

In "Modern topics in mathematical, quantum and statistical physics",
 Proceedings of the 2025 CIF-UnB conferences. Chapter 2, pp.: 17-46.
 Alexandre Dodonov, Aleksandr Pinzul, Ismael Segundo da Silva Carrasco,
 Carolina Matte Gregory, Arsen Melikyan (editors).
 ISBN: 978-65-5563-767-0. São Paulo: LF Editorial, (2026).

Chapter 2

The Algebraic Formulation of Quantum Theories

João Carlos Alves Barata^{1,a}

¹ Departamento de Física Matemática, Instituto de Física da USP, São Paulo, SP, Brazil

^a jbarata@if.usp.br

2.1 Abstract

We present some ingredients common to physical theories that justify the use of the ideas and of the language of Operator Algebras in the context of Quantum Physics.

2.2 Introduction and Summary

The algebraic formulation of Quantum Physics provides a powerful unifying language that can be employed in the description and understanding of numerous phenomena of the quantum world. In this note, we intend to present some of the ideas that guide this language, centred on physical and probabilistic principles. The illustrative cases of Classical Mechanics, Quantum Mechanics, and relativistic Quantum Field Theories are discussed, and some results are presented, such as Heisenberg's and Schrödinger's uncertainty relations. This text does not intend to be an exhaustive discussion of the various topics presented here, but aims to motivate the reader in future studies. Therefore, a detailed, albeit incomplete, bibliography is presented.

2.3 General Description of Physical Theories

Physics is composed of various theories, related to each other in diverse ways, and which, to a greater or lesser degree of approximation, describe the observable world. We can list Classical Mechanics, Thermodynamics, Quantum Mechanics, Relativistic Quantum Field Theory, the Theory of General Relativity, and Statistical Mechanics. These diverse

theories, however, share a series of common ingredients. Every physical theory should be able to specify:

- Physically observable quantities and their mathematical description, the relationships between these observables, such as compatibility relations, algebraic relations, etc.
- The set of possible outcomes of individual measurements of these observables.
- The association between physical systems, observables, and the probability distributions that describe measurements of these observables in states.
- The set of pure states (see [1] for a general discussion of this notion).
- The dynamics of observables and states.
- The symmetries of the described physical systems and their implementations in states and observables.

These questions are not necessarily independent. The dynamics, for example, can be associated with symmetry by temporal evolution, and the set of pure states can be fixed by the algebra of observables.

Let's try to discuss some of the points above in more detail.

Observables, measurements and probability distributions

Each physical theory has its own set of observable quantities (in Heisenberg's¹ words, "*it is the theory that defines what is observable*" [2]) and one of its main objectives is to describe the result of measurements of these observables in physical systems. Let A be an observable physical quantity and $\mathcal{C}(A)$ the set of possible values resulting from measurements of A (in any state). It is an experimental fact that repeated measurements of an observable A , under the same conditions, that is, in the same physical state E of the system studied, do not necessarily provide the same value in $\mathcal{C}(A)$, having a random character.

It is an observational fact that an ideally infinite succession of experimental measurements of A , *all under the same physical conditions of the system in question*, should produce a statistical distribution in $\mathcal{C}(A)$ defined by a probability measure.

Let us generically denote these physical conditions by E (which can concretely represent a set of physical parameters of the system) and by $\mu_{E,A}$ the probability measure in question.

This probability measure $\mu_{E,A}$ is a function of both the set of conditions E that specifies the system and the observable A considered. This probability measure $\mu_{E,A}$ is called the *state* (or *physical state*) of the system in question with respect to the observable A . Since all information about the properties of the physical system, with respect to the observable A , must be the result of the statistical analysis of the experimental measurements of A in the system, we conclude that the probability measure $\mu_{E,A}$, that is, the physical state of the system, contains within itself all the available information about these properties.

¹Werner Karl Heisenberg (1901–1976).

Herein lies embedded a physical (or philosophical, if you will) principle according to which only objective reality derived from experimentation allows inferences about a physical system, and this reality manifests itself in the form of statistical distributions in the sets $\mathcal{C}(A)$ for the various observables A with which we study the system. In other words, the reality of a physical system is only achieved based on experimentation, and inferences about it must be statistical inferences based on experimental data. It is only based on these inferences that one can determine general patterns (if any) that lead to the elaboration of physical laws and theories to explain them based on simpler principles (physical postulates) and mathematical inference. Allow us a historical-philosophical comment. It is a general belief of physicists, first expressed by Galilei² in the 16th-17th centuries, but with deeper roots, that the formulation of physical theories based on mathematical ideas, a construct of the human mind, is possible. That such a thing is true, which is corroborated by the history of Physics up to now, is perhaps the greatest enigma of all Science.

There are three possible origins for the randomness, which we mentioned above, observed in the measurement of an observable in a physical system, origins that can occur simultaneously: it can stem from experimental measurement errors, it can stem from an incomplete knowledge of the system studied, or it can be intrinsic to the described system, a fact first identified in Atomic Physics.

Normally, in the elaboration of physical theories, the ideal situation is considered, in which experimental inaccuracies are neglected. However, the other two sources of randomness remain, which then must be duly considered in the theoretical framework. Later we will remember how this is done in some cases.

The fact we want to emphasise is that physical theories must be able to associate with each physical state of a system and with each observable a probability distribution that describes a succession of measurements of that observable in that state. Note that this does not exclude deterministic theories, such as Classical Mechanics, since deterministic situations can also be described by probability distributions, such as Dirac delta probability measures.

In the statistical analysis of the results of measurements of an observable A of a physical system in a given state, several quantities play a role. One of them is the so-called average value of the measurements of A in this distribution, that is, its expectation or expected value, which will be denoted here by $\langle A \rangle_E$. Other relevant quantities are the momenta $\langle A^n \rangle_E$, $n \in \mathbb{N}$. It is a well-known mathematical fact (a consequence, incidentally, of Weierstrass's Theorem. See the discussion on the "Hamburger moment problem" in [3]) that if $\mathcal{C}(A)$ is a compact set, then the probability measure $\mu_{E,A}$ can be recovered from the set of momenta $\langle A^n \rangle_E$, $n \in \mathbb{N}$.³

²Galileo di Vincenzo Bonaiuti de Galilei (1564–1642): *"The book of nature cannot be read until we learn its language and become familiar with the symbols in which it is written. And it is written in mathematical language, and its letters are triangles, circles, and other geometric figures, without which it is humanly impossible to understand a single word and there is only a wandering lost in a dark labyrinth."* Il Saggiatore, 1623.

³Hence the importance of considering observables A that are *bounded*, that is, for which $\mathcal{C}(A)$ is compact. As we will discuss, in Quantum Physics $\mathcal{C}(A)$ is identified with $\sigma(A)$, the spectrum of a selfadjoint operator A . $\sigma(A)$ is compact (closed and bounded) if A is a selfadjoint and bounded operator. In the so-called algebraic formulation of Quantum Field Theories, the entire treatment is done considering observables that are selfadjoint and bounded operators, in Hilbert spaces or algebras. See [4] or [5].

Another important stochastic quantity is the so-called *variance*, defined by

$$\text{Var}_E(A) := \langle A^2 \rangle_E - \langle A \rangle_E^2 = \left\langle (A - \langle A \rangle_E)^2 \right\rangle_E \geq 0, \quad (2.1)$$

which provides a qualitative indication of how much the values of the A measurements deviate from their average value. The quantity

$$\Delta_E(A) := \sqrt{\text{Var}_E(A)} \quad (2.2)$$

is called *standard deviation* of the observable A in the state E . It provides a qualitative indication of how much the values of the A measurements deviate from their average value.

Although not the only stochastic quantity that provides this type of qualitative information, the variance is a useful quantity. As we will see, the famous Uncertainty Relations of Heisenberg in Quantum Mechanics can be expressed as statements about the variance of two non-commuting observables (for example, momentum and position in the same Cartesian direction: $\text{Var}(p_x) \text{Var}(x) \geq \hbar^2/4$).

The *correlation*, or *covariance*, between two observables A and B , with respect to a state E , is defined by

$$\text{Cov}_E(A, B) := \left\langle (A - \langle A \rangle_E)(B - \langle B \rangle_E) \right\rangle_E \quad (2.3)$$

and one has, as one easily sees,

$$\text{Cov}_E(A, B) = \langle AB \rangle_E - \langle A \rangle_E \langle B \rangle_E.$$

In words, $\text{Cov}_E(A, B)$ “measures” how much the average deviation of A from its average value $\langle A \rangle_E$ is statistically related to the average deviation of B from its average value $\langle B \rangle_E$. If A and B are stochastically independent, then $\text{Cov}_E(A, B) = 0$. The converse is not generally true.

It is also relevant to note that, by definition,

$$\text{Var}_E(A) = \text{Cov}_E(A, A).$$

In Probability Theory, the expected value (or “expectation”) of a measurable function (“random variable”) A defined on a sample space Ω and its variance with respect to a probability measure μ on Ω are given by

$$\begin{aligned} \mathbb{E}_\mu(A) &\equiv \langle A \rangle_\mu := \int_\Omega A \, d\mu, \\ \text{Var}_\mu(A) &:= \int_\Omega (A - \langle A \rangle_\mu)^2 \, d\mu = \mathbb{E}_\mu(A^2) - \mathbb{E}_\mu(A)^2, \end{aligned}$$

respectively. The correlation, or covariance, of two random variables A and B defined in a sample space Ω , in relation to a probability measure μ in Ω is given by

$$\begin{aligned} \text{Cov}_E(A, B) &= \int_\Omega \left((A - \langle A \rangle_\mu)(B - \langle B \rangle_\mu) \right) d\mu \\ &= \int_\Omega AB \, d\mu - \left(\int_\Omega A \, d\mu \right) \left(\int_\Omega B \, d\mu \right) = \mathbb{E}_\mu(AB) - \mathbb{E}_\mu(A)\mathbb{E}_\mu(B). \end{aligned}$$

Variance and pure states

In probability theory, a probability measure μ in a sample space is said to be *pure* if it cannot be written as a convex linear combination of two other distinct probability measures of μ from the same sample space, that is, if it cannot be written in the form $\mu = \alpha\mu_1 + (1 - \alpha)\mu_2$ where μ_1 and μ_2 are also distinct probability measures of μ and $0 < \alpha < 1$. It is an easy exercise to show that if $\mu = \alpha\mu_1 + (1 - \alpha)\mu_2$, then

$$\langle A \rangle_\mu = \alpha \langle A \rangle_{\mu_1} + (1 - \alpha) \langle A \rangle_{\mu_2}$$

and

$$\text{Var}_\mu(A) = \alpha \text{Var}_{\mu_1}(A) + (1 - \alpha) \text{Var}_{\mu_2}(A) + \alpha(1 - \alpha) \left[\langle A \rangle_{\mu_1} - \langle A \rangle_{\mu_2} \right]^2.$$

From this, we conclude

$$\text{Var}_\mu(A) \geq \alpha \text{Var}_{\mu_1}(A) + (1 - \alpha) \text{Var}_{\mu_2}(A) \geq \min \{ \text{Var}_{\mu_1}(A), \text{Var}_{\mu_2}(A) \}.$$

Thus, the variance $\text{Var}_\mu(A)$ in the non-pure measure μ is always greater than or equal to the smaller of the two variances $\text{Var}_{\mu_1}(A)$ or $\text{Var}_{\mu_2}(A)$. We understand, therefore, that if we restrict the measures μ to a certain set of measures $\mathfrak{M}$ on the sample space, then the smallest possible values of the variances $\text{Var}_\mu(A)$ of a fixed function A are achieved when μ is found in the subset of the pure probability measures of $\mathfrak{M}$. In this sense, the pure probability measures represent those with the smallest possible deviation of the magnitude represented by A from its mean value.

We say that a physical system is in a *pure state* for a given observable A if $\mu_{E,A}$ is pure. Thus, the pure states of a physical system represent those states with the smallest fluctuations of the observable quantity A . We understand, therefore, that determining which are the pure states of a physical system and what are the variances of observables in these pure states provide important information about the smallest possible fluctuations that can be observed in that system. This is important information about the degree of intrinsic randomness (*i.e.*, not resulting from experimental errors or incomplete knowledge) of the underlying physical theory that describes the system in question.

As we will discuss later, another reason for the importance of pure states lies in the fact that in both Classical Mechanics and Quantum Mechanics it can be stated that the knowledge of the expected values of an observable in all pure states of a system uniquely determines that observable.

2.4 The Case of Classical Mechanics

In Classical Mechanics, all basic experimental measurement processes involve measurements of position and velocity, which can be performed simultaneously and independently, so that, in principle, any functions involving the coordinates and momenta of a system are observable physical quantities. It is possible to create new observables by performing simple algebraic operations with other observables, such as linear combinations, products, etc. It is, therefore, convenient to consider the algebra of all functions defined in the phase space $\mathbb{F}$ of the system considered as constituting the collection of all observable physical quantities of that system. Since the result of a physical measurement is always a real

number, the observable physical quantities must be functions of the phase space $\mathbb{F}$ in real numbers $\mathbb{R}$. For technical reasons it is convenient to take only the algebra of functions defined in the phase space that are *measurable* with respect to the measure of Liouville⁴ $dqdp$, thus avoiding mathematical pathologies.

An important characteristic of classical systems is the possibility of simultaneous and independent measurement of any distinct observables. This characteristic is called the *compatibility* of observables. A consequence of the compatibility of classical observables, which ended up implicitly embedded in the observations above, is that they form a commutative algebra.

Given an observable thus abstractly defined as a function $f(q, p)$, we can ask ourselves what values we will obtain when making a measurement of this observable at a certain instant in time? The answer is somewhat disappointingly obvious: if the coordinates of the system considered are at that instant in time q_0 and its moment p_0 , then the measured value of f will be $f(q_0, p_0)$. The collection $\mathcal{C}(f)$ of all possible results of measurements of f is, therefore, the image of f as a function of $\mathbb{F}$ in $\mathbb{R}$.

In Classical Mechanics, physical states are described by probability distributions in phase space, such that average values of an observable f are given by

$$\langle f \rangle = \int_{\mathbb{F}} f(q, p) \rho(q, p) dqdp, \quad (2.4)$$

with $\rho(q, p) \geq 0$ e $\int_{\mathbb{F}} \rho(q, p) dqdp = 1$. In this sense, we can identify the function (or measure) ρ with the state of the system itself, since from it we uniquely obtain the probability distributions on the sets $\mathcal{C}(f)$, which we identify with the image of the functions $f : \mathbb{F} \rightarrow \mathbb{R}$.

Dirac delta measures $\rho_{q_0, p_0}(q, p) = \delta(q - q_0)\delta(p - p_0)$ with

$$\langle f \rangle_{q_0, p_0} = \int_{\mathbb{F}} f(q, p) \rho_{q_0, p_0}(q, p) dqdp = f(q_0, p_0)$$

represent pure states of the system being treated and can be interpreted as states with maximal information. For states like $\rho_{q_0, p_0}(q, p) = \delta(q - q_0)\delta(p - p_0)$ we have certainty regarding the positions and momenta of the system's constituents and the variance of the distribution of f is zero, as are the other fluctuations, because

$$\text{Var}_{q_0, p_0}(f) = \langle f^2 \rangle_{q_0, p_0} - \langle f \rangle_{q_0, p_0}^2 = f(q_0, p_0)^2 - f(q_0, p_0)^2 = 0.$$

In such states, measurements of the observable f provide one and only one value, namely, $f(q_0, p_0)$. No randomness occurs, therefore, in the measurement of any observables when the system is in a pure classical state. The belief that it is always possible to fix all the parameters of a system in order to completely fix its state and eliminate all randomness in measurements of observables is sometimes called "*realism*". Classical Mechanics, as well as all Classical Physics in general, is in this sense realistic. This characteristic is not found in Quantum Physics, where pure states can produce non-zero variances.

In Classical Mechanics, not only pure states are of interest. In Classical Statistical Mechanics, for example, one can consider states with distributions of the type

$$\rho(q, p) = \frac{1}{V(\mathcal{E})} \delta(H(q, p) - \mathcal{E}) \quad (2.5)$$

⁴Joseph Liouville (1809–1882).

in the so-called *microcanonical ensemble* with energy $\mathcal{E}$, where $H(q, p)$ is the Hamiltonian of the system and $V(\mathcal{E})$ is the normalisation constant $V(\mathcal{E}) = \int_{\mathbb{F}} \delta(H(q, p) - \mathcal{E}) dqdp$ (assumed finite). In the so-called *canonical ensemble* the so-called *Gibbs state*⁵ is adopted:

$$\rho(q, p) = \frac{1}{Z(\beta)} e^{-\beta H(q, p)}, \quad (2.6)$$

with the normalisation constant $Z(\beta) = \int_{\mathbb{F}} e^{-\beta H(q, p)} dqdp$ assumed finite, β being the inverse of the temperature.

The dynamics of observables in a classical mechanical system is defined by the Hamiltonian flux in phase space, which is characterised by Hamilton's⁶ equations:

$$\dot{q} = \partial_p H(q, p), \quad \dot{p} = -\partial_q H(q, p),$$

where the Hamiltonian H is a differentiable function defined in phase space and satisfying suitable conditions to guarantee uniqueness and the existence of solutions (preferably global) for the above equations from initial conditions $q(0)$ and $p(0)$. If q_t and p_t are solutions of Hamilton's equations, the evolution of an observable f is expressed by $f_t(q, p) := f(q_t, p_t)$. Thus, by (2.4),

$$\langle f \rangle_t := \langle f_t \rangle = \int_{\mathbb{F}} f(q_t, p_t) \rho(q, p) dqdp = \int_{\mathbb{F}} f(q, p) \rho(q_{-t}, p_{-t}) dq_{-t} dp_{-t}.$$

Since the Liouville measure $dqdp$ is invariant under a Hamiltonian flow (Liouville's Theorem), we conclude that $\langle f \rangle_t = \int_{\mathbb{F}} f(q, p) \rho_t(q, p) dqdp$, where $\rho_t(q, p) := \rho(q_{-t}, p_{-t})$ represents the temporal evolution of the state described by ρ . This relation teaches us how the evolution of observables in Classical Mechanics is reflected in the evolution of states.

By (2.5) and (2.6), it is evident that the measurements of the microcanonical and canonical ensembles are invariant under temporal evolution (a requirement for them to describe equilibrium states), since $H(q_t, p_t) = H(q, p)$ for all t .

2.5 The Framework of Quantum Physics and the Relevance of the Spectral Theorem

If $\psi \in \mathcal{H}$ is a non-zero vector in a Hilbert⁷ space $\mathcal{H}$ and $g : \mathcal{B}_l(\sigma(A)) \rightarrow \mathbb{C}$ is a bounded Borelian function defined on the spectrum of a selfadjoint and bounded operator A , we know from the Spectral Theorem (see [6, 7]) that an expression like $\langle \psi, g(A)\psi \rangle$ can be represented as an integral over the spectrum of A :

$$\langle \psi, g(A)\psi \rangle = \int_{\sigma(A)} g d\mu_{\psi, A},$$

where $\sigma(A)$ is the spectrum of A , which is a closed and bounded set of A . The measure $\mu_{\psi, A}$ is a positive measure in $\sigma(A)$ and if $\|\psi\| = 1$ we also know that

$$\int_{\sigma(A)} d\mu_{\psi, A} = 1.$$

⁵Josiah Willard Gibbs (1839–1903).

⁶Sir William Rowan Hamilton (1805–1865)

⁷David Hilbert (1862–1943).

These two facts tell us that $\mu_{\psi, A}$ is a *probability measure in $\sigma(A)$* . This simple mathematical fact has a significant consequence in the context of Quantum Physics, which is at the root of its axiomatisation and formalisation in terms of Hilbert spaces and operators acting on Hilbert spaces. To better understand this fact, let's make some general considerations.

The Quantum Physics Framework

In Quantum Physics, it is no longer true that experimental measurement processes involve measurements of position and velocity, since these cannot be done independently and simultaneously. The compatibility property of some observables is therefore lost. As is well known, the historical development of Quantum Mechanics led to the proposition that observables must be represented by selfadjoint operators acting in a Hilbert space. One of the adopted postulates states that individual measurements of an observable represented by an operator A must be elements of the spectrum of that operator.

According to the postulates of Quantum Mechanics, the physical states of the quantum system with a finite number of degrees of freedom (*i.e.*, describing a finite number of particles) are described by “density matrices”⁸ positive selfadjoint operators ρ acting on a Hilbert space $\mathcal{H}$ with $\text{Tr}(\rho) = 1$ such that the average value of an ideally infinite set of measurements of the observable A in the state described by ρ are given by $\langle A \rangle_\rho = \text{Tr}(\rho A)$.

The choice of selfadjoint operators for the role of observables is motivated by two properties: first, the spectrum of a selfadjoint operator is a subset of the real line, a fact consistent with the postulate that states that individual measurements of an observable must be elements of the spectrum of the associated operator. According to the spectral theorem, selfadjoint operators can be represented by sums (or integrals) like $A = \sum_{\lambda \in \sigma(A)} \lambda P_\lambda$. Here we assumed for simplicity that A has a discrete spectrum $\sigma(A)$ of eigenvalues, in which case P_λ designates the projector onto the subspace of eigenvectors of A with eigenvalue λ .

The importance of the Spectral Theorem in the formalisation of quantum theories is enormous, since it is through it that we can obtain the probability distributions associated with measurements of an observable A in a given state. In fact, by the above prescription and by the Spectral Theorem, we have:

$$\langle A \rangle = \text{Tr}(\rho A) = \sum_{\lambda \in \sigma(A)} \lambda p_\lambda, \quad (2.7)$$

where $p_\lambda = \text{Tr}(\rho P_\lambda)$. Now, it is clear that $p_\lambda \geq 0$ and

$$\sum_{\lambda \in \sigma(A)} p_\lambda = \text{Tr} \left(\rho \sum_{\lambda \in \sigma(A)} P_\lambda \right) = \text{Tr}(\rho) = 1.$$

These two facts together with (2.7) lead to the interpretation that p_λ represents the probability measure in $\sigma(A)$ that describes distributions of measures of the observable values A in the state described by ρ . In this sense, we can identify ρ with the state of the

⁸It should be mentioned that a good part of the mathematical interpretation of Quantum Physics that we will present in a summarised way below originates from the contributions of von Neumann. János Neumann (1903-1957) also adopted the names Johann von Neumann and John von Neumann.

system itself, since from it we uniquely obtain the probability distributions on the sets $\mathcal{C}(A)$, which we identify with the spectra $\sigma(A)$ of the selfadjoint operators A .

The observations above show that the usual interpretation of observables in Quantum Physics in terms of selfadjoint operators acting on Hilbert spaces is consistent with the basic purpose of describing experimental measurements of observables and their probability distributions.

We comment in passing that the scheme above can be further generalised and abstracted in the following sense: the algebras of observables of quantum systems can be taken as abstract C^* -algebras and the physical states correspond to states on these algebras, that is, positive and normalised linear functionals. In this context, it is equally possible to recover the probabilistic description that we outlined above. The great advantage of this description manifests itself in the treatment of quantum systems with an infinite number of degrees of freedom, as in Quantum Statistical Mechanics and Quantum Field Theory. Because it is independent of a underlying Hilbert spaces, the description of observables in terms of algebras allows us to describe certain phenomena typically found in systems with an infinite number of degrees of freedom, such as super-selection rules and phase transitions. For applications in Physics, we refer to the references [4, 5, 8].

The temporal evolution of observables in a system with a finite number of degrees of freedom is characterised by a strongly continuous unitary representation of the additive group $\mathbb{R}$ (representing the symmetry of temporal evolution, for time-independent systems): $\mathbb{R} \ni t \mapsto U(t)$, where $U(0) = \mathbb{1}$, $U(t)U(t') = U(t+t')$ and $U(t)^{-1} = U(t)^*$ for all $t, t' \in \mathbb{R}$. If A is an observable, its evolution will be given by $A_t := U(t)AU(t)^*$. Thus, $\langle A \rangle_t := \langle A_t \rangle = \text{Tr}(\rho A_t) = \text{Tr}(\rho U(t)AU(t)^*)$ and, by the cyclic property of the trace, we obtain $\langle A \rangle_t = \text{Tr}(\rho_t A)$ where $\rho_t := U(t)^* \rho U(t)$. This expression shows how the evolution of observables is reflected in the evolution of states. The fact that the evolution $U(t)$ is strongly continuous guarantees, by Stone's Theorem (see [6]), that there exists a selfadjoint operator (not necessarily bounded) H such that $U(t) = e^{-iHt/\hbar}$ for all $t \in \mathbb{R}$. With this we can (except for technicalities related to domains) transform by differentiation the relation $A_t := U(t)AU(t)^*$ into the Heisenberg equation: $i\hbar \partial_t A_t = [H, A_t]$. For the states we will have, analogously, $i\hbar \partial_t \rho_t = -[H, \rho_t]$.

In Quantum Physics, the question of the compatibility of two observables is directly linked to the commutativity of the associated operators: two observables can only be measured simultaneously if the corresponding operators commute with each other. This question is particularly important in relativistic quantum field theories, where the so-called *Einstein locality principle* must be respected.

This principle, one of the most central in all of Physics, states that events separated by space-like intervals cannot be causally related. This principle must be translated into relativistic quantum field theories by the imposition that observables associated with points or regions separated by space-like intervals must commute with each other. The consequences of this imposition on the structure of relativistic quantum field theories are enormous, but it is not our place to discuss all its consequences here (see, for example, [4, 5]). Returning to (2.7), pure states of quantum systems described in a Hilbert space $\mathcal{H}$ correspond to the situation in which ρ is a projector onto a one-dimensional subspace of $\mathcal{H}$: $\rho = P_\psi$, that is, in Dirac notation $\rho = |\psi\rangle\langle\psi|$, where $\psi \in \mathcal{H}$ is a normalised vector $\|\psi\| = 1$. Thus, for a pure state with $\rho = P_\psi$ and $\|\psi\| = 1$ we will have $\langle A \rangle_\psi = \langle \psi, A\psi \rangle$.

The equivalent of the Gibbs state (2.6) at the inverse temperature β for a quantum system with a finite number of particles is $\rho_\beta = e^{-\beta H} / \text{Tr}(e^{-\beta H})$, if the Hamiltonian

operator is such that $\text{Tr}(e^{-\beta H}) < \infty$ (which is typically the case if the system is restricted to a finite spatial volume). Such operators ρ_β commute with H and are, therefore, invariant under time evolution, as desired for equilibrium states.

An important fact is that pure states can exhibit non-zero variance for mean values of measurements of certain observables, what does not occur in Classical Mechanics:

$$\langle A^2 \rangle_\psi - \langle A \rangle_\psi^2 = \langle \psi, A^2 \psi \rangle - \langle \psi, A \psi \rangle^2 \neq 0,$$

unless ψ is an eigenvector of A . In fact, for selfadjoint A and $\|\psi\| = 1$, one has

$$\langle \psi, A^2 \psi \rangle - \langle \psi, A \psi \rangle^2 = \left\langle \psi, \left[A - \langle \psi, A \psi \rangle \mathbb{1} \right]^2 \psi \right\rangle = \left\| \left[A - \langle \psi, A \psi \rangle \mathbb{1} \right] \psi \right\|^2,$$

showing that $\langle A^2 \rangle_\psi - \langle A \rangle_\psi^2 = 0$ if and only if $[A - \langle \psi, A \psi \rangle \mathbb{1}] \psi = 0$, that is, if and only if $A\psi = \lambda\psi$ with $\lambda = \langle \psi, A \psi \rangle$.

2.6 An Overview of the Mathematical Formulation of Quantum Field Theories

Relativistic Quantum Field Theory was born almost simultaneously with Quantum Mechanics, with the first important works emerging in the 1930s (for a historical overview of the first developments, see, *e.g.*, [9]). Its first applications were in the field of Atomic Physics, in the study of the spectrum of atoms under the action of electromagnetic radiation, but from early on it revealed itself as the fundamental theory for the description of the then nascent Elementary Particle Physics, its main area to this day, and which places it, along with the Theory of General Relativity, among the most fundamental physical theories.

In simplified terms, three ingredients comprise Relativistic Quantum Field Theories: the quantum nature of observables and states, relativistic covariance (through Poincaré transformations), and, not dissociated from the latter, the idea that observations and influences separated by space-like intervals do not influence each other, regardless of the particular states considered, and should therefore be described by commuting operators belonging to an appropriate algebra of observables. This fundamental principle is often referred to as *Einstein's causality*.

Quantum theories of relativistic fields, particularly the so-called Quantum Electrodynamics (QED), have developed rapidly and spectacularly, becoming one of the most accurate physical theories in its experimental predictions. However, from very early on, voices of dissatisfaction arose regarding its conceptual and mathematical foundations and the recipe-like nature of some of its inference methods.

Among the problems pointed out, we list:

- The formulation centred on perturbation theory, sometimes obscuring the underlying conceptual foundations. One must say, however, that perturbation theory played a very important historical role in the development of Particle Physics, specially due to its heuristic power in the discovery of “true” physical models from experimental data.

- The need for regularisation/renormalisation and its consequent mathematical difficulties.
- Problems with the convergence of perturbative series, even after the renormalisation process. In many important cases, perturbative series in a coupling constant g apparently behave as $\sum_{n=0}^{\infty} c_n g^n$ with the coefficients c_n increasing as $c_n \approx n!$ for $n \in \mathbb{N}$ “large”, which leads to a series with zero radius of convergence. It should be noticed that there is only one mathematically complete demonstration of this behaviour (for scalar field models in two dimensions. See [10]), but the general belief (based, for example, on counting Feynman diagrams) is that this behaviour is the same in physically more realistic models in any dimension, like the QED. In fact, no interactive QFT models are known for which the *convergence* of the perturbative series has been demonstrated⁹.
- Fundamental problems concerning the nature of Quantum Field Theories: what objects do they describe? Do they describe particles? How to deal with super-selection sectors?
- Problems with the mathematical justification of the so-called Feynman Functional Integration in Minkowski spacetime (and in other locally or globally hyperbolic spacetimes), an instrument frequently employed, especially in the construction of models of interest to Elementary Particle Physics and in didactic texts.
- Problems with defining models beyond the limits of perturbation theory.

Faced with these questions, essentially two mathematical formulations of Quantum Field Theories emerged: the Wightman formulation, also known as Axiomatic Field Theory (a rather inappropriate name, since Physics is an empirical science, not an axiomatic one), where Quantum Field Theories in the vacuum sector are formulated in terms of certain distributions, the so-called Wightman distributions, and the Algebraic Formulation of Quantum Field Theories, also known as Local Quantum Field Theory, formulated by several authors, such as Haag, Kastler, Araki and others (see, *e.g.*, [4, 5, 12]), which formulates Quantum Field Theories in the language of C^* algebras and von Neumann algebras, being in several aspects more general than the other formulation, better suited to the treatment of super-selection sectors and the formulation of QFT in curved spacetimes.

The main ingredient of Wightman theories is a family of tempered distributions $W_n(x_1, \dots, x_n)$ in $\mathbb{R}^{dn}$, for all $n \in \mathbb{N}$, d being the space-time dimension, to be interpreted as vacuum expectation values of products of scalar fields together with some additional assumptions about positivity, Poincaré covariance, Einstein causality, etc. Einstein causality, for instance, means in this context (bosonic case): $W_2(x_1, x_2) = W_2(x_2, x_1)$ etc, for space-like separated points. Wightman’s reconstruction theorem [13] allows the construction of a Hilbert space of physical states $\mathcal{H}$ provided with a singled normalised vector Ω representing the vacuum state, invariant under the Poincaré group. It also allows the association of a quantum field Φ through the linear map

$$f \longmapsto \Phi(f) \equiv \int f(x) \Phi(x) dx$$

⁹It should be noticed that the existence of divergent perturbative series is not exclusive to Quantum Field Theories, as it happens, in certain special cases, in Classical Mechanics and in Quantum Mechanics, as in the computation of the energy of the ground state of the anharmonic oscillator [11].

as an (unbounded) operator acting on $\mathcal{H}$ (operator valued distributions). The function f is taken as a function in Schwartz space. The Wightman distributions can be expressed as vacuum expectations as $W_n(x_1, \dots, x_n) = \langle \Omega, \Phi(x_1) \cdots \Phi(x_n) \Omega \rangle$. Einstein causality implies $\Phi(f)\Phi(g) = \Phi(g)\Phi(f)$ if $\text{supp}(f)$ and $\text{supp}(g)$ are space-like. The Poincaré group is implemented in the Hilbert space by unitary transformations.

The most important results of Wightman's formulation are:

- PCT Theorem [14, 15].
- Spin and Statistics Theorem [15].
- Reeh-Schlieder Theorem [4, 5, 16].
- Bisognano-Wichmann Theorem [4, 17].
- Haag-Ruelle scattering theory [18–20].
- LSZ reduction formulas [21–26].

Wightman's formulation opened the possibility for *constructing* Quantum Field theoretical models satisfying all required postulates by means of Euclidean methods based ON the Feynman-Kac path integral formulation. See, *e.g.*, [27, 28]. It has, though, some shortcomings: a limited treatment of super-selection sectors, a limited treatment of mixed states, a limited treatment of thermal states and a limited treatment of gauge fields.

2.6.1 The Algebraic Formulation of Quantum Fields

As mentioned above, the Algebraic Formulation of Quantum Fields (AQFT) started with works of Rudolf Haag, Daniel Kastler and Huzihiro Araki, but many other names should be added, as Sergio Doplicher, John Roberts, Bert Schroer, Detlev Buchholz and Klaus Fredenhagen. In order to present the main ideas of AQFT we have to briefly describe some algebraic notions, that of C^* -algebras and of von Neumann algebras.

C^* and von Neumann-algebras. What are they and why they are useful for Physics

In mathematical terms, a C^* -algebra is a normed, complete, associative algebra over $\mathbb{C}$ with an antilinear involution $A \mapsto A^*$, such that

- $\|AB\| \leq \|A\| \|B\|$,
- $\|A^*\| = \|A\|$,
- $\|A^*A\| = \|A\|^2$,

for all A, B elements of the algebra. There are many examples of C^* -algebras, the most basic ones is the algebra $\mathcal{B}(\mathcal{H})$ of bounded (*i.e.*, continuous) operators acting on a Hilbert space $\mathcal{H}$ (see [6, 7, 29]).

Examples of C^* -algebra are:

- $\mathcal{B}(\mathcal{H})$, the algebra of bounded operators acting on a Hilbert space $\mathcal{H}$,
- The algebra of compact operators acting on Hilbert spaces,

- CAR and CCR algebras,
- AF-algebras,
- Cuntz algebras and Cuntz-Krieger algebras,

and many more. For a more complete account of the theory of C^* -algebras, see, *e.g.*, [8, 29–32].

A C^* -algebra may or may not have a unit $\mathbb{1}$. In most physical applications we always assume there is one. If $\mathcal{A}$ is a C^* -algebra with unit, we say that $\lambda \in \mathbb{C}$ is an element of the spectrum of A if the inverse $(\lambda\mathbb{1} - A)^{-1}$ does not exist in $\mathcal{A}$. The spectrum of $A \in \mathcal{A}$ is denoted by $\sigma(A)$. The spectrum $\sigma(A)$ is always closed and bounded and $\sigma(A) \subset \mathbb{R}$ if A is selfadjoint. For A selfadjoint one has

$$\sigma(A) \subset [-\|A\|, \|A\|] .$$

C^* -algebras are abstract algebras and do not necessarily act on vector spaces. However, they can be represented as operator algebras acting on Hilbert spaces.

Moreover, C^* -algebras also admit a probabilistic interpretation for expectation values. To understand that we have to introduce the important notion of *state*.

A state ω on a C^* -algebra $\mathcal{A}$ is a map $\omega : \mathcal{A} \rightarrow \mathbb{C}$ that is

- linear: $\omega(\alpha A + \beta B) = \alpha\omega(A) + \beta\omega(B)$,
- positive: $\omega(A^*A) \geq 0$,
- normalised: $\omega(\mathbb{1}) = 1$.

It follows that

- $\omega(A^*) = \overline{\omega(A)}$,
- ω is continuous: if $\|A_n - A\| \xrightarrow{n \rightarrow \infty} 0$, then $\lim_{n \rightarrow \infty} \omega(A_n) = \omega(A)$,
- For each selfadjoint $A \in \mathcal{A}$ there is a probability measure $\mu_{A, \omega}$ on $\sigma(A)$ such that

$$\omega(A) = \int_{\sigma(A)} \lambda \, d\mu_{A, \omega}(\lambda) .$$

Hence, states can be interpreted as expectation values.

The following result is quite important, since it says that C^* -algebras can be represented in Hilbert spaces, provided a state is singled out on it. This result is known as the *GNS construction* or *GNS representation*, due to its original discoveres Gelfand¹⁰, Naimark¹¹ e Segal¹².

Theorem [GNS Representation] *Let ω be a state of an algebra C^* which we will denote by $\mathcal{A}$. It is possible to construct a Hilbert space $\mathcal{H}_\omega$ and a representation π_ω of $\mathcal{A}$ by bounded operators acting on $\mathcal{H}_\omega$ such that*

$$\pi_\omega(A^*) = \pi_\omega(A)^*$$

¹⁰Israel Moiseevic Gelfand (1913–2009).

¹¹Mark Aronovich Naimark (or Neumark) (1909–1978).

¹²Irving Ezra Segal (1918–1998).

for all $A \in \mathcal{A}$ (a representation with this property is said to be a $*$ -representation). Furthermore, if the algebra $\mathcal{A}$ has a unit, then there exists in $\mathcal{H}_\omega$ a vector Ω with the property that

$$\omega(A) = \langle \Omega, \pi_\omega(A)\Omega \rangle_{\mathcal{H}_\omega}.$$

This vector Ω is a cyclic vector for the representation π_ω , that is, $\{\pi_\omega(A)\Omega, A \in \mathcal{A}\}$ is a dense set in $\mathcal{H}_\omega$. $\square$

By a representation of the algebra one means a map π_ω from $\mathcal{A}$ into the bounded linear operators acting on $\mathcal{H}_\omega$ such that

$$\pi_\omega(\alpha A + \beta B) = \alpha \pi_\omega(A) + \beta \pi_\omega(B) \quad \text{and} \quad \pi_\omega(AB) = \pi_\omega(A)\pi_\omega(B),$$

for all $A, B \in \mathcal{A}$ and for all $\alpha, \beta \in \mathbb{C}$.

For the sake of the reader we provide a simplified proof of the statement above (see also, e.g., [7, 29]). Consider $\mathcal{A}$ as a vector space and identify $\Omega \equiv \mathbb{1}$. Define $\pi_\omega(A)\Omega := A\mathbb{1} = A$. For two vectors $A, B \in \mathcal{A}$, define a scalar product¹³ $\langle A, B \rangle_\omega := \omega(A^*B)$, that means

$$\langle \pi_\omega(A)\Omega, \pi_\omega(B)\Omega \rangle_\omega = \omega(A^*B).$$

Taking $A = \mathbb{1}$, this is particular says that

$$\langle \Omega, \pi_\omega(B)\Omega \rangle_\omega = \omega(B).$$

By these definitions, one has $\|\pi_\omega(A)\Omega\|^2 = \omega(A^*A)$. Now, complete the set $\{\pi_\omega(A)\Omega, A \in \mathcal{A}\}$ in this norm, producing a Hilbert space $\mathcal{H}_\omega$.

By definition, $\{\pi_\omega(A)\Omega, A \in \mathcal{A}\}$ is a dense set in $\mathcal{H}_\omega$. Moreover,

$$\begin{aligned} \|\pi_\omega(A)\pi_\omega(B)\Omega\|_\omega^2 &= \langle \pi_\omega(A)\pi_\omega(B)\Omega, \pi_\omega(A)\pi_\omega(B)\Omega \rangle_\omega = \omega(B^*A^*AB) \\ &\leq \|A^*A\| \omega(B^*B) = \|A^*A\| \|\pi_\omega(B)\Omega\|_\omega^2, \end{aligned}$$

and, therefore

$$\|\pi_\omega(A)\|^2 \leq \|A^*A\| = \|A\|^2,$$

showing that $\pi_\omega(A)$ are bounded operators acting on $\mathcal{H}_\omega$. Finally, for $A, B, C \in \mathcal{A}$ one has

$$\pi_\omega(A)\pi_\omega(B)\pi_\omega(C)\Omega = ABC\mathbb{1} = \pi_\omega(AB)\pi_\omega(C)\Omega$$

and, hence, $\pi_\omega(A)\pi_\omega(B) = \pi_\omega(AB)$, since $\{\pi_\omega(C)\Omega, C \in \mathcal{A}\}$ is a dense set in $\mathcal{H}_\omega$.

A remarkable fact about the GNS construction is that the purity of the state ω reflects on algebraic properties of the representation π_ω , namely if ω is a pure state, π_ω is irreducible [1, 29]. What happens if ω is not a pure state? Take, for instance, $\omega = \lambda\omega_1 + (1 - \lambda)\omega_2$, with ω_1 and ω_2 pure. Then, $\pi_\omega(A)$ has the form of a direct sum

$$\pi_\omega(A) = \begin{pmatrix} \pi_{\omega_1}(A) & 0 \\ 0 & \pi_{\omega_2}(A) \end{pmatrix},$$

that means $\pi_\omega = \pi_{\omega_1} \oplus \pi_{\omega_2}$ and $\mathcal{H}_\omega$ also splits into two orthogonal subspaces where π_{ω_1} and π_{ω_2} act irreducibly.

¹³In many cases this is just a bilinear form, but it can be restricted to a scalar product by taking quotients.

There is a very beautiful interpretation of these facts: if ω is not pure, the decomposition into a direct sum $\pi_\omega = \pi_{\omega_1} \oplus \pi_{\omega_2}$ indicates that π describes a system decomposed into independent subsystems, reflecting the fact that ω is composed by a statistical mixture of independent states.

We now turn to the notion of von Neumann algebras. Let $\mathcal{A}$ be a C^* -subalgebra of some $\mathcal{B}(\mathcal{H})$. We denote by $\mathcal{A}'$ the commutant of $\mathcal{A}$:

$$\mathcal{A}' := \left\{ B \in \mathcal{B}(\mathcal{H}) \mid BA = AB \text{ for all } A \in \mathcal{A} \right\}.$$

By definition, one has

$$\mathcal{A} \subset \mathcal{A}''.$$

A C^* -subalgebra $\mathcal{N}$ of some $\mathcal{B}(\mathcal{H})$ is said to be a *von Neumann algebra* if

$$\mathcal{N} = \mathcal{N}''.$$

Equivalently (by von Neumann's bicommutant theorem, see, *e.g.*, [7, 29]), $\mathcal{N}$ is a von Neumann algebra if it is *weakly closed*: if for any sequence $A_n \in \mathcal{N}$ one has

$$\lim_{n \rightarrow \infty} \langle \psi, A_n \phi \rangle = \langle \psi, A \phi \rangle$$

for all $\psi, \phi \in \mathcal{H}$, then $A \in \mathcal{N}$.

Von Neumann algebras are C^* -algebras that are realised as concrete algebras of bounded operators acting in some concrete Hilbert space.

The Algebraic Formulation of QFT

In Minkowski space-time, the Algebraic Formulation of QFT is based on the Haag¹⁴-Kastler¹⁵ postulates [4, 5, 12]. The main ingredient of this formulation is a net of von Neumann algebras acting on a common Hilbert space $\mathcal{H}$, that means, an association $\mathcal{O} \rightarrow \mathcal{A}(\mathcal{O})$ that assigns to every open subset $\mathcal{O}$ of Minkowski space with compact closure a von Neumann algebra $\mathcal{A}(\mathcal{O})$ under some additional requirements:

- Isotony: $\mathcal{A}(\mathcal{O}_1) \subset \mathcal{A}(\mathcal{O}_2)$ if $\mathcal{O}_1 \subset \mathcal{O}_2$. This allows to define the C^* -algebra

$$\mathfrak{A} := \bigcup_{\mathcal{O}} \mathcal{A}(\mathcal{O})$$

as an inductive limit.

- Einstein causality:

$$\mathcal{A}(\mathcal{O}_1) \subset \mathcal{A}(\mathcal{O}_2)'$$

if $\mathcal{O}_1$ and $\mathcal{O}_2$ are space-like. Here $\mathcal{A}(\mathcal{O}_2)'$ denotes the commutant of $\mathcal{A}(\mathcal{O}_2)$: the set of all bounded operators acting on $\mathcal{H}$ that commute with all the elements of $\mathcal{A}(\mathcal{O}_2)$.

This property is compatible with the choice of the algebras $\mathcal{A}(\mathcal{O})$ as von Neumann algebras.

¹⁴Rudolf Haag (1922–2016).

¹⁵Daniel Kastler (1926–2015).

- Poincaré covariance: for g in the Poincaré group

$$\mathcal{A}(g\mathcal{O}) = U(g)^* \mathcal{A}(\mathcal{O}) U(g) ,$$

$U(g)$ unitary.

- Spectrum condition: the joint spectrum of the generators of translations is contained in the closed forward light cone.
- Vacuum vector: existence of vector $\Omega \in \mathcal{H}$ that is cyclic, Poincaré-invariant and normalised.

The algebras $\mathcal{A}(\mathcal{O})$ are to be interpreted as algebras of observables associated to a quantum field model. The requirements above correspond to a minimal set of assumptions, all physically grounded, that should be found in any relativistic Quantum Field Theory model in Minkowski space-time. Generalisations to more general curved space-times can be presented [33], but will not be considered here.

The postulates above are the beginning of a fully developed research field dealing with the foundations of quantum field theories and the reader can find in [4, 5] excellent starting points for an additional study. We shall limit ourselves to comment on some remarkable results with striking consequences.

Reeh-Schlieder Theorem and Bisognano-Wichmann Theorem

We start with some basic notions. For a Haag-Kastler net as described above, as before, the inductive limit of C^* -algebras

$$\mathfrak{A} := \bigcup_{\mathcal{O}} \mathcal{A}(\mathcal{O}) ,$$

where the union is taken over all open sets $\mathcal{O}$ in Minkowski space-times with compact closure. We say that *weak additivity* holds for a state φ on $\mathfrak{A}$ if

$$\pi_{\varphi}(\mathfrak{A})'' = \left(\bigcup_{x \in \mathbb{M}} \pi_{\varphi}(\mathcal{A}(\mathcal{O} + x)) \right)'' .$$

Let $\mathcal{N}$ be a von Neumann algebra acting on a Hilbert space $\mathcal{H}$. A vector $\Omega \in \mathcal{H}$ is said to be

- cyclic for $\mathcal{N}$ if the set of vectors $\{A\Omega, A \in \mathcal{N}\}$ is dense in $\mathcal{H}$.
- separating for $\mathcal{N}$ if $A\Omega = 0$ for some $A \in \mathcal{N}$ only if $A = 0$.

This means that Ω has no annihilators in $\mathcal{N}$.

Then we have the following

Theorem. Reeh-Schlieder Theorem [16]: Assume weak additivity for the vacuum state ω . Then, the vacuum vector Ω obtained from the GNS construction from ω is

- a **cyclic** vector for $\pi_{\omega}(\mathcal{A}(D))$ for any open domain D .

- a **cyclic** and **separating** vector for $\pi_\omega(\mathcal{A}(D))$ for any open domain D with $D' \neq \emptyset$, where D' stands for the set of all points in Minkowski space-times that are space-like with respect to all points of D . $\square$

A proof of this theorem in the algebraic setting can be found for instance in [5].

This a remarkable result, among other things, because it claims (due to cyclicity) that all states in the Hilbert space can be *approximated* by acting on the vacuum vector with local observables contained in any domain D , a quite counter-intuitive result!

This results has some resemblance with a well-known result from Functional Analysis, Wiener's¹⁶ Theorem in $L^2(\mathbb{R}, dx)$ [7, 34], that claims that for any finite interval $I \equiv (a, b) \subset \mathbb{R}$ the set of all translates of a Gaussian by elements of I is a total set, *i.e.*, a set whose orthogonal complement in $L^2(\mathbb{R}, dx)$ consists on the zero vector only.

Another remarkable fact about Reeh-Schlieder's Theorem is that it is closely related with a well established theorem from the theory of von Neumann algebras, the Tomita-Takesaki Theorem (see, *e.g.*, [29]).

Theorem. Tomita-Takesaki Theorem. Let $\mathcal{N}$ be a von Neumann algebra and Ω be a cyclic and separating vector for $\mathcal{N}$. Then,

- There is a antilinear operator J such that

$$J\mathcal{N}J = \mathcal{N}'$$

and with $J^2 = \mathbb{1}$.

- These is a positive and selfadjoint operator Δ such that

$$\Delta^{it}\mathcal{N}\Delta^{-it} = \mathcal{N}$$

for all $t \in \mathbb{R}$.

- Ω is a KMS state (a temperature state) with $\beta = 1/2$ for the dynamics defined in $\mathcal{N}$ by Δ^{it} , $t \in \mathbb{R}$. $\square$

A proof of this theorem can be found for instance in [29, 32].

As we see, the claims of the Reeh-Schlieder Theorem are the hypotheses of the Tomita-Takesaki Theorem. Therefore, applications to Quantum Field Theory are immediate. This was first pointed out in a celebrated article by Bisognano and Wichmann [17], later extended by Geoffrey Sewell [35], showing, among other things, that an observer in constant acceleration in Minkowski space-times “sees” the vacuum as a temperature state, an effect also known as the Unruh effect, a similar effect to Hawking's radiation and Hawking's temperature in Black Hole Physics. The original treatments of Unruh and Hawking, though, where perturbative and semiclassical, but Bisognano and Wichmann's Theorem is general and of an structural nature, thus showing the power of the algebraic treatment of Quantum Field Theories.

2.7 Uncertainty Relations

In this section we will discuss inequalities that reveal notable distinctions between Quantum Physics and Classical Physics. They represent an important contribution of the operatorial formalism to some fundamental questions.

¹⁶Norbert Wiener (1894–1964).

As we will see, one of the basic reason behind the distinction between the Quantum Physics and Classical Physics lies in the non-commutative nature of the algebra of observables in Quantum Physics, a fact that manifests itself, for example, with the operators of momentum and position (in one dimension), for which we have Heisenberg's canonical commutation relation $[p, x] = -i\hbar$ (in the one-dimensional case).

Here we will present Heisenberg's uncertainty relations and an improvement on these, Schrödinger's¹⁷ uncertainty relations. It is noteworthy that most quantum mechanics textbooks focus on the treatment of Heisenberg relations and ignore Schrödinger relations (a notable exception being [36]). However, Schrödinger relations have become important in recent decades, especially in the context of quantum optics and the study of so-called squeezed states.

In our presentation, we will first address the case of observables described by bounded operators, and then the mathematically more elaborate case of observables described by unbounded operators acting in a Hilbert space.

We note, in passing, that uncertainty relations such as Heisenberg's below also exist in the theory of Fourier Transforms, as presented in [7, 37].

Some Stochastic Definitions

In the description of quantum systems, following all our previous discussion, observables will be selfadjoint elements of an algebra with unity $\mathcal{A}$ and expected values will be described by states in this algebra, that is, by positive linear functionals (i.e., $\omega(A^*A) \geq 0$ for all $A \in \mathcal{A}$) and normalised (i.e., $\omega(1) = 1$). In the case of pure states, we can describe them by vector states in a Hilbert space, where $\mathcal{B}(\mathcal{H})$ is the algebra of the observables.

We will make repeated use of the Cauchy-Schwarz inequality, which in the present algebraic context states that, for operators A and B and for a state ω , the following holds:

$$|\omega(A^*B)| \leq \sqrt{\omega(A^*A)\omega(B^*B)} \quad (2.8)$$

Detailed proofs of this relation can be found, for instance, in [7, 29].

Note that if A and B are selfadjoint operators, (2.8) says that

$$|\omega(AB)| \leq \sqrt{\omega(A^2)\omega(B^2)} . \quad (2.9)$$

Let's recall some of the definitions of stochastic quantities that we have already presented, adapted to the context of Quantum Physics. The average value of an observable¹⁸ B in a state ω is given by

$$\langle B \rangle_\omega := \omega(B) . \quad (2.10)$$

For example, if ω is a pure state associated with a ψ in a Hilbert space $\mathcal{H}$ (with $\|\psi\|_{\mathcal{H}} = 1$) we have $\langle B \rangle_\omega := \omega(B) = \langle \psi, B\psi \rangle_{\mathcal{H}}$. If ω is described by a density matrix ρ (a selfadjoint, positive operator with trace equal to 1), we will have $\langle B \rangle_\rho := \omega(B) = \text{Tr}(\rho B)$.

¹⁷Erwin Rudolf Josef Alexander Schrödinger (1887–1961).

¹⁸In this section we will consider, unless explicitly stated otherwise, only observables described by bounded operators acting in a Hilbert space. However, all relevant results also apply to unbounded operators, provided that appropriate restrictions are made on the pure states considered to the domains of these operators.

The variance, or quadratic deviation, of the same observable in the same state is given by

$$\text{Var}_\omega(B) := \left\langle (B - \langle B \rangle_\omega)^2 \right\rangle_\omega = \langle B^2 \rangle_\omega - \langle B \rangle_\omega^2 = \omega(B^2) - \omega(B)^2, \quad (2.11)$$

and represents a measure of how much the measurements of the observable B deviated from its average value in the state ω .

Another statistically significant quantity is the so-called *covariance*, or *correlation*, between observables. If A and B are observables and ω is a state, the *covariance* of these observables is defined by

$$\begin{aligned} \text{Cov}_\omega(A, B) &:= \frac{1}{2}\omega\left((A - \omega(A))(B - \omega(B)) + (B - \omega(B))(A - \omega(A))\right) \\ &= \frac{1}{2}\omega(AB + BA) - \omega(A)\omega(B). \end{aligned} \quad (2.12)$$

This definition is similar to that presented in (2.3) and takes into account the fact that A and B do not necessarily commute. Since A and B are selfadjoint, $\frac{1}{2}(AB + BA)$ is also selfadjoint, and therefore, $\text{Cov}_\omega(A, B)$ is real, but not necessarily positive. $\text{Cov}_\omega(A, B)$ measures how much A and B jointly deviate on average from their mean values. Note also that

$$\text{Var}_\omega(A) = \text{Cov}_\omega(A, A). \quad (2.13)$$

Another relevant point to note is the inequality

$$|\text{Cov}_\omega(A, B)|^2 \leq \text{Var}_\omega(A)\text{Var}_\omega(B). \quad (2.14)$$

It is a simple consequence of the definition (2.12) and the Cauchy-Schwarz inequality:

$$\begin{aligned} \left| \omega\left((A - \omega(A))(B - \omega(B))\right) \right|^2 &\leq \omega\left((A - \omega(A))^2\right)\omega\left((B - \omega(B))^2\right) \\ &= \text{Var}_\omega(A)\text{Var}_\omega(B). \end{aligned}$$

Remember that A and B are selfadjoint.

In the case of bounded operators, all the statistical notions above are perfectly well defined. However, care must be taken when extending them to include the case of unbounded selfadjoint operators, a case of interest in many discussions about Quantum Mechanics, for example, involving the uncertainty relations of position and momentum operators. The case of unbounded operators will be studied in Section 2.7.3.

As an exercise, note that (2.14) means that

$$\det(\Sigma_\omega(A, B)) \geq 0, \quad \text{where} \quad \Sigma_\omega(A, B) := \begin{pmatrix} \text{Cov}_\omega(A, A) & \text{Cov}_\omega(A, B) \\ \text{Cov}_\omega(B, A) & \text{Cov}_\omega(B, B) \end{pmatrix} \quad (2.15)$$

is called the covariance matrix of the observables A and B in the state ω . This matrix will play a role later on.

A historical comment

The so-called Uncertainty Principle was first formulated by Heisenberg in 1927, in the early days of Quantum Mechanics, in a work that aimed to clarify the conceptual content

of this newly born theory in contrast to the view offered by Classical Physics. In that work, Heisenberg considered a quantum particle (an electron) moving in one dimension, with p and q being the momentum and position observables associated with it, and considered a process for measuring the position observable through the interaction of the particle with a gamma-ray microscope (the so called “*Heisenberg microscope*”). In his analysis, Heisenberg showed that if this measurement process provides the position with an error of $\varepsilon(q)$ (at the resolution of the microscope), then the momentum of the particle is altered by an amount of $\delta(p)$ and so that the (mathematically informal) relation $\delta(p)\varepsilon(q) \sim \hbar$ holds.

This relationship exposes the idea that each of the quantities $\delta(p)$ and $\varepsilon(q)$ cannot be made arbitrarily small without the other becoming large, thus expressing the impossibility of determining with arbitrarily small precision the position of a particle without a drastically large transfer of momentum to it, suggesting, therefore, that the state of a quantum particle does not allow the simultaneous knowledge of these two quantities. Heisenberg pointed out that this circumstance is an essential characteristic of Quantum Mechanics, as it does not manifest itself in Classical Mechanics. This idea became known as Heisenberg’s Uncertainty Principle and is a fundamental conceptual piece of Quantum Mechanics, discussed in all textbooks that deal with the subject.

Following on from Heisenberg’s conceptual work, in the same year of 1927, Kennard¹⁹ provided more mathematical and conceptual precision to Heisenberg’s relation, expressing it in the form of an inequality involving the variances of p and q :

$$\text{Var}(p)\text{Var}(q) \geq \hbar^2/4. \quad (2.16)$$

(A side note: possibly due to a calculation error, Kennard actually obtained $\text{Var}(p)\text{Var}(q) \geq \hbar^2$).

The inequality (2.16), called Heisenberg’s uncertainty principle, states in mathematically precise terms the idea that, in a given quantum state, each of the stochastic quantities $\text{Var}(p)$ and $\text{Var}(q)$ cannot be made arbitrarily small without the other becoming large, thus expressing the intrinsic impossibility of a quantum state allowing the knowledge of both quantities with arbitrary precision, regardless of what occurs in a measurement process.

En passant, (2.16) also informs that neither p nor q can have eigenstates, since the variance of each cannot be zero.

Kennard’s result was also obtained around the same time by Weyl²⁰ and presented in one of his classic books on applications of Group Theory to Quantum Physics [39]. In that book, however, Weyl attributes his observations to Pauli²¹. In fact, a demonstration similar to Weyl’s can be found in Pauli’s book [40]. The Weyl-Pauli deduction is much simpler than Kennard’s and can also be found in quantum mechanics textbooks, such as in [41] and [40].

In contrast to Heisenberg’s, the Kennard-Weyl-Pauli analysis does not involve a specific measurement process and describes an intrinsic stochastic relationship between the variances of p and q , independent of what occurs in a particular measurement process of these observables.

¹⁹Earle H. Kennard (1885-1968). The work in question is [38].

²⁰Hermann Klaus Hugo Weyl (1885-1955).

²¹Wolfgang Ernst Pauli (1900-1958)

Thus, as someone has already stated, the uncertainty principle does not claim that we observe the world through an opaque window, but that this opacity is inherent in Nature.

The formulations of Heisenberg and Kennard are not the same, and it is in Kennard's version, and not in Heisenberg's, that the uncertainty principle is currently understood and presented.

The next historical step was an extension, obtained by Robertson²² in 1929, of inequality (2.16). Robertson's result involves an inequality for the product of variances of certain operators more general than those of position and momentum. In truth, Robertson's results were not yet ideal, as they were valid for a restricted class of observables. What we will present in Section 2.7.1 is an even more general version of Robertson's inequality, with fewer restrictions on the type of operators considered.

One of the important aspects of the demonstration of uncertainty relations that we will present is to highlight the role of non-commutativity of observables in Quantum Physics, which manifests itself, for example, in the Heisenberg commutation relations $[p, q] = -i\hbar$ between the momentum and position observables of a particle in one-dimensional motion.

The last historical step we want to consider was an important improvement on the uncertainty relations obtained by Erwin Schrödinger in 1930 [42]. This Schrödinger uncertainty relation, still rarely discussed in textbooks today, has recently become important in developments in Quantum Optics and Quantum Information Theory. We will dedicate Section 2.7.2 to it.

2.7.1 Heisenberg's Uncertainty Principle

The first inequality we will discuss is known as Heisenberg's Uncertainty Relation, which we will present first in its most general form, attributed, as mentioned above, to Robertson. In Section 2.7.2, we will deal with the version of the uncertainty relations that was obtained by Schrödinger.

Deducing Heisenberg's Uncertainty Relation

To arrive at the desired result, we need two simple mathematical steps. The first is the simple observation that, for an operator A (not necessarily selfadjoint), we have

$$\omega(A^*A) \geq 0, \quad (2.17)$$

due to the positivity of the state.

The second step for proving the uncertainty relation is the following. Let λ be a real number and let

$$p(\lambda) = a\lambda^2 + b\lambda + c, \quad (2.18)$$

with $a \neq 0$, a second-degree polynomial in λ with real coefficients a , b , and c . Suppose we know that $p(\lambda) \geq 0$ for all $\lambda \in \mathbb{R}$. This is obviously only possible if $a > 0$ and implies, in particular, that $p(\lambda_0) \geq 0$, where λ_0 is the point where $p(\lambda)$ assumes a minimum. It is easy to calculate that

$$\lambda_0 = -\frac{b}{2a} \quad \text{and that} \quad p(\lambda_0) = c - \frac{b^2}{4a}. \quad (2.19)$$

²²Howard Percy Robertson (1902–1961).

Therefore, under the hypothesis that $p(\lambda)$ is always greater than or equal to zero, it follows that

$$c - \frac{b^2}{4a} \geq 0. \quad (2.20)$$

Let A and B be two selfadjoint operators and let ω be a state. For a real λ , we have by (2.17) that

$$\omega\left((A + i\lambda B)^*(A + i\lambda B)\right) \geq 0. \quad (2.21)$$

Writing the left-hand side explicitly, we have:

$$\omega(B^2) \lambda^2 + \omega(i[A, B]) \lambda + \omega(A^2) \geq 0. \quad (2.22)$$

Note that the coefficients of the polynomial on the left-hand side are real numbers, since $i[A, B]$ is selfadjoint if A and B are.

We can then use (2.20) and, assuming without loss of generality (see below) that $\omega(B^2) \neq 0$, we will have,

$$\omega(A^2) - \frac{\omega(i[A, B])^2}{4\omega(B^2)} \geq 0, \quad (2.23)$$

i.e.,

$$\omega(A^2)\omega(B^2) \geq \frac{1}{4}\omega(i[A, B])^2. \quad (2.24)$$

(Note, *en passant*, that this inequality is also valid if $\omega(B^2) = 0$ because, in that case,

$$\left|\omega([A, B])\right| \leq |\omega(AB)| + |\omega(BA)| \leq 2\sqrt{\omega(A^2)\omega(B^2)} = 0,$$

where the Cauchy-Schwarz inequality (2.8) was used and the fact that A and B are selfadjoints. Thus, if B is selfadjoint and $\omega(B^2) = 0$, both sides of (2.24) cancel each other out, preserving the truth of that expression).

Relation (2.24) holds for any observables A and B . Let us take in its place, in particular, the observables $A' = A - \omega(A)$ and $B' = B - \omega(B)$. Since $[A', B'] = [A, B]$, the relation (2.24) tells us that

$$\left\langle (A - \langle A \rangle_\omega)^2 \right\rangle_\omega \left\langle (B - \langle B \rangle_\omega)^2 \right\rangle_\omega \geq \frac{1}{4} \langle i[A, B] \rangle_\omega^2, \quad (2.25)$$

that is, that

$$\text{Var}_\omega(A) \text{Var}_\omega(B) \geq \frac{1}{4} \langle i[A, B] \rangle_\omega^2. \quad (2.26)$$

In terms of the standard deviations of A and B , this is written as

$$\Delta_\omega(A) \Delta_\omega(B) \geq \frac{1}{2} \left| \langle [A, B] \rangle_\omega \right|. \quad (2.27)$$

Note that the right-hand side of the equation may be non-zero if the observables A and B do not commute with each other.

Relation (2.26) is the *Heisenberg uncertainty relation* for the observables A and B in the state ω . The generalisation to any observables that we presented above was obtained (by other means) by Robertson.

Some authors refer to (2.26) as the *Heisenberg-Robertson Uncertainty Relation* and call only the particular case of position and momentum operators, which we will discuss later, the *Heisenberg Uncertainty Relation*.

2.7.2 Schrödinger's Uncertainty Principle

The inequality expressed in (2.26) has an important refinement, which is due to Schrödinger [42]. We again assume A and B are selfadjoint and bounded. For simplicity, let us provisionally consider $A' := A - \omega(A)\mathbb{1}$ and $B' := B - \omega(B)\mathbb{1}$.

These operators A' and B' are equally bounded and selfadjoint. We have, trivially,

$$A'B' = \frac{1}{2}(A'B' + B'A') + \frac{1}{2}[A', B']$$

and therefore

$$\omega(A'B') = \frac{1}{2}\omega(A'B' + B'A') - \frac{i}{2}\omega(i[A', B']) . \quad (2.28)$$

The two quantities $\omega(A'B' + B'A')$ and $\omega(i[A', B'])$ are real numbers, since $A'B' + B'A'$ and $i[A', B']$ are selfadjoint. Therefore, we can consider (2.28) as the decomposition of the complex number $\omega(A'B')$ into real and imaginary parts and, with this, state that

$$|\omega(A'B')|^2 = \frac{1}{4}\left(\omega(A'B' + B'A')\right)^2 + \frac{1}{4}\left(\omega(i[A', B'])\right)^2 . \quad (2.29)$$

Now, by the Cauchy-Schwarz inequality (2.8), $|\omega(A'B')|^2 \leq \omega((A')^2)\omega((B')^2)$. Thus, returning to (2.29) we obtain

$$\omega((A')^2)\omega((B')^2) \geq \frac{1}{4}\left(\omega(A'B' + B'A')\right)^2 + \frac{1}{4}\left(\omega(i[A', B'])\right)^2 . \quad (2.30)$$

We have $\omega((A')^2) = \omega((A - \omega(A))^2) = \text{Var}_\omega(A)$ and $\omega((B')^2) = \omega((B - \omega(B))^2) = \text{Var}_\omega(B)$, by definition. Moreover,

$$\begin{aligned} \frac{1}{2}\left(\omega(A'B' + B'A')\right) &= \frac{1}{2}\left((A - \omega(A))(B - \omega(B)) + (B - \omega(B))(A - \omega(A))\right) \\ &= \text{Cov}_\omega(A, B) . \end{aligned}$$

Finally, $[A', B'] = [A, B]$. Bringing these facts together, (2.30) can be rewritten as

$$\text{Var}_\omega(A)\text{Var}_\omega(B) \geq (\text{Cov}_\omega(A, B))^2 + \frac{1}{4}\left(\omega(i[A, B])\right)^2 , \quad (2.31)$$

that is,

$$\text{Var}_\omega(A)\text{Var}_\omega(B) \geq \text{Cov}_\omega(A, B)^2 + \frac{1}{4}\left\langle i[A, B] \right\rangle_\omega^2 . \quad (2.32)$$

It is clearly perceived, due to the positivity of the first term on the right-hand side, that (2.26) is a consequence of (2.32) and that this inequality is more precise, that is, more restrictive, than (2.26). Relation (2.32) is sometimes referred to as the *Schrödinger Uncertainty Relation*, or the *Modified Uncertainty Relation*.

If we recall (2.13) and the fact that $\text{Cov}_\omega(A, B) = \text{Cov}_\omega(B, A)$, the Schrödinger Uncertainty Relation (2.32) can be rewritten in a more compact and elegant way as

$$\det \begin{pmatrix} \text{Cov}_\omega(A, A) & \text{Cov}_\omega(A, B) \\ \text{Cov}_\omega(B, A) & \text{Cov}_\omega(B, B) \end{pmatrix} \geq \frac{1}{4}\left\langle i[A, B] \right\rangle_\omega^2 . \quad (2.33)$$

The matrix

$$\Sigma_\omega(A, B) := \begin{pmatrix} \text{Cov}_\omega(A, A) & \text{Cov}_\omega(A, B) \\ \text{Cov}_\omega(B, A) & \text{Cov}_\omega(B, B) \end{pmatrix} \quad (2.34)$$

is called the *covariance matrix* of the observables A and B in the state ω . We had already established in (2.15) that $\det(\Sigma_\omega(A, B)) \geq 0$. Inequality (2.32) can be seen as a refinement of this.

Finally, we mention that for the case already considered of the observables x and p_x in the Hilbert space $L^2(\mathbb{R}, dx)$ the Schrödinger uncertainty relation (2.33) assumes the form

$$\det \begin{pmatrix} \text{Cov}_\omega(x, x) & \text{Cov}_\omega(x, p_x) \\ \text{Cov}_\omega(p_x, x) & \text{Cov}_\omega(p_x, p_x) \end{pmatrix} \geq \frac{\hbar^2}{4}. \quad (2.35)$$

We will see that this expression is invariant under symplectic linear transformations in phase space.

Invariance groups of uncertainty relations

A relevant point to be addressed regarding the relation (2.33) is its covariance under certain linear transformations. Let's first study how the covariance matrix (2.34) of two observables A and B is transformed by linear transformations $\begin{pmatrix} A' \\ B' \end{pmatrix} = G \begin{pmatrix} A \\ B \end{pmatrix}$, with $G = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{Mat}(\mathbb{R}, 2)$, where $a, b, c, d \in \mathbb{R}$. The transformations considered are real so that the property of the observables being selfadjoint is not lost.

Since $A' = aA + bB$ and $B' = cA + dB$, it is elementary to verify that

$$\text{Cov}_\omega(A', A') = a^2 \text{Cov}_\omega(A, A) + 2ab \text{Cov}_\omega(A, B) + b^2 \text{Cov}_\omega(B, B), \quad (2.36)$$

$$\begin{aligned} \text{Cov}_\omega(A', B') &= \text{Cov}_\omega(B', A') \\ &= ac \text{Cov}_\omega(A, A) + (ad + cb) \text{Cov}_\omega(A, B) + bd \text{Cov}_\omega(B, B) \end{aligned} \quad (2.37)$$

and

$$\text{Cov}_\omega(B', B') = c^2 \text{Cov}_\omega(A, A) + 2cd \text{Cov}_\omega(A, B) + d^2 \text{Cov}_\omega(B, B). \quad (2.38)$$

Thus, we obtain the following linear transformation rule for the covariance matrix:

$$\Sigma_\omega(A', B') = G \Sigma_\omega(A, B) G^T. \quad (2.39)$$

From this we conclude that

$$\det(\Sigma_\omega(A', B')) = (\det G)^2 \det(\Sigma_\omega(A, B)). \quad (2.40)$$

Secondly, for the same linear transformations, we have for the commutator

$$[A', B'] = [aA + bB, cA + dB] = (ad - bc)[A, B] = (\det G)[A, B]. \quad (2.41)$$

From this we extract the important observation that if $\det G$ is non-zero, both sides of (2.33) are multiplied by the same positive factor $(\det G)^2$ when performing a linear transformation such as the one considered here. We can capture these considerations in the following proposition:

Proposition. The Schrödinger uncertainty relation (2.33) is preserved by linear transformations of the observables $\begin{pmatrix} A' \\ B' \end{pmatrix} = G \begin{pmatrix} A \\ B \end{pmatrix}$, with $G \in \text{GL}(\mathbb{R}, 2)$, the group of real non-singular 2×2 matrices. $\square$

The group of linear transformations that preserve Schrödinger's uncertainty relations is larger than that corresponding to the uncertainty relations of Heisenberg. Due to (2.13) and (2.36)–(2.38), it is easy to see that only linear transformations such as $A' = \alpha A$ and $B' = \beta B$, with $\alpha, \beta \in \mathbb{R} \setminus \{0\}$, do not alter (2.26).

2.7.3 Uncertainty Relations for Unbounded Operators

We will now discuss the problems that arise when we try to generalise the uncertainty relations to situations in which we deal with unbounded selfadjoint operators. Before that, however, we will present an important theorem, due to Wielandt²³, on Heisenberg's commutation relations. After that, we will discuss, as motivation, a pathological example of the application of Heisenberg's uncertainty relations, that of a particle moving in a circle.

Wielandt's Theorem

Before we proceed, let's make an observation about the Heisenberg's commutation relations. Let's assume that q and p are selfadjoint operators on a common domain in a Hilbert space $\mathcal{H}$ and satisfy $[p, q] = -i\hbar$. We claim that at least one of the two cannot be a bounded operator. This statement is sometimes called Wielandt's Theorem [43].

Let's assume, by contradiction, that $\|p\|$ and $\|q\|$ are finite. Then, we have $\hbar = \|[p, q]\| \leq 2\|p\| \|q\|$. It is easy to prove by induction that it holds $[p, q^n] = -in\hbar q^{n-1}$ for all $n \in \mathbb{N}$. Thus, we have $n\hbar \|q^{n-1}\| \leq \|[p, q^n]\| \leq 2\|p\| \|q^n\|$. Now, since q is selfadjoint, it holds $\|q^n\| = \|q\|^n$ for all $n \in \mathbb{N}$. Thus, we have the inequality $n\hbar \|q\|^{n-1} \leq 2\|p\| \|q\|^n$. Since, naturally, we assume that $\|q\| \neq 0$, we can cancel a factor $\|q\|^{n-1}$ from both sides and obtain $n\hbar \leq 2\|p\| \|q\|$.

Now, this inequality cannot be satisfied for all $n \in \mathbb{N}$, unless $\hbar = 0$ (a trivial case) or that p or q (or both) is/are unbounded, proving what we wanted.

It should be noticed that there are, indeed, situations in which one of the operators q or p is bounded. This is the case, for example, when considering a particle in one-dimensional motion in a finite interval, for example, in the finite interval $[0, L]$, with $L > 0$. The Hilbert space to be considered is $L^2([0, L], dx)$ and here $\|q\| = L < \infty$.

Particle moving on a line

Expressions (2.26) and (2.32) were obtained under the hypothesis that A and B are bounded operators. In the case where the state ω is a pure state associated with a normalised vector ψ in a Hilbert space, that is, $\omega(\cdot) = \langle \psi, \cdot \psi \rangle$, it is possible to reproduce exactly the same proofs that led to (2.26) and (2.32) (since they make use only of the Cauchy-Schwarz inequality) even for unbounded selfadjoint operators A and B , provided that ψ belongs to the domain of the operators $A, A^2, B, B^2, AB - BA$ and $AB + BA$.

²³Helmut Wielandt (1910–2001).

These conditions, however, are very restrictive in certain cases of interest, but in the we will not present the mathematically adequate treatment of these issues here and refer the reader, as an example, to references such as [44].

The most important illustration of the relation (2.26) occurs in Hilbert space $\mathcal{H} = L^2(\mathbb{R}, dx)$ when we take $A = p_x = -i\hbar \frac{\partial}{\partial x}$ and $B = x$ (the multiplication operator by x) and consider a pure state associated with a normalised vector $\psi \in \mathcal{H}$ satisfying the domain conditions mentioned above. This is the case, for example, if ψ is an element of the Schwartz space $\mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R}, dx)$, which is kept invariant by p_x and x and is dense in $L^2(\mathbb{R}, dx)$. Given that $[p_x, x] = -i\hbar$ (in a domain such as $\mathcal{S}(\mathbb{R})$), the relation (2.26) then states that

$$\text{Var}_\psi(p_x) \text{Var}_\psi(x) \geq \frac{\hbar^2}{4}. \quad (2.42)$$

What does this result mean? The expression (2.42), as well as (2.26), contains statements of a statistical nature. Both refer to average quantities (such as variance), where the average in question is taken over a set of systems, all being described by the same state ψ .

Thus, when we state that for any physical system, say a hydrogen atom, the average value of an observable, say, of the energy H , in a state ψ is $\langle \psi, H\psi \rangle_{\mathcal{H}}$, we are stating that if we make several measurements of the energy of *a large set of hydrogen atoms, all in the state ψ* , with only one measurement for each atom, then, by taking the average of the values thus obtained, we will have as a result the value $\langle \psi, H\psi \rangle_{\mathcal{H}}$.

Similarly, the expression (2.42) tells us that if we make measurements of the momentum component x of, say, an electron moving around a proton for a large set of identical hydrogen atoms, all in a state ψ , and we make measurements of the position x , also of an electron moving around a proton for a large set of identical hydrogen atoms, all in a state ψ , then the product of the variances of momentum and position satisfies the inequality (2.42).

We want to emphasise that (2.42), as well as (2.26), refer to measurements on "large" sets of identical systems in the same state, not on individual successive measurements of position and momentum in a single system.

Bearing this in mind, what the relation (2.42) tells us is that in Quantum Mechanics there is no state ψ for which the variance of the x -component of the momentum operator and the variance of the position x can be arbitrarily small in both. This fact is radically contrary to what occurs in Classical Mechanics, because there, at any instant of time, a material particle has well-defined position and momentum.

Particle in a circle. An example to keep in mind

To illustrate the difficulties that arise when we naively deal with Heisenberg's uncertainty relation (2.26) in the case of unbounded operators, let us consider the situation of a quantum particle moving in a circle of radius 1, which we will denote as usually done by $\mathbb{S}^1$. The Hilbert space to be considered is $\mathcal{H} = L^2(\mathbb{S}^1, d\varphi) \equiv L^2([-\pi, \pi], d\varphi)$, the position operator can be adopted as the multiplication operator by the angle $\varphi \in [-\pi, \pi]$, used as a coordinate in the circle $\mathbb{S}^1$. We will denote this operator by $\hat{\varphi}$. Consider the angular momentum operator, which in this case is given by $\hat{\ell} = -i\hbar \frac{\partial}{\partial \varphi}$. The operator $\hat{\varphi}$ is bounded in $\mathcal{H}$, but $\hat{\ell}$ is not.

Formally, we have the commutation relations

$$[\hat{\ell}, \hat{\varphi}] = -i\hbar.$$

These commutation relations are identical to Heisenberg's commutation relations.

If we now naively apply Heisenberg's uncertainty relation (2.26) we would conclude that for $\psi \in \mathcal{H}$, with $\|\psi\| = 1$, the inequality would hold:

$$\text{Var}_\psi(\hat{\ell})\text{Var}_\psi(\hat{\varphi}) \geq \frac{\hbar^2}{4}, \quad (2.43)$$

where, proceeding naively with the definition (2.11),

$$\begin{aligned} \text{Var}_\psi(\hat{\ell}) &:= \langle \psi, \hat{\ell}^2 \psi \rangle - \langle \psi, \hat{\ell} \psi \rangle^2 \quad \text{and} \\ \text{Var}_\psi(\hat{\varphi}) &:= \langle \psi, \hat{\varphi}^2 \psi \rangle - \langle \psi, \hat{\varphi} \psi \rangle^2. \end{aligned} \quad (2.44)$$

However, there is a problem with the inequality (2.43). For $m \in \mathbb{Z}$, the wave function $\psi_m(\varphi) := \frac{1}{\sqrt{2\pi}} e^{im\varphi}$ is an element of $\mathcal{H}$ that is an eigenfunction of $\hat{\ell}$ (with eigenvalue m) and therefore $\text{Var}_{\psi_m}(\hat{\ell}) = 0$, while $\text{Var}_{\psi_m}(\hat{\varphi}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi^2 d\varphi = \pi^2/3$ is a finite quantity. We see, therefore, that the inequality (2.43) is invalid in this situation, because the left side is zero!

Where does the problem lie? This question has been lingering in the literature for some time and has produced numerous works, not all of them correct and satisfactory (and even some that, in the well-known words of Pauli, “are not even wrong”). As we will point, one of the problems lies in the fact that the operator that we naively defined by $\hat{\ell} = -i\hbar \frac{\partial}{\partial \varphi}$ is not a bounded operator in $\mathcal{H}$ and must be defined in a domain in which it is a selfadjoint operator²⁴. As we will see, this domain is not invariant under the action of the multiplication operator $\hat{\varphi}$, which makes the definition of the commutator $[\hat{\ell}, \hat{\varphi}] = \hat{\ell}\hat{\varphi} - \hat{\varphi}\hat{\ell}$ very problematic because if ψ is in a domain where $\hat{\ell}$ is selfadjoint, then $\hat{\varphi}\psi$ is not necessarily in such domain, making the term $\hat{\ell}\hat{\varphi}\psi$ meaningless.

In general, the definition of commutators between unbounded selfadjoint operators is subtle and problematic, and we refer the interested student to references such as [3, 6, 45, 46], where this issue is treated in more depth.

A second problem lies in the very definition of the variances $\text{Var}_\psi(\hat{\ell})$ and $\text{Var}_\psi(\hat{\varphi})$. By the definition we have adopted so far, offered in (2.44), for bounded operators, $\text{Var}_\psi(\hat{\ell})$ and $\text{Var}_\psi(\hat{\varphi})$ are only defined if ψ is simultaneously an element of the selfadjoint domain of $\hat{\ell}$, of $\hat{\varphi}$, of $\hat{\ell}^2$ and of $\hat{\varphi}^2$, which can only have the null vector in common!²⁵

We learn from this that the naive use of Heisenberg's uncertainty relations (2.26) in this case leads to a paradoxical situation and that much more care is needed in dealing with uncertainty relations in the case of unbounded operators. A mathematically adequate treatment of these issues can be found in references such as [44].

²⁴For the theory of unbounded operators, see references such as [3, 6, 45, 46]

²⁵In the case of Schrödinger uncertainty relations, covariances must also be considered, which require additionally taking into account the domains of the operators $\hat{\ell}\hat{\varphi} + \hat{\varphi}\hat{\ell}$.

Acknowledgements

I would like to thank the organisers of the XIV EFRAS for the very stimulating atmosphere.

2.8 References

- [1] J. C. A. Barata, M. Brum, V. Chabu, R. Correa da Silva. “Pure and mixed states”. *Brazilian Journal of Physics*, volume 51, pages 244–262 (2021). <https://doi.org/10.1007/s13538-020-00808-0>
- [2] Werner Heisenberg. “Der Teil und das Ganze: Gespräche im Umkreis der Atomphysik”. R. Piper & Co. Verlag (1969). ISBN 3-492-22297-8.
English edition: Werner Heisenberg. “Physics and Beyond: Encounters and Conversations”. Harper & Row. ISBN 9780049250086.
Portuguese translation: Werner Heisenberg. “Diálogos sobre Física Atômica”. Editorial Verbo (1971).
- [3] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics. Vol. 2: Fourier Analysis, Self-Adjointness*. Academic Press. New York. (1972–1979).
- [4] Rudolf Haag. *Local Quantum Physics*. Second edition. Berlin Heidelberg: Springer Verlag (1996). ISBN 3-540-53610-8. ISBN 0-387-53610-6.
- [5] Huzihiro Araki. *Mathematical Theory of Quantum Fields*. Oxford Science Publications. (1999). ISBN 0 19 851773 4.
- [6] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics. Vol. 1: Functional Analysis*. Academic Press. New York. (1972–1979).
- [7] JCAB. *Notas para um curso de Física-Matemática* (título provisório). <https://fma.if.usp.br/~jbarata/>
- [8] Ola Bratteli and Derek W. Robinson. *Operator Algebras and Quantum Statistical Mechanics II*. Springer Verlag. (1979).
- [9] Silvan S. Schweber. *QED and the men who made it: Dyson, Feynman, Schwinger and Tomonaga*. Princeton University Press. (1994). ISBN 0-691-03685-3.
- [10] Arthur Jaffe, “Divergence of Perturbation Theory for Bosons”. *Commun. Math. Phys.* **1**, 127-149 (1965).
- [11] Barry Simon and A. Dicke, “Coupling constant analyticity for the anharmonic oscillator”. *Annals of Physics*, Volume **58**, Issue 1, p. 76-136 (1970). DOI: 10.1016/0003-4916(70)90240-X.
- [12] R. Haag and D. Kastler, “An Algebraic Approach to Quantum Field Theory”. *J. Math. Phys.* **5**, 848-861 (1964).
- [13] A. S. Wightman. “Quantum Field Theory in Terms of Vacuum Expectation Values”. *Physical Review*, **101**, n. 2, p. 860-866 (1956).
- [14] R. Jost. *The General Theory of Quantized Fields*. American Mathematical Society (1965).
- [15] R. F. Streater and A. S. Wightman. *PCT, Statistics and All That*. Terceira Edição. Princeton Univ. Press. (1980).
- [16] H. Reeh, S. Schlieder, “Bemerkungen zur Unitäräquivalenz von Lorentzinvarianten Feldern”. *Il Nuovo Cimento*. **22**, 1059-1068 (1961).
- [17] J. Bisognano and E. H. Wichmann, “On the duality condition for a Hermitian scalar field”. *J. Math. Phys.* **16**, 985-1007 (1975).

- [18] R. Haag, “Quantum Field Theories with Composite Particles and Asymptotic Conditions”. Phys. Rev. **112**, 669-673 (1958).
- [19] R. Haag, “The Framework of Quantum Field Theory”. Supplemento al volume XIV, serie X del Nuovo Cimento, 1563 (1959).
- [20] D. Ruelle, “On the Asymptotic Condition in Quantum Field Theory”. Helv. Phys. Acta **35**, 147-163 (1962).
- [21] H. Araki, K. Hepp and D. Ruelle, “On the Asymptotic Behavior of Wightman Functions in Space-Like Directions”. Helv. Phys. Acta, **35**, 164 (1962).
- [22] K. Hepp, “On the Connection Between The LSZ and the Wightman Quantum Field Theory”. Commun. Math. Phys. **1**, 95-111 (1965).
- [23] K. Hepp, “On the Connection Between The LSZ and the Wightman Quantum Field Theory”. In: Axiomatic Field Theory. Brandeis (1965). M. Chretien and S. Deser, eds., New York: Gordon and Breach.
- [24] K. Hepp, “On the Analytical Properties of the Scattering Amplitude in Relativistic Quantum Field Theory”. Helv. Phys. Acta **37**, 639-658 (1964).
- [25] K. Hepp, “One-Particle Singularities of the S-Matrix in Quantum Field Theory”. J. Math. Phys. **6**, 1762-1767 (1965).
- [26] K. Hepp, “Spacial Cluster Decomposition of the S-matrix”. Helv. Phys. Acta **37**, 659-662 (1964).
- [27] James Glimm and Arthur Jaffe. *Quantum Physics. A Functional Integral Point of View*. Second Edition. Springer-Verlag. (1987).
- [28] J. Glimm and A. Jaffe. *Functional Integral Methods in Quantum Field Theory*. In: New Developments in Quantum Field Theory and Statistical Mechanics, Cargèse (1976). Edited by Maurice Levy and Pronob Mitter. Plenum Press, New York and London. ISBN 978-1-4615-8920-4. DOI 10.1007/978-1-4615-8918-1.
- [29] Ola Bratteli and Derek W. Robinson. *Operator Algebras and Quantum Statistical Mechanics I*. Springer Verlag. (1979).
- [30] W. B. Arveson. *An Invitation to C^* -Algebras*. Springer Verlag, NewYork, Heidelberg, Berlin (1998).
- [31] Kenneth R. Davidson. *C^* -Algebras by Example*. Fields Institute Monographs. American Mathematical Society. (1996). ISBN 0-8218-0599-1.
- [32] Richard V. Kadison and John R. Ringrose. *Fundamentals of the Theory of Operator Algebras. Vol. I. Elementary Theory*. American Mathematical Society (1997).
- [33] Romeo Brunetti, Klaus Fredenhagen and Rainer Verch, “The Generally Covariant Locality Principle – A New Paradigm for Local Quantum Field Theory”. Commun. Math. Phys. **237**, 31-68 (2003).
- [34] Walter Rudin. *Functional Analysis*. McGraw-Hill International Editions. (1991). ISBN: 0-07-054236-8.
- [35] Geoffrey L. Sewell, “Quantum Fields on Manifolds: PCT and Gravitationally Induced Thermal States”. Annals of Physics **141**, 201-224 (1982).
- [36] A. F. R. de Toledo Piza. *Mecânica Quântica*. Edusp, Editora da Universidade de São Paulo. (2003). ISBN: 85-314-0748-6.
- [37] Elias M. Stein and Rami Shakarchi. *Fourier Analysis. An Introduction*. Princeton Univ. Press. (2003).
- [38] E. H. Kennard, “Zur Quantenmechanik einfacher Bewegungstypen”. Zeitschrift für Physik **44**, (4-5): 326-352 (1927).
- [39] Hermann Weyl. *The Theory of Groups and Quantum Mechanics*. Dover Publications,

- Inc. (1950). Original: *Gruppentheorie und Quantummechanik*, 1928 edition.
- [40] Wolfgang Pauli. *Wave Mechanics*. Pauli Lectures on Physics, Volume 5. Dover Publications, Inc. (2000). ISBN: 0-486-41462-0.
 - [41] Lev Davidovitch Landau et Evgeny Mikhailovich Lifshitz. *Mécanique Quantique*. Edition Mir, Moscou (1966).
 - [42] E. Schrödinger, “Zum Heisenbergschen Unschärfepprinzip”, Proceedings of the Prussian Academy of Sciences, Physics-Mathematical Section. *Berliner Berichte* 296-303 (1930). Findable in [47], Volume 3, pp. 348-356.
For an English translation: A. Angelow and M.-C. Batoni, “About Heisenberg Uncertainty Relation”, arXiv: quant-ph/9903100v3 23 May 2008, and published in *Bulg. J. Phys.* **26**, 193 (1999).
 - [43] Philippe Blanchard and Erwin Brüning. *Mathematical Methods in Physics. Distributions, Hilbert Space Operators and Variational Methods*. Ed. Birkhäuser (2003).
 - [44] Eric D. Chisolm, “Generalizing the Heisenberg Uncertainty Relation”. *American Journal of Physics* 69, 368 (2001).
 - [45] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics. Vol. 3: Scattering Theory*. Academic Press. New York. (1972–1979).
 - [46] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics. Vol. 4: Analysis of Operators*. Academic Press. New York. (1972–1979).
 - [47] Erwin Schrödinger, *Gesammelte Abhandlungen*, Verlag de Österreichischen Akademie der Wissenschaften, Wien (1984).