q$ where N_Q is a Hilbert basis of $Q(\mathcal{H})$ and N'_Q of $Q(\mathcal{H})^\perp$, the notion of orthogonal being referred to the Hilbert space $P(\mathcal{H})$. From $Q = \sum_{z \in N_Q} \langle z | \cdot \rangle z$ and $P = Q + \sum_{z \in N'_Q} \langle z | \cdot \rangle z$ we have $PQ = Q$. The converse implication is obvious.

As preannounced, it turns out that $(\mathcal{L}(\mathcal{H}), \geq)$ is a lattice and, in particular, it enjoys the following properties (e.g., see [6]) whose proof is direct.

Proposition 98. *Let \mathcal{H} be a complex separable Hilbert space and, if $P \in \mathcal{L}(\mathcal{H})$, define $\neg P := I - P$ (the orthogonal projector onto $P(\mathcal{H})^\perp$). With this definition, $(\mathcal{L}(\mathcal{H}), \geq, \neg)$ turns out to be bounded, orthocomplemented, σ -complete lattice which is not distributive if $\dim(\mathcal{H}) \geq 2$.*

More precisely,

(i) $P \vee Q$ is the orthogonal projector onto $\overline{P(\mathcal{H}) + Q(\mathcal{H})}$.

The analogue holds for a countable set $\{P_n\}_{n \in \mathbb{N}} \subset \mathcal{P}(\mathcal{H})$, $\bigvee_{n \in \mathbb{N}} P_n$ is the orthogonal projector onto $\overline{\bigoplus_{n \in \mathbb{N}} P_n(\mathcal{H})}$.

(ii) $P \wedge Q$ is the orthogonal projector on $P(\mathcal{H}) \cap Q(\mathcal{H})$.

The analogue holds for a countable set $\{P_n\}_{n \in \mathbb{N}} \subset \mathcal{P}(\mathcal{H})$, $\bigwedge_{n \in \mathbb{N}} P_n$ is the orthogonal projector onto $\bigcap_{n \in \mathbb{N}} P_n(\mathcal{H})$.

(iii) The bottom and the top are respectively 0 and I .

Referring to (i) and (ii), it turns out that

$$\bigvee_{n \in \mathbb{N}} P_n = \lim_{k \rightarrow +\infty} \bigvee_{n \leq k} P_n \quad \text{and} \quad \bigwedge_{n \in \mathbb{N}} P_n = \lim_{k \rightarrow +\infty} \bigwedge_{n \leq k} P_n$$

with respect to the strong operator topology.

Remark 99. The fact that the distributive property does not hold is evident from the following elementary counterexample in \mathbb{C}^2 (so that it is valid for every dimension > 1). Let $\{e_1, e_2\}$ be the standard basis of \mathbb{C}^2 and define the subspaces $\mathcal{H}_1 := \text{span}(e_1)$, $\mathcal{H}_2 := \text{span}(e_2)$, $\mathcal{H}_3 := \text{span}(e_1 + e_2)$. Finally P_1, P_2, P_3 respectively denote the orthogonal projectors onto these spaces. By direct inspection one sees that $P_1 \wedge (P_2 \vee P_3) = P_1 \wedge I = P_1$ and $(P_1 \wedge P_2) \vee (P_1 \wedge P_3) = 0 \vee 0 = 0$, so that $P_1 \wedge (P_2 \vee P_3) \neq (P_1 \wedge P_2) \vee (P_1 \wedge P_3)$. ■

The crucial observation is that, nevertheless $(\mathcal{L}(\mathcal{H}), \geq, \neg)$ includes lots of Boolean σ algebras, and precisely the maximal sets of pairwise compatible projectors [6].

Proposition 100. Let \mathcal{H} be a complex separable Hilbert space and consider the lattice $(\mathcal{L}(\mathcal{H}), \geq, \neg)$. If $\mathcal{L}_0 \subset \mathcal{L}(\mathcal{H})$ is a maximal subset of pairwise commuting elements, then \mathcal{L}_0 contains 0, I is \neg -closed and, if equipped with the restriction of the lattice structure of $(\mathcal{L}(\mathcal{H}), \geq, \neg)$, turns out to be a Boolean σ -algebra.

In particular, if $P, Q \in \mathcal{L}_0$,

$$(i) \quad P \vee Q = P + Q - PQ,$$

$$(ii) \quad P \wedge Q = PQ.$$

Proof. \mathcal{L}_0 includes both 0 and I because \mathcal{L}_0 is maximally commutative. Having (i) and (ii), due to (iii) in proposition 98, the sup and the inf of a sequence of projectors of \mathcal{L}_0 commute with the elements of \mathcal{L}_0 , maximality implies that they belong to \mathcal{L}_0 . Finally (i) and (ii) prove by direct inspection that \vee and \wedge are mutually distributive. Let us prove (ii) and (i) to conclude. If $PQ = QP$, PQ is an orthogonal projector and $PQ(\mathcal{H}) = QP(\mathcal{H}) \subset P(\mathcal{H}) \cap Q(\mathcal{H})$. On the other hand, if $x \in P(\mathcal{H}) \cap Q(\mathcal{H})$ then $Px = x$ and $x = Qx$ so that $PQx = x$ and thus $P(\mathcal{H}) \cap Q(\mathcal{H}) \subset PQ(\mathcal{H})$ and (ii) holds. To prove (i) observe that $\overline{\langle P(\mathcal{H}), Q(\mathcal{H}) \rangle}^\perp = P(\mathcal{H})^\perp \cap Q(\mathcal{H})^\perp$. Using (ii), this can be rephrased as $I - P \vee Q = (I - P)(I - Q)$ which entails (i) immediately. ■

Remark 101.

(a) Every set of pairwise commuting orthogonal projectors can be completed to a maximal set as an elementary application of Zorn's lemma. However, since the commutativity property is *not* transitive, there are many possible maximal subsets of pairwise commuting elements in $\mathcal{L}(\mathcal{H})$ with non-empty intersection.

(b) As a consequence of the stated proposition, the symbols \vee , \wedge and \neg have the same properties in \mathcal{L}_0 as the corresponding symbols of classical logic *OR*, *AND* and *NOT*. Moreover $P \geq Q$ can be interpreted as “*Q IMPLIES P*”.

(c) There were and are many attempts to interpret \vee and \wedge as connectives of a new non-distributive logic when dealing with the whole $\mathcal{L}(\mathcal{H})$: a *quantum logic*. The first noticeable proposal was due to Birkhoff and von Neumann [14]. Nowadays there are lots of quantum logics [15,16] all regarded with suspicion by physicists. Indeed, the most difficult issue is the physical operational interpretation of these connectives taking into account the fact that they put together incompatible propositions, which cannot be measured simultaneously. An interesting interpretative attempt, due to Jauch, relies upon a result by von Neumann (e.g., [6])

$$(P \wedge Q)x = \lim_{n \rightarrow +\infty} (PQ)^n x \quad \text{for every } P, Q \in \mathcal{L}(\mathcal{H}) \text{ and } x \in \mathcal{H}.$$

Notice that the result holds in particular if P and Q do not commute, so they are incompatible elementary observables. The right hand side of the identity above can be interpreted as the consecutive and alternated measurement of an infinite sequence of elementary observables P and Q . As

$$\|(P \wedge Q)x\|^2 = \lim_{n \rightarrow +\infty} \|(PQ)^n x\|^2 \quad \text{for every } P, Q \in \mathcal{L}(\mathcal{H}) \text{ and } x \in \mathcal{H},$$

the probability that $P \wedge Q$ is true for a state represented by the unit vector $x \in \mathcal{H}$ is the probability that the infinite sequence of consecutive alternated measurements of P and Q produce is true at each step. ■

We are in a position to clarify why, in this context, observables are PVMs. Exactly as in CM, an observable A is a collection of elementary observables $\{P_E\}_{E \in \mathcal{B}(\mathbb{R})}$ labelled on the Borel sets E of \mathbb{R} . Exactly as for classical quantities, (73) we can say that the meaning of P_E is

$$P_E = \text{“the value of the observable belongs to } E\text{”} \quad (74)$$

We expect that all those elementary observables are pairwise compatible and that they satisfy the same properties (Fi)-(Fiii) as for classical quantities. We can complete $\{P_E\}_{E \in \mathcal{B}(\mathbb{R})}$ to a maximal set of compatible elementary observables. Taking Proposition 100 into account (Fi)-(Fiii) translate into

- (i) $P_{\mathbb{R}} = I$,
- (ii) $P_E P_F = P_{E \cap F}$,
- (iii) If $N \subset \mathbb{N}$ and $\{E_k\}_{k \in N} \subset \mathcal{B}(\mathbb{R})$ satisfies $E_j \cap E_k = \emptyset$ for $k \neq j$, then

$$\sum_{j \in N} P_{E_j} x = P_{\cup_{j \in N} E_j} x \quad \text{for every } x \in \mathcal{H}.$$

(The presence of x is due to the fact that the convergence of the series if N is infinite is in the strong operator topology as declared in the last statement of Proposition 98.) *In other words we have just found Definition 62, specialized to PVM on \mathbb{R} : Observables in QM are PVM over \mathbb{R} !*

We know that all PVM over \mathbb{R} are one-to-one associated to all selfadjoint operators in view of the results presented in the previous section (see (e) in remark 75). We conclude that, adopting von Neumann's framework, in QM observables are naturally described by selfadjoint operators, whose spectra coincide with the set of values attained by the observables.

3.3. Recovering the Hilbert space structure

A reasonable question to ask is whether there are better reasons for choosing to describe quantum systems via a lattice of orthogonal projectors, other than the kill-off argument "it works". To tackle the problem we start by listing special properties of the lattice of orthogonal projectors, whose proof is elementary.

Theorem 102. *The bounded, orthocomplemented, σ -complete lattice $\mathcal{L}(\mathcal{H})$ of Propositions 98 and 100 satisfies these additional properties:*

(i) **separability** (for \mathcal{H} separable): *if $\{P_a\}_{a \in A} \subset \mathcal{L}(\mathcal{H})$ satisfies $P_i P_j = 0$, $i \neq j$, then A is at most countable;*

(ii) **atomicity and atomisticity**: *there exist elements in $A \in \mathcal{L}(\mathcal{H}) \setminus \{0\}$, called **atoms**, for which $0 \leq P \leq A$ implies $P = 0$ or $P = A$; for any $P \in \mathcal{L}(\mathcal{H}) \setminus \{0\}$ there exists an atom A with $A \leq P$ ($\mathcal{L}(\mathcal{H})$ is then called **atomic**); For every $P \in \mathcal{L}(\mathcal{H}) \setminus \{0\}$, P is the sup of the set of atoms $A \leq P$ ($\mathcal{L}(\mathcal{H})$ is then called **atomistic**);*

(iii) **orthomodularity**: *$P \leq Q$ implies $Q = P \vee ((\neg P) \wedge Q)$;*

(iv) **covering property**: *if $A, P \in \mathcal{L}(\mathcal{H})$, with A an atom, satisfy $A \wedge P = 0$, then (1) $P \leq A \vee P$ with $P \neq A \vee P$, and (2) $P \leq Q \leq A \vee P$ implies $Q = P$ or $Q = A \vee P$;*

(v) **irreducibility**: *only 0 and I commute with every element of $\mathcal{L}(\mathcal{H})$.*

The orthogonal projectors onto one-dimensional spaces are the only atoms of $\mathcal{L}(\mathcal{H})$.

Irreducibility can easily be proved observing that if $P \in \mathcal{L}(\mathcal{H})$ commutes with all projectors along one-dimensional subspaces, $Px = \lambda_x x$ for every $x \in \mathcal{H}$. Thus $P(x + y) = \lambda_{x+y}(x + y)$ but also $Px + Py = \lambda_x x + \lambda_y y$ and thus $(\lambda_x - \lambda_{x+y})x = (\lambda_{x+y} - \lambda_y)y$, which entails $\lambda_x = \lambda_y$ if $x \perp y$. If $N \subset \mathcal{H}$ is a Hilbert basis, $Pz = \sum_{x \in N} \langle x|z \rangle \lambda_x = \lambda z$ for some fixed $\lambda \in \mathbb{C}$. Since $P = P^* = PP$, we conclude that either $\lambda = 0$ or $\lambda = 1$, i.e., either $P = 0$ or $P = I$, as wanted. Orthomodularity is a weaker version of distributivity of \vee with respect to \wedge that we know to be untenable in $\mathcal{P}(\mathcal{H})$.

Actually each of the listed properties admits a physical operational interpretation (e.g. see [15].) So, based on the experimental evidence of quantum systems, we could try to prove, in the absence of any Hilbert space, that elementary propositions with

experimental outcome in $\{0, 1\}$ form a poset. More precisely, we could attempt to find a bounded, orthocomplemented σ -complete lattice that verifies conditions (i)–(v) above, and then prove this lattice is described by the orthogonal projectors of a Hilbert space.

The partial order relation of elementary propositions can be defined in various ways. But it will always correspond to the logical implication, in some way or another. Starting from [17] a number of approaches (either of essentially physical nature, or of formal character) have been developed to this end: in particular, those making use of the notion of (quantum) *state*, which we will see in a short while for the concrete case of propositions represented by orthogonal projectors. The object of the theory is now [17] the pair $(\mathcal{O}, \mathcal{S})$, where \mathcal{O} is the class of observables and \mathcal{S} the one of states. The elementary propositions form a subclass \mathcal{L} of \mathcal{O} equipped with a natural poset structure (\mathcal{L}, \leq) (also satisfying a weaker version of some of the conditions (i)–(v)). A state $s \in \mathcal{S}$, in particular, defines the probability $m_s(P)$ that P is true for every $P \in \mathcal{L}$ [17]. As a matter of fact, if $P, Q \in \mathcal{L}$, $P \leq Q$ means by definition that the probability $m_s(P) \leq m_s(Q)$ for every state $s \in \mathcal{S}$. More difficult is to justify that the poset thus obtained is a lattice, i.e. that it admits a greatest lower bound $P \vee Q$ and a least upper bound $P \wedge Q$ for every P, Q . There are several proposals, very different in nature, to introduce this lattice structure (see [15] and [16] for a general treatise) and make the physical meaning explicit in terms of measurement outcome. See Aerts in [16] for an abstract but operational viewpoint and [15, §21.1] for a summary on several possible ways to introduce the lattice structure on the partially ordered set of abstract elementary propositions \mathcal{L} . If we accept the lattice structure on elementary propositions of a quantum system, then we may define the operation of orthocomplementation by the familiar logical/physical negation. Compatible propositions can then be defined in terms of commuting propositions, i.e. commuting elements of a orthocomplemented lattice as follows.

Definition 103. Let $(\mathcal{L}, \geq, \neg)$ an orthocomplemented lattice. Two elements $a, b \in \mathcal{L}$ are said to be:

orthogonal written $a \perp b$, if $\neg a \geq b$ (or equivalently $\neg b \geq a$);

commuting, if $a = c_1 \vee c_3$ and $b = c_2 \vee c_3$ with $c_i \perp c_j$ if $i \neq j$. ■

These notions of orthogonality and compatibility make sense because, *a posteriori*, they turn out to be the usual ones when propositions are interpreted via projectors. As the reader may easily prove, two elements $P, Q \in \mathcal{L}(\mathcal{H})$ are orthogonal in accordance with Definition 103 if and only if $PQ = QP = 0$ (in other words they project onto mutually orthogonal subspaces), and commute in accordance with Definition 103 if and only if $PQ = QP$. (If $P = P_1 + P_3$ and $Q = P_2 + P_3$ where the orthogonal projectors satisfy $P_i \perp P_j = 0$ for $i \neq j$, we trivially have $PQ = QP$. If conversely, $PQ = QP$, the said decomposition arises for $P_3 := PQ$, $P_1 := P(I - Q)$, $P_2 := Q(I - P)$.)

Now fully-fledged with an orthocomplemented lattice and the notion of compati-

ble propositions, we can attach a physical meaning (an interpretation backed by experimental evidence) to the requests that the lattice be orthocomplemented, complete, atomistic, irreducible and that it have the covering property [15]. Under these hypotheses and assuming there exist at least 4 pairwise-orthogonal atoms, Piron ([18,19], [15, §21], Aerts in [16]) used projective geometry techniques to show that the lattice of quantum propositions can be canonically identified with the closed (in a generalised sense) subsets of a generalised Hilbert space of sorts. In the latter: (a) the field is replaced by a division ring (usually not commutative) equipped with an involution, and (b) there exists a certain non-singular Hermitian form associated with the involution. It has been conjectured by many people (see [15]) that if the lattice is also orthomodular and separable, the division ring can only be picked among \mathbb{R} , \mathbb{C} or \mathbb{H} (quaternion algebra). More recently Solèr^m, Hollandⁿ and Aerts–van Steirteghem^o have found sufficient hypotheses, in terms of the existence of infinite orthogonal systems, for this to happen. Under these hypotheses, if the ring is \mathbb{R} or \mathbb{C} , we obtain precisely the lattice of orthogonal projectors of the separable Hilbert space. In the case of \mathbb{H} , one gets a similar generalised structure. In all these arguments the assumption of irreducibility is not really crucial: if property (v) fails, the lattice can be split into irreducible sublattices [20,15]. Physically-speaking this situation is natural in the presence of *superselection rules*, of which more soon. It is worth stressing that the covering property in Theorem 102 is a crucial property. Indeed there are other lattices relevant in physics verifying all the remaining properties in the afore-mentioned theorem. Remarkably the family of the so-called *causally closed sets* in a general spacetime satisfies all the said properties but the covering one^p. This obstruction prevents one from endowing a spacetime with a natural (generalized) Hilbert space structure, while it suggests some ideas towards a formulation of quantum gravity.

3.4. States as measures on $\mathcal{L}(\mathcal{H})$: Gleason's Theorem

Let us introduce an important family of operators. This family will play a decisive rôle in the issue concerning a possible justification of the fact that quantum states are elements of the projective space $P\mathcal{H}$.

^mSolèr, M. P.: Characterization of Hilbert spaces by orthomodular spaces. *Communications in Algebra*, **23**, 219-243 (1995).

ⁿHolland, S.S.: Orthomodularity in infinite dimensions; a theorem of M. Solèr. *Bulletin of the American Mathematical Society*, **32**, 205-234, (1995).

^oAerts, D., van Steirteghem B.: Quantum Axiomatics and a theorem of M.P. Solèr. *International Journal of Theoretical Physics*. **39**, 497-502, (2000).

^pSee H. Casini, *The logic of causally closed spacetime subsets*, *Class.Quant.Grav.* **19**, 2002, 6389-6404

3.4.1. Trace class operators

Definition 104. If \mathcal{H} is a complex Hilbert space, $\mathfrak{B}_1(\mathcal{H}) \subset \mathfrak{B}(\mathcal{H})$ denotes the set of **trace class** or **nuclear** operators, i.e. the operators $T \in \mathfrak{B}(\mathcal{H})$ satisfying

$$\sum_{z \in N} \langle z | |T| z \rangle < +\infty \quad (75)$$

for some Hilbertian basis $N \in \mathcal{H}$ and where $|T| := \sqrt{T^*T}$ defined via functional calculus. \blacksquare

Remark 105. Notice that, above, T^*T is selfadjoint and $\sigma(T^*T) \in [0, +\infty)$ because of exercise 73 so that $\sqrt{T^*T}$ is well defined as a function of T^*T . \blacksquare

Trace class operators enjoy several remarkable properties [6]. Here we only mention the ones relevant for these lecture notes.

Proposition 106. *Let \mathcal{H} a complex Hilbert space, $\mathfrak{B}_1(\mathcal{H})$ satisfy the following properties.*

(a) *If $T \in \mathfrak{B}_1(\mathcal{H})$ and $N \subset \mathcal{H}$ is any Hilbertian basis, then (75) holds and thus*

$$\|T\|_1 := \sum_{z \in N} \langle z | |T| z \rangle$$

is well defined.

(b) $\mathfrak{B}_1(\mathcal{H})$ *is a subspace of $\mathfrak{B}(\mathcal{H})$ which is moreover a two-sided *-ideal, namely*

(i) $AT, TA \in \mathfrak{B}_1(\mathcal{H})$ *if $T \in \mathfrak{B}_1(\mathcal{H})$ and $A \in \mathfrak{B}(\mathcal{H})$,*

(ii) $T^* \in \mathfrak{B}_1(\mathcal{H})$ *if $T \in \mathfrak{B}_1(\mathcal{H})$.*

(c) $\|\cdot\|_1$ *is a norm on $\mathfrak{B}_1(\mathcal{H})$ making it a Banach space and satisfying*

(i) $\|TA\|_1 \leq \|A\| \|T\|_1$ *and $\|AT\|_1 \leq \|A\| \|T\|_1$ if $T \in \mathfrak{B}_1(\mathcal{H})$ and $A \in \mathfrak{B}(\mathcal{H})$,*

(ii) $\|T\|_1 = \|T^*\|_1$ *if $T \in \mathfrak{B}_1(\mathcal{H})$.*

(d) *If $T \in \mathfrak{B}_1(\mathcal{H})$, the **trace** of T ,*

$$\text{tr } T := \sum_{z \in N} \langle z | T z \rangle \in \mathbb{C}$$

is well defined, does not depend on the choice of the Hilbertian basis N and the sum converges absolutely (so can be arbitrarily re-ordered).

Remark 107.

(1) Obviously we have $\text{tr } |T| = \|T\|_1$ if $T \in \mathfrak{B}_1(\mathcal{H})$.

(2) The trace just possesses the properties one expects from the finite dimensional case. In particular, [6],

(i) it is linear on $\mathfrak{B}_1(\mathcal{H})$,

(ii) $\text{tr } T^* = \overline{\text{tr } T}$ if $T \in \mathfrak{B}_1(\mathcal{H})$,

(iii) the trace satisfies the **cyclic property**,

$$\text{tr}(T_1 \cdots T_n) = \text{tr}(T_{\pi(1)} \cdots T_{\pi(n)}) \quad (76)$$

if at least one of the T_k belongs to $\mathfrak{B}_1(\mathcal{H})$, the remaining ones are in $\mathfrak{B}(\mathcal{H})$, and $\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ is a *cyclic* permutation. \blacksquare

The trace of $T \in \mathfrak{B}_1(\mathcal{H})$ can be computed on a basis of eigenvectors in view of the following further result [6]. Actually (d) and (e) easily follow from (a),(b),(c), (d), and the spectral theory previously developed.

Proposition 108. *Let \mathcal{H} a complex Hilbert space and $T^* = T \in \mathfrak{B}_1(\mathcal{H})$. The following facts hold.*

- (a) $\sigma(T) \setminus \{0\} = \sigma_p(T) \setminus \{0\}$. If $0 \in \sigma(T)$ it may be either the unique element of $\sigma_c(T)$ or an element of $\sigma_p(T)$.
- (b) Every eigenspace \mathcal{H}_λ has finite dimension d_λ provided $\lambda \neq 0$.
- (c) $\sigma_p(T)$ is made of at most countable number of reals such that
 - (i) 0 is unique possible accumulation point,
 - (ii) $\|T\| = \max_{\lambda \in \sigma_p(T)} |\lambda|$.
- (d) There is a Hilbert basis of eigenvectors $\{x_{\lambda,a}\}_{\lambda \in \sigma_p(T), a=1,2,\dots,d_\lambda}$ (d_0 may be infinite) and

$$\operatorname{tr}(T) = \sum_{\lambda \in \sigma_p(T)} d_\lambda \lambda,$$

where the sum converges absolutely (and thus can be arbitrarily re-ordered).

(e) Referring to the basis presented in (d), the spectral decomposition of T reads

$$T = \sum_{\lambda \in \sigma_p(T)} \lambda P_\lambda$$

where $P_\lambda = \sum_{a=1,2,\dots,d_\lambda} \langle x_{\lambda,s} | \cdot \rangle x_{\lambda,a}$ and the sum is computed in the strong operator topology and can be re-ordered arbitrarily. The convergence holds in the uniform topology too if the set of eigenspaces are suitably ordered in the count.

Corollary 109. $\operatorname{tr} : \mathfrak{B}_1(\mathcal{H}) \rightarrow \mathbb{C}$ is continuous with respect to the norm $\|\cdot\|_1$ because $|\operatorname{tr} T| \leq \operatorname{tr}|T| = \|T\|_1$ if $T \in \mathfrak{B}_1(\mathcal{H})$.

Proof. If $T \in \mathfrak{B}(\mathcal{H})$, we have the polar decomposition $T = U|T|$ (see, e.g., [6]) where $U \in \mathfrak{B}(\mathcal{H})$ is isometric on $\operatorname{Ker}(T)^\perp$ and $\operatorname{Ker}(U) = \operatorname{Ker}(T) = \operatorname{Ker}(|T|)$ so that, in particular $\|U\| \leq 1$. Let N be a Hilbertian basis of \mathcal{H} made of eigenvectors of $|T|$ (it exists for the previous theorem since $|T|$ is trace class). We have

$$|\operatorname{tr} T| = \left| \sum_{u \in N} \langle u | U |T| u \rangle \right| = \left| \sum_{u \in N} \langle u | U u \rangle \lambda_u \right| \leq \sum_{u \in N} |\lambda_u| |\langle u | U u \rangle|.$$

Next observe that $|\lambda_u| = \lambda_u$ because $|T| \geq 0$ and $|\langle u | U u \rangle| \leq \|u\| \|U u\| \leq 1 \|U u\| \leq \|u\| = 1$ and thus, $|\operatorname{tr} T| \leq \sum_{u \in N} \lambda_u = \sum_{u \in N} \langle u | |T| u \rangle = \operatorname{tr}|T| = \|T\|_1$. \square

3.4.2. The notion of quantum state and the crucial theorem by Gleason

As commented in (a) in remark 84, the probabilistic interpretation of quantum states is not well defined because there is no a true probability measure in view of the fact that there are incompatible observables. The idea is to re-define the notion of probability in the bounded, orthocomplemented, σ -complete lattice like $\mathcal{L}(\mathcal{H})$ instead of on a σ -algebra. Exactly as in CM, where the generic states are probability measures on Boolean lattice $\mathcal{B}(\Gamma)$ of the elementary properties of the system (Sect. 3.1), we can think of states of a quantum system as σ -additive probability measures over the non-Boolean lattice of the elementary observables $\mathcal{L}(\mathcal{H})$.

Definition 110. Let \mathcal{H} be a complex Hilbert space. A **quantum state** in \mathcal{H} is a map $\rho : \mathcal{L}(\mathcal{H}) \rightarrow [0, 1]$ such that the following requirements are satisfied.

(1) $\rho(I) = 1$.

(2) If $\{P_n\}_{n \in N} \subset \mathcal{L}(\mathcal{H})$, for N at most countable satisfies $P_k(\mathcal{H}) \perp P_h(\mathcal{H})$ when $h \neq k$ for $h, k \in N$, then

$$\rho(\bigvee_{k \in N} P_k) = \sum_{k \in N} \rho(P_k). \quad (77)$$

The set of the states in \mathcal{H} will be denoted by $\mathfrak{S}(\mathcal{H})$. ■

Remark 111.

(a) The condition $P_k(\mathcal{H}) \perp P_h(\mathcal{H})$ is obviously equivalent to $P_k P_h = 0$. Since (taking the adjoint) we also obtain $P_h P_k = 0 = P_k P_h$, we conclude that we are dealing with pairwise compatible elementary observables. Therefore Proposition 100 permits us to equivalently re-write the σ -additivity (2) as follows.

(2) If $\{P_n\}_{n \in N} \subset \mathcal{L}(\mathcal{H})$, for N at most countable satisfies $P_k P_h = 0$ when $h \neq k$ for $h, k \in N$, then

$$\rho\left(\sum_{k \in N} P_k\right) = \sum_{k \in N} \rho(P_k), \quad (78)$$

the sum on the left hand side being computed with respect to the strong operator topology if N is infinite.

(b) Requirement (2), taking (1) into account implies $\rho(0) = 0$.

(c) Quantum states do exist. It is immediately proved that, in fact, $\psi \in \mathcal{H}$ with $\|\psi\| = 1$ defines a quantum state ρ_ψ as

$$\rho_\psi(P) = \langle \psi | P \psi \rangle \quad P \in \mathcal{L}(\mathcal{H}). \quad (79)$$

This is in nice agreement with what we already know and proves that these types of quantum states are one-to-one with the elements of $P\mathcal{H}$ as well known.

However these states do not exhaust $\mathfrak{S}(\mathcal{H})$. In fact, it immediately arises from Definition 110 that the set of the states is convex: If $\rho_1, \dots, \rho_n \in \mathfrak{S}(\mathcal{H})$ then $\sum_{j=1}^n p_j \rho_j \in \mathfrak{S}(\mathcal{H})$ if $p_j \geq 0$ and $\sum_{k=1}^n p_k = 1$. These convex combinations of states generally do not have the form ρ_ψ .

(d) Restricting ourselves to a maximal set \mathcal{L}_0 of pairwise commuting projectors, which in view of Proposition 100 has the abstract structure of a σ -algebra, a quantum state ρ reduces thereon to a standard probability measure. In this sense the “quantum probability” we are considering extends the classical notion. Differences show up just when one deals with conditional probability involving incompatible elementary observables. ■

An interesting case of (c) in the remark above is a convex combination of states induced by unit vectors as in (79), where $\langle \psi_k | \psi_h \rangle = \delta_{hk}$,

$$\rho = \sum_{k=1}^n p_k \rho_{\psi_k} .$$

By direct inspection, completing the finite orthonormal system $\{\psi_k\}_{k=1,\dots,n}$ to a full Hilbertian basis of \mathcal{H} , one quickly proves that, defining

$$T = \sum_{k=1}^n p_k \langle \psi_k | \cdot \rangle \psi_k \tag{80}$$

$\rho(P)$ can be computed as

$$\rho(P) = \text{tr}(TP) \quad P \in \mathcal{L}(\mathcal{H})$$

In particular it turns out that T is in $\mathfrak{B}_1(\mathcal{H})$, satisfies $T \geq 0$ (so it is selfadjoint for (3) in exercise 46) and $\text{tr} T = 1$. As a matter of fact, (80) is just the spectral decomposition of T , whose spectrum is $\{p_k\}_{k=1,\dots,n}$. This result is general [6]

Proposition 112. *Let \mathcal{H} be a complex Hilbert space and let $T \in \mathfrak{B}_1(\mathcal{H})$ satisfy $T \geq 0$ and $\text{Tr} T = 1$, then the map*

$$\rho_T : \mathcal{L}(\mathcal{H}) \ni P \mapsto \text{tr}(TP)$$

is well defined and $\rho_T \in \mathfrak{S}(\mathcal{H})$.

The very remarkable fact is that these operators exhaust $\mathfrak{S}(\mathcal{H})$ if \mathcal{H} is separable with dimension $\neq 2$, as established by Gleason in a celebrated theorem we restate re-adapting it to these lecture notes (see [6] for the original statement and [21] for a general treatise on the subject).

Theorem 113 (Gleason’s Theorem). *Let \mathcal{H} be a complex Hilbert space of finite dimension $\neq 2$, or infinite-dimensional and separable. If $\rho \in \mathfrak{S}(\mathcal{H})$ there exists a unique operator $T \in \mathfrak{B}_1(\mathcal{H})$ with $T \geq 0$ and $\text{tr} T = 1$ such that $\text{tr}(TP) = \rho(P)$ for every $P \in \mathcal{L}(\mathcal{H})$.*

Concerning the existence of T , Gleason’s proof works for real Hilbert spaces too. If the Hilbert space is complex, the operator T associated to ρ is unique for the following reason. Any other T' of trace class such that $\rho(P) = \text{tr}(T'P)$ for any $P \in \mathcal{L}(\mathcal{H})$ must also satisfy $\langle x | (T - T')x \rangle = 0$ for any $x \in \mathcal{H}$. If $x = 0$ this is clear, while if $x \neq 0$ we may complete the vector $x/\|x\|$ to a basis, in which

$\text{tr}((T - T')P_x) = 0$ reads $\|x\|^{-2}\langle x|(T - T')x\rangle = 0$, where P_x is the projector onto $\text{span}(x)$. By (3) in exercise 46, we obtain $T - T' = 0^{\mathfrak{q}}$.

Remark 114.

(a) Imposing $\dim \mathcal{H} \neq 2$ is mandatory, a well known counterexample can be found, e.g. in [6].

(b) Particles with spin 1/2, like electrons, admit a Hilbert space – in which the observable spin is defined – of dimension 2. The same occurs to the Hilbert space in which the polarisation of light is described (cf. helicity of photons). When these systems are described in full, however, for instance including degrees of freedom relative to position or momentum, they are representable on a separable Hilbert space of infinite dimension.

(c) Gleason’s characterization of states has an important consequence known as the *Kochen-Specker theorem*. It proves that in QM there are no states assigning probability 1 to some elementary observables and 0 to the remaining ones, differently to what happens in CM.

Theorem 115 (Kochen-Specker Theorem). *Let \mathcal{H} be a complex Hilbert space of finite dimension $\neq 2$, or infinite-dimensional and separable. There is no quantum state $\rho : \mathcal{L}(\mathcal{H}) \rightarrow [0, 1]$, in the sense of Def. 110, such that $\rho(\mathcal{L}(\mathcal{H})) = \{0, 1\}$*

Proof. Define $\mathbb{S} := \{x \in \mathcal{H} \mid \|x\| = 1\}$ endowed with the topology induced by \mathcal{H} , and let $T \in \mathfrak{B}_1(\mathcal{H})$ be the representative of ρ using Gleason’s theorem. The map $f_\rho : \mathbb{S} \ni x \mapsto \langle x|Tx\rangle = \rho(\langle x| \cdot \rangle x) \in \mathbb{C}$ is continuous because T is bounded. We have $f_\rho(\mathbb{S}) \subset \{0, 1\}$, where $\{0, 1\}$ is equipped with the topology induced by \mathbb{C} . Since \mathbb{S} is connected its image must be connected also. So either $f_\rho(\mathbb{S}) = \{0\}$ or $f_\rho(\mathbb{S}) = \{1\}$. In the first case $T = 0$ which is impossible because $\text{tr}T = 1$, in the second case $\text{tr}T \neq 2$ which is similarly impossible. □

This negative result produces no-go theorems in some attempts to explain QM in terms of CM introducing *hidden variables* [3]. ■

Remark 116. *In view of Proposition 112 and Theorem 113, assuming that \mathcal{H} has finite dimension or is separable, we henceforth identify $\mathfrak{S}(\mathcal{H})$ with the subset of $\mathfrak{B}_1(\mathcal{H})$ of positive operators with unit trace. We simply disregard the states in \mathcal{H} with dimension 2 which are not of this form especially taking (b) in remark 114 into account.* ■

We are in a position to state some definitions of interest for physicists, especially the distinction between pure and mixed states, so we proceed to analyse the structure of

^qIn a real Hilbert space $\langle x|Ax\rangle = 0$ for all x does not imply $A = 0$. Think of real anti symmetric matrices in \mathbb{R}^n equipped with the standard scalar product.

the space of the states. To this end, we remind the reader that, if C is a convex set in a vector space, $e \in C$ is called **extreme** if it cannot be written as $e = \lambda x + (1 - \lambda)y$, with $\lambda \in (0, 1)$, $x, y \in C \setminus \{e\}$.

We have the following simple result whose proof can be found in [6].

Proposition 117. *Let \mathcal{H} be a complex separable Hilbert space.*

(a) $\mathfrak{S}(\mathcal{H})$ is a convex closed subset in $\mathfrak{B}_1(\mathcal{H})$ whose extreme points are those of the form: $\rho_\psi := \langle \psi | \cdot \rangle \psi$ for every vector $\psi \in \mathcal{H}$ with $\|\psi\| = 1$. (This sets up a bijection between extreme states and elements of $P\mathcal{H}$.)

(b) A state $\rho \in \mathfrak{S}(\mathcal{H})$ is extreme if and only if $\rho\rho = \rho$. (All the elements of $\mathfrak{S}(\mathcal{H})$ however satisfy $\langle x | \rho\rho x \rangle \leq \langle x | \rho x \rangle$ for all $x \in \mathcal{H}$.)

(c) Any state $\rho \in \mathfrak{S}(\mathcal{H})$ is a linear combination of extreme states, including infinite combinations in the strong operator topology. In particular there is always a decomposition

$$\rho = \sum_{\phi \in N} p_\phi \langle \phi | \cdot \rangle \phi,$$

where N is an eigenvector basis for ρ , $p_\phi \in [0, 1]$ for any $\phi \in N$, and

$$\sum_{\phi \in N} p_\phi = 1.$$

The stated proposition allows us to introduce some notions and terminology relevant in physics. First of all, extreme elements in $\mathfrak{S}(\mathcal{H})$ are usually called **pure states** by physicists. We shall denote their set is denoted $\mathfrak{S}_p(\mathcal{H})$. Non-extreme states are instead called **mixed states**, **mixtures** or **non-pure states**. If

$$\psi = \sum_{i \in I} a_i \phi_i,$$

with I finite or countable (and the series converges in the topology of \mathcal{H} in the second case), where the vectors $\phi_i \in \mathcal{H}$ are all non-null and $0 \neq a_i \in \mathbb{C}$, physicists say that the state $\langle \psi | \cdot \rangle \psi$ is called an **coherent superposition** of the states $\langle \phi_i | \cdot \rangle \phi_i / \|\phi_i\|^2$.

The possibility of creating pure states by non-trivial combinations of vectors associated to other pure states is called, in the jargon of QM, **superposition principle of (pure) states**

There is however another type of superposition of states. If $\rho \in \mathfrak{S}(\mathcal{H})$ satisfies:

$$\rho = \sum_{i \in I} p_i \rho_i$$

with I finite, $\rho_i \in \mathfrak{S}(\mathcal{H})$, $0 \neq p_i \in [0, 1]$ for any $i \in I$, and $\sum_i p_i = 1$, the state ρ is called **incoherent superposition** of states ρ_i (possibly pure).

If $\psi, \phi \in \mathcal{H}$ satisfy $\|\psi\| = \|\phi\| = 1$ the following terminology is very popular: The complex number $\langle \psi | \phi \rangle$ is the **transition amplitude** or **probability amplitude** of the state $\langle \phi | \cdot \rangle \phi$ on the state $\langle \psi | \cdot \rangle \psi$, moreover the non-negative real number $|\langle \psi | \phi \rangle|^2$ is the **transition probability** of the state $\langle \phi | \cdot \rangle \phi$ on the state $\langle \psi | \cdot \rangle \psi$.

We make some comments about these notions. Consider the pure state $\rho_\psi \in \mathfrak{S}_p(\mathcal{H})$, written $\rho_\psi = \langle \psi | \cdot \rangle \psi$ for some $\psi \in \mathcal{H}$ with $\|\psi\| = 1$. What we want to emphasise is that this pure state is also an orthogonal projector $P_\psi := \langle \psi | \cdot \rangle \psi$, so it must correspond to an elementary observable of the system (an *atom* using the terminology of Theorem 102). The naïve and natural interpretation^r of that observable is this: “*the system’s state is the pure state given by the vector ψ* ”. We can therefore interpret the square modulus of the transition amplitude $\langle \phi | \psi \rangle$ as follows. If $\|\phi\| = \|\psi\| = 1$, as the definition of transition amplitude imposes, $\text{tr}(\rho_\psi P_\phi) = |\langle \phi | \psi \rangle|^2$, where $\rho_\psi := \langle \psi | \cdot \rangle \psi$ and $P_\phi = \langle \phi | \cdot \rangle \phi$. Using (4) we conclude:

$|\langle \phi | \psi \rangle|^2$ is the probability that the state, given (at time t) by the vector ψ , following a measurement (at time t) on the system becomes determined by ϕ .

Notice $|\langle \phi | \psi \rangle|^2 = |\langle \psi | \phi \rangle|^2$, so the probability transition of the state determined by ψ on the state determined by ϕ coincides with the analogous probability where the vectors are swapped. This fact is, *a priori*, highly non-evident in physics.

Since we have introduced a new notion of state the axiom concerning the collapse of the state (Sect. 2.8) must be improved in order to encompass all states of $\mathfrak{S}(\mathcal{H})$. The standard formulation of QM assumes the following axiom (introduced by von Neumann and generalised by Lüders) about what occurs to the physical system, in state $\rho \in \mathfrak{S}(\mathcal{H})$ at time t , when subjected to the measurement of an elementary observable $P \in \mathcal{L}(\mathcal{H})$, if the latter is true (so in particular $\text{tr}(\rho P) > 0$, prior to the measurement). We are referring to *non-destructive* testing, also known as *indirect measurement* or *first-kind measurement*, where the physical system examined (typically a particle) is not absorbed/annihilated by the instrument. They are idealised versions of the actual processes used in labs, and only in part they can be modelled in such a way.

Collapse of the state revisited. If the quantum system is in state $\rho \in \mathfrak{S}(\mathcal{H})$ at time t and proposition $P \in \mathcal{L}(\mathcal{H})$ is true after a measurement at time t , the system’s state immediately afterwards is:

$$\rho_P := \frac{P\rho P}{\text{tr}(\rho P)}.$$

In particular, if ρ is pure and determined by the unit vector ψ , the state immediately after measurement is still pure, and determined by:

$$\psi_P = \frac{P\psi}{\|P\psi\|}.$$

Obviously, in either case ρ_P and ψ_P define states. In the former, in fact, ρ_P is

^rWe cannot but notice how this interpretation muddles the semantic and syntactic levels. Although this could be problematic in a formulation within formal logic, the use physicists make of the interpretation eschews the issue.

positive of trace class, with unit trace, while in the latter $\|\psi_P\| = 1$.

Remark 118.

(a) Measuring a property of a physical quantity goes through the interaction between the system and an instrument (supposed to be macroscopic and obeying the laws of classical physics). Quantum Mechanics, in its standard formulation, does not establish what a measuring instrument is, it only says they exist; nor is it capable of describing the interaction of instrument and quantum system set out in the von Neumann Lüders' postulate quoted above. Several viewpoints and conjectures exist on how to complete the physical description of the measuring process; these are called, in the slang of QM, **collapse**, or **reduction, of the state** or **of the wavefunction** (see [6] for references).

(b) Measuring instruments are commonly employed to *prepare a system in a certain pure state*. Theoretically-speaking the preparation of a *pure* state is carried out like this. A finite collection of *compatible* propositions P_1, \dots, P_n is chosen so that the projection subspace of $P_1 \wedge \dots \wedge P_n = P_1 \cdots P_n$ is *one-dimensional*. In other words $P_1 \cdots P_n = (\psi |)\psi$ for some vector with $\|\psi\| = 1$. The existence of such propositions is seen in practically all quantum systems used in experiments. (From a theoretical point of view these are *atomic* propositions) Then propositions P_i are simultaneously measured on several identical copies of the physical system of concern (e.g., electrons), whose initial states, though, are unknown. If for one system the measurements of all propositions are successful, the post-measurement state is determined by the vector ψ , and the system was **prepared** in that particular pure state.

Normally each projector P_i belongs to the PVM $P^{(A)}$ of an observable A_i whose spectrum is made of isolated points (thus a pure point spectrum) and $P_i = P_{\{\lambda_i\}}^{(A)}$ with $\lambda_i \in \sigma_p(A_i)$.

(c) Let us finally explain how to practically obtain non-pure states from pure ones. Consider q_1 identical copies of system S prepared in the pure state associated to ψ_1 , q_2 copies of S prepared in the pure state associated to ψ_2 and so on, up to ψ_n . If we mix these states each one will be in the non-pure state: $\rho = \sum_{i=1}^n p_i \langle \psi_i |)\psi_i$, where $p_i := q_i / \sum_{i=1}^n q_i$. In general, $\langle \psi_i | \psi_j \rangle$ is not zero if $i \neq j$, so the above expression for ρ is not the decomposition with respect to an eigenvector basis for ρ . This procedure hints at the existence of two different types of probability, one intrinsic and due to the quantum nature of state ψ_i , the other epistemic, and encoded in the probability p_i . But this is not true: once a non-pure state has been created, as above, there is no way, within QM, to distinguish the states forming the mixture. For example, the same ρ could have been obtained mixing other pure states than those determined by the ψ_i . In particular, one could have used those in the decomposition of ρ into a basis of its eigenvectors. For physics, no kind of measurement would distinguish the two mixtures. ■

Another delicate point is that, dealing with mixed states, definitions (64) and (66) for, respectively the expectation value $\langle A \rangle_\psi$ and the standard deviation ΔA_ψ of an observable A referred to the pure state $\langle \psi | \cdot | \psi \rangle$ with $\|\psi\| = 1$ are no longer valid. We just say that extended natural definitions can be stated referring to the probability measure associated to both the mixed state $\rho \in \mathfrak{B}_1(\mathcal{H})$ (with $\rho \geq 0$ and $\text{tr } \rho = 1$) and the observable,

$$\mu_\rho^{(A)} : \mathcal{B}(\mathbb{R}) \ni E \mapsto \text{tr}(\rho P_E^{(A)}) .$$

We refer the reader to [6] for a technical discussion on these topics.

3.5. von Neumann algebra of observables, superselection rules

The aim of this section is to focus on the class of observables of a quantum system, described in the complex Hilbert space \mathcal{H} , exploiting some elementary results of the theory of *von Neumann algebras*. Up to now, we have tacitly supposed that *all* selfadjoint operators in \mathcal{H} represent observables, *all* orthogonal projectors represent elementary observables, *all* normalized vectors represent pure states. This is not the case in physics due to the presence of the so-called *superselection rules*. Within the Hilbert space approach the modern tool to deal with this notion is the mathematical structure of a *von Neumann algebra*. For this reason we spend the initial part of this section to introduce this mathematical tool.

3.5.1. von Neumann algebras

Before we introduce it, let us define first the *commutant* of an operator algebra and state an important preliminary theorem. If $\mathfrak{M} \subset \mathfrak{B}(\mathcal{H})$ is a subset in the algebra of bounded operators on the complex Hilbert space $\mathfrak{B}(\mathcal{H})$, the **commutant** of \mathfrak{M} is:

$$\mathfrak{M}' := \{T \in \mathfrak{B}(\mathcal{H}) \mid TA - AT = 0 \text{ for any } A \in \mathfrak{M}\} . \quad (81)$$

If \mathfrak{M} is closed under the adjoint operation (i.e. $A^* \in \mathfrak{M}$ if $A \in \mathfrak{M}$) the commutant \mathfrak{M}' is certainly a *-algebra with unit. In general: $\mathfrak{M}'_1 \subset \mathfrak{M}'_2$ if $\mathfrak{M}_2 \subset \mathfrak{M}_1$ and $\mathfrak{M} \subset (\mathfrak{M}')'$, which imply $\mathfrak{M}' = ((\mathfrak{M}')')'$. Hence we cannot reach beyond the second commutant by iteration.

The continuity of the product of operators in the uniform topology says that the commutant \mathfrak{M}' is closed in the uniform topology, so if \mathfrak{M} is closed under the adjoint operation, its commutant \mathfrak{M}' is a C^* -algebra (C^* -subalgebra) in $\mathfrak{B}(\mathcal{H})$.

\mathfrak{M}' has other pivotal topological properties in this general setup. It is easy to prove that \mathfrak{M}' is both strongly and weakly closed. This holds, despite the product of operators is not continuous with respect to the strong operator topology, because separate continuity in each variable is sufficient.

In the sequel we shall adopt the standard convention used for von Neumann algebras and write \mathfrak{M}'' in place of $(\mathfrak{M}')'$ etc. The next crucial result is due to von Neumann (see e.g. [6]).

Theorem 119 (von Neumann's double commutant theorem). *If \mathcal{H} is a complex Hilbert space and \mathfrak{A} a unital $*$ -subalgebra in $\mathfrak{B}(\mathcal{H})$, the following statements are equivalent.*

- (a) $\mathfrak{A} = \mathfrak{A}''$.
- (b) \mathfrak{A} is weakly closed.
- (c) \mathfrak{A} is strongly closed.

At this juncture we are ready to define von Neumann algebras.

Definition 120. Let \mathcal{H} be a complex Hilbert space. A **von Neumann algebra** in $\mathfrak{B}(\mathcal{H})$ is a $*$ -subalgebra of $\mathfrak{B}(\mathcal{H})$, with unit, that satisfies any of the equivalent properties appearing in von Neumann's theorem 119. ■

In particular \mathfrak{M}' is a von Neumann algebra provided \mathfrak{M} is a $*$ -closed subset of $\mathfrak{B}(\mathcal{H})$, because $(\mathfrak{M}')'' = \mathfrak{M}'$ as we saw above. Note how, by construction, a von Neumann algebra in $\mathfrak{B}(\mathcal{H})$ is a C^* -algebra with unit, or better, a C^* -subalgebra with unit of $\mathfrak{B}(\mathcal{H})$.

It is not hard to see that the intersection of von Neumann algebras is a von Neumann algebra. If $\mathfrak{M} \subset \mathfrak{B}(\mathcal{H})$ is closed under the adjoint operation, \mathfrak{M}'' turns out to be the smallest (set-theoretically) von Neumann algebra containing \mathfrak{M} as a subset [13]. Thus \mathfrak{M}'' is called the **von Neumann algebra generated** by \mathfrak{M} .

Since in QM it is natural to deal with unbounded selfadjoint operators, the definition of commutant is extended to the case of a set of generally unbounded selfadjoint operators, exploiting the fact that these operators admit spectral measures made of bounded operators.

Definition 121. If \mathfrak{N} is a set of (generally unbounded) selfadjoint operators in the complex Hilbert space \mathcal{H} , the **commutant** \mathfrak{N}' of \mathfrak{N} , is defined as the commutant in the sense of (81) of the set of all the spectral measures $P^{(A)}$ of every $A \in \mathfrak{N}$. The von Neuman algebra \mathfrak{N}'' generated by \mathfrak{N} is defined as $(\mathfrak{N}')'$, where the external prime is the one of definition (81). ■

Remark 122. Notice that, if the selfadjoint operators are all bounded, \mathfrak{N}' obtained this way coincides with the one already defined in (81) as a consequence of (ii) and (iv) of Proposition 86 (for a bounded selfadjoint operator A). Thus \mathfrak{N}' is well-defined and gives rises to a von Neumann algebra because the set of spectral measures is $*$ -closed. \mathfrak{N}'' is a von Neumann algebra too for the same reason. ■

We are in a position to state a technically important result which concerns both the spectral theory and the notion of von Neumann algebra [6].

Proposition 123. *Let $\mathfrak{N} = \{A_1, \dots, A_n\}$ be a finite collection of self-adjoint operators in the separable Hilbert space \mathcal{H} whose spectral measures commute. The von*

Neumann algebra \mathfrak{N}'' coincides with the collection of operators

$$f(A_1, \dots, A_n) := \int_{\text{supp}(P^{(\mathbf{A})})} f(x_1, \dots, x_n) dP^{(\mathbf{A})},$$

with $f : \text{supp}(P^{(\mathbf{A})}) \rightarrow \mathbb{C}$ measurable and bounded.

3.5.2. Lattices of von Neumann algebras

To conclude this elementary mathematical survey, we will say some words about von Neumann algebras and their associated lattices of orthogonal projectors.

Consider a von Neumann algebra \mathfrak{R} on the complex Hilbert space \mathcal{H} . It is easy to prove that the set $\mathcal{L}_{\mathfrak{R}}(\mathcal{H}) \subset \mathfrak{R}$ of the orthogonal projectors included in \mathfrak{R} form a lattice, which is bounded by 0 and I , orthocomplemented with respect to the orthocomplementation operation of $\mathcal{L}(\mathcal{H})$ and σ -complete (because this notion involves only the strong topology ((iii) in Proposition 98) and \mathfrak{R} is closed with respect to that topology in view of Theorem 119. Moreover $\mathcal{L}_{\mathfrak{R}}(\mathcal{H})$ is orthomodular, and separable like the whole $\mathcal{L}(\mathcal{H})$, assuming that \mathcal{H} is separable. It is interesting to note that, as expected, $\mathcal{L}_{\mathfrak{R}}(\mathcal{H})$ contains all information about \mathfrak{R} itself since the following result holds.

Proposition 124. *Let \mathfrak{R} be a von Neumann algebra on the complex Hilbert space \mathcal{H} and consider the lattice $\mathcal{L}_{\mathfrak{R}}(\mathcal{H}) \subset \mathfrak{R}$ of the orthogonal projectors in \mathfrak{R} . then the equality $\mathcal{L}_{\mathfrak{R}}(\mathcal{H})'' = \mathfrak{R}$ holds.*

Proof. Since $\mathcal{L}_{\mathfrak{R}}(\mathcal{H}) \subset \mathfrak{R}$, we have $\mathcal{L}_{\mathfrak{R}}(\mathcal{H})' \supset \mathfrak{R}'$ and $\mathcal{L}_{\mathfrak{R}}(\mathcal{H})'' \subset \mathfrak{R}'' = \mathfrak{R}$. Let us prove the other inclusion. $A \in \mathfrak{R}$ can always be decomposed as a linear combination of two self adjoint operators of \mathfrak{R} , $A + A^*$ and $i(A - A^*)$. So we can restrict ourselves to the case of $A^* = A \in \mathfrak{R}$, proving that $A \in \mathcal{L}_{\mathfrak{R}}(\mathcal{H})''$ if $A \in \mathfrak{R}$. The PVM of A belongs to \mathfrak{R} because of (ii) and (iv) of Proposition 86: $P^{(A)}$ commutes with every bounded operator B which commutes with A . So $P^{(A)}$ commutes, in particular, with the elements of \mathfrak{R}' because $\mathfrak{R} \ni A$. We conclude that every $P_E^{(A)} \in \mathfrak{R}'' = \mathfrak{R}$. Finally, there is a sequence of simple functions s_n uniformly converging to id in a compact $[-a, a] \supset \sigma(A)$ (e.g, see [6]). By construction $\int_{\sigma(A)} s_n dP^{(A)} \in \mathcal{L}_{\mathfrak{R}}(\mathcal{H})''$ because it is a linear combination of elements of $P^{(A)}$ and $\mathcal{L}_{\mathfrak{R}}(\mathcal{H})''$ is a linear space. Finally $\int_{\sigma(A)} s_n dP^{(A)} \rightarrow A$ for $n \rightarrow +\infty$ uniformly, and thus strongly, as seen in (2) of example 88. Since $\mathcal{L}_{\mathfrak{R}}(\mathcal{H})''$ is closed with respect to the strong topology, we must have $A \in \mathcal{L}_{\mathfrak{R}}(\mathcal{H})''$, proving that $\mathcal{L}_{\mathfrak{R}}(\mathcal{H}) \supset \mathfrak{R}$ as wanted. \square

3.5.3. General algebra of observables and its centre

Let us pass to physics and we apply these notions and results. Relaxing the hypothesis that all selfadjoint operators in the separable Hilbert space \mathcal{H} associated to a quantum system represent observables, there are many reasons to assume that the observables of a quantum system are represented (in the sense we are going to

illustrate) by the selfadjoint elements of an algebra of von Neumann, we hereafter indicated by \mathfrak{R} , called the **von Neumann algebra of observables** (though only the selfadjoint elements are observables). Including non-selfadjoint elements $B \in \mathfrak{R}$ is armless, as they can always be one-to-one decomposed into a pair of selfadjoint elements

$$B = B_1 + iB_2 = \frac{1}{2}(B + B^*) + i\frac{1}{2i}(B - B^*).$$

The fact that the elements of \mathfrak{R} are bounded does not seem a physical problem. If $A = A^*$ is unbounded and represents an observable it does not belong to \mathfrak{R} . Nevertheless the associated *class* of bounded selfadjoint operators $\{A_n\}_{n \in \mathbb{N}}$ where

$$A_n := \int_{[-n, n] \cap \sigma(A)} \lambda dP^{(A)}(\lambda),$$

embodies the same information as A itself. A_n is bounded due to Proposition 76 because the support of its spectral measures is included in $[-n, n]$. Physically speaking, we can say that A_n is nothing but the observable A when it is measured with an instrument unable to produce outcomes larger than $[-n, n]$. All real measurement instruments are similarly limited. We can safely assume that every A_n belongs to \mathfrak{R} . Mathematically speaking, the whole (unbounded) observable A is recovered as the limit in the *strong operator topology* $A = \lim_{n \rightarrow +\infty} A_n$ ((1) in examples 88). Moreover the union of the spectral measures of all the A_n is that of A . Finally the spectral measure of A belongs to \mathfrak{R} since the spectral measure of every $A_n \in \mathfrak{R}$ does, as has been established in the proof of Proposition 124 above.

Within this framework the orthogonal projectors $P \in \mathfrak{R}$ represent all elementary observables of the system. The lattice of these projectors, $\mathcal{L}_{\mathfrak{R}}(\mathcal{H})$, encompass the amount of information about observables as established Proposition 124. As said above $\mathcal{L}_{\mathfrak{R}}(\mathcal{H}) \subset \mathfrak{R}$ is bounded, orthocomplemented, σ -complete, orthomodular and separable like the whole $\mathcal{L}(\mathcal{H})$ (assuming that \mathcal{H} is separable) but there is no guarantee for the validity of the other properties listed in Theorem 102. The natural question is whether \mathfrak{R} is $*$ -isomorphic to $\mathfrak{B}(\mathcal{H}_1)$ for a suitable complex Hilbert space \mathcal{H}_1 , which would automatically imply that also the remaining properties were true. In particular there would exist atomic elements in $\mathcal{L}_{\mathfrak{R}}(\mathcal{H})$ and the covering property would be satisfied. A necessary condition is that, exactly as it happens for $\mathfrak{B}(\mathcal{H}_1)$, there are no non-trivial elements in $\mathfrak{R} \cap \mathfrak{R}'$, since $\mathfrak{B}(\mathcal{H}_1) \cap \mathfrak{B}(\mathcal{H}_1)' = \mathfrak{B}(\mathcal{H}_1)' = \{cI\}_{c \in \mathbb{C}}$.

Definition 125. *A von Neumann algebra \mathfrak{R} is a **factor** when its **centre**, the subset $\mathfrak{R} \cap \mathfrak{R}'$ of elements commuting with the whole algebra, is trivial: $\mathfrak{R} \cap \mathfrak{R}' = \{cI\}_{c \in \mathbb{C}}$.*

■

Remark 126. It is possible to prove that a von Neumann algebra is always a direct sum or a direct integral of factors. Therefore factors play a crucial role. The classification of factors, started by von Neumann and Murray, is one of the key

chapters in the theory of operator algebras, and has enormous consequences in the algebraic theory of quantum fields. The factors isomorphic to $\mathfrak{B}(\mathcal{H}_1)$ for some complex Hilbert space \mathcal{H}_1 , are called of *type I*. These factors admit atoms, fulfil the covering property (orthomodularity and irreducibility are always true). Regarding separability, it depends on separability of \mathcal{H}_1 and requires a finer classification in factors of *type I_n* where n is a cardinal number. There are however factors of type *II* and *III* which do not admit atoms and are not important in elementary QM. ■

The centre of the von Neumann algebra of observables enters the physical theory in a nice way. A common situation dealing with quantum systems is the existence of a **maximal set of compatible observables**, i.e. a finite maximal class $\mathfrak{A} = \{A_1, \dots, A_n\}$ of pairwise compatible observables. The notion of maximality here means that, if a (bounded) selfadjoint operator commutes with all the observables in \mathfrak{A} , then it is a *function* of them. In particular it is an observable as well. In view of proposition 123 the existence of a maximal set of compatible observables is equivalent to say that there is a finite set of observables \mathfrak{A} such that $\mathfrak{A}' = \mathfrak{A}''$. We have the following important consequence

Proposition 127. *If a quantum physical system admits a maximal set of compatible observables, then the commutant \mathfrak{A}' of the von Neumann algebra of observables \mathfrak{A} is Abelian and coincides with the center of \mathfrak{A} .*

Proof. As the spectral measures of each $A \in \mathfrak{A}$ belong to \mathfrak{A} , it must be (i) $\mathfrak{A}'' \subset \mathfrak{A}$. Since $\mathfrak{A}' = \mathfrak{A}''$, (i) yields $\mathfrak{A}' \subset \mathfrak{A}$ and thus, taking the commutant, (ii) $\mathfrak{A}'' \supset \mathfrak{A}'$. Comparing (i) and (ii) we have $\mathfrak{A}' \subset \mathfrak{A}$. In other words $\mathfrak{A}' = \mathfrak{A}' \cap \mathfrak{A}$. In particular, \mathfrak{A}' must be Abelian. □

Example 128.

(1) Considering a quantum particle without spin and referring to the rest space \mathbb{R}^3 of an inertial reference frame, $\mathcal{H} = L^2(\mathbb{R}^3, d^3x)$. A maximal set of compatible observables is the set of the three position operators $\mathfrak{A}_1 = \{X_1, X_2, X_3\}$ or the the set of the three momenta operators $\mathfrak{A}_2 = \{P_1, P_2, P_3\}$. \mathfrak{A} is the von Neumann algebra generated by $\mathfrak{A}_1 \cup \mathfrak{A}_2$. It is possible to prove that the commutant (which coincides with the centre) of this von Neumann algebra is trivial (as it includes a unitary irreducible representation of the Weyl-Heisenberg group) so that $\mathfrak{A} = \mathfrak{B}(\mathcal{H})$ (see also Theorem 167).

(2) If adding the spin space (for instance dealing with an electron “without charge”), we have $\mathcal{H} = L^2(\mathbb{R}^3, d^3x) \otimes \mathbb{C}^2$. Referring to (11) a maximal set of compatible observables is, for instance, $\mathfrak{A}_1 = \{X_1 \otimes I, X_2 \otimes I, X_3 \otimes I, I \otimes S_z\}$, another is $\mathfrak{A}_2 = \{P_1 \otimes I, P_2 \otimes I, P_3 \otimes I, I \otimes S_x\}$. As before $(\mathfrak{A}_1 \cup \mathfrak{A}_2)''$ is the von Neumann algebra of observables of the system (changing the component of the spin passing from \mathfrak{A}_1 to \mathfrak{A}_2 is crucial for this result). Also in this case, it turns out that the commutant of the von Neumann algebra of observables is trivial yielding $\mathfrak{A} = \mathfrak{B}(\mathcal{H})$. ■

3.5.4. Superselection charges and coherent sectors

We must have accumulated enough formalism to successfully investigate the structure of the Hilbert space (always supposed to be separable) and the algebra of the observables when not all selfadjoint operators represent observables and not all orthogonal projectors are interpreted as elementary observables. Re-adapting the approach by Wightman [22] to our framework, we make two assumptions generally describing the so called *superselection rules* for QM formulated in a (separable) Hilbert space where \mathfrak{R} denotes the von Neumann algebra of observables.

(SS1) There is a maximal set of compatible observables in \mathfrak{R} , so that $\mathfrak{R}' = \mathfrak{R}' \cap \mathfrak{R}$.

(SS2) $\mathfrak{R}' \cap \mathfrak{R}$ contains a finite class of observables $\Omega = \{Q_1, \dots, Q_n\}$, with $\sigma(Q_k) = \sigma_p(Q_k)$, $k = 1, 2, \dots, n$, generating the centre: $\Omega'' \supset \mathfrak{R}' \cap \mathfrak{R}$. (If the Q_k are unbounded, $\Omega \subset \mathfrak{R}' \cap \mathfrak{R}$ means that the PVM of the Q_j are included in $\mathfrak{R}' \cap \mathfrak{R}$.)

The Q_k are called **superselection charges**.

As the reader can easily prove, the joint spectral measure $P^{(\Omega)}$ in \mathbb{R}^n has support given exactly by $\times_{k=1}^n \sigma_p(Q_k)$ and, if $E \subset \mathbb{R}^n$,

$$P_E^{(\Omega)} = \sum_{(q_1, \dots, q_n) \in \times_{k=1}^n \sigma_p(Q_k) \cap E} P_{\{q_1\}}^{(Q_1)} \cdots P_{\{q_n\}}^{(Q_n)} \quad (82)$$

We have the following remarkable result where we occasionally adopt the notation $\mathbf{q} := (q_1, \dots, q_n)$ and $\sigma(\Omega) := \times_{k=1}^n \sigma_p(Q_k)$.

Proposition 129. *Let \mathcal{H} be a complex separable Hilbert and suppose that the von Neumann algebra \mathfrak{R} in \mathcal{H} satisfies **(SS1)** and **(SS2)**. The following facts hold.*

(a) \mathcal{H} admits the following direct decomposition into closed pairwise orthogonal subspaces, called **superselection sectors** or **coherent sectors**,

$$\mathcal{H} = \bigoplus_{\mathbf{q} \in \sigma(\Omega)} \mathcal{H}_{\mathbf{q}} \quad (83)$$

where

$$\mathcal{H}_{\mathbf{q}} := P_{\mathbf{q}}^{(\Omega)} \mathcal{H}.$$

and each $\mathcal{H}_{\mathbf{q}}$ is invariant and irreducible under \mathfrak{R} .

(b) An analogous direct decomposition occurs for \mathfrak{R} .

$$\mathfrak{R} = \bigoplus_{\mathbf{q} \in \sigma(\Omega)} \mathfrak{R}_{\mathbf{q}} \quad (84)$$

where

$$\mathfrak{R}_{\mathbf{q}} := \{ A|_{\mathcal{H}_{\mathbf{q}}} \mid A \in \mathfrak{R} \}$$

is a von Neumann algebra on $\mathcal{H}_{\mathbf{q}}$ considered as Hilbert space in its own right. Finally,

$$\mathfrak{R}_{\mathbf{q}} = \mathfrak{B}(\mathcal{H}_{\mathbf{q}})$$

(c) Each map

$$\mathfrak{R} \ni A \mapsto A|_{\mathcal{H}_{\mathbf{q}}} \in \mathfrak{R}_{\mathbf{q}}$$

is a *-algebra representation of \mathfrak{R} (Def.33). Representations associated with different values of \mathbf{q} are (unfithful and) unitarily inequivalent: In other words there is no isometric surjective map $U : \mathcal{H}_{\mathbf{q}} \rightarrow \mathcal{H}_{\mathbf{q}'}$ such that

$$UA|_{\mathcal{H}_{\mathbf{q}}}U^{-1} = A|_{\mathcal{H}_{\mathbf{q}'}}$$

when $\mathbf{q} \neq \mathbf{q}'$

Proof. (a) Since $P_{\mathbf{q}}^{(\Omega)}P_{\mathbf{s}}^{(\Omega)} = 0$ if $\mathbf{q} \neq \mathbf{s}$ and $\sum_{\mathbf{q} \in \sigma_p(\Omega)} P_{\mathbf{q}}^{(\Omega)} = I$, \mathcal{H} decomposes as in (83). Since $P_{\mathbf{q}}^{(\Omega)}$ belongs to the centre of \mathfrak{R} , the subspaces of the decomposition are invariant under the action of each element of \mathfrak{R} . Let us pass to the irreducibility. If $P \in \mathfrak{R}' \cap \mathfrak{R}$ is an orthogonal projector it must be a function of the Q_k by hypotheses: $P = \int_{\mathbb{R}^n} f(x)dP^{(\Omega)}(x)$ since $P = PP \geq 0$ and $P = P^*$, exploiting the measurable functional calculus, we easily find that $f(x) = \chi_E(x)$ for some $E \subset \text{supp}(P^{(\Omega)})$. In other words P is an element of the joint PVM of Ω : that PVM exhausts all orthogonal projectors in $\mathfrak{R}' \cap \mathfrak{R}$. Now, if $\{0\} \neq \mathcal{K} \subset \mathcal{H}_{\mathbf{s}}$ is an invariant closed subspace for \mathfrak{R} , its orthogonal projector $P_{\mathcal{K}}$ must commute with \mathfrak{R} , so it must belong to the centre for (SS2) and thus it belongs to $P^{(\Omega)}$ for (SS1) and, more precisely it must be of the form $P_{\mathcal{K}} = P_{\mathbf{s}}^{(\Omega)}$ because $P_{\mathcal{K}} \leq P_{\mathbf{s}}^{(\Omega)}$ by hypothesis, but there are no projectors smaller than $P_{\mathbf{s}}^{(\Omega)}$ in the PVM of Ω . So $\mathcal{K} = \mathcal{H}_{\mathbf{s}}$.

(b) $\mathfrak{R}_{\mathbf{q}} := \{A|_{\mathcal{H}_{\mathbf{q}}} \mid A \in \mathfrak{R}\}$ is a von Neumann algebra on $\mathcal{H}_{\mathbf{s}}$ considered as a Hilbert space in its own right as it arises by direct inspection. (84) holds by definition. Since $\mathcal{H}_{\mathbf{q}}$ is irreducible for $\mathfrak{R}_{\mathbf{q}}$, we have $\mathfrak{R}_{\mathbf{s}} = \mathfrak{R}'_{\mathbf{s}} = \mathfrak{B}(\mathcal{H}_{\mathbf{s}})$. Each map $\mathfrak{R} \ni A \mapsto A|_{\mathcal{H}_{\mathbf{q}}} \in \mathfrak{R}_{\mathbf{q}}$ is a representation of *-algebras as follows by direct check. If $\mathbf{q} \neq \mathbf{q}'$ –for instance $q_1 \neq q'_1$ – there is no isometric surjective map $U : \mathcal{H}_{\mathbf{q}} \rightarrow \mathcal{H}_{\mathbf{q}'}$ such that

$$UA|_{\mathcal{H}_{\mathbf{q}}}U^{-1} = A|_{\mathcal{H}_{\mathbf{q}'}}$$

If such an operator existed one would have, contrarily to our hypothesis $q_1 \neq q'_1$, $q_1 I_{\mathcal{H}_{\mathbf{q}'}} = UQ_1|_{\mathcal{H}_{\mathbf{q}}}U^{-1} = Q_1|_{\mathcal{H}_{\mathbf{q}'}} = q'_1 I_{\mathcal{H}_{\mathbf{q}'}}$ so that $q_1 = q'_1$. \square

We have found that, in the presence of superselection charges, the Hilbert space decomposes into pairwise orthogonal subspaces which are invariant and irreducible with respect to the algebra of the observables, giving rise to inequivalent representations of the algebra itself. Restricting ourselves to each such subspace, QM takes its standard form as all orthogonal projectors are representatives of elementary observables, differently from what happens in the whole Hilbert space where there are orthogonal projectors which cannot represent observables: These are the projectors

which do not commute with $P^{(\Omega)}$.

There are several superselection structures as the one pointed out in physics. The three most known are of very different nature: The superselection structure of the *electric charge*, the superselection structure of *integer/semi integers values of the angular momentum*, and the one related to the mass in non-relativistic physics, i.e., *Bargmann's superselection rule*.

Example 130. The electric charge is the typical example of superselection charge. For instance, referring to an electron, its Hilbert space is $L^2(\mathbb{R}^3, d^3x) \otimes \mathcal{H}_s \otimes \mathcal{H}_e$. The space of the electric charge is $\mathcal{H}_e = \mathbb{C}^2$ and therein $Q = e\sigma_z$ (see (12)). Many other observables could exist in \mathcal{H}_e in principle, but the electric charge superselection rule imposes that the only possible observables are functions of σ_z . The centre of the algebra of observables is $I \otimes I \otimes f(\sigma_3)$ for every function $f : \sigma(\sigma_z) = \{1, 1\} \rightarrow \mathbb{C}$. We have the decomposition in coherent sectors

$$\mathcal{H} = (L^2(\mathbb{R}^3, d^3x) \otimes \mathcal{H}_s \otimes \mathcal{H}_+) \oplus (L^2(\mathbb{R}^3, d^3x) \otimes \mathcal{H}_s \otimes \mathcal{H}_-),$$

where \mathcal{H}_{\pm} are respectively the eigenspaces of Q with eigenvalue $\pm e$. ■

Remark 131.

(a) A fundamental requirement is that the superselection charges have punctual spectrum. If instead $\mathfrak{R} \cap \mathfrak{R}'$ includes an operator A with a continuous part in its spectrum (A may also be the strong limit on $D(A)$ of a sequence of elements in $\mathfrak{R} \cap \mathfrak{R}'$), the established proposition does not hold. \mathcal{H} cannot be decomposed into a direct sum of closed subspaces. In this case it decomposes into a direct integral and we find a much more complicated structure whose physical meaning seems dubious.

(b) The representations $\mathfrak{R} \ni A \mapsto A|_{\mathcal{H}_{\mathbf{q}}} \in \mathfrak{R}_{\mathbf{q}}$ are not faithful (injective), because both I and $P_s^{(\Omega)}$ have the same image under the representation.

(c) The discussed picture is not the most general one though we only deal with it in these notes. There are quantum physical systems such that their \mathfrak{R}' is not Abelian (think of chromodynamics where \mathfrak{R}' includes a faithful representation of $SU(3)$) so that the centre of \mathfrak{R} does not contain the full information about \mathfrak{R}' . In this case, the non-Abelian group of the unitary operators in \mathfrak{R}' is called the **gauge group** of the theory. The existence of a gauge group is compatible with the presence of superselection rules which are completely described by the centre $\mathfrak{R}' \cap \mathfrak{R}$. The only difference is that now $\mathfrak{R}_{\mathbf{q}} = \mathfrak{B}(\mathcal{H}_{\mathbf{q}})$ cannot be possible for every coherent subspace otherwise we would have $\mathfrak{R}' = \mathfrak{R} \cap \mathfrak{R}'$. ■

3.5.5. States in the presence of superselection rules

Let us come to the problem to characterize the states when a superselection structure is assumed on a complex separable Hilbert space \mathcal{H} in accordance with (SS1) and (SS2). In principle we can extend Definition 110 already given for the case of

\mathfrak{A} with trivial centre. As usual $\mathcal{L}_{\mathfrak{A}}(\mathcal{H})$ indicates the lattice of orthogonal projectors in \mathfrak{A} , which we know to be bounded by 0 and I , orthocomplemented, σ -complete, orthomodular and separable, but not atomic and it does not satisfy the covering property in general. The atoms are one-dimensional projectors exactly as pure states, so we may expect some difference at that level when $\mathfrak{A} \neq \mathfrak{B}(\mathcal{H})$.

Definition 132. *Let \mathcal{H} be a complex separable Hilbert space. A **quantum state** in \mathcal{H} , for a quantum system with von Neumann algebra of observables \mathfrak{A} , is a map $\rho : \mathcal{L}_{\mathfrak{A}}(\mathcal{H}) \rightarrow [0, 1]$ such that the following requirements are satisfied.*

(1) $\rho(I) = 1$.

(2) If $\{Q_n\}_{n \in N} \subset \mathcal{L}_{\mathfrak{A}}(\mathcal{H})$, for N at most countable satisfies $Q_k \wedge Q_h = 0$ when $h, k \in N$, then

$$\rho(\bigvee_{k \in N} Q_k) = \sum_{k \in N} \rho(Q_k). \quad (85)$$

The set of the states will be denoted by $\mathfrak{S}_{\mathfrak{A}}(\mathcal{H})$. ■

If there is a superselection structure we have the decompositions we re-write down into a simpler version,

$$\mathcal{H} = \bigoplus_{k \in K} \mathcal{H}_k, \quad \mathfrak{A} = \bigoplus_{k \in K} \mathfrak{A}_k, \quad \mathfrak{A}_k = \mathfrak{B}(\mathcal{H}_k), \quad k \in K \quad (86)$$

where K is some finite or countable set. The lattice $\mathcal{L}_{\mathfrak{A}}(\mathcal{H})$, as a consequence of (85), decomposes as (the notation should be obvious)

$$\mathcal{L}_{\mathfrak{A}}(\mathcal{H}) = \bigvee_{k \in K} \mathcal{L}_{\mathfrak{A}_k}(\mathcal{H}_k) = \bigvee_{k \in K} \mathcal{L}(\mathcal{H}_k) \quad (87)$$

where

$$\mathcal{L}_{\mathfrak{A}_k}(\mathcal{H}_k) \wedge \mathcal{L}_{\mathfrak{A}_h}(\mathcal{H}_h) = \{0\} \quad \text{if } k \neq h.$$

In other words $Q \in \mathcal{L}_{\mathfrak{A}}(\mathcal{H})$ can uniquely be written as $Q = \bigvee_{k \in K} Q_k$ where $Q_k \in \mathcal{L}(\mathfrak{B}(\mathcal{H}_k))$. In fact $Q_k = P_k Q_k$, where P_k is the orthogonal projector onto \mathcal{H}_k .

In this framework, it is possible to readapt Gleason's result simply observing that a state ρ on $\mathcal{L}_{\mathfrak{A}}(\mathcal{H})$ as above defines a state ρ_k on $\mathcal{L}_{\mathfrak{A}_k}(\mathcal{H}_k) = \mathcal{L}(\mathcal{H}_k)$ by

$$\rho_k(P) := \frac{1}{\rho(P_k)} \rho(P), \quad P \in \mathcal{L}(\mathcal{H}_k).$$

If $\dim(\mathcal{H}_k) \neq 2$ we can exploit Gleason's theorem.

Theorem 133. *Let \mathcal{H} be a complex separable Hilbert space and assume that the von Neumann algebra \mathfrak{A} in \mathcal{H} satisfies **(SS1)** and **(SS2)**, so that the decomposition (86) in coherent sectors is valid where we suppose $\dim \mathcal{H}_k \neq 2$ for every $k \in K$. The following facts hold.*

(a) If $T \in \mathfrak{B}_1(\mathcal{H})$ satisfies $T \geq 0$ and $\text{tr } T = 1$ then

$$\rho_T : \mathcal{L}_{\mathfrak{A}}(\mathcal{H}) \ni P \mapsto \text{tr}(TP)$$

is an element of $\mathfrak{S}_{\mathfrak{R}}(\mathcal{H})$ that is a state on $\mathcal{L}_{\mathfrak{R}}(\mathcal{H})$.

(b) For $\rho \in \mathfrak{S}_{\mathfrak{R}}(\mathcal{H})$ there is a $T \in \mathfrak{B}_1(\mathcal{H})$ satisfies $T \geq 0$ and $\text{tr } T = 1$ such that $\rho = \rho_T$.

(c) If $T_1, T_2 \in \mathfrak{B}_1(\mathcal{H})$ satisfy same hypotheses as T in (a), then $\rho_{T_1} = \rho_{T_2}$ is valid if and only if $P_k T_1 P_k = P_k T_2 P_k$ for all $k \in K$, P_k being the orthogonal projector onto \mathcal{H}_k .

(d) A unit vector $\psi \in \mathcal{H}$ defines a pure state only if belongs to a coherent sector. More precisely, a state $\rho \in \mathfrak{S}_{\mathfrak{R}}(\mathcal{H})$ is pure, that is extremal, if and only if there is $k_0 \in K$, $\psi \in \mathcal{H}_{k_0}$ with $\|\psi\| = 1$ such that

$$\rho(P) = 0 \quad \text{if } P \in \mathcal{L}(\mathcal{H}_k), k \neq k_0 \quad \text{and} \quad \rho(P) = \langle \psi | P \psi \rangle \quad \text{if } P \in \mathcal{L}(\mathcal{H}_{k_0})$$

Proof. (a) is obvious from Proposition 112, as restricting a state ρ on $\mathcal{L}(\mathcal{H})$ to $\mathcal{L}_{\mathfrak{R}}(\mathcal{H})$ we still obtain a state as one can immediately verify. Let us prove (b). Evidently, every $\rho|_{\mathcal{L}(\mathcal{H}_k)}$ is a positive measure with $0 \leq \rho(P_k) \leq 1$. We can apply Gleason's theorem finding $T_k \in \mathfrak{B}(\mathcal{H}_k)$ with $T_k \geq 0$ and $\text{Tr } T_k = \rho(P_k)$ such that $\rho(Q) = \text{tr}(T_k Q)$ if $Q \in \mathcal{L}(\mathcal{H}_k)$. Notice also that $\|T_k\| \leq \rho(P_k)$ because

$$\|T_k\| = \sup_{\lambda \in \sigma_p(T_k)} |\lambda| = \sup_{\lambda \in \sigma_p(T_k)} \lambda \leq \sum_{\lambda \in \sigma_p(T_k)} d_\lambda \lambda = \text{Tr } T_k = \rho(P_k).$$

If $Q \in \mathcal{L}_{\mathfrak{R}}(\mathcal{H})$, $Q = \sum_k Q_k$, where $Q_k := P_k Q \in \mathcal{L}(\mathcal{H}_k)$, $Q_k Q_h = 0$ if $k \neq h$ and thus, by σ -additivity,

$$\rho(Q) = \sum_k \rho(Q_k) = \sum_k \text{tr}(T_k Q_k)$$

since $\mathcal{H}_k \perp \mathcal{H}_h$, this identity can be rewritten as

$$\rho(Q) = \text{tr}(TQ)$$

provided $T := \oplus_k T_k \in \mathfrak{B}_1(\mathcal{H})$. It is clear that $T \in \mathfrak{B}(\mathcal{H})$ because, if $x \in \mathcal{H}$ and $\|x\| = 1$ then, as $x = \sum_k x_k$ with $x_k \in \mathcal{H}_k$, $\|Tx\| \leq \sum_k \|T_k\| \|x_k\| \leq \sum_k \|T_k\| \mathbf{1} \leq \sum_k \rho(P_k) = 1$. In particular $\|T\| \leq 1$. $T \geq 0$ because each $T_k \geq 0$. Hence $|T| = \sqrt{T^* T} = \sqrt{T T} = T$ via functional calculus, and also $|T_k| = T_k$. Moreover, using the spectral decomposition of T , whose PVM commutes with each P_k , one easily has $|T| = \oplus_k |T_k| = \oplus_k T_k$. The condition

$$1 = \rho(I) = \sum_k \rho(P_k) = \sum_k \text{tr}(T_k P_k) = \sum_k \text{tr}(|T_k| P_k)$$

is equivalent to say that $\text{tr } |T| = 1$ using a Hilbertian basis of \mathcal{H} made of the union of bases in each \mathcal{H}_k . We have obtained, as wanted, that $T \in \mathfrak{B}_1(\mathcal{H})$, $T \geq 0$, $\text{tr } T = 1$ and $\rho(Q) = \text{tr}(TQ)$ for all $Q \in \mathcal{L}_{\mathfrak{R}}(\mathcal{H})$.

(c) The proof straightforwardly follows from $\mathcal{L}_{\mathfrak{R}_k}(\mathcal{H}_k) = \mathcal{L}(\mathfrak{B}(\mathcal{H}_k))$ because $\mathfrak{R}_k = \mathfrak{B}(\mathcal{H}_k)$ and, evidently, $\rho_{T_1} = \rho_{T_2}$ if and only if $\rho_{T_1}|_{\mathcal{L}(\mathfrak{B}(\mathcal{H}_k))} = \rho_{T_2}|_{\mathcal{L}(\mathfrak{B}(\mathcal{H}_k))}$ for all $k \in K$. Regarding (d) it is clear that if ρ encompasses more than one component $\rho|_{\mathcal{L}(\mathcal{H}_k)} \neq 0$ cannot be extremal because is, by construction, a convex combination of other states which vanishes in some of the given coherent subspace. Therefore

only states such that only one restriction $\rho|_{\mathcal{L}(\mathcal{H}_{k_0})}$ does not vanish may be extremal. Now (a) of Proposition 117 implies that, among these states, the extremal ones are precisely those of the form said in (d) of the thesis. \square

Remark 134.

(a) Take $\psi = \sum_{k \in K} c_k \psi_k$ where the $\psi_k \in \mathcal{H}_k$ are unit vectors and also suppose that $\|\psi\|^2 = \sum_k |c_k|^2 = 1$. This vector induces a state ρ_ψ on \mathfrak{A} by means of the standard procedure (which is nothing but the trace procedure with respect to $T_\psi := \langle \psi | \cdot \rangle \psi$)

$$\rho_\psi(P) = \langle \psi | P \psi \rangle \quad P \in \mathcal{L}_{\mathfrak{A}}(\mathcal{H}) .$$

In this case however, since $PP_k = P_kP$ and $\psi_k = P_k\psi_k$ we have

$$\begin{aligned} \rho_\psi(P) &= \sum_k \sum_h \bar{c}_k c_h \langle \psi_k | P_k P P_h \psi_k \rangle = \sum_k \sum_h \bar{c}_k c_h \langle \psi_k | P P_k P_h \psi_k \rangle \\ &= \sum_k \sum_h \bar{c}_k c_h \langle \psi | P P_k \psi \rangle \delta_{kh} = \sum_k |c_k|^2 \langle \psi_k | P \psi_k \rangle = \text{tr}(T'_\psi P) \end{aligned}$$

where

$$T'_\psi = \sum_{k \in K} |c_k|^2 \langle \psi_k | \cdot \rangle \psi_k$$

We conclude that the apparent pure state ψ and the mixed state T'_ψ cannot be distinguished, just because the algebra \mathfrak{A} is too small to make a difference. Actually they define the same state at all and this is an elementary case of (c) in the above theorem with $T_1 = \langle \psi | \cdot \rangle \psi$ and $T_2 = T'_\psi$.

This discussion, in the language of physicist is often stated as follows:

No coherent superpositions $\psi = \sum_{k \in K} c_k \psi_k$ of pure states $\psi_k \in \mathcal{H}_k$ of different coherent sectors are possible, only incoherent superpositions $\sum_{k \in K} |c_k|^2 \langle \psi_k | \cdot \rangle \psi_k$ are allowed.

(b) It should be clear that the one-to-one correspondence between pure states and atomic elementary observables (one-dimensional projectors) here does not work. Consequently, notions like *probability amplitude* must be handled with great care. In general, however, everything goes right if staying in a fixed superselection sector \mathcal{H}_k where the said correspondence exists. \blacksquare

3.6. Quantum Symmetries: unitary projective representations

The notion of symmetry in QM is quite abstract. Actually there are three distinct ideas, respectively by Wigner, Kadison and Segal [23]. Here we focus on the first pair only. Physically speaking, a *symmetry* is an active transformation on the quantum system changing its state. It is supposed that this transformation preserves some properties of the physical system and here we have to distinguish between the two afore-mentioned cases. However in both cases the transformation is required to be reversible (injective) and to cover (surjective) the space of the states. Symmetries

are supposed to mathematically describe some concrete transformation acting on the physical system. Sometimes their action, in practice, can be cancelled simply changing the reference frame. This is not the general case however, even if this class of symmetries plays a relevant role in physics.

3.6.1. Wigner and Kadison theorems, groups of symmetries

Consider a quantum system described in the complex Hilbert space \mathcal{H} with dimension $\neq 2$, separable whenever infinite dimensional. We assume that either \mathcal{H} is the whole Hilbert space in the absence of superselection charges or it denotes a single coherent sector. Let $\mathfrak{S}(\mathcal{H})$ and $\mathfrak{S}_p(\mathcal{H})$ respectively indicate the convex body of the quantum states and the set of pure states, referred to the sector \mathcal{H} if it is the case.

Definition 135. If \mathcal{H} is a complex Hilbert space with dimension $\neq 2$, separable if infinite dimensional, we have the following definitions.

(a) A **Wigner symmetry** is a bijective map

$$s_W : \mathfrak{S}_p(\mathcal{H}) \ni \langle \psi | \cdot \rangle \psi \rightarrow \langle \psi' | \cdot \rangle \psi' \in \mathfrak{S}_p(\mathcal{H})$$

which preserves the probabilities of transition. In other words

$$|\langle \psi_1 | \psi_2 \rangle|^2 = |\langle \psi'_1 | \psi'_2 \rangle|^2 \quad \text{if } \psi_1, \psi_2 \in \mathcal{H} \quad \text{with } \|\psi_1\| = \|\psi_2\| = 1.$$

(b) A **Kadison symmetry** is a bijective map

$$s_K : \mathfrak{S}(\mathcal{H}) \ni \rho \rightarrow \rho' \in \mathfrak{S}(\mathcal{H})$$

which preserves the convex structure of the space of the states. In other words

$$(p\rho_1 + q\rho_2)' = p\rho'_1 + q\rho'_2 \quad \text{if } \rho_1, \rho_2 \in \mathfrak{S}(\mathcal{H}) \quad \text{and } p, q \geq 0 \quad \text{with } p + q = 1. \quad \blacksquare$$

We observe that the first definition is well-posed even if unit vectors define pure states just up to a phase, as the reader can immediately prove, because transition probabilities are not affected by that ambiguity.

Though the definitions are evidently of different nature, they lead to the same mathematical object, as established in a pair of famous characterization theorems we quote into a unique statement. We need a preliminary definition.

Definition 136. Let \mathcal{H} a complex Hilbert space. A map $U : \mathcal{H} \rightarrow \mathcal{H}$ is said to be an **antiunitary operator** if it is surjective, isometric and $U(ax + by) = \bar{a}Ux + \bar{b}Uy$ when $x, y \in \mathcal{H}$ and $a, b \in \mathbb{C}$. \blacksquare

We come to the celebrated theorem. The last statement is obvious, the difficult parts are (a) and (b) (see, e.g., [6]).

Theorem 137 (Wigner and Kadison theorems). *Let \mathcal{H} be a complex Hilbert space with dimension $\neq 2$, separable if infinite dimensional. The following facts hold.*

(a) For every Wigner symmetry s_W there is an operator $U : \mathcal{H} \rightarrow \mathcal{H}$, which can be either unitary or anti unitary (depending on s_w) such that

$$s_w : \langle \psi | \cdot \rangle \psi \rightarrow \langle U\psi | \cdot \rangle U\psi, \quad \forall \langle \psi | \cdot \rangle \psi \in \mathfrak{S}_p(\mathcal{H}). \quad (88)$$

U and U' are associated to the same s_W if and only if $U' = e^{ia}U$ for $a \in \mathbb{R}$.

(b) For every Kadison symmetry s_K there is an operator $U : \mathcal{H} \rightarrow \mathcal{H}$, which can be either unitary or anti unitary (depending on s_K) such that

$$s_w : \rho \rightarrow U\rho U^{-1}, \quad \forall \rho \in \mathfrak{S}(\mathcal{H}). \quad (89)$$

U and U' are associated to the same s_K if and only if $U' = e^{ia}U$ for $a \in \mathbb{R}$.

(c) $U : \mathcal{H} \rightarrow \mathcal{H}$, either unitary or antiunitary, simultaneously defines a Wigner and a Kadison symmetry by means of (88) and (89) respectively.

Remark 138.

(a) It is worth stressing that the Kadison notion of symmetry is an extension of the Wigner one, after the result above. In fact, a Kadison symmetry $\rho \mapsto U\rho U^{-1}$ restricted to one-dimensional projectors preserves the probability transitions, as immediately follows from the identity $|\langle \psi | \phi \rangle|^2 = \text{tr}(\rho_\psi \rho_\phi)$ and the cyclic property of the trace, where we use the notation $\rho_\chi = \langle \chi | \cdot \rangle \chi$. In particular we can use the same operator U to represent also the found Wigner symmetry.

(b) If superselection rules are present, in general quantum symmetries are described in a similar way with unitary or antiunitary operators acting in a single coherent sector or also swapping different sectors [6]. ■

If a unitary or antiunitary operator V represents a symmetry s , it has an action on observables, too. If A is an observable (a selfadjoint operator on \mathcal{H}), we define the **transformed observable** along the action of s as

$$s^*(A) := VAV^{-1}. \quad (90)$$

Obviously $D(s^*(A)) = V(D(A))$. It is evident that this definition is not affected by the ambiguity of the arbitrary phase in the choice of V when s is given.

According with (i) in Proposition 82 the spectral measure of $s^*(A)$ is

$$P_E^{(s^*(A))} = VP_E^{(A)}V^{-1} = s^*(P_E^{(A)})$$

as expected.

The meaning of $s^*(A)$ should be evident: The probability that the observable $s^*(A)$ produces the outcome E when the state is $s(\rho)$ (namely $\text{tr}(P_E^{(s^*(A))} s(\rho))$) is the same as the probability that the observable A produces the outcome E when the state is ρ (that is $\text{tr}(P_E^{(A)} \rho)$). Changing simultaneously and coherently observables and states nothing changes. Indeed

$$\begin{aligned} \text{tr}(P_E^{(s^*(A))} s(\rho)) &= \text{tr}(VP_E^{(A)}V^{-1}V\rho V^{-1}) = \text{tr}(VP_E^{(A)}\rho V^{-1}) \\ &= \text{tr}(P_E^{(A)}\rho V^{-1}V) = \text{tr}(P_E^{(A)}\rho). \end{aligned}$$

Example 139.

(1) Fixing an inertial reference frame, the pure state of a quantum particle is defined, up to phases, as a unit norm element ψ of $L^2(\mathbb{R}^3, d^3x)$, where \mathbb{R}^3 stands for the rest three space of the reference frame. The group of isometries $IO(3)$ of \mathbb{R}^3 equipped with the standard Euclidean structure acts on states by means of symmetries the sense of Wigner and Kadison. If $(R, t) : x \mapsto Rx + t$ is the action of the generic element of $IO(3)$, where $R \in O(3)$ and $t \in \mathbb{R}^3$, the associated quantum (Wigner) symmetry $s_{(R,t)}(\langle \psi | \cdot \rangle \psi) = \langle U_{(R,t)} \psi | \cdot \rangle U_{(R,t)} \psi$ is completely fixed by the unitary operators $U_{(R,t)}$. They are defined as

$$(U_{(R,t)}\psi)(x) := \psi((R, t)^{-1}x), \quad x \in \mathbb{R}^3, \psi \in L^2(\mathbb{R}^3, d^3x), \quad \|\psi\| = 1.$$

The fact that the Lebesgue measure is invariant under $IO(3)$ immediately proves that $U_{(R,t)}$ is unitary. It is furthermore easy to prove that, with the given definition

$$U_{(I,0)} = I, \quad U_{(R,t)}U_{(R',t')} = U_{(R,t) \circ (R',t')}, \quad \forall (R, t), (R', t') \in IO(3). \quad (91)$$

(2) The so called *time reversal* transformation classically corresponds to invert the sign of all the velocities of the physical system. It is possible to prove [6] (see also (3) in exercise 157 below) that, in QM and for systems whose energy is bounded below but not above, the time reversal symmetry cannot be represented by unitary transformations, but only antiunitary. In the most elementary situation as in (1), the time reversal is defined by means of the anti unitary operator

$$(T\psi)(x) := \overline{\psi(x)}, \quad x \in \mathbb{R}^3, \psi \in L^2(\mathbb{R}^3, d^3x), \quad \|\psi\| = 1.$$

(3) According to the example in (1), let us focus on the subgroup of $IO(3)$ of displacements along x_1 parametrized by $u \in \mathbb{R}$,

$$\mathbb{R}^3 \ni x \mapsto x + u\mathbf{e}_1,$$

where \mathbf{e}_1 denotes the unit vector in \mathbb{R}^3 along x_1 . For every value of the parameter u , we indicate by s_u the corresponding (Wigner) quantum symmetry, $s_u(\langle \psi | \cdot \rangle \psi) = \langle U_u \psi | \cdot \rangle U_u \psi$ with

$$(U_u \psi)(x) = \psi(x - u\mathbf{e}_1), \quad u \in \mathbb{R},$$

The action of this symmetry on the observable X_k turns out to be

$$s_u^*(X_k) = U_u X_k U_u^{-1} = X_k + u\delta_{k1}I, \quad u \in \mathbb{R}.$$

■

3.6.2. Groups of quantum symmetries

As in (1) in the example above, very often in physics one deals with groups of symmetries. In other words, there is a certain group G , with unit element e and group product \cdot , and one associates each element $g \in G$ to a symmetry s_g (if Kadison or Wigner is immaterial here, in view of the above discussion). In turn, s_g is associated to an operator U_g , unitary or antiunitary.

Remark 140. In the rest of this section, we assume that all the U_g are unitary. ■

It would be nice to fix these operators U_g in order that the map $G \ni g \mapsto U_g$ be a **unitary representation** of G on \mathcal{H} , that is

$$U_e = I, \quad U_g U_{g'} = U_{g \cdot g'} \quad g, g' \in G \quad (92)$$

The identities (91) found in (1) in example 139 shows that it is possible at least in certain cases. In general the requirement (92) does not hold. What we know is that $U_{g \cdot g'}$ equals $U_g U_{g'}$ just *up to phases*:

$$U_g U_{g'} U_{g \cdot g'}^{-1} = \omega(g, g') I \quad \text{with } \omega(g, g') \in U(1) \text{ for all } g, g' \in G. \quad (93)$$

For $g = e$ this identity gives in particular

$$U_e = \omega(e, e) I. \quad (94)$$

The numbers $\omega(g, g')$ are called **multipliers**. They cannot be completely arbitrary, indeed associativity of composition of operators $(U_{g_1} U_{g_2}) U_{g_3} = U_{g_1} (U_{g_2} U_{g_3})$ yields the identity

$$\omega(g_1, g_2) \omega(g_1 \cdot g_2, g_3) = \omega(g_1, g_2 \cdot g_3) \omega(g_2, g_3), \quad g_1, g_2, g_3 \in G \quad (95)$$

which also implies

$$\omega(g, e) = \omega(e, g) = \omega(g', e), \quad \omega(g, g^{-1}) = \omega(g^{-1}, g), \quad g, g' \in G. \quad (96)$$

Definition 141. If G is a group, a map $G \ni g \mapsto U_g$ – where the U_g are unitary operators in the complex Hilbert space \mathcal{H} – is named a **unitary projective representation** of G on \mathcal{H} if (93) holds (so that also (94) and (95) are valid). Moreover,

(i) two unitary projective representation $G \ni g \mapsto U_g$ and $G \ni g \mapsto U'_g$ are said to be **equivalent** if $U'_g = \chi_g U_g$, where $\chi_g \in U(1)$ for every $g \in G$. That is the same as requiring that there are numbers $\chi_h \in U(1)$, if $h \in G$, such that

$$\omega'(g, g') = \frac{\chi_g \chi_{g'}}{\chi_{g \cdot g'}} \omega(g, g') \quad \forall g, g' \in G \quad (97)$$

with obvious notation;

(ii) a unitary projective representation with $\omega(e, e) = \omega(g, e) = \omega(e, g) = 1$ for every $g \in G$ is said to be **normalized**. ■

Remark 142.

(a) It is easily proved that every unitary projective representation is always equivalent to a normalized representation.

(b) It is clear that two projective unitary representations are equivalent if and only if they are made of the same Wigner (or Kadison) symmetries.

(c) In case of superselection rules, continuous symmetries representing a connected topological group do not swap different coherent sectors when acting on pure states [6].

(d) One may wonder if it is possible to construct a group representation $G \ni g \mapsto V_g$ where the operators V_g may be both unitary or antiunitary. If every $g \in G$ can be written as $g = h \cdot h$ for some h depending on g – and this is the case if G is a connected Lie group – all the operators U_g must be unitary because $U_g = U_h U_h$ is necessarily linear no matter if U_h is linear or anti linear. The presence of arbitrary phases does not change the result. ■

Given a unitary projective representation, a technical problem is to check if it is equivalent to a unitary representation, because unitary representations are much simpler to handle. This is a difficult problem [11,6] which is tackled especially when G is a *topological group* (or *Lie group*) and the representation satisfies the following natural *continuity property*

Definition 143. A unitary projective representation of the topological group G , $G \ni g \mapsto U_g$ on the Hilbert space \mathcal{H} is said to be **continuous** if the map

$$G \ni g \mapsto |\langle \psi | U_g \phi \rangle|$$

is continuous for every $\psi, \phi \in \mathcal{H}$. ■

The notion of continuity defined above is natural as it regards continuity of probability transitions. A well known co-homological condition assuring that a unitary projective representation of Lie groups is equivalent to a unitary one is due to Bargmann [24,6].

Theorem 144 (Bargmann’s criterion). *Let G be a connected and simply connected (real finite dimensional) Lie group with Lie algebra \mathfrak{g} . Every continuous unitary projective representation of G in a complex Hilbert space is equivalent to a strongly continuous unitary representation of G if, for every bilinear antisymmetric map $\Theta : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{R}$ such that*

$$\Theta([u, v], w) + \Theta([v, w], u) + \Theta([w, u], v) = 0, \quad \forall u, v, w \in \mathfrak{g}$$

there is a linear map $\alpha : \mathfrak{g} \rightarrow \mathbb{R}$ such that $\Theta(u, v) = \alpha([u, v])$, for all $u, v \in \mathfrak{g}$.

Remark 145. The condition is equivalent to require that the second cohomology group $H^2(G, \mathbb{R})$ is trivial. $SU(2)$ for instance satisfies the requirement. ■

However, non-unitarisable unitary projective representations do exist and one has to deal with them. There is nevertheless a way to circumvent the technical problem. Given a unitary projective representation $G \ni g \mapsto U_g$ with multipliers ω , let us put on $U(1) \times G$ the group structure arising by the product \circ

$$(\chi, g) \circ (\chi', g') = (\chi \chi' \omega(g, g'), g \cdot g')$$

and indicate by \hat{G}_ω the obtained group. The map

$$\hat{G}_\omega \ni (\chi, g) \mapsto \chi U_g =: V_{(\chi, g)}$$

is a *unitary representation* of \hat{G}_ω . If the initial representation is normalized, \hat{G}_ω is said to be a **central extension of G** by means of $U(1)$ [11,6]. Indeed, the elements (χ, e) , $\chi \in U(1)$, commute with all the elements of \hat{G}_ω and thus they belong to the centre of the group.

Remark 146. These types of unitary representations of central extensions play a remarkable role in physics. Sometimes \hat{G}_ω with a particular choice for ω is seen as the *true* group of symmetries at quantum level, when G is the *classical group* of symmetries. There is a very important case. If G is the *Galileian group* – the group of transformations between inertial reference frames in classical physics, viewed as *active* transformations – as clarified by Bargmann [6] the only physically relevant unitary projective representations in QM are just the ones which are *not* equivalent to unitary representations! The multipliers embody the information about the *mass* of the system. This phenomenon gives also rise to a famous superselection structure in the Hilbert space of quantum systems admitting the Galileian group as a symmetry group, known as *Bargmann’s superselection rule* [6]. ■

To conclude we just state a technically important result [6] which introduces the one-parameter strongly continuous unitary groups as crucial tool in QM.

Theorem 147. *Let $\gamma : \mathbb{R} \ni r \mapsto U_r$ be a continuous unitary projective representation of the additive topological group \mathbb{R} on the complex Hilbert space \mathcal{H} . The following facts hold.*

- (a) γ is equivalent to a strongly continuous unitary representation $\mathbb{R} \ni r \mapsto V_r$ of the same topological additive group on \mathcal{H} .
- (b) A strongly continuous unitary representation $\mathbb{R} \ni r \mapsto V'_r$ is equivalent to γ if and only if

$$V'_r = e^{icr} V_r$$

for some constant $c \in \mathbb{R}$ and all $r \in \mathbb{R}$.

The above unitary representation can also be defined as *strongly continuous one-parameter unitary group*.

Definition 148. If \mathcal{H} is a Hilbert space, $V : \mathbb{R} \ni r \mapsto V_r \in \mathfrak{B}(\mathcal{H})$, such that

- (i) V_r is unitary for every $r \in \mathbb{R}$
- (ii) $V_r V_s = V_{r+s}$ for all $r, s \in \mathbb{R}$,

is called **one-parameter unitary group**. It is called **strongly continuous one-parameter unitary group** if in addition to (i) and (ii) we also have

- (iii) V is continuous referring to the strong operator topology. In other words $V_r \psi \rightarrow V_{r_0} \psi$ for $r \rightarrow r_0$ and every $r_0 \in \mathbb{R}$ and $\psi \in \mathcal{H}$. ■

Remark 149.

- (a) It is evident that, in view of the group structure, a one-parameter unitary group $\mathbb{R} \ni r \mapsto V_r \in \mathfrak{B}(\mathcal{H})$ is strongly continuous if and only if it is strongly continuous for $r = 0$.

(b) It is a bit less evident but true that a one-parameter unitary group $\mathbb{R} \ni r \mapsto V_r \in \mathfrak{B}(\mathcal{H})$ is strongly continuous if and only if it is *weakly continuous* at $r = 0$. Indeed, if V is weakly continuous at $r = 0$, for every $\psi \in \mathcal{H}$, we have

$$\|U_r\psi - \psi\|^2 = \|U_r\psi\|^2 + \|\psi\|^2 - \langle \psi | U_r\psi \rangle - \langle U_r\psi | \psi \rangle = 2\|\psi\|^2 - \langle \psi | U_r\psi \rangle - \langle U_r\psi | \psi \rangle \rightarrow 0$$

for $r \rightarrow 0$. ■

3.6.3. One-parameter strongly continuous unitary groups: von Neumann and Stone theorems

Theorem 147 establishes that, dealing with continuous unitary projective representation of the additive topological group \mathbb{R} , one can always reduce to work with proper strongly continuous one-parameter unitary groups. So, for instance, the action on a quantum system of rotations around an axis can always be described by means of strongly continuous one-parameter unitary groups. There is a couple of technical results of very different nature which are very useful in QM. The former is due to von Neumann [6] and proves that the one-parameter unitary group which are not strongly continuous are not so many in separable Hilbert spaces.

Theorem 150. *If \mathcal{H} is a separable complex Hilbert space and $V : \mathbb{R} \ni r \mapsto V_r \in \mathfrak{B}(\mathcal{H})$ is a one parameter unitary group, it is strongly continuous if and only if the maps $\mathbb{R} \ni r \mapsto \langle \psi | U_r\phi \rangle$ are Borel measurable for all $\psi, \phi \in \mathcal{H}$.*

The second proposition we quote [6] is a celebrated result due to Stone (and later extend to the famous *Hille-Yosida theorem* in Banach spaces). We start by noticing that, if A is a selfadjoint operator in a Hilbert space, $U_t := e^{itA}$, for $t \in \mathbb{R}$, defines a strongly continuous one-parameter unitary group as one easily proves using the functional calculus. The result is remarkably reversible.

Theorem 151 (Stone theorem). *Let $\mathbb{R} \ni t \mapsto U_t \in \mathfrak{B}(\mathcal{H})$ be a strongly continuous one-parameter unitary group in the complex Hilbert space \mathcal{H} . The following facts hold.*

(a) *There exists a unique selfadjoint operator, called the **generator** of the group, $A : D(A) \rightarrow \mathcal{H}$ in \mathcal{H} , such that*

$$U_t = e^{-itA}, \quad t \in \mathbb{R}. \quad (98)$$

(b) *The generator is determined as*

$$A\psi = i \lim_{t \rightarrow 0} \frac{1}{t} (U_t - I)\psi \quad (99)$$

and $D(A)$ is made of the vectors $\psi \in \mathcal{H}$ such that the right hand side of (99) exists in \mathcal{H} .

(c) *$U_t(D(A)) \subset D(A)$ for all $t \in \mathbb{R}$ and*

$$AU_t\psi = U_tA\psi \quad \text{if } \psi \in D(A) \text{ and } t \in \mathbb{R}.$$

Remark 152.

(a) For a selfadjoint operator A , the expansion

$$e^{-itA}\psi = \sum_{n=0}^{+\infty} \frac{(-it)^n}{n!} A^n \psi$$

generally does *not* work for $\psi \in D(A)$. It works in two cases however: (i) if ψ is an *analytic* vector of A (Def. 51 and this result is due to Nelson), (ii) if $A \in \mathfrak{B}(\mathcal{H})$ which is equivalent to say that $D(A) = \mathcal{H}$. In the latter case, one more strongly finds $e^{-itA} = \sum_{n=0}^{+\infty} \frac{(-it)^n}{n!} A^n$, referring to the uniform operator topology. [6].

(b) One parameter unitary group generated by selfadjoint operators can be used to check if the associated observables are compatible in view of the following nice result [6].

Proposition 153. *If A and B are selfadjoint operators in the complex Hilbert space \mathcal{H} , the identity holds*

$$e^{-itA}e^{-isB} = e^{-isB}e^{-itA} \quad \forall t, s \in \mathbb{R}$$

if and only if the spectral measures of A and B commute. ■

3.6.4. Time evolution, Heisenberg picture and quantum Noether theorem

Consider a quantum system described in the Hilbert space \mathcal{H} when an inertial reference frame is fixed. Suppose that, physically speaking, the system is either isolated or interacts with some external stationary environment. With these hypotheses, the time evolution of states is axiomatically described by a continuous symmetry, more precisely, by a continuous one-parameter group of unitary projective operators $\mathbb{R} \ni t \mapsto V_t$. In view of Theorems 147 and 151, this group is equivalent to a strongly continuous one-parameter group of unitary operators $\mathbb{R} \ni t \mapsto U_t$ and, up to additive constant, there is a unique selfadjoint operator H , called the **Hamiltonian operator** such that (notice the sign in front of the exponent)

$$U_t = e^{-itH}, \quad t \in \mathbb{R}. \tag{100}$$

The observable represented by H is usually identified with *the energy of the system* in the considered reference frame.

Within this picture, if $\rho \in \mathfrak{S}(\mathcal{H})$ is the state of the system at $t = 0$, as usual described by a positive trace-class operator with unit trace, the state at time t is $\rho_t = U_t \rho U_t^{-1}$. If the initial state is pure and represented by the unit vector $\psi \in \mathcal{H}$, the state at time t is $\psi_t := U_t \psi$. In this case, if $\psi \in D(H)$ we have that $\psi_t \in D(H)$ for every $t \in \mathbb{R}$ in view of (c) in Theorem 151 and furthermore, for (b) of the same theorem

$$-iH\psi_t = \frac{d\psi_t}{dt}. \tag{101}$$

where the derivative is computed with respect to the topology of \mathcal{H} . One recognises in Eq. (101) the general form of **Schödinger equation**.

Remark 154. It is possible to study quantum systems interacting with some external system which is not stationary. In this case the Hamiltonian observable depends parametrically on time as already introduced in remark 8. In these cases a Schrödinger equation is assumed to describe the time evolution of the system giving rise to a groupoid of unitary operators [6]. We shall not enter into the details of this technical issue here. ■

Adopting the above discussed framework, observables do not evolve and states do. This framework is called **Schrödinger picture**. There is however another approach to describe time evolution called **Heisenberg picture**. In this representation states do not evolve in time but observables do. If A is an observable at $t = 0$, its evolution at time t is the observable

$$A_t := U_t^{-1} A U_t .$$

Obviously $D(A_t) = U_t^{-1}(D(A)) = U_{-t}(D(A)) = U_t^*(D(A))$. According with (i) in Proposition 82 the spectral measure of A_t is

$$P_E^{(A_t)} = U_t^{-1} P_E^{(A)} U_t$$

as expected. The probability that, at time t , the observable A produces the outcome E when the state is ρ at $t = 0$, can equivalently be computed both using the standard picture, where states evolve as $\text{tr}(P_E^{(A)} \rho_t)$, or Heisenberg picture where observables do obtaining $\text{tr}(P_E^{(A_t)} \rho)$. Indeed

$$\text{tr}(P_E^{(A)} \rho_t) = \text{tr}(P_E^{(A)} U_t^{-1} \rho U_t) = \text{tr}(U_t P_E^{(A)} U_t^{-1} \rho) = \text{tr}(P_E^{(A_t)} \rho) .$$

The two pictures are completely equivalent to describe physics. Heisenberg picture permits to give the following important definition

Definition 155. In the complex Hilbert space \mathcal{H} equipped with a strongly continuous unitary one-parameter group representing the time evolution $\mathbb{R} \ni t \mapsto U_t$, an observable represented by the selfadjoint operator A is said to be a **constant of motion** with respect to U , if $A_t = A_0$.

The meaning of the definition should be clear: Even if the state evolve, the probability to obtain an outcome E , measuring a constant of motion A , remains stationary. Also expectation values and standard deviations do not change in time.

We are now in a position to state the equivalent of the *Noether theorem* in QM.

Theorem 156 (Noether quantum theorem). *Consider a quantum system described in the complex Hilbert space \mathcal{H} equipped with a strongly continuous unitary one-parameter group representing the time evolution $\mathbb{R} \ni t \mapsto U_t$. If A is an observable represented by a (generally unbounded) selfadjoint operator A in \mathcal{H} , the following facts are equivalent.*

- (a) A is a constant of motion: $A_t = A_0$ for all $t \in \mathbb{R}$.

(b) The one-parameter group of symmetries generated by A , $\mathbb{R} \ni s \mapsto e^{-isA}$ is a **group of dynamical symmetries**: It commutes with time evolution

$$e^{-isA}U_t = U_t e^{-isA} \quad \text{for all } s, t \in \mathbb{R}. \quad (102)$$

In particular transforms evolutions of pure states into evolutions of (other) pure states, i.e., $e^{-isA}U_t\psi = U_t e^{-isA}\psi$.

(c) The action on observables (90) of the one-parameter group of symmetries generated by A , $\mathbb{R} \ni s \mapsto e^{isA}$ leaves H invariant. That is

$$e^{-isA}He^{isA} = H, \quad \text{for all } s \in \mathbb{R}.$$

Proof. Suppose that (a) holds. By definition $U_t^{-1}AU_t = A$. By (i) in Proposition 82 we have that $U_t^{-1}e^{-isA}U_t = e^{-isA}$ which is equivalent to (b). If (b) is true, we have that $e^{-isA}e^{-itH}e^{isA} = e^{-itH}$. Here an almost direct application of Stone theorem yields $e^{-isA}He^{isA} = H$. Finally suppose that (c) is valid. Again (i) in Proposition 82 produces $e^{-isA}U_t e^{isA} = U_t$ which can be rearranged into $U_t^{-1}e^{-isA}U_t = e^{-isA}$. Finally Stone theorem leads to $U_t^{-1}AU_t = A$ which is (a), concluding the proof. \square

Remark 157.

(a) In physics textbooks the above statements are almost always stated using time derivatives and commutators. This is useless and involves many subtle troubles with domains of the involved operators.

(b) The theorem can be extended to observables $A(t)$ parametrically depending on time already in the Schrödinger picture [6]. In this case (a) and (b) are equivalent too. With this more general situation, (102) in (b) has to be re-written as

$$e^{-isA(t)}U_t = U_t e^{-isA(0)} \quad \text{for all } s, t \in \mathbb{R}$$

and Heisenberg evolution considered in (a) encompasses both time dependences

$$A_t = U_t^{-1}A(t)U_t.$$

At this juncture, (c) can similarly be stated but, exactly as it happens in Hamiltonian classical mechanics, it has a more complicated interpretation [6].

An example is the generator of the boost one-parameter subgroup along the axis \mathbf{n} of transformations of the Galileian group $\mathbb{R}^3 \ni x \mapsto x + tv\mathbf{n} \in \mathbb{R}^3$, where the speed $v \in \mathbb{R}$ is the parameter of the group. The generator is [6] the unique self adjoint extension of

$$K_{\mathbf{n}}(t) = \sum_{j=1}^3 n_j(mX_j|_D - tP_j|_D), \quad (103)$$

the constant $m > 0$ denoting the mass of the system and D being the Gårding or the Nelson domain of the representation of (central extension of the) Galileian group as we will discuss later.

(c) In QM there are symmetries described by operators which are simultaneously selfadjoint and unitary, so they are also observables and can be measured. The

parity is one of them: $(\mathcal{P}\psi)(x) := \psi(-x)$ for a particle described in $L^2(\mathbb{R}^3, d^3x)$. These are constants of motion ($U_t^{-1}\mathcal{P}U_t = \mathcal{P}$) if and only if they are dynamical symmetries ($\mathcal{P}U_t = \mathcal{P}U_t$). This phenomenon has no classical corresponding.

(d) The *time reversal symmetry*, when described by an anti unitary operator T is supposed to satisfy: $THT^{-1} = H$. However, since it is antilinear gives rise to the identity (exercise) $Te^{-itH}T^{-1} = e^{-itTHT^{-1}}$, so that $TU_t = U_{-t}T$ as physically expected. There is no conserved quantity associated with this operator because it is not selfadjoint.

Exercise 158.

(1) *Prove that if the Hamiltonian observable does not depend on time is a constant of motion.*

Solution. In this case the time translation is described by $U_t = e^{itH}$ and trivially it commutes with U_s . Noether theorem implies the thesis \square

(2) *Prove that for the free particle in \mathbb{R}^3 the momentum along x_1 is a constant of motion as consequence of translational invariance along that axis. Assume that the unitary group representing translations along x_1 is U_u with $(U_u\psi)(x) = \psi(x - u\mathbf{e}_1)$ if $\psi \in L^2(\mathbb{R}^3, d^3x)$.*

Solution. The Hamiltonian is $H = \frac{1}{2m} \sum_{j=1}^3 P_j^2$. It commutes with the one-parameter unitary group describing displacements along x_1 , because as one can prove the said groups is generated by P_1 itself: $U_u := e^{-iuP_1}$. Theorem 156 yields the thesis. \square

(3) *Prove that if $\sigma(H)$ is bounded below but not above, the time reversal symmetry cannot be unitary.*

Solution. We look for an operator, unitary or antiunitary such that $TU_t = U_{-t}T$ for all $t \in \mathbb{R}$. If the operator is unitary, the said identity easily implies $THT^{-1} = -H$ and therefore, with obvious notation, $\sigma(THT^{-1}) = -\sigma(H)$. (e) in remark 57 immediately yields $\sigma(H) = -\sigma(H)$ which is false if $\sigma(H)$ is bounded below but not above. \square

3.6.5. *Strongly continuous unitary representations of Lie groups, Nelson theorem*

Topological and Lie groups are intensively used in QM [24]. More precisely they are studied in terms of their strongly continuous unitary representations. The reason to consider strongly continuous representations is that they immediately induce continuous representations of the group in terms of quantum symmetries (Def. 143). In the rest of the section we consider only the case of a *real Lie group*, G , whose Lie algebra is indicated by \mathfrak{g} endowed with the Lie bracket or commutator $\{ , \}$.

Definition 159. If G is a Lie group, a **strongly continuous unitary representation** of G over the complex Hilbert space \mathcal{H} is a group homomorphism $G \ni g \mapsto U_g \in \mathfrak{B}(\mathcal{H})$ such that every U_g is unitary and $U_g \rightarrow U_{g_0}$, in the strong

operator topology, if $g \rightarrow g_0$. ■

We leave to the reader the elementary proof that strong continuity is equivalent to strong continuity at the unit element of the group and in turn, this is equivalent to weak continuity at the unit element of the group.

A fundamental technical fact is that the said unitary representations are associated with representations of the Lie algebra of the group in terms of (anti)selfadjoint operators. These operators are often physically interpreted as constants of motion (generally parametrically depending on time) when the Hamiltonian of the system belongs to the representation of the Lie algebra. We want to study this relation between the representation of the group on the one hand and the representation of the Lie algebra on the other hand. First of all we define the said operators representing the Lie algebra.

Definition 160. Let G be a real Lie group and consider a strongly continuous unitary representation U of G over the complex Hilbert space \mathcal{H} .

If $A \in \mathfrak{g}$ let $\mathbb{R} \ni t \mapsto \exp(tA) \in G$ be the generated one-parameter Lie subgroup. The **self-adjoint generator associated with A**

$$A : D(A) \rightarrow \mathcal{H}$$

is the generator of the strongly continuous one-parameter unitary group

$$\mathbb{R} \ni t \mapsto U_{\exp\{tA\}} = e^{-isA}$$

in the sense of Theorem 151. ■

The expected result is that these generators (with a factor $-i$) define a representation of the Lie algebra of the group. The utmost reason is that they are associated to the unitary one-parameter subgroups exactly as the elements of the Lie algebra are associated to the Lie one-parameter subgroups. In particular we expect that the Lie parenthesis correspond to the commutator of operators. The technical problem is that the generators A may have different domains. Thus we look for a common invariant (because the commutator must be defined thereon) domain, where all them can be defined. This domain should embody all the amount of information about the operators A themselves, disregarding the fact that they are defined in larger domains. In other words we would like that the domain be a *core* ((3) in Def. 35) for each generator. There are several candidates for this space, one of the most appealing is the so called *Gårding space*.

Definition 161. Let G be a (finite-dimensional real) Lie group and consider a strongly continuous unitary representation U of G over the complex Hilbert space \mathcal{H} . If $f \in C_0^\infty(G; \mathbb{C})$ and $x \in \mathcal{H}$, define

$$x[f] := \int_G f(g)U_g x dg \tag{104}$$

where dg denotes the Haar measure over G and the integration is defined in a weak sense exploiting Riesz' lemma: Since the map $\mathcal{H} \ni x \mapsto \int_G f(g) \langle y | U_g x \rangle dg$ is continuous (the proof being elementary), $x[f]$ is the unique vector in \mathcal{H} such that

$$\langle y | x[f] \rangle = \int_G f(g) \langle y | U_g x \rangle dg, \quad \forall y \in \mathcal{H}.$$

The complex span of all vectors $x[f] \in \mathcal{H}$ with $f \in C_0^\infty(G; \mathbb{C})$ and $x \in \mathbb{H}$ is called **Gårding space** of the representation and is denoted by $D_G^{(U)}$. ■

The subspace $D_G^{(U)}$ enjoys very remarkable properties we state in the next theorem. In the following $L_g : C_0^\infty(G; \mathbb{C}) \rightarrow C_0^\infty(G; \mathbb{C})$ denotes the standard left-action of $g \in G$ on complex valued smooth compactly supported functions defined on G :

$$(L_g f)(h) := f(g^{-1}h) \quad \forall h \in G, \quad (105)$$

and, if $A \in \mathfrak{g}$, $X_A : C_0^\infty(G; \mathbb{C}) \rightarrow C_0^\infty(G; \mathbb{C})$ is the smooth vector field over G (a smooth differential operator) defined as:

$$(X_A(f))(g) := \lim_{t \rightarrow 0} \frac{f(\exp\{-tA\}g) - f(g)}{t} \quad \forall g \in G. \quad (106)$$

so that that map

$$\mathfrak{g} \ni A \mapsto X_A \quad (107)$$

defines a representation of \mathfrak{g} in terms of vector fields (differential operators) on $C_0^\infty(G; \mathbb{C})$. We conclude with the following theorem [24], establishing that the Gårding space has all the expected properties.

Theorem 162. *Referring to Definitions 160 and 161, the Gårding space $D_G^{(U)}$ satisfies the following properties.*

(a) $D_G^{(U)}$ is dense in \mathcal{H} .

(b) If $g \in G$, then $U_g(D_G^{(U)}) \subset D_G^{(U)}$. More precisely, if $f \in C_0^\infty(G)$, $x \in \mathcal{H}$, $g \in G$, it holds

$$U_g x[f] = x[L_g f]. \quad (108)$$

(c) If $A \in \mathfrak{g}$, then $D_G^{(U)} \subset D(A)$ and furthermore $A(D_G^{(U)}) \subset D_G^{(U)}$. More precisely

$$-iAx[f] = x[X_A(f)] \quad (109)$$

(d) The map

$$\mathfrak{g} \ni A \mapsto -iA|_{D_G^{(U)}} =: U(A) \quad (110)$$

is a Lie algebra representation in terms of anti symmetric operators defined on the common dense invariant domain $D_G^{(U)}$. In particular if $\{ , \}$ is the Lie commutator of \mathfrak{g} we have:

$$[U(A), U(A')] = U(\{A, A'\}) \quad \text{if } A, A' \in \mathfrak{g}.$$

(e) $D_G^{(U)}$ is a core for every selfadjoint generator A with $A \in \mathfrak{g}$, that is

$$A = \overline{A|_{D_G^{(U)}}}, \quad \forall A \in \mathfrak{g}. \quad (111)$$

Now we tackle the inverse problem: We suppose to have a certain representation of a Lie algebra \mathfrak{g} in terms of symmetric operators defined in common invariant domain of a complex Hilbert space \mathcal{H} . We are interested in lifting this representation to a whole strongly continuous representation of the unique simply connected Lie group G admitting \mathfrak{g} as Lie algebra. This is a much more difficult problem solved by Nelson.

Given a strongly continuous representation U of a (real) Lie group G , there is another space $D_N^{(U)}$ with similar features to $D_G^{(U)}$. Introduced by Nelson [24], it turns out to be more useful than the Gårding space to *recover* the representation U by exponentiating the Lie algebra representation.

By definition $D_N^{(U)}$ consists of vectors $\psi \in \mathcal{H}$ such that $G \ni g \mapsto U_g \psi$ is *analytic* in g , i.e. expansible in power series in (real) analytic coordinates around every point of G . The elements of $D_N^{(U)}$ are called **analytic vectors of the representation** U and $D_N^{(U)}$ is the **space of analytic vectors of the representation** U . It turns out that $D_N^{(U)}$ is invariant for every U_g , $g \in G$.

A remarkable relationship exists between analytic vectors in $D_N^{(U)}$ and analytic vectors according to Definition 51. Nelson proved the following important result [24], which implies that $D_N^{(U)}$ is dense in \mathcal{H} , as we said, because analytic vectors for a self-adjoint operator are dense (exercise 89). An operator is introduced, called *Nelson operator*, that sometimes has to do with the *Casimir operators* [24] of the represented group.

Proposition 163. *Let G be a (finite dimensional real) Lie group and $G \ni g \mapsto U_g$ a strongly continuous unitary representation on the Hilbert space \mathcal{H} . Take $A_1, \dots, A_n \in \mathfrak{g}$ a basis and define **Nelson's operator** on $D_G^{(U)}$ by*

$$\Delta := \sum_{k=1}^n U(A_k)^2,$$

where the $U(A_k)$ are, as before, the selfadjoint generators A_k restricted to the Gårding domain $D_G^{(U)}$. Then

- (a) Δ is essentially selfadjoint on $D_G^{(U)}$.
- (b) Every analytic vector of the selfadjoint operator $\overline{\Delta}$ is analytic an element of $D_N^{(U)}$, in particular $D_N^{(U)}$ is dense.
- (c) Every vector in $D_N^{(U)}$ is analytic for every self-adjoint operator $\overline{U(A_k)}$, which is thus essentially selfadjoint in $D_N^{(U)}$ by Nelson's criterion.

We finally state the well-known theorem of Nelson that enables to associate representations of the only simply connected Lie group with a given Lie algebra to representations of that Lie algebra.

Theorem 164 (Nelson theorem). Consider a real n -dimensional Lie algebra V of operators $-iS$ – with each S symmetric on the Hilbert space \mathcal{H} , defined on a common invariant subspace \mathcal{D} dense in \mathcal{H} and V -invariant – with the usual commutator of operators as Lie bracket.

Let $-iS_1, \dots, -iS_n \in V$ be a basis of V and define Nelson's operator with domain \mathcal{D} :

$$\Delta := \sum_{k=1}^n S_k^2.$$

If Δ is essentially self-adjoint, there exists a strongly continuous unitary representation

$$G_V \ni g \mapsto U_g$$

on \mathcal{H} , of the unique simply connected Lie group G_V with Lie algebra V .

U is completely determined by the fact that the closures \overline{S} , for every $-iS \in V$, are the selfadjoint generators of the representation of the one-parameter subgroups of G_V in the sense of Def. 160.

In particular, the symmetric operators S are essentially selfadjoint on \mathcal{D} , their closure being selfadjoint.

Exercise 165. Let A, B be selfadjoint operators in the complex Hilbert space \mathcal{H} with a common invariant dense domain D where they are symmetric and commute. Prove that if $A^2 + B^2$ is essentially self adjoint on D , then the spectral measures of A and B commute.

Solution. Apply Nelson's theorem observing that A, B define the Lie algebra of the additive Abelian Lie group \mathbb{R}^2 and that D is a core for A and B , because they are essentially selfadjoint therein again by Nelson theorem. ■

Example 166.

(1) Exploiting spherical polar coordinates, the Hilbert space $L^2(\mathbb{R}^3, d^3x)$ can be factorised as $L^2([0, +\infty), r^2 dr) \otimes L^2(\mathbb{S}^2, d\Omega)$, where $d\Omega$ is the natural rotationally invariant Borel measure on the sphere \mathbb{S}^2 with unit radius in \mathbb{R}^3 , with $\int_{\mathbb{S}^2} 1 d\Omega = 4\pi$. In particular a Hilbertian basis of $L^2(\mathbb{R}^3, d^3x)$ is therefore made of the products $\psi_n(r)Y_m^l(\theta, \phi)$ where $\{\psi_n\}_{n \in \mathbb{N}}$ is any Hilbertian basis in $L^2([0, +\infty), r^2 dr)$ and $\{Y_m^l \mid l = 0, 1, 2, \dots, m = 0, \pm 1, \pm 2, \dots, \pm l\}$ is the standard Hilbertian basis of spherical harmonics of $L^2(\mathbb{S}^2, d\Omega)$ [24]. Since the function Y_m^l are smooth on \mathbb{S}^2 , it is possible to arrange the basis of ψ_n made of compactly supported smooth functions whose derivatives in 0 vanish at every order, in order that $\mathbb{R}^3 \ni x \mapsto (\psi_n \cdot Y_m^l)(x)$ are elements of $\mathcal{C}^\infty(\mathbb{R}^n; \mathbb{C})$ (and therefore also of $\mathcal{S}(\mathbb{R}^3)$). Now consider the three symmetric operators defined on the common dense invariant domain $\mathcal{S}(\mathbb{R}^3)$

$$\mathcal{L}_k = \sum_{i,j=1}^3 \epsilon_{kij} X_i P_j |_{\mathcal{S}(\mathbb{R}^3)}$$

where ϵ_{ijk} is completely antisymmetric in ijk and $\epsilon_{123} = 1$. By direct inspection one sees that

$$[-i\mathcal{L}_k, -i\mathcal{L}_h] = \sum_{r=1}^3 \epsilon_{khr} (-i\mathcal{L}_r)$$

so that the finite real span of the operators $i\mathcal{L}_k$ is a representation of the Lie algebra of the simply connected real Lie group $SU(2)$ (the universal covering of $SO(3)$). Define the Nelson operator $\mathcal{L}^2 := -\sum_{k=1}^3 \mathcal{L}_k^2$ on $\mathcal{S}(\mathbb{R}^3)$. Obviously this is a symmetric operator. A well known computation proves that

$$\mathcal{L}^2 \psi_n(r) Y_m^l = l(l+1) \psi_n(r) Y_m^l .$$

We conclude that \mathcal{L}^2 admits a Hilbertian basis of eigenvectors. Corollary 53 implies that \mathcal{L}^2 is essentially self adjoint. Therefore we can apply Theorem 164 concluding that there exists a strongly continuous unitary representation $SU(2) \ni M \mapsto U_M$ of $SU(2)$ (actually it can be proved to be also of $SO(3)$). The three selfadjoint operators $L_k := \overline{\mathcal{L}_k}$ are the generators of the one-parameter of rotations around the corresponding three orthogonal Cartesian axes x_k , $k = 1, 2, 3$. The one-parameter subgroup of rotations around the generic unit vector \mathbf{n} , with components n_k , admits the selfadjoint generator $L_{\mathbf{n}} = \sum_{k=1}^3 n_k \mathcal{L}_k$. The observable $L_{\mathbf{n}}$ has the physical meaning of the *\mathbf{n} -component of the angular momentum* of the particle described in $L^2(\mathbb{R}^3, d^3x)$. It turns out that, for $\psi \in L^2(\mathbb{R}^3, d^3x)$,

$$(U_M \psi)(x) = \psi(\pi(M)^{-1}x) , \quad M \in SU(2) \quad , x \in \mathbb{R}^3 \quad (112)$$

where $\pi : SU(2) \rightarrow SO(3)$ is the standard covering map. (112) is the action of the rotation group on pure states in terms of quantum symmetries. This representation is, in fact, a subrepresentation of the unitary representation of $IO(3)$ already found in (1) of example 139.

(2) Given a quantum system, a quite general situation is the one where the quantum symmetries of the systems are described by a strongly continuous representation $V : G \ni g \mapsto V_g$ on the Hilbert space \mathcal{H} of the system, and the time evolution is the representation of a one-parameter Lie subgroup with generator $\mathbf{H} \in \mathfrak{g}$. So that

$$V_{\exp(t\mathbf{H})} = e^{-it\mathbf{H}} =: U_t .$$

This is the case, for instance, of relativistic quantum particles, where G is the *special orthochronous Lorentz group*, $SO(1, 3)_+$, (or its universal covering $SL(2, \mathbb{C})$). Describing non-relativistic quantum particles, the relevant group G is an $U(1)$ central extension of the universal covering of the (connected orthochronous) Galileian group.

In this situation, every element of \mathfrak{g} determines a constant of motion. Actually there are two cases.

(i) If $\mathbf{A} \in \mathfrak{g}$ and $\{\mathbf{H}, \mathbf{A}\} = 0$, then the Lie subgroups $\exp(t\mathbf{H})$ and $\exp(s\mathbf{A})$ commute as, for example, follows from *Baker-Campbell-Hausdorff* formula (see [24,6], for instance). Consequently \mathbf{A} is a constant of motion because $V_{\exp(t\mathbf{H})} = e^{-it\mathbf{H}}$ and

$V_{\exp(sA)} = e^{-isA}$ commute as well and Theorem 156 is valid. In this case e^{-isA} defines a dynamical symmetry in accordance with the afore-mentioned theorem. This picture applies in particular, referring to a free particle, to $A = J_{\mathbf{n}}$, the observable describing total angular momentum along the unit vector \mathbf{n} computed in an inertial reference frame.

(ii) A bit more complicated is the case of $A \in \mathfrak{g}$ with $\{H, A\} \neq 0$. However, even in this case A defines a constant of motion in terms of selfadjoint operators (observables) belonging to the representation of the Lie algebra of G . The difference with respect to the previous case is that, now, the constant of motion *parametrically depend on time*. We therefore have a class of observables $\{A(t)\}_{t \in \mathbb{R}}$ in the Schrödinger picture, in accordance with (b) in remark 157, such that $A_t := U_t^{-1}A(t)U_t$ are the corresponding observables in the Heisenberg picture. The equation stating that we have a constant of motion is therefore $A_t = A_0$.

Exploiting the natural action of the Lie one-parameters subgroups on \mathfrak{g} , let us define the time parametrised class of elements of the Lie algebra

$$A(t) := \exp(tH)A \exp(-tH) \in \mathfrak{g}, \quad t \in \mathbb{R}.$$

If $\{A_k\}_{k=1, \dots, n}$ is a basis of \mathfrak{g} , it must consequently hold

$$A(t) = \sum_{k=1}^n a_k(t)A_k \quad (113)$$

for some real-valued smooth functions $a_k = a_k(t)$. By construction, the corresponding class of selfadjoint generators $A(t)$, $t \in \mathbb{R}$, define a parametrically time dependent constant of motion. Indeed, since (exercise)

$$\exp(s \exp(tH)A \exp(-tH)) = \exp(tH) \exp(sA) \exp(-tH),$$

we have

$$\begin{aligned} -iA(t) &= \frac{d}{ds} \Big|_{s=0} V_{\exp(tH)A \exp(-tH)} = \frac{d}{ds} \Big|_{s=0} V_{\exp(tH) \exp(sA) \exp(-tH)} \\ &= \frac{d}{ds} \Big|_{s=0} V_{\exp(tH)} V_{\exp(sA)} V_{\exp(-tH)} = -iU_t A U_t^{-1} \end{aligned}$$

Therefore

$$A_t = U_t^{-1}A(t)U_t = U_t^{-1}U_t A U_t^{-1}U_t = A = A_0.$$

In view of Theorem 162, as the map $\mathfrak{g} \ni A \mapsto A|_{D_G^{(V)}}$ is a Lie algebra isomorphism, we can recast (113) for selfadjoint generators

$$A(t)|_{D_G^{(V)}} = \sum_{k=1}^n a_k(t)A_k|_{D_G^{(V)}} \quad (114)$$

(where $D_G^{(V)}$ may be replaced by $D_N^{(V)}$ as the reader can easily establish, taking advantage of Proposition 163 and Theorem 164). Since $D_G^{(V)}$ (resp. $D_N^{(V)}$) is a core

for $A(t)$, it also hold

$$A(t) = \overline{\sum_{k=1}^n a_k(t) A_k|_{D_G^{(V)}}}, \quad (115)$$

the bar denoting the closure of an operator as usual. (The same is true replacing $D_G^{(V)}$ for $D_N^{(V)}$.) An important case, both for the non-relativistic and the relativistic case is the selfadjoint generator $K_{\mathbf{n}}(t)$ associated with the boost transformation along the unit vector $\mathbf{n} \in \mathbb{R}^3$, the rest space of the inertial reference frame where the boost transformation is viewed as an active transformation. In fact, referring to the Lie generators of (a $U(1)$ central extension of the universal covering of the connected orthochronous) Galileian group, we have $\{h, k_{\mathbf{n}}\} = -p_{\mathbf{n}} \neq 0$, where $p_{\mathbf{n}}$ is the generator of spatial translations along \mathbf{n} , corresponding to the observable momentum along the same axis when passing to selfadjoint generators. The non-relativistic expression of $K_{\mathbf{n}}(t)$, for a single particle, appears in (103). For a more extended discussion on the non-relativistic case see [6]. A pretty complete discussion including the relativistic case is contained in [24]. ■

3.6.6. Selfadjoint version of Stone - von Neumann - Mackey Theorem

A remarkable consequence of Nelson's theorem is a selfadjoint operator version of Stone-von Neumann theorem usually formulated in terms of unitary operators [1,6], proving that the CCRs always give rise to the standard representation in $L^2(\mathbb{R}^n, d^n x)$. We state and prove this version of the theorem, adding a last statement which is the selfadjoint version of Mackey completion to Stone von Neumann statement [6].

Theorem 167 (Stone - von Neumann - Mackey Theorem). *Let \mathcal{H} be a complex Hilbert space and suppose that there are $2n$ selfadjoint operators in \mathcal{H} we indicate with Q_1, \dots, Q_n and M_1, \dots, M_n such the following requirements are valid.*

(1) *There is a common dense invariant subspace $D \subset \mathcal{H}$ where the CCRs hold*
 $[Q_h, M_k]\psi = i\hbar\delta_{hk}\psi, \quad [Q_h, Q_k]\psi = 0, \quad [M_h, M_k]\psi = 0 \quad \psi \in D, \quad h, k = 1, \dots, n.$ (116)

(2) *The representation is irreducible, in the sense that there is no closed subspace $\mathcal{K} \subset \mathcal{H}$ such that $Q_k(\mathcal{K} \cap D(Q_k)) \subset \mathcal{K}$ and $M_k(\mathcal{K} \cap D(M_k)) \subset \mathcal{K}$ for $k = 1, \dots, n$.*

(3) *The operator $\sum_{k=1}^n Q_k^2|_D + M_k^2|_D$ is essentially self adjoint.*
Under these conditions, there is a Hilbert space isomorphism, that is a surjective isometric map, $U : \mathcal{H} \rightarrow L^2(\mathbb{R}^n, d^n x)$ such that

$$UQ_kU^{-1} = X_k \quad \text{and} \quad UM_kU^{-1} = P_k \quad k = 1, \dots, n \quad (117)$$

where X_k and P_k respectively are the standard position (35) and momentum (36) selfadjoint operators in $L^2(\mathbb{R}^n, d^n x)$. In particular \mathcal{H} results to be separable.

If (1), (2) and (3) are valid with the exception that the representation is not reducible, then \mathcal{H} decomposes into an orthogonal Hilbertian sum $\mathcal{H} = \bigoplus_{r \in R} \mathcal{H}_r$ where R is finite or countable if \mathcal{H} is separable, the $\mathcal{H}_r \subset \mathcal{H}$ are closed subspaces with

$$Q_k(\mathcal{H}_r \cap D(Q_k)) \subset \mathcal{H}_r \quad \text{and} \quad M_k(\mathcal{H}_r \cap D(M_k)) \subset \mathcal{H}_r$$

for all $r \in R$, $k = 1, \dots, n$ and the restrictions of all the Q_k and M_k to each \mathcal{H}_r satisfy (117) for suitable surjective isometric maps $U_r : \mathcal{H}_r \rightarrow L^2(\mathbb{R}^n, d^n x)$

Proof. If (1) holds, the restrictions to D of the selfadjoint operators Q_k , M_k define symmetric operators (since they are selfadjoint and D is dense and included in their domains), also their powers are symmetric since D is invariant. If also (2) is valid, in view of Nelson theorem (since evidently the symmetric operator $I|_D^2 + \sum_{k=1}^n Q_k^2|_D + M_k^2|_D$ is essentially selfadjoint if $\sum_{k=1}^n Q_k^2|_D + M_k^2|_D$ is), there is a strongly continuous unitary representation $W \ni g \mapsto V_g \in \mathfrak{B}(\mathcal{H})$ of the simply connected $2n + 1$ -dimensional Lie group W whose Lie algebra is defined by (116) (correspondingly re-stated for the operators $-iI, -iQ_k, -iM_k$) together with $[-iQ_k, -iI] = [-iM_k, -iI] = 0$, where the operator $-iI$ restricted to D is the remaining Lie generator. W is the *Weyl-Heisenberg* group [6]. The selfadjoint generators of this representation are just the operators Q_k and P_k (and I), since they coincide with the closure of their restrictions to D , because they are selfadjoint (so they admit unique selfadjoint extensions) and D is a core. If furthermore the Lie algebra representation is irreducible, the unitary representation is irreducible, too: If \mathcal{K} were an invariant subspace for the unitary operators, Stone theorem would imply that \mathcal{K} be also invariant under the selfadjoint generators of the one parameter Lie subgroups associated to each Q_k and P_k . This is impossible if the Lie algebra representation is irreducible as we are assuming. The standard version of Stone-von Neumann theorem [6] implies that there is isometric surjective operator $U : \mathcal{H} \rightarrow L^2(\mathbb{R}^n, d^n x)$ such that $W \ni g \mapsto UV_gU^{-1} \in \mathfrak{B}(L^2(\mathbb{R}^n, d^n x))$ is the standard unitary representation of the group W in $L^2(\mathbb{R}^n, d^n x)$ generated by X_k and P_k (and I) [6]. Again, Stone theorem immediately yields (117). The last statement easily follows from the standard form of Mackey's theorem completing Stone-von Neumann result [6]. \square

Remark 168.

(a) The result *a posteriori* gives, in particular, a strong justification of the requirement that the Hilbert space of an elementary quantum system, like a particle, must be separable.

(b) Physical Hamiltonian operators have spectrum bounded from below to avoid thermodynamical instability. This fact prevents the definition of a “time operator” canonically conjugated with H following the standard way. This result is sometime quoted as **Pauli theorem**. As a consequence, the meaning of Heisenberg relations $\Delta E \Delta T \geq \hbar/2$ is different from the meaning of the analogous relations for posi-

tion and momentum. It is however possible to define a sort of time observable just extending the notion of PVM to the notion of POVM (positive valued operator measure) [1,6]. POVMs are exploited to describe concrete physical phenomena related to measurement procedures, especially in quantum information theory [30,31].

■

Corollary 169. *If the Hamiltonian operator $\sigma(H)$ of a quantum system is bounded below, there is no selfadjoint operator (time operator) T satisfying the standard CCR with H and the hypotheses (1), (2), (3) of Theorem 167.*

Proof. The couple H, T should be mapped to a corresponding couple X, P in $L^2(\mathbb{R}, dx)$, or a direct sum of such spaces, by means of a Hilbert space isomorphism. In all cases the spectrum of H should therefore be identical to the one of X , namely is \mathbb{R} . This fact is false by hypotheses. □

4. Just few words about the Algebraic Approach

The fundamental theorem 3.6.6 of Stone-von Neumann and Mackey is stated in the jargon of theoretical physics as follows:

“all irreducible representations of the CCRs with a finite, and fixed, number of degrees of freedom are unitarily equivalent.”

The expression *unitarily equivalent* refers to the existence of the Hilbert-space isomorphism U , and the finite number of degrees of freedom is the dimension of the Lie algebra spanned by the generators I, X_k, P_k .

What happens then in infinite dimensions? This is the case when dealing with *quantum fields*, where the $2n + 1$ generators I, X_k, P_k ($k = 1, 2, \dots, n$), are replaced by a *continuum* of generators, the so-called *field operators at fixed time* and the **conjugated momentum at fixed time**: $I, \Phi(f), \Pi(g)$ which are smeared by arbitrary functions $f, g \in C_0^\infty(\mathbb{R}^3)$. Here \mathbb{R}^3 is the rest space of a given reference frame in the spacetime. Those field operators satisfy commutation relations similar to the ones of X_k and P_k (e.g., see [26,27,28]). Then the Stone-von Neumann theorem no longer holds. In this case, theoretical physicists would say that

“there exist irreducible non-equivalent CCR representations with an infinite number of degrees of freedom”.

What happens in this situation, in practice, is that one finds two *isomorphic* *-algebras of field operators, the one generated by $\Phi(f), \Pi(g)$ in the Hilbert space \mathcal{H} and the other generated by $\Phi'(f), \Pi'(g)$ in the Hilbert space \mathcal{H}' that admit *no Hilbert space isomorphism* $U : \mathcal{H}' \rightarrow \mathcal{H}$ satisfying:

$$U\Phi'(f)U^{-1} = \Phi(f), \quad U\Pi'(g)U^{-1} = \Pi(g) \quad \text{for any pair } f, g \in C_0^\infty(\mathbb{R}^3).$$

Pairs of this kind are called (*unitarily*) *non-equivalent*. Jumping from the finite-dimensional case to the infinite-dimensional one corresponds to passing from Quantum Mechanics to Quantum Field Theory (possibly relativistic, and on curved

spacetime [28]). The presence of non-equivalent representations of one single physical system shows that a formulation in a fixed Hilbert space is fully inadequate, at least because it insists on a fixed Hilbert space, whereas the physical system is characterized by a more abstract object: An algebra of observables which may be represented in different Hilbert spaces in terms of operators. These representations are not unitarily equivalent and none can be considered more fundamental than the remaining ones. We must abandon the structure of Hilbert space in order to lay the foundations of quantum theories in broader generality.

This programme has been widely developed (see e.g., [13,25,26,27]), starting from the pioneering work of von Neumann himself, and is nowadays called *algebraic formulation of quantum (field) theories*. Within this framework it was possible to formalise, for example, field theories in curved spacetime in relationship to the quantum phenomenology of black-hole thermodynamics.

4.1. Algebraic formulation

The algebraic formulation prescind, anyway, from the nature of the quantum system and may be stated for systems with finitely many degrees of freedom as well [25]. The new viewpoint relies upon two assumptions [26,27,25,29,6].

AA1. A physical system S is described by its **observables**, viewed now as selfadjoint elements in a certain C^* -algebra \mathfrak{A} with unit $\mathbb{1}$ associated to S .

AA2. An **algebraic state** on \mathfrak{A}_S is a linear functional $\omega : \mathfrak{A}_S \rightarrow \mathbb{C}$ such that:

$$\omega(a^*a) \geq 0 \quad \forall a \in \mathfrak{A}_S, \quad \omega(\mathbb{1}) = 1 ,$$

that is, *positive* and *normalised to 1*.

We have to stress that \mathfrak{A} is not seen as a concrete C^* -algebra of operators (a von Neumann algebra for instance) on a given Hilbert space, but remains an abstract C^* -algebra. Physically, $\omega(a)$ is the *expectation value* of the observable $a \in \mathfrak{A}$ in state ω .

Remark 170.

(a) \mathfrak{A} is usually called *the algebra of observables of S* though, properly speaking, the observables are the selfadjoint elements of \mathfrak{A} only.

(b) Differently from the Hilbert space formulation, the algebraic approach can be adopted to describe *both classical and quantum systems*. The two cases are distinguished on the base of commutativity of the algebra of observables \mathfrak{A}_S : A commutative algebra is assumed to describe a classical system whereas a non-commutative one is supposed to be associated with a quantum systems.

(c) The notion of **spectrum** of an element a of a C^* -algebra \mathfrak{A} , with unit element $\mathbb{1}$, is defined analogously to the operatorial case [6]. $\sigma(a) := \mathbb{C} \setminus \rho(a)$ where

we have introduced the **resolvent** set:

$$\rho(a) := \{\lambda \in \mathbb{C} \mid \exists (a - \lambda \mathbb{1})^{-1} \in \mathfrak{A}\}.$$

When applied to the elements of $\mathfrak{B}(\mathcal{H})$, this definition coincides with the one previously discussed for operators in view of (2) in exercise 58. It turns out that if $a^*a = aa^*$, namely $a \in \mathfrak{A}$ is normal, then

$$\|a\| = \sup_{\lambda \in \sigma(a)} |\lambda|.$$

The right hand side of the above identity is called **spectral radius** of a . If a is not normal, a^*a is selfadjoint and thus normal in any cases. Therefore the C^* -property of the norm $\|a\|^2 = \|a^*a\|$ permits us to write down $\|a\|$ in terms of the spectrum of a^*a . As the spectrum is a completely algebraic property, we conclude that it is impossible to change the norm of a C^* -algebra preserving the C^* -algebra property of the new norm. A unital $*$ -algebra admits at most one C^* -norm.

(d) Unital C^* -algebras are very rigid structures. In particular, every $*$ -homomorphism $\pi : \mathfrak{A} \rightarrow \mathfrak{B}$ (which is a pure algebraic notion) between two unital C^* -algebras is necessarily [6] norm decreasing ($\|\pi(a)\| \leq \|a\|$) thus continuous. Its image, $\pi(\mathfrak{A})$, is a C^* -subalgebra of \mathfrak{B} . Finally π is injective if and only if it is isometric. The spectra satisfy a certain permanence property [6], with obvious meaning of the symbols

$$\sigma_{\mathfrak{B}}(\pi(a)) = \sigma_{\pi(\mathfrak{A})}(a) \subset \sigma_{\mathfrak{A}}(a), \quad \forall a \in \mathfrak{A},$$

where the last inclusion becomes an equality if π is injective. ■

The most evident *a posteriori* justification of the algebraic approach lies in its powerfulness [26]. However there have been a host of attempts to account for assumptions **AA1** and **AA2** and their physical meaning in full generality (see the study of [32], [27] and [25,29] and especially the work of I. E. Segal [33] based on so-called *Jordan algebras*). Yet none seems to be definitive [34].

An evident difference with respect to the standard QM, where states are measures on the lattice of elementary propositions, is that we have now a complete identification of the notion of state with that of expectation value. This identification would be natural within the Hilbert space formulation, where the class of observables includes the elementary ones, represented by orthogonal projectors, and corresponding to “Yes-No” statements. The expectation value of such an observable coincides with the probability that the outcome of the measurement is “Yes”. The set of all those probabilities defines, in fact, a quantum state of the system as we know. However, the analogues of these elementary propositions generally do not belong to the C^* -algebra of observables in the algebraic formulation. Nevertheless, this is not an insurmountable obstruction. Referring to a completely general physical system and following [27], the most general notion of state, ω , is the assignment of all probabilities, $w_{\omega}^{(A)}(a)$, that the outcome of the measurement of the observable A is a , for all observables A and all of values a . On the other hand, it is known [25] that

all experimental information on the measurement of an observable A in the state ω – the probabilities $w_\omega^{(A)}(a)$ in particular – is recorded in the expectation values of the polynomials of A . Here, we should think of $p(A)$ as the observable whose values are the values $p(a)$ for all values a of A . This characterization of an observable is theoretically supported by the various solutions to the *moment problem* in probability measure theory. To adopt this paradigm we have thus to assume that the set of observables must include at least all real polynomials $p(A)$ whenever it contains the observable A . This is in agreement with the much stronger requirement **AA1**.

4.1.1. The GNS reconstruction theorem

The set of algebraic states on \mathfrak{A}_S is a convex subset in the dual \mathfrak{A}'_S of \mathfrak{A}_S : if ω_1 and ω_2 are positive and normalised linear functionals, $\omega = \lambda\omega_1 + (1 - \lambda)\omega_2$ is clearly still the same for any $\lambda \in [0, 1]$.

Hence, just as we saw for the standard formulation, we can define *pure algebraic states* as extreme elements of the convex body.

Definition 171. An algebraic state $\omega : \mathfrak{A} \rightarrow \mathbb{C}$ on the C^* -algebra with unit \mathfrak{A} is called a **pure algebraic state** if it is extreme in the set of algebraic states. An algebraic state that is not pure is called **mixed**. ■

Surprisingly, most of the entire abstract apparatus introduced, given by a C^* -algebra and a set of states, admits elementary Hilbert space representations when a reference algebraic state is fixed. This is by virtue of a famous procedure that Gelfand, Najmark and Segal came up with, and that we prepare to present [26,27, 25,6].

Theorem 172 (GNS reconstruction theorem). *Let \mathfrak{A} be a C^* -algebra with unit $\mathbb{1}$ and $\omega : \mathfrak{A} \rightarrow \mathbb{C}$ a positive linear functional with $\omega(\mathbb{1}) = 1$. Then the following holds.*

(a) *There exist a triple $(\mathcal{H}_\omega, \pi_\omega, \Psi_\omega)$, where \mathcal{H}_ω is a Hilbert space, the map $\pi_\omega : \mathfrak{A} \rightarrow \mathfrak{B}(\mathcal{H}_\omega)$ a \mathfrak{A} -representation over \mathcal{H}_ω and $\Psi_\omega \in \mathcal{H}_\omega$, such that:*

- (i) Ψ_ω is cyclic for π_ω . In other words, $\pi_\omega(\mathfrak{A})\Psi_\omega$ is dense in \mathcal{H}_ω ,
- (ii) $\langle \Psi_\omega | \pi_\omega(a)\Psi_\omega \rangle = \omega(a)$ for every $a \in \mathfrak{A}$.

(b) *If (\mathcal{H}, π, Ψ) satisfies (i) and (ii), there exists a unitary operator $U : \mathcal{H}_\omega \rightarrow \mathcal{H}$ such that $\Psi = U\Psi_\omega$ and $\pi(a) = U\pi_\omega(a)U^{-1}$ for any $a \in \mathfrak{A}$.*

Remark 173. The GNS representation $\pi_\omega : \mathfrak{A} \rightarrow \mathfrak{B}(\mathcal{H}_\omega)$ is a $*$ -homomorphism and thus (c) in remark 170 applies. In particular π_ω is norm decreasing and continuous. Moreover, again referring to the same remark, if π_ω is faithful – i.e., injective – it is isometric and preserves the spectra of the elements. If $a \in \mathfrak{A}$ is selfadjoint $\pi_\omega(a)$ is a selfadjoint operator and its spectrum has the well-known quantum meaning. This meaning, in view of the property of permanence of the spectrum, can be directly attributed to the spectrum of $a \in \mathfrak{A}$: *If $a \in \mathfrak{A}$ represents an abstract*

observable, $\sigma(a)$ is the set of the possible values attained by a . ■

As we initially said, it turns out that different algebraic states ω, ω' give generally rise to unitarily inequivalent GNS representations $(\mathcal{H}_\omega, \pi_\omega, \Psi_\omega)$ and $(\mathcal{H}_{\omega'}, \pi_{\omega'}, \Psi_{\omega'})$: There is no isometric surjective operator $U : \mathcal{H}_{\omega'} \rightarrow \mathcal{H}_\omega$ such that

$$U\pi_{\omega'}(a)U^{-1} = \pi_\omega(a) \quad \forall a \in \mathfrak{A} .$$

The fact that one may simultaneously deal with all these inequivalent representations is a representation of the power of the algebraic approach with respect to the Hilbert space framework.

However one may also focus on states referred to a fixed GNS representation. If ω is an algebraic state on \mathfrak{A} , every statistical operator on the Hilbert space of a GNS representation of ω – i.e. every positive, trace-class operator with unit trace $T \in \mathfrak{B}_1(\mathcal{H}_\omega)$ – determines an algebraic state

$$\mathfrak{A} \ni a \mapsto \text{tr}(T\pi_\omega(a)) ,$$

evidently. This is true, in particular, for $\Phi \in \mathcal{H}_\omega$ with $\|\Phi\|_\omega = 1$, in which case the above definition reduces to

$$\mathfrak{A} \ni a \mapsto \langle \Phi | \pi_\omega(a) \Phi \rangle_\omega .$$

Definition 174. If ω is an algebraic state on the C^* -algebra with unit \mathfrak{A} , every algebraic state on \mathfrak{A} obtained either from a density operator or a unit vector, in a GNS representation of ω , is called **normal state** of ω . Their set $Fol(\omega)$ is the **folium** of the algebraic state ω . ■

Note that in order to determine $Fol(\omega)$ one can use a fixed GNS representation of ω . In fact, as the GNS representation of ω varies, normal states do not change, as implied by part (b) of the GNS theorem.

4.1.2. Pure states and irreducible representations

To conclude we would like to explain how pure states are characterised in the algebraic framework. To this end we have the following simple result (e.g., see [26,27,25,6]).

Theorem 175 (Characterisation of pure algebraic states). *Let ω be an algebraic state on the C^* -algebra with unit \mathfrak{A} and $(\mathcal{H}_\omega, \pi_\omega, \Psi_\omega)$ a corresponding GNS triple. Then ω is pure if and only if π_ω is irreducible.*

The algebraic notion of pure state is in nice agreement with the Hilbert space formulation result where pure states are represented by unit vectors (in the absence of superselection rules). Indeed we have the following proposition which make a comparison between the two notions.

Proposition 176. *Let ω be a pure state on the C^* -algebra with unit \mathfrak{A} and $\Phi \in \mathbf{H}_\omega$ a unit vector. Then*

(a) *the functional*

$$\mathfrak{A} \ni a \mapsto \langle \Phi | \pi_\omega(a) \Phi \rangle_\omega,$$

defines a pure algebraic state and $(\mathbf{H}_\omega, \pi_\omega, \Phi)$ is a GNS triple for it. In that case, GNS representations of algebraic states given by non-zero vectors in \mathcal{H}_ω are all unitarily equivalent.

(b) *Unit vectors $\Phi, \Phi' \in \mathbf{H}_\omega$ give the same (pure) algebraic state if and only if $\Phi = c\Phi'$ for some $c \in \mathbb{C}$, $|c| = 1$, i.e. if and only if Φ and Φ' belong to the same ray.*

The correspondence pure (algebraic) states vs. state vectors, automatic in the standard formulation, holds in Hilbert spaces of GNS representations of pure algebraic states, but in general not for mixed algebraic states. The following exercise focusses on this apparent problem.

Exercise 177. *Consider, in the standard (not algebraic) formulation, a physical system described on the Hilbert space \mathcal{H} and a mixed state $\rho \in \mathfrak{S}(\mathcal{H})$. The map $\omega_\rho : \mathfrak{B}(\mathcal{H}) \ni A \mapsto \text{tr}(\rho A)$ defines an algebraic state on the C^* -algebra $\mathfrak{B}(\mathcal{H})$. By the GNS theorem, there exist another Hilbert space \mathcal{H}_ρ , a representation $\pi_\rho : \mathfrak{B}(\mathcal{H}) \rightarrow \mathfrak{B}(\mathcal{H}_\rho)$ an unit vector $\Psi_\rho \in \mathbf{H}_\rho$ such that*

$$\text{tr}(\rho A) = \langle \Psi_\rho | \pi_\rho(A) \Psi_\rho \rangle$$

for $A \in \mathfrak{B}(\mathcal{H})$. Thus it seems that the initial mixed state has been transformed into a pure state! How is this fact explained?

Solution. There is no transformation from mixed to pure state because the mixed state is represented by a vector, Ψ_ρ , in a different Hilbert space, \mathcal{H}_ρ . Moreover, there is no Hilbert space isomorphism $U : \mathcal{H} \rightarrow \mathcal{H}_\rho$ with $UAU^{-1} = \pi_\rho(A)$, so that $U^{-1}\Psi_\rho \in \mathcal{H}$. In fact, the representation $\mathfrak{B}(\mathcal{H}) \ni A \mapsto A \in \mathfrak{B}(\mathcal{H})$ is irreducible, whereas π_ρ cannot be irreducible (as it would be if U existed), because the state ρ is not an extreme point in the space of non-algebraic states, and so it cannot be extreme in the larger space of algebraic states. ■

Acknowledgments

The author is grateful to S. Mazzucchi who carefully read this manuscript, and to G. Marmo, S. Mazzucchi, R. Picken and M. Sánchez for useful discussions about the content of these lectures.

References

- [1] E. Ercolessi, *A short course on quantum mechanics and methods of quantization* (Lectures held at the XXIII International Workshop on Geometry and Physics)
Int. J. Geom. Meth. Mod. Phys. (2015) online version 1560008

- [2] G. Ghirardi, *Sneaking a Look at God's Cards: Unraveling the Mysteries of Quantum Mechanics* Princeton University Press; Revised edition (March 25, 2007)
- [3] Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* <http://plato.stanford.edu/>
- [4] P.A.M. Dirac, *The principles of Quantum Mechanics*. Oxford University Press, Oxford (1930)
- [5] I. M. Gelfand and N. J. Vilenkin, *Generalized Functions*, vol. 4: Some Applications of Harmonic Analysis. Rigged Hilbert Spaces. Academic Press, New York, 1964.
- [6] V. Moretti, *Spectral Theory and Quantum Mechanics: Mathematical Foundations of Quantum Theories, Symmetries and Introduction to the Algebraic Formulation*, Springer, 2018
- [7] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*. Springer-Verlag, Berlin (1932)
- [8] W. Rudin, *Functional Analysis* 2nd edition, Mc Graw Hill, 1991
- [9] K. Schmüdgen, *Unbounded Self-adjoint Operators on Hilbert Space*, Springer, 2012
- [10] G. K. Pedersen, *Analysis Now*, Graduate Texts in Mathematics, Vol. 118 (Springer-Verlag, New York, 1989);
- [11] V.S. Varadarajan, *Geometry of Quantum Theory*, Second Edition, Springer, Berlin (2007)
- [12] Ph. Blanchard, D. Giulini D., E. Joos, C. Kiefer, I.-O. Stamatescu (Eds.): *Decoherence: Theoretical, Experimental, and Conceptual Problems*. Lecture Notes in Physics. Springer-Verlag, Berlin, (2000)
- [13] O. Bratteli, D.W. Robinson, *Operator Algebras and Quantum Statistical Mechanics* (Vol I and II, Second Edition). Springer, Berlin (2002)
- [14] G. Birkhoff and J. von Neumann, *The logic of quantum mechanics*, Ann. of Math. (2) **37**(4) (1936) 823843;
- [15] E.G. Beltrametti, G. Cassinelli, *The logic of quantum mechanics*. Encyclopedia of Mathematics and its Applications, vol. 15, Addison-Wesley, Reading, Mass., (1981)
- [16] K. Engesser, D. M. Gabbay, D. Lehmann (Eds), *Handbook of Quantum Logic and Quantum Structures*. Elsevier, Amsterdam (2009)
- [17] G. Mackey, *The Mathematical Foundations of Quantum Mechanics*. Benjamin, New York (1963)
- [18] C. Piron, *Axiomatique Quantique* Helv. Phys. Acta **37** 439-468 (1964)
- [19] J.M., Jauch and C. Piron, *On the structure of quantal proposition system* Helv. Phys. Acta **42**, 842 (1969)
- [20] J.M., Jauch, *Foundations of Quantum Mechanics* Addison-Wesley Publishing Company, Reading USA, (1978)
- [21] A. Dvurecenskij, *Gleasons theorem and its applications*. Kluwer academic publishers, Dordrecht (1992)
- [22] A. S. Wightman, *Superselection rules; old and new* Nuovo Cimento B **110**, 751-769, (1995)
- [23] B. Simon, *Quantum dynamics: From automorphism to Hamiltonian*. Studies in Mathematical Physics, Essays in Honor of Valentine Bargmann (ed. E.H. Lieb, B. Simon and A.S. Wightman), Princeton University Press, Princeton, 327-349 (1976)
- [24] A.O. Barut, R. Raczka, *Theory of group representations and applications*, World Scientific, 1984
- [25] F. Strocchi, *An Introduction To The Mathematical Structure Of Quantum Mechanics: A Short Course For Mathematicians*, F. Strocchi. World Scientific, Singapore (2005)
- [26] R. Haag, *Local Quantum Physics* (Second Revised and Enlarged Edition). Springer, Berlin (1996)

- [27] H. Araki, *Mathematical Theory of Quantum Fields*. Oxford University Press, Oxford, (2009)
- [28] I. Khavkine, V. Moretti, *Algebraic QFT in Curved Spacetime and quasifree Hadamard states: an introduction*. Advances in Algebraic Quantum Field Theory by Springer 2015 (Eds R. Brunetti, C. Dappiaggi, K. Fredenhagen, and J. Yngvason)
- [29] F. Strocchi, *The Physical Principles of Quantum Mechanics*. European Physics Journal Plus **127**, 12 (2012)
- [30] P. Busch, *Quantum states and generalized observables: a simple proof of Gleason's theorem*. Physical Review Letters **91**, 120403 (2003)
- [31] P. Busch, M. Grabowski, P.J. Lahti, *Operational Quantum Physics*. Springer, Berlin (1995)
- [32] G.G. Emch, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory*. Wiley-Interscience, New York (1972)
- [33] I. Segal, *Postulates for general quantum mechanics*, Annals of Mathematics (2), **48**, 930-948 (1947)
- [34] R.F., Streater, *Lost Causes in and beyond Physics*, Springer-Verlag, Berlin (2007)