AN INTRODUCTION TO QUANTUM MEASUREMENTS
WITH A HISTORICAL MOTIVATION

Leonardo Andreta de Castro[*]**, Olímpio Pereira de Sá Neto**[†]**, Carlos Alexandre Brasil**[1][‡]

[*] *Department of Electrical and Computer Engineering, Duke University, Durham*
NC, 27708, USA
[†] *Coordenação de Ciência da Computação, Universidade Estadual do Piauí*
CEP 64202-220, Parnaíba, PI, Brazil
[‡] *Department of Physics and Materials Science (FCM), São Carlos Institute of Physics (IFSC)*
University of São Paulo (USP), PO Box 369, CEP 13560-970, São Carlos, SP, Brazil
[‡] *Departamento Acadêmico de Ciências da Natureza (DACIN), Campus Cornélio Procópio (CP)*
Universidade Tecnológica Federal do Paraná (UTFPR), Avenida Alberto Carazzai
1640, CEP 86300-000, Cornélio Procópio, PR, Brazil

We provide an introduction to the theory of quantum measurements that is centered on the pivotal role played by John von Neumann's model. This introduction is accessible to students and researchers from outside the field of foundations of quantum mechanics and presented within a historical context. We first explain the origins and the meaning of the measurement problem in quantum theory, and why it is not present in classical physics. We perform a chronological review of the quantization of action and explain how this led to successive restrictions on what could be measured in atomic phenomena, until the consolidation of the orthodox interpretation of quantum mechanics. The clear separation between quantum system and classical apparatus that causes these restrictions is subverted in von Neumann's paradigmatic model of quantum measurements, a subject whose concepts we explain, while also providing the mathematical tools necessary to apply it to new problems. We show how this model was important in discussing the interpretations of quantum mechanics and how it is still relevant in modern applications. In particular, we explain in detail how it can be used to describe weak measurements and the surprising results they entail. We also discuss the limitations of von Neumann's model of measurements, and explain how they can be overcome with POVMs and Kraus operators. We provide the mathematical tools necessary to work with these generalized measurements and to derive master equations from them. Finally, we demonstrate how these can be applied in research problems by calculating the Quantum Zeno Effect.

---

[1]E-mail address: carlosbrasil.physics@gmail.com

Contents

1   Introduction

Quantum mechanics caused a profound change to our understanding of nature, giving it an inherently probabilistic character. Previously, classical physics favored the view of Pierre-Simon Laplace (1749-1827) that the future of the universe could be exactly predicted from its current conditions [1]. The only obstacles would be our lack of complete knowledge of every particle that exists, and the difficulty in solving all the necessary differential equations. The conflict between the new quantum physics and the previous paradigm motivated the famous dictum attributed to Albert Einstein (1879-1955) that *"God does not throw dice"* [2].

Today, the non-deterministic character of quantum theory is not a mere curiosity, but a fact that has impact in society. For example, the developing technology of quantum computation uses the superposition states of quantum mechanics as an edge over classical computers [3]. Techniques such as measurement-based quantum computation [4, 5] make it even clearer that being acquainted with the theory of measurement is now important even for applied research.

The significance of this subject contrasts with how its discussion was for a long time discouraged among physicists and relegated to the domain of philosophy, something that is reflected in some textbooks still in use [6–10]. Only more recent textbooks began to feature these foundational questions prominently, such as Ballentine [11], Peres [12], and Griffiths [13].

In this article, we provide an introduction to the quantum theory of measurements that is accessible to students and researchers from outside the field of foundations quantum mechanics, while also giving a historical context. Our main innovation is to center our exposition on the model of measurements proposed by John von Neumann (1903-1957), whose long-lasting influence is not often emphasized enough. Indeed, his work is still important for current topics such as weak measurements, quantum algorithms, and even the many-worlds interpretation of quantum mechanics.

We begin the article with a qualitative overview of the problem of measurement in quantum mechanics. In Sec. 2, we present a broad discussion of measurements in classical physics, which will begin to be subverted by the initial developments of quantum mechanics, presented in Sec. 3. In Sec. 4, we explain the introduction of the wave function, which would lead to a statistical interpretation of quantum mechanics, which is the subject of Sec. 5. The current complete orthodox formulation of the measurement problem is given in Sec. 6 and Sec. 7, which deal with the principle of complementarity and uncertainty relations, respectively.

After this qualitative introduction, we provide all the conceptual and mathematical tools to apply his description of measurements. In Sec. 8 we explain the concepts involved in von Neumann's model and the required mathematical calculations. Sec. 9 deals with an important concept that is implicit in von Neumann's model—entanglement—and the apparent paradoxes it implies. Further discussions of these foundational questions using von Neumann's model led to the many-worlds interpretation and the much-researched decoherence program, as shown in Sec. 10. As an example of a more recent development, in Sec. 11 we talk about the inclusion of the time variable in the measurement process and the surprising phenomena related to weak measurements.

Despite its wide range of applications, von Neumann's model sometimes is not sufficient to describe a quantum measurement. After this qualitative introduction, we present a typology of measurements in Sec. 12 and explain why a complete measurement described by von Neumann's model is called projective. We generalize these into POVMs in Sec. 13, and present Kraus

operators as a means of describing their dynamics. These are used as tools to derive master equations for finite-time measurements in Sec. 14, which are subsequently used in Sec. 15 to describe an example of application, the Quantum Zeno Effect.

2   Measurements before quantum physics

Measurements that are part of our everyday life are assumed to be simple and intuitive processes, where the measurement apparatus—the timer, the ruler, the scale—remains distinct from the measured object. Sometimes, however, the interaction between the apparatus and the object has to be taken into account. For example, when we use a spring scale, we have to make sure that it will not deform excessively and violate Hooke's law.

There are a few much more fundamental cases in classical physics where this interaction is important:

1. If we wish to determine experimentally the electric field of a charge distribution by means of a probe charge [14], the latter must be as small as possible so that its interaction with the distribution will not cause a displacement of its constituent charges and significantly alter the system. To avoid these problems, the definition of the electric field assumes that the charges of the distribution are fixed.

2. When we measure the temperature of a body with a thermometer, we implicitly use the zeroth law of thermodynamics [15]: we put the thermometer in contact with the system whose temperature we wish to determine and wait until they both reach thermal equilibrium. However, the thermometer must be chosen so that the temperature of the system does not suffer great variation before the equilibrium is reached.

Moreover, in the theory of relativity [9, 16], even measurements of space and time have to take into account the relative movement of the system with respect to the measurer, in stark contrast with Newtonian physics.

However, in all the cases mentioned above, the effects of the measurement apparatus over the measured system can be theoretically calculated and minimized with the improvement of the equipments and the techniques. This is not the case in quantum physics, where the measurement process is often considered intrinsically probabilistic. Measurements no longer find a result that can be predicted with absolute certainty, except for very specific cases. We will explain from the beginning how this situation came to be.

3   The Old Quantum Theory

The first step towards overcoming the classical view of measurements happened with the emergence of the *Old Quantum Theory* [17–21] a heuristic approach to atomic phenomena that made explicit use of elements from classical physics together with *ad hoc* postulates. Despite this apparently precarious character, it provided an explanation for the black body spectrum, the specific heat of solids, the photoelectric effect, and the atomic spectrum [19, 22–27]. However, each of these developments also played a role in restricting what could be measured in an atomic system.

3.1   Quantization of energy

The Old Quantum Theory began with the hypothesis of quantization proposed by Max Planck (1858-1947) in 1900 as a solution to the problem of *black body* radiation [28–30]. This was a term coined by Gustav Kirchhoff (1824-1887) to designate a body that absorbs all the radiation incident upon it. Many previous attempts had been made to explain the spectrum of such a body based on classical physics, to no avail [19, 31]. Planck's success was characterized by himself as "*an act of desperation [done because] a theoretical explanation [to the black–body–radiation spectrum] had to be supplied at all cost, whatever the price*" [19].

What Planck was willing to sacrifice to obtain his solution was nothing less than the continuity of the values of the exchanged energy inside the black body. Planck's theory restricted the radiation in the interior of the cavity to discrete amounts (quanta) of energy proportional to its frequency v multiplied by a constant h (now known as Planck's constant):

$E = hv.$

Within the next decade, this hypothesis would lead to the recognition of discontinuities in atomic physical processes in general [31]. A particularly important contributor to this development was Albert Einstein, who would apply the quantization to other phenomena, such as the specific heat of solids [32] and the photoelectric effect [32, 33]. About the latter, what Einstein proposed was that:

> "[T]he energy of a light ray spreading out from a point source is not continuously distributed over an increasing space but consists of a finite number of energy quanta which are localized at points in space, which move without dividing, and which can only be produced and absorbed as complete units." [33]

While for Planck the quantization was restricted to cavities, Einstein extrapolated it to free electromagnetic radiation [31], thus introducing the quantum of light, or *photon*. In this way, quantities that could be found in a continuous range of values in classical physics became restricted to discrete amounts in quantum mechanics.

3.2   The quantized atom

In the beginning of the 20th century, matter was explained by different atomic models based on classical physics [19, 23]. A well-known example is the atom proposed by Joseph John Thomson (1856-1940) [34, 35], which in his own words was "*built up of large numbers of negatively*

electrified corpuscles revolving around the centre of a sphere filled with uniform positive electri-fication" [34]. Ernest Rutherford (1871-1937) was responsible for the next famous development:

> "I supposed that the atom consisted of a positively charged nucleus of small dimensions in which practically all the mass of the atom was concentrated. The nucleus was supposed to be surrounded by a distribution of electrons to make the atom electrically neutral, and extending to distances from the nucleus comparable with the ordinary accepted radius of the atom." [36]

Rutherford's model was preceded by the similar *Saturnian model* by Hantaro Nagaoka (1865-1950) [37–39]. There was also the model by Arthur Erich Haas (1884-1941), who made the first attempt to apply Planck's quantum of action to the constitution of the atom [18], and with this managed to calculate the *Rydberg constant* that characterizes many atomic spectra—albeit with a wrong numerical factor [18, 19, 31].

A more successful attempt at quantizing the atom was the one made in 1913 by Niels Bohr (1885-1962). An assistant to Rutherford since 1912, Bohr grew concerned that there were no means to conciliate the atomic model proposed by his mentor with Newton's mechanics and Maxwell's electrodynamics. He then introduced what Werner Heisenberg (1901-1976) would call "*the most direct expression of discontinuity in all atomic processes*" [40] by restricting the electron orbits to quantized *stationary stable states*. Bohr stated that it made no sense to ask where the electron was to be found during the transition between these orbits, the so-called *quantum jumps* [41].

An important tool used by Bohr was the *correspondence principle*, formulated in distinct forms by him, Planck and Heisenberg [19, 23, 31, 42–44]. This principle states that when we take the limits of some basic parameters of quantum mechanics we must find the classical results again [19, 43–46]. Bohr used this to find the dependence between the electronic frequencies of emission/absorption and translation around the nucleus, obtaining the correct value for Rydberg's constant [47].

An improvement over Bohr's model would be attained through phase-space analysis, something that is not surprising, since Planck's constant h [17, 22, 23, 30, 48, 49] has dimensions of action. The works of Arnold Sommerfeld (1868-1951) [50–53], William Wilson (1875-1965) [54, 55], Jun Ishiwara (1881-1947) [56, 57], and Bohr [58] led to the restriction of the stationary states of atomic systems to those that satisfied [22, 48]:

$$\oint p \, \mathrm{d}q = nh, \tag{3.1}$$

where q is a position coordinate, p its conjugated momentum, and n is an integer. The imposition of Eq. (3.1) to the hydrogen atom allowed the explanation of the fine structure of its spectrum [53]. According to Planck, this was "*an achievement fully comparable with that of the famous discovery of the planet Neptune whose existence and orbit was calculated by Leverrier before the human eye had seen it*" [59]. The theory was being successful, but the idea of measuring the trajectory of the particle between its discrete allowed states had to be abandoned.

3.3   Wave-particle duality

The old controversy about whether light is a particle or a wave was rekindled when the x-rays were discovered in 1895 by Wilhelm Conrad Röntgen (1845-1923). Some facts suggested x-rays

had a particle-like character: they were emitted with the acknowledgedly corpuscular α and β radiations, and energetic considerations led some to believe that they were particles [19]. However, diffraction experiments suggesting also a wave-like character prompted the experimentalist William Henry Bragg (1862-1942) to say:

> "The problem becomes, as it seems to me, not to decide between two theories of x-rays, but to find, as I have said elsewhere, one theory which possesses the capacity of both." [19]

Among the experimentalists who worked with x-rays, Maurice de Broglie (1875-1960) was responsible for one of the first rigorous determinations of the charge of the electron [23]. Maurice had already considered an analogy between electrons and x-rays, implying a possible fundamental unity of radiation and matter [60] which influenced his younger brother Louis de Broglie (1892-1987). Louis had worked with telegraphy and communication during World War I, becoming familiarized with wave mechanics [60,61]. After the war, he came to work together with his brother on the analysis of the properties of the x-rays and read the works of Einstein, Bohr, and Sommerfeld [60] about the quanta of action. This led him to propose wave properties for matter in the Ph.D. thesis he defended in 1924 [62,63], which some consider "*the most influential and successful doctoral dissertation in the history of physics*" [64].

De Broglie postulated that material particles should be associated to a plane and infinitely extended wave, whose energy and momentum were connected with the frequency and wave length by the same relations that Einstein had used. He associated the momentum p of a particle to its wave length λ,

$$p = \frac{h}{\lambda},$$

which would be justified by the electron diffraction experiments and allowed an explanation of Bohr's theory of electronic orbits [19, 60]. In his theory, when the orbit length is equal to an integer number of electron wavelengths, the orbit is stable. To de Broglie, the wave guided the electron's trajectory, *without probabilistic considerations*—wave and particles were two aspects of the system coexisting at the same level, independent of the measurement apparatus [60, 65].

Einstein had contact with de Broglie's ideas via Paul Langevin (1872-1946), who participated in de Broglie's thesis examination before meeting Einstein at the Fourth Solvay Conference in 1924 [19]. He would put this hypothesis to practice right in the next year, stating in an article about an ideal gas:

> "The way we can associate a (scalar) undulatory field to a material particle or to a system of material particles was shown to us in a remarkable work by Mr. L. de Broglie." [66]

Thus Einstein associated a wave character to gas particles, following the reverse path of his 1905 article that attributed corpuscular features to light.

Alongside the works of de Broglie and Einstein, the last attempt to explain electronic transitions and solve the "paradox" of the wave-particle duality would be made by Bohr, together with Hendrik Anthony Kramers (1894-1952) and John Clark Slater (1900-1976), in what would be known as the BKS theory [67]. The authors associated the electromagnetic wave inside the

atom with the probability of electronic transitions, and hypothesized a *virtual radiation field* that would be responsible for communication between far away atoms and inducing spontaneous transitions. This unprecedented concept of a *probability wave* [68] would become a central feature in quantum mechanics. However, the same would not be true for the BKS theory. When applied to the Compton scattering, BKS predicted that the recoil direction of the electron would exhibit a statistical distribution, something that was soon refuted experimentally [19].

To conclude, Old Quantum Mechanics was a heuristic theory strongly based on classical physics that explained a few recently-observed phenomena using hypotheses such as the quantization of action, the wave-particle duality, and the suggestion of the existence of a field of probabilistic character. These facts imposed a few constraints on which measurable quantities had meaning in quantum theory: some observables only had discrete values, and there was no traceable transition between them. It was the prelude to the problem of measurement that would emerge with the complete quantum theory.

4 The rise of the wave function

De Broglie's waves had a lasting influence in modern quantum mechanics after they were transformed into the modern concept of wave function by Erwin Schrödinger (1887-1961). Schrödinger used the analogy proposed by William Rowan Hamilton (1805-1865) between the motion of a point mass under the influence of a potential $V(x, y, z)$ and the propagation of light rays in a medium with refractive index $n(x, y, z)$ [19, 69–71]. Hamilton showed [23, 72] that the optical path and the mechanical *action S* obey the same variational principle. In terms of a Hamiltonian function $H(q_k, p_k, t)$, which depends on the coordinates q_k and their associated momenta p_k, the principle can be expressed as the Hamilton-Jacobi equation:

$$H(q_k, p_k, t) = -\frac{\partial S}{\partial t}, \tag{4.1}$$

where the momenta are given by:

$$p_k = \frac{\partial S}{\partial q_k}. \tag{4.2}$$

Hamilton's formalism was treated solely as a formal analogy, lacking deeper physical meaning [19]. Nevertheless, Schrödinger was convinced that it represented not only a robust mathematical theory, but something real in nature. Reading Einstein's 1925 article led him to find de Broglie's thesis, which he proceeded to study systematically [60]. Schrödinger then used the optical-mechanical analogy [73] as a heuristic argument to derive a wave equation for matter, introducing the *wave function* ψ, related to the action via [74]:

$$S = K \ln \psi, \tag{4.3}$$

where K is a constant to be determined. If one replaces Eq. (4.3) in Eq. (4.1), one obtains the Schrödinger equation, to this day one of the pinnacles of quantum mechanics:

$$H(q_k, p_k, t)\psi = -K\frac{\partial \psi}{\partial t}, \tag{4.4}$$

where the momenta are found replacing Eq. (4.3) in Eq. (4.2):

$$p_k \psi = K\frac{\partial \psi}{\partial q_k}. \tag{4.5}$$

Currently, we choose $K = -i\hbar$, where $\hbar = h/2\pi$, but Schrödinger initially chose $K = \hbar$ because he at first considered ψ to be real, before having to review this idea. In the beginning, he did not ascribe any physical interpretation to ψ, but he did remark that:

> "It is, of course, strongly suggested that we should try to connect the function ψ with some *vibration process* in the atom, which would more nearly approach reality than the electronic orbits, the real existence of which is being very much questioned today." [75]

The problem of interpreting the function ψ persisted despite the success of the Schrödinger equation in calculating the values of energies in experimental situations. To Schrödinger, the wave function was as much a real entity as the electromagnetic waves described by the Maxwell's equations [41, 65]. However, this realistic interpretation faced some obstacles, such as the fact that ψ is a complex entity that depends on the set of observables chosen for its representation (the most evident example, at that time, was the case of position and momentum representations).

In 1926, Schrödinger argued for the necessity of attributing electromagnetic meaning to ψ, because the electron, in transitions, emits electromagnetic waves whose energy is equal to the difference between two eigenvalues. Schrödinger then established that $|\psi|^2$ obeys a continuity equation (known today as *probability current continuity* [76]) and proposed that it multiplied by the electric charge provides the *density of charge*. In this view, charge would be distributed in space, unlike de Broglie's view that it should be located at a given point [60]. This new diffuse character of the charge of a particle would open space for a probabilistic interpretation where the particle itself would not be properly localized.

5 The probabilistic interpretation

The introduction of the wave function gave a diffuse character to the quantum particles, associating with them also wave–like features. This was the first step towards making their measurable quantities no longer as well defined as in classical physics, but rather probabilistic phenomena. This change was due to the contribution of two influential physicists: Max Born (1882-1970) and Paul Dirac (1902-1984).

5.1 Born's rule

Max Born took a position at the University of Göttingen in 1921, where he established a research program of elaborating quantum theories for those simpler atomic problems that had a great amount of empirical data available [77]. Analyzing the problem of an electron beam colliding with an atom, Born provided a probabilistic interpretation of ψ in a work of just a few pages and a mainly qualitative argument, consisting of the asymptotic analysis of the elastic collision of an electron with an atomic core [78, 79].

Born considers a beam with F particles per unit time colliding with an atom. The number of particles dn that are deflected in the infinitesimal solid angle $d\Omega$ is obtained via the *differential scattering cross section*, $\sigma(\theta, \phi)$, which depends on the potential of the atom:

$$dn = F\sigma(\theta, \phi)\, d\Omega. \tag{5.1}$$

Initially, the beam is not affected by the interaction with the atom and its wave function is assumed to be the plane wave e^{ikz}, where z is the direction of propagation. After the scattering, and when the beam is distant from the target, it will be composed of a part proportional to the initial wave function and a part containing a scattering contribution:

$$\psi_{\text{scatter}} = f_k(\theta, \phi)\, \frac{e^{ikr}}{r}.$$

The scattering term has the form of a spherical wave, and is inversely proportional to the distance in order to conserve the flux of $|\psi|^2$, a condition Schrödinger had shown to be necessary. Likewise, the part of this flux that is deviated from the main beam to a certain direction is given by the cross section $\sigma(\theta, \phi)$ (calculations can be seen in detail in Ref. [76]):

$$\sigma(\theta, \phi) = |f_k(\theta, \phi)|^2,$$

so that Eq. (5.1) becomes:

$$dn = F\, |f_k(\theta, \phi)|^2\, d\Omega.$$

Therefore, the number of particles detected in the solid angle by unit of time will be proportional to the squared modulus of the scattered wave function:

$$dn \propto |\psi_{\text{scatter}}(r)|^2. \tag{5.2}$$

For a single particle, Born interpreted the squared modulus of the wave function as the probability of measuring it in that direction after the scattering. He summarized how this interpretation inserted uncertainty into the nature of measurement as follows:

"One gets no answer to the question, 'what is the state after the collision', but only to the question 'how probable is a specified outcome of the collision'." [79]

Born published a considerably longer paper in the same year [80] where he stated the current paradigm of the evolution of the wave function in quantum mechanics:

"[T]he motion of the particles follows laws of probability, but the probability itself propagates in harmony with the causal law." [81]

In this paper he also took a step further and attributed the probabilistic interpretation not to the transition, but to the stationary states. Born started with the time-independent Schrödinger equation, considering that ψ in Eq. (4.4) is an eigenfunction of the time derivative:

$$\hat{H}\psi_n(\mathbf{x}) = E_n\psi_n(\mathbf{x}). \tag{5.3}$$

Here, the system is considered non-degenerate and discrete, and ψ_n are orthonormal eigenfunctions of the Hamiltonian:

$$\int_V d^3x\, \psi_n(\mathbf{x})\,\psi_m^*(\mathbf{x}) = \delta_{n,m}, \tag{5.4}$$

Born assumes the system is submitted to a potential $U(\mathbf{x})$, so that Eq. (5.3) takes the following expanded form (adapting the notation):

$$\frac{\hbar^2}{2m}\nabla^2\psi_n(\mathbf{x}) + [E_n - U(\mathbf{x})]\,\psi_n(\mathbf{x}) = 0, \tag{5.5}$$

where m is the particle mass. Now, multiplying by ψ_m^* and integrating over the volume V, we obtain:

$$\int_V d^3x\left\{\frac{\hbar^2}{2m}\psi_m^*(\mathbf{x})\nabla^2\psi_n(\mathbf{x}) + [E_n - U(\mathbf{x})]\,\psi_m^*(\mathbf{x})\psi_n(\mathbf{x})\right\} = 0. \tag{5.6}$$

In the first term on the left-hand side, we use one of *Green's identities* [82, 83]:

$$\int_V d^3x\,\psi_m^*(\mathbf{x})\nabla^2\psi_n(\mathbf{x}) = \int_S da\,\psi_m^*(\mathbf{x})\hat{\mathbf{n}}\cdot\boldsymbol{\nabla}\psi_n(\mathbf{x}) - \int_V d^3x\,\boldsymbol{\nabla}\psi_m^*(\mathbf{x})\cdot\boldsymbol{\nabla}\psi_n(\mathbf{x}), \tag{5.7}$$

where S is a surface that involves the volume V. We will consider that the wave functions vanish at this surface—which could be as distant as necessary—so that when we replace Eq. (5.7) in Eq. (5.6), we obtain the expression:

$$\int_V d^3x\left\{-\frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi_m^*(\mathbf{x})\cdot\boldsymbol{\nabla}\psi_n(\mathbf{x}) + [E_n - U(\mathbf{x})]\,\psi_m^*(\mathbf{x})\psi_n(\mathbf{x})\right\} = 0. \tag{5.8}$$

Using the normalization from Eq. (5.4), we find:

$$\int_V d^3x\left\{\frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi_m^*(\mathbf{x})\cdot\boldsymbol{\nabla}\psi_n(\mathbf{x}) + U(\mathbf{x})\psi_m^*(\mathbf{x})\psi_n(\mathbf{x})\right\} = E_n\int_V d^3x\,\psi_m^*(\mathbf{x})\psi_n(\mathbf{x})$$
$$= E_n\delta_{n,m}. \tag{5.9}$$

The terms on the left-hand side can be interpreted, respectively, as the volumetric integrals of the kinetic and potential energies, so that "*each energy level [E_n] can therefore be regarded as a space integral of the energy density of the eigenvibrations*" [81].

Following a similar reasoning for any normalized wave function ψ, the energy E of the particle associated to it should be the sum of the integrals of the kinetic and potential energy densities:

$$\int_{\mathbb{R}^3} \mathrm{d}^3x \left\{ \frac{\hbar^2}{2m} |\nabla\psi|^2 + U(\mathbf{x}) |\psi(\mathbf{x})|^2 \right\} = E. \tag{5.10}$$

This wave function can always be decomposed into a superposition of eigenfunctions $\psi_n(\mathbf{x})$:

$$\psi(\mathbf{x}) = \sum_n c_n \psi_n(\mathbf{x}), \tag{5.11}$$

which can be substituted into Eq. (5.10) and compared to Eq. (5.9), yielding:

$$E = \sum_r |c_n|^2 E_n. \tag{5.12}$$

According to Eq. (5.12), the total energy is obtained by weighing the energy of each eigenfunction by $|c_n|^2$. Hence, the $|c_n|^2$ are probabilities associated with finding the system in the eigenstate ψ_n.

The interpretation of the square modulus of the wave function as the probability of finding the particle in a region of space, as in Eq. (5.2), or with the probability of finding it in a certain eigenstate, as in Eq. (5.12), is called *Born's rule*. This is still an important part of quantum theory, although it seems to contradict the classical idea of measurement as the process of acquiring a value that already exists in nature. Despite that, Born still saw the particle as a point mass, possessing at each instant a definite position and a definite momentum, like a classical entity [65]. The wave function ψ would represent our knowledge about the physical system, and not the system itself.

5.2   Dirac's contribution

While theoretical quantum mechanics was being intensely developed in Göttingen by Born, Heisenberg, Kramers, and Pauli, who were in continuous communication with Bohr in Copenhagen [84], England had not contributed much. This started to change in July 1925, after Heisenberg gave a seminar in Cambridge that was attended by Ralph Howard Fowler (1889-1944), who asked for more details and forwarded the received material to a young researcher under his supervision, Paul Dirac, with the question: "*What do you think of this?*". Dirac was on holiday in Bristol, his homeland, and would declare later:

> "[I]t needed about ten days or so before I was really able to master it. And I suddenly became convinced that this would provide the key to understanding the atom."

When his first article on the subject reached Germany, there was such a great outburst that Born would declare afterwards it had been *"one of the great surprises of my scientific life"* [84].

Dirac's long-lasting influence is evident in the fact that he introduced many of the notations and conventions still in use in modern quantum mechanics—for example, bras and kets. His

articles, extremely clear even today, contrast with the difficult texts of the Göttingen group, published in German with rare English translations.

For our discussion of the measurement problem, Dirac made the significant contribution of modifying Eq. (5.11) by adding an extra time-varying term to the potential for $t > 0$, so that the coefficients c_n would vary in time for $t > 0$. He called these new time-dependent coefficients $a_n(t)$:

$$\psi = \sum_r a_n(t)\,\psi_n. \tag{5.13}$$

Replacing Eq. (5.13) in Eq. (4.4), Dirac found the differential equations for each of the $a_n(t)$. In his own words,

> "We shall consider the general solution [Eq. (5.11)] to represent an assembly of the undisturbed atoms in which $|c_n|^2$ is the number of atoms in the nth state, and shall assume that [Eq. (5.13)] represents in the same way an assembly of the disturbed atoms, $|a_n(t)|^2$ being the number in the nth state at any time t. We take $|a_n|^2$ instead of any other function of a_n because, as will be shown later, this makes the total number of atoms remain constant." [85]

Indeed, Dirac found that the number of atoms in the nth state is $N_n = a_n a_n^*$ and demonstrated that $\dot{N} = \sum_n \dot{N}_n = 0$. Therefore, using an approach independent from the one used by Born, Dirac presented a similar interpretation of the coefficients of the expansion of the wave function in its eigenstates, although he did not use the term *probability*.

6 The complementarity principle

As the probabilistic interpretation was adequate for experimental analyses, it became the hegemonic orthodox interpretation. In this interpretation, the act of measurement influences the observed system [23, 41, 86, 87], disturbing it in a way that cannot be completely predicted by quantum theory [23, 86, 88, 89]. Therefore, the wave functions are mere symbolic entities used to make probabilistic predictions about what is observed in the lab, under specific experimental conditions [88]. The complete character of the theory would be based on Bohr's pragmatic criterion: the description of the phenomena offered by quantum theory would encompass everything that is *possible* to be experimentally measured [23, 41, 86, 90], as encapsulated in von Neumann's statement:

> "Although we believe that after having specified [the wave function] we know the state completely, nevertheless, only statistical statements can be made on the values of the physical quantities involved." [91]

Yet, von Neumann considered the wave function just a part of a theoretical scheme and not an entity that represents a real aspect of nature [91].

The orthodox interpretation is sometimes called *Copenhagen interpretation*, although it was neither created by Niels Bohr nor developed in Copenhagen. The term originated [92] in the 1955 text "*The development of the interpretation of the quantum theory*" [93] by Heisenberg, who would in a later book state that this interpretation actually emerged at the 1927 Solvay conference in Brussels [68]. One of Bohr's closest collaborators, Leon Rosenfeld (1904-1974), even said that "*...we in Copenhagen do not like at all [the phrase 'Copenhagen interpretation']*" [90].

Despite this misnomer, Bohr's influence is sufficiently great for his view of measurement to be considered a paradigm in the orthodox interpretation. To Bohr, we cannot treat the measurement process without necessarily taking into account the interaction between the system and the measurement apparatus [94]—this is the concept of *indivisibility* [95]. As the measurement apparatus is a macroscopic system that is classical by nature [19, 42, 65], it must be necessarily described by classical physics [19, 96]. From this need to describe the experimental apparatus in classical language, Bohr derived his concept of complementarity [97].

Complementarity was so vital to Bohr's thinking that he would even suggest the expansion of it to other areas such as biology and anthropology [97]. In quantum mechanics, Bohr derived the concept of complementarity from the necessity to use the mutually-exclusive ideas of "wave" and "particle" [19], which he saw as mere abstractions when applied to quantum systems [23, 86]. A system can manifest its corpuscular or wave-like character *depending on the experimental set-up used to observe it*. This does not mean that it is a wave in certain occasions and a particle in others; the system is an entity of "undefined" nature that can reveal wave-like features when interacting with a given measurement apparatus, and particle-like features for another measurement apparatus—the results would be mutually complementary, not exclusive [19, 23].

As such, the complementarity principle does not apply to the isolated system, but only to the complete set including the measurement apparatus [19, 41, 42, 86, 88, 96, 98]. As opposed to some later interpretations, Bohr saw the measurement as a consequence of the experimental set-up, having nothing to do with the consciousness of the observer. Moreover, Bohr saw the wave packet collapse (the reduction of the wave function to the eigenstate corresponding to the measured eigenvalue) as a mere "*artifact of the formalism*" [92, 99]. Despite this, Bohr

(a) Classical particle

(b) Old Quantum Mechanics

(c) De Broglie's guiding wave

(d) Schroedinger's original interpretation of the wave function

(e) Born's probabilistic interpretation
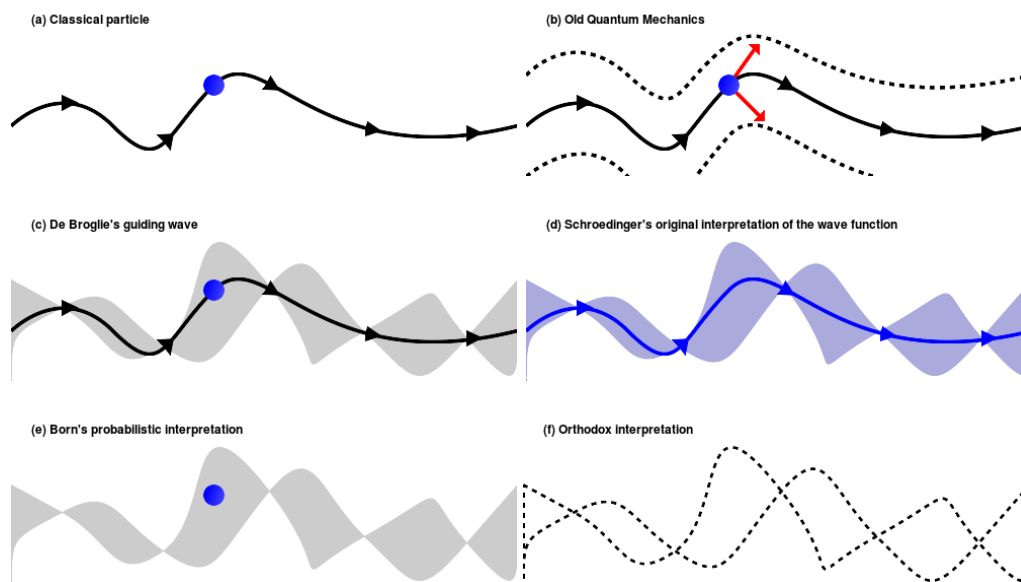
(f) Orthodox interpretation

Fig. 6.1: Summary of the concepts of subatomic particles: (a) Classically, they could move in a continuous trajectory in a potential. (b) In the Old Quantum Theory, only certain stable trajectories were allowed, with quantum jumps occurring between them. (c) De Broglie's waves would guide the particle in its trajectory. (d) Schrödinger's original interpretation of the wave function would have the particle's charge distributed throughout it. (e) To Born, the wave function would encapsulate the probabilities associated with the particle, which would have precise position and momentum at every instant. (f) In the established orthodox interpretation, the wave function is an abstraction representing our maximum knowledge: position and momentum of the particle are only defined if associated with an experimental apparatus.

believed that quantum mechanics could be applied to individual particles, rather than being just an ensemble theory:

> "To my mind, there is no other alternative than to admit that, in this field of experience, we are dealing with individual phenomena and that our possibilities of handling the measuring instruments allow us only to make a choice between the different complementary types of phenomena we want to study." [97]

Many authors [65,90] emphasize that Bohr's writings were not always clear enough. Einstein, for example, said Bohr's clear thinking became obscure when written [65]. However, his position was summarized in his own words during an enlightening debate between him and Heisenberg:

> "That is the whole paradox of quantum theory. On the one hand, we establish laws that differ from those of classical physics; on the other, we apply the concepts of classical physics quite unreservedly whenever we make observations, or take measurements or photographs. And we have to do just that because, when all is said and done, we are forced to use language if we are to communicate our results to other people." [100]

7   The uncertainty principle

Any discussion of the measurement problem would be incomplete without mentioning Heisenberg's uncertainty principle. Heisenberg was a doctoral student under the supervision of Arnold Sommerfeld, who had a singular view of physics and who sought to explain spectroscopic phenomena based solely on experimentally-measured physical quantities [101, 102]. This point of view was inherited by Heisenberg, who built his theory taking into account the way the parameters are determined experimentally—this assumption is called *rule of restriction* [103]. For example, he considered that the concept of electron orbit was inadequate due to the spatial dimensions involved, small enough that any observation would require a wavelength with such an energy that it would remove the electron from its orbit [40]. In his words:

> "[O]ne ought to ignore the problem of electron orbits inside the atom, and treat the frequencies and amplitudes associated with the line intensities as perfectly good substitutes. In any case, these magnitudes could be observed directly, and ... physicists must consider none but observable magnitudes when trying to solve the atomic puzzle." [100]

Based on this, Heisenberg concluded that the quantum theory demanded a revision of kinematic and mechanical concepts [79, 104]. His analysis was based on the canonical commutation relation:

$$[\hat{p}, \hat{q}] = -i\hbar. \tag{7.1}$$

To illustrate his point, Heisenberg used an example that became famous [19, 23, 40, 91]: an electron interacting with radiation of a sufficiently short wavelength λ to be observed under an optical microscope. This wavelength will define the uncertainty of the position observed by the microscope, $\Delta x \sim \lambda$. But the electron will also suffer a Compton recoil [6, 9], that implies an uncertainty in the momentum of magnitude $\Delta p \sim \hbar/\lambda$. Thus, Heisenberg obtained the famous formula:

$$\Delta x \Delta p \sim \hbar, \tag{7.2}$$

which constitutes a formulation of the *uncertainty principle*. What Eq. (7.2) is expressing is that the measurement apparatus causes uncontrollable perturbations on the system being measured, making it impossible to obtain simultaneously the values of two conjugated variables with arbitrary precision [40, 68, 103]. Despite this formulation of the uncertainty principle not being very rigorous—for example, Heisenberg omits the angular aperture of the microscope lens—the general idea was confirmed by new uncertainty relations obtained shortly later by Ditchburn [105], Robertson [106–108], and Heisenberg himself [40].

More evidently than Bohr's complementarity principle, the uncertainty relations represent an essential distinction between quantum theory and classical physics [23]. Until then, we assumed that, despite the experimental apparatus being able to affect the system and the result of the measurement, we could improve the experimental set-up and take precautions to make the distortion of the result arbitrarily small. But, after this point, a minimum perturbation became impossible to eliminate when two conjugate variables are measured. Even if the imperfections

Year	Theoretical Development	Restrictions to Measurements	Refs.
1900	Planck quantizes the energy exchanges in a black body	Energy inside a cavity only could be exchanged in discrete amounts	(28), (29), [30]
1905	Einstein quantizes free electromagnetic radiation	Light only could be detected in discrete amounts	(109), [33]
1913	Bohr's atomic model	Trajectory of the electron between orbits cannot be measured	(47), (110), (111)
1924	de Broglie attributes a wave-like character to matter	—	(62), [63]
1926	Schrödinger introduces wave functions	The charge of a particle may not be localized	(74), (73), [75]
1926	Born's rule	Measurements become probabilistic phenomena	(78), (80), [79]
1927	Heisenberg introduces the uncertainty principle	Measurements of non-commuting observables have limited precision	(104), [79]
1928	Bohr introduces the complementarity principle	Measurements results depend on the experimental set-up	(94)

Tab. 7.1: Summary of the early developments of quantum mechanics and how each step affected how we see measurements. We also list in the last column which of our references correspond to the original papers (in parenthesis) and to their English translations [in square brackets].

were removed by a highly trained and competent observer with a perfectly-tuned equipment, a fundamental limitation would still exist [23].

However, it is important to know what the uncertainty principles still allow us to do. We will remark on two important subtle aspects [40]:

1. The relations do not rule out the arbitrarily precise knowledge of one observable alone. For example, we can know completely, at a specific instant, the position of a particle, because we can have $\Delta x \to 0$ while $\Delta p \to \infty$.

2. If we know the value of the position at an instant t_1, and we measure the value of the momentum at an instant t_2, with $t_2 > t_1$, then we can calculate both values for $t_1 < t < t_2$ with greater precision than the allowed by the uncertainty relations. However, we cannot use that information to predict future measurements, as the second measurement of the momentum affected the position of the particle. In Heisenberg's words, *"[i]t is a matter of personal belief whether such a calculation concerning the past history of the electron can be ascribed any physical reality or not"* [40].

To Bohr, the uncertainty relations were a confirmation of his view: "*in quantum mechanics,*

we are not dealing with an arbitrary renunciation of a more detailed analysis of atomic phenomena, but with a recognition that such an analysis is in principle excluded" [97]. Complementing, Heisenberg stated:

> "The demand to 'describe what happens' in the quantum-theoretical process between two successive observations is a contradiction *in adjecto*, since the word 'describe' refers to the use of the classical concepts, while these concepts cannot be applied in the space between the observations; they can only be applied at the points of observation." [68]

Like Bohr, Heisenberg believed the problem began with language itself, given that we have only the language of classical physics to describe both quantum and non-quantum phenomena. The wave function would represent our language to deal with the experimental conditions, and for this reason would change discontinuously during the measurement process, when our knowledge about the systems changes. But, contrary to Bohr, Heisenberg believed the quantum theory was an ensemble theory [19,65]:

> "The probability function does—unlike the common procedure in Newtonian mechanics— not describe a certain event but, at least during the process of observation, a whole ensemble of possible events." [68]

In short, this new prevailing view of the measurement process divided the world in a quantum part, totally inaccessible to the experimentalist, and a classical part, which we can access via the experimental record. The interaction between the system and the classical apparatus exerts a fundamental role in obtaining the results, preventing measurements more precise than a certain fundamental limit.

8   von Neumann's model of measurements

So far, we have described quantum measurements in a mostly qualitative fashion, as the orthodox interpretation would see them as an artifact of language, their precise description lying outside the scope of scientific inquiry. Despite that, a model by John von Neumann, where quantum measurements are derived from the explicit interaction between the quantum system and the macroscopic apparatus, has had a lasting influence.

With a PhD in mathematics and a degree in chemical engineering, John von Neumann began his work with David Hilbert (1862-1943) at the University of Göttingen in 1926. He used his mentor's discussions about the mathematical foundations of quantum mechanics as a starting point for a series of papers that would become the basis of his 1932 book [19]. There, von Neumann distinguished the two ways in which a wave function can evolve:

- **Process 1:** Reduction, when a measurement is performed.

- **Process 2:** Unitary evolution according to the Schrödinger equation.

An important feature of Process 1 is its violation of causality. Planck had already stated that causality was a "*a heuristic principle*" because "*it is never possible to predict a physical occurrence with unlimited precision*" [112]. In the specific domain of quantum mechanics, Heisenberg had concluded that his uncertainty principle would only allow causal laws to be defined for isolated systems, which are not disturbed by an external classical observer [40]. Therefore, the presence of an observation would be the source of randomness in Process 1.

To avoid this problem, von Neumann's model uses as much as possible of Process 2's determinism to describe the interaction between the quantum system and the measurement apparatus. To see how this works, suppose that the quantum system initially interacts with a single atom of the apparatus. This interaction will be quantum in nature, as well as the interaction between other microscopic parts of the apparatus with this first atom. In principle, then, we can model the beginning of the measurement, before Process 1 takes place, as a unitary evolution. But in order for it to be a true measurement, this evolution will have to satisfy a few conditions, as we will see bellow.

8.1   The dynamics of the measurement

The measurement model introduced by von Neumann requires the following four assumptions, stated clearly by David Bohm (1917–1992) [113]:

1. "*One obtains information by studying the interaction of the system of interest, which we denote hereafter by [S], with the observing apparatus, which we denote by [M]*".

2. "*Before the experiment begins, the observing apparatus [M] and the system [S] under observation are, in general, not coupled*".

3. "*After the interaction has taken place, the state of the apparatus [M] must be correlated to the state of the system [S] in a reproducible and reliable way*".

4. "*It is, in principle, always possible to design an apparatus that measures any given variable without changing that variable during the course of measurement*".

Assumption 1 simply says that we are interested in describing the evolution of a measurement apparatus M that is observing a quantum system S, although we are not explaining why the measurement of an observable $\hat{S}^{(S)}$ yields a specific result.

According to Assumption 2, the quantum system S can initially be described by its own state vector $\left|\phi^{(S)}(0)\right\rangle$, which can be expanded in terms of the eigenstates $|s_n\rangle$ of $\hat{S}^{(S)}$:

$$\left|\phi^{(S)}(0)\right\rangle = \sum_n c_n |s_n\rangle, \quad \sum_n |c_n|^2 = 1, \tag{8.1}$$

where the c_n are constant complex scalars, and where we are assuming discrete eigenstates.

We are also tacitly assuming that the measurement apparatus M can be initially associated to a wave function $|m_0\rangle$ of its own. This would not be permissible under orthodox quantum mechanics, where the measurement apparatus necessarily had to be described by classical physics [65], but Bohm argues that von Neumann's assumption is valid because quantum theory "*should be able to describe the process of observation itself in terms of the wave functions of the observing apparatus and those of the system under observation*", otherwise "*quantum theory could not be regarded as a complete logical system*" [113]. In this case, the joint initial state is *von Neumann's pre-measurement state*:

$$\left|\phi^{(S+M)}(0)\right\rangle = \sum_n c_n |s_n\rangle |m_0\rangle. \tag{8.2}$$

Next, to satisfy Assumption 3, we associate a state vector $|m_n\rangle$ to the state of the measurement apparatus when the eigenvalue s_n has been observed. If we assume that the states $\{|m_n\rangle\}$ are orthonormal, they will not to be confused with each other, and we can propose the following unitary evolution operator via Process 2:

$$\hat{\Delta} = \sum_{n,p} |s_n\rangle \left|m_{n+p}\right\rangle \langle s_n| \left\langle m_p\right|,$$

where the sum in the subscript is calculated *modulo* the total number of states of S. This interaction means that:

$$\hat{\Delta} |s_n\rangle \left|m_p\right\rangle = |s_n\rangle \left|m_{n+p}\right\rangle, \tag{8.3}$$

so that when we apply $\hat{\Delta}$ to Eq. (8.2), we find:

$$\hat{\Delta} \left|\phi^{(S+M)}(0)\right\rangle = \sum_n c_n |s_n\rangle |m_n\rangle. \tag{8.4}$$

We call Eq. (8.4) *von Neumann's measurement states*. They tell us that each state $|m_n\rangle$ of the measurer is correlated to a state $|s_n\rangle$ of the system, and has a probability $|c_n|^2$ of being found—according to the Born–Dirac rule.

This joint evolution of the system and the measurement apparatus already contains all the necessary elements for a measurement, up to the moment when Process 1 takes place. However, the exact dynamics before and after this point depend on the Hamiltonian.

8.2   A specific interaction Hamiltonian

The unitary evolution $\hat{\Delta}$ representing Process 2 is governed by a Hamiltonian $\hat{H}$:

$$\hat{H} = \hat{H}^{(S)} + \hat{H}^{(M)} + \hat{H}^{(S+M)},$$

where $\hat{H}^{(S)}$ contains the terms that act only on the system, $\hat{H}^{(M)}$ has the terms that act only on the measurement apparatus, and the *interaction Hamiltonian* $\hat{H}^{(S+M)}$ encompasses the terms that act on both. von Neumann assumes that this interaction term predominates.
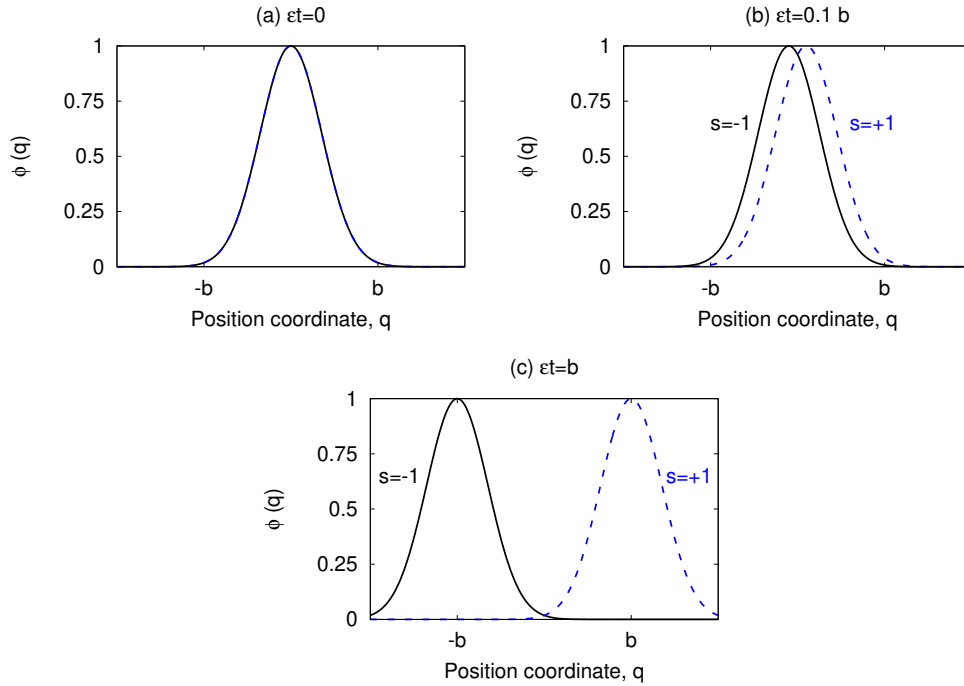
Fig. 8.1: Evolution of the wave function of the measurement apparatus, originally with a compact support of approximate extent $[-b, b]$. (a) At $t = 0$ there is only one wave function. (b) A small period of time after the beginning of the measurement, the wave functions corresponding to different measurements of s become different, but still overlap. (c) After sufficient time, the wave functions cease to overlap and the measured states become orthogonal.

According to Assumption 4, this measurement can be repeated any number of times in sequence and still yield the same result. As we demonstrate in Appendix A, this is true if we impose the condition:

$$\left[\hat{H}^{(S+M)}, \hat{S}^{(S)} \right] = \hat{0}, \tag{8.5}$$

as long as the term $\hat{H}^{(S)}$ is irrelevant. To satisfy Eq. (8.5), we choose:

$$\hat{H}^{(S+M)} = \varepsilon \hat{S}^{(S)} \hat{M}^{(M)}, \tag{8.6}$$

where ε is a real constant, and $\hat{M}^{(M)}$ is a Hermitian operator that acts only on the measurement apparatus. For $\hat{M}^{(M)}$, von Neumann chose a momentum operator $\hat{P}^{(M)}$ conjugated to a certain position operator $\hat{Q}^{(M)}$, which could represent where is the pointer that indicates the measurement result. In this case, the time evolution operator for a measurement that lasts for a period τ will be:

$$\hat{\Delta} = \exp\left\{-i\frac{\tau}{\hbar}\hat{H}\right\} = \exp\left\{-i\frac{\varepsilon}{\hbar}\tau \hat{S}^{(S)} \hat{P}^{(M)}\right\}.$$

Writing the initial state of the apparatus in terms of the eigenbasis of $\hat{Q}^{(M)}$, which we represent by $\{|q\rangle\}$:

$$|m_0\rangle = \int_{-\infty}^{\infty} \mathrm{d}q \, \phi_0(q) |q\rangle,$$

we can evolve the initial state of the system from Eq. (8.2) into:

$$\hat{\Delta} \left|\phi^{(S+M)}(0)\right\rangle = \sum_n c_n \int_{-\infty}^{\infty} \mathrm{d}q \, \exp\left\{-i\frac{\varepsilon}{\hbar}\tau s_n \frac{\hbar}{i}\frac{\partial}{\partial q}\right\} \phi_0(q) |s_n\rangle |q\rangle. \tag{8.7}$$

If we expand the exponential into a power series, we see a Taylor expansion of $\phi_0(m)$:

$$\exp\left\{-i\frac{\varepsilon}{\hbar}\tau s_n \frac{\hbar}{i}\frac{\partial}{\partial q}\right\} \phi_0(q) = \sum_{k=0}^{\infty} \frac{1}{k!} (-\varepsilon\tau s_n)^k \frac{\partial^k}{\partial q^k} \phi_0(q) = \phi_0\left(q - \varepsilon\tau s_n\right). \tag{8.8}$$

Hence, the measurement state given in Eq. (8.4) can be written in the form:

$$\hat{\Delta} \left|\phi^{(S+M)}(0)\right\rangle = \sum_n c_n \int_{-\infty}^{\infty} \mathrm{d}q \, \phi_0\left(q - \varepsilon\tau s_n\right) |s_n\rangle |q\rangle.$$

Therefore, the final state of the measurement apparatus corresponding to the measurement result s_n will be the original wave function displaced by an amount $\varepsilon\tau s_n$:

$$|m_n\rangle = \int_{-\infty}^{\infty} \mathrm{d}q \, \phi_0\left(q - \varepsilon\tau s_n\right) |q\rangle.$$

As long as $\phi_0(q)$ is a function with a compact support smaller than $\varepsilon\tau\Delta s$, where Δs is the minimum distance between the eigenvalues, the states $|m_n\rangle$ will be orthogonal, as demanded by Assumption 3. In this way, the apparatus has distinguishable states for each outcome, a fact illustrated in Fig. 8.1.

8.3   The general interaction Hamiltonian

Despite the lasting influence of von Neumann's model of measurements, his description is not as clear as its reformulation published in Bohm's 1951 textbook [113]. This book was written

as part of Bohm's efforts to understand quantum theory and to explain it to beginners while he delivered a course on the subject at Princeton [114]. In fact, a major portion of it is dedicated to explaining experiments and arguments using mathematics that would be accessible even for a first-year physics student. The book was soon adopted by many universities and even received praise from Einstein, who would consider it the clearest possible exposition of the orthodox interpretation of quantum mechanics [114].

Like von Neumann, Bohm considered that only the interaction term $\hat{H}^{(S+M)}$ is relevant, but went further, justifying this fact with the assumption that τ is short enough for this measurement to be *impulsive*. Bohm also generalized the form of the Hamiltonian from Eq. (8.6). He only assumed that, being a Hermitian operator, $\hat{M}^{(M)}$ must be diagonalizable:

$$\hat{M}^{(M)} = \int_{-\infty}^{\infty} dm \, f(m) \, |m\rangle \langle m| \,.$$

Here, $\{|m\rangle\}$ are orthonormal eigenvalues and $f(m)$ is a real function that yields the continuous set of eigenvalues. If the eigenvalues are discrete, then $f(m)$ has the form:
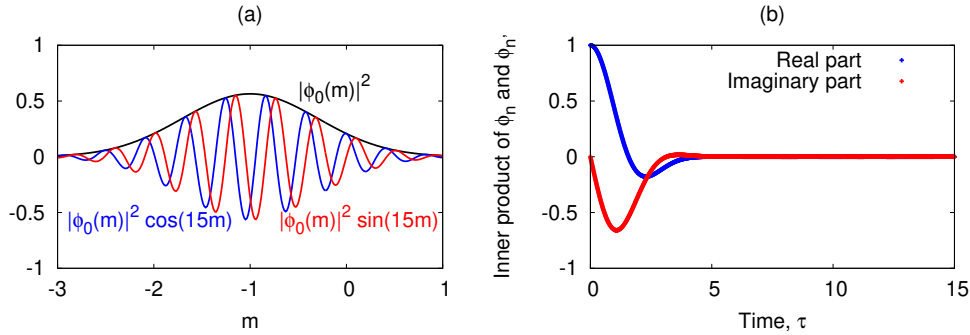
$$f(m) = \sum_n f(m_n)\delta(m - m_n).$$

Fig. 8.2: How a probability distribution $|\phi_0(m)|^2$, chosen as a Gaussian significantly different from zero only in a region $\Delta m = 4$, vanishes when multiplied by rapidly oscillating trigonometric functions. Here, we chose $f(m) = m$ and $\varepsilon(s_n - s_{n'})/\hbar = 1$. Fig. 2(a) shows how the multiplication of $|\phi_0(m)|^2$ by $\cos(15m)$ and $\sin(15m)$ creates an approximately equal number of positive and negative regions. Fig. 2(b) shows the numerical calculation of the real and imaginary parts of the integrals on the right-hand side of Eq. (8.10), which vanish when the trigonometric functions oscillate faster. Hence, the functions become practically orthogonal when $\tau > \Delta m = 4$.

We can re-write the pre-measurement state from Eq. (8.2) in terms of this basis $|m\rangle$:

$$\left|\phi^{(S+M)}(0)\right\rangle = \sum_n c_n \, |s_n\rangle \int_{-\infty}^{\infty} dm \, \phi_0(m) \, |m\rangle \,, \tag{8.9}$$

where $\phi_0(m)$ is the initial wave function of the apparatus. The time evolution due to the Hamiltonian from Eq. (8.6) after the short time interval τ is given by:

$$\hat{\Delta} = \exp\left\{-i\frac{\varepsilon}{\hbar}\tau\hat{S}^{(S)}\hat{M}^{(M)}\right\}.$$

Applied to Eq. (8.9), this $\hat{\Delta}$ results in the measurement state from Eq. (8.4):

$$\hat{\Delta}\left|\phi^{(S+M)}(0)\right\rangle = \sum_n c_n \int_{-\infty}^{\infty} dm\, \phi_n(m)\,|s_n\rangle\,|m\rangle,$$

where $\phi_n(m)$ are the wave functions associated with each outcome of the measurement:

$$\phi_n(m) \equiv e^{-i\varepsilon\tau s_n f(m)/\hbar}\phi_0(m).$$

The inner product between the wave functions $\phi_n(m)$ corresponding to different measurement outcomes is:

$$\int_{-\infty}^{\infty} dm\, \phi_n(m)\,\phi_{n'}^*(m) = \int_{-\infty}^{\infty} dm\, |\phi_0(m)|^2 \cos\left[\frac{\varepsilon}{\hbar}\tau(s_n - s_{n'})f(m)\right]$$
$$- i \int_{-\infty}^{\infty} dm\, |\phi_0(m)|^2 \sin\left[\frac{\varepsilon}{\hbar}\tau(s_n - s_{n'})f(m)\right]. \tag{8.10}$$

Now, suppose the probability distribution $|\phi_0(m)|^2$ has a compact support of length Δm. In this case, the strength $\varepsilon\tau/\hbar$ of the impulsive interaction will make the sines and cosines oscillate so fast whenever $s_n \neq s_{n'}$ that, when multiplied by the comparatively slowly-varying positive function $|\phi_0(m)|^2$, the resulting integrals in Eq. (8.10) will vanish, as can be seen in Fig. 8.2. Hence, Bohm argues that the states of the apparatus corresponding to different eigenstates of the observable "*are very nearly orthogonal*" [113]:

$$|m_n\rangle = \int_{-\infty}^{\infty} dm\, \phi_n(m)\,|m\rangle, \quad \langle m_n|\, m_{n'}\rangle \approx \delta_{n,n'}.$$

In both Bohm's and von Neumann's versions of this measurement, while the complete state of the system can be represented by a single state vector, the state of the individual system S can be any of the set $\{|s_n\rangle\}$. In order to represent this situation, we will need to go beyond wave functions and use the density matrix formalism.

8.4   Density matrices

Density matrices were introduced during the development of quantum statistical mechanics, which required the description of systems whose state might not be completely known. They were independently [42] introduced in 1928 by John von Neumann [91] and Lev Davidovich Landau (1908–1968) [115]. Here, we will show Landau's reasoning to derive this mathematical object—von Neumann's derivation is more elaborate and can be seen in modern notation in Appendix B.

Suppose we have a general state that involves both S and M:

$$\left|\Psi^{(S+M)}\right\rangle = \sum_{n,p} c_{n,p}\,|s_n\rangle\,|m_p\rangle, \quad \sum_{n,p} |c_{n,p}|^2 = 1,$$

but we are only interested in observables of the system S. Landau notes that any such observable $\hat{S}^{(S)}$ will have the following expectation value, according to Born's rule:

$$\left\langle \hat{S}^{(S)} \right\rangle = \left\langle \Psi^{(S+M)} \middle| \hat{S}^{(S)} \middle| \Psi^{(S+M)} \right\rangle = \sum_{n,p} \sum_{n',p'} c^*_{n',p'} c_{n,p} \left\langle s_{n'} \middle| \left\langle m_{p'} \middle| \hat{S}^{(S)} \middle| s_n \right\rangle \middle| m_p \right\rangle.$$

As $\hat{S}^{(S)}$ only acts on the S system and the $\left\{ \middle| m_p \right\rangle \right\}$ are orthonormal, we have:

$$\left\langle \hat{S}^{(S)} \right\rangle = \sum_{n,p} \sum_{n'} c_{n,p} c^*_{n',p} \left\langle s_{n'} \middle| \hat{S}^{(S)} \middle| s_n \right\rangle. \tag{8.11}$$

We then define the reduced density matrix of the system S as:

$$\hat{\rho}^{(S)} \equiv \sum_{n,n'} \rho_{n,n'} \left| s_n \right\rangle \left\langle s_{n'} \right|, \quad \rho_{n,n'} \equiv \sum_p c_{n,p} c^*_{n',p}, \tag{8.12}$$

so that Eq. (8.11) can be re-written as a trace (sum of diagonal elements of a matrix):

$$\left\langle \hat{S}^{(S)} \right\rangle = \sum_{n,n'} \left\langle s_{n'} \middle| \rho_{n,n'} \hat{S}^{(S)} \middle| s_n \right\rangle = \sum_n \left\langle s_n \middle| \hat{\rho}^{(S)} \hat{S}^{(S)} \middle| s_n \right\rangle \equiv \mathrm{Tr} \left\{ \hat{\rho}^{(S)} \hat{S}^{(S)} \right\}. \tag{8.13}$$

This matrix can be used to find the expectation values of all observables of S and thus contains all the information we can know about this system, without including the unnecessary information from M.

States that can be represented by either density matrices or state vectors are called *pure states*. Simply choose $\rho_{n,n'} = c_n c^*_{n'}$ in Eq. (8.12), so that:

$$\hat{\rho}^{(S)} = \left| \Psi \right\rangle \left\langle \Psi \right|, \quad \left| \Psi \right\rangle = \sum_n c_n \left| s_n \right\rangle. \tag{8.14}$$

However, the density matrix can also be a convex combination of pure states, in which case we have a *mixture*:

$$\hat{\rho}^{(S)} = \sum_i p_i \left| \Psi_i \right\rangle \left\langle \Psi_i \right|, \quad \left| \Psi_i \right\rangle \equiv \frac{1}{\sqrt{p_i}} \sum_n c_{n,i} \left| s_n \right\rangle, \quad p_i \equiv \sum_n \left| c_{n,i} \right|^2. \tag{8.15}$$

Eq. (8.15) represents an ensemble where there is a probability p_i of finding each state $\left| \Psi_i \right\rangle$. We can see this from the fact that the expectation value of the mixture is the weighed average of the expectation values of the individual states:

$$\left\langle \hat{S}^{(S)} \right\rangle = \mathrm{Tr} \left\{ \hat{\rho}^{(S)} \hat{S}^{(S)} \right\} = \sum_i p_i \left\langle \Psi_i \middle| \hat{S}^{(S)} \middle| \Psi_i \right\rangle. \tag{8.16}$$

The different states in the mixture do not interfere in a quantum manner with each other, making this an *"incoherent superposition"* [116].

In the case of the measurement state from Eq. (8.4), we have $c_{n,p} = \delta_{n,p} c_n$, which means that the density matrix only has diagonal elements, called *populations*. All the off-diagonal terms,

called *coherences*, vanished after the measurement:

$$
\begin{pmatrix}
\rho_{1,1} & \rho_{1,2} & \cdots & \rho_{1,n} \\
\rho_{2,1} & \rho_{2,2} & \cdots & \rho_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
\rho_{n,1} & \rho_{n,2} & \cdots & \rho_{n,n}
\end{pmatrix}
\longrightarrow
\begin{pmatrix}
\rho_{1,1} & 0 & \cdots & 0 \\
0 & \rho_{2,2} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \rho_{n,n}
\end{pmatrix}.
\tag{8.17}
$$

By the end of the measurement, the initial state becomes an incoherent mixture, a process known as *decoherence*. We will explore this further in Sec. 10.

9 Entanglement and paradoxes

Implicit in the discussion of von Neumann's model of measurements is the idea that the measurement apparatus M and the observed quantum system S end up correlated in a particularly quantum manner, called entanglement. This concept was developed in the 1930s and is still one of the most surprising features of quantum mechanics, responsible for three apparent paradoxes that we will discuss in this section: Wigner's friend, the EPR problem, and Schrödinger's cat.

9.1 Wigner's friend

Previously, we considered two systems S and M interacting. Suppose we now add a new system E, whose basis $\{|e_n\rangle\}$ becomes correlated to the states of M in the same way M became correlated to S:

$$|s_n\rangle |m_0\rangle |e_0\rangle \longrightarrow |s_n\rangle |m_n\rangle |e_0\rangle \longrightarrow |s_n\rangle |m_n\rangle |e_n\rangle. \tag{9.1}$$

Should we consider E to be the measurer of the joint system $S + M$? Or is this process invalid because M was already what we had defined as the measurer? And what if we append another system after E? Is there no end to the number of candidate systems that can claim to be the "real" measurer?

This is the question of *von Neumann's chain* [42]. Heisenberg had considered this problem before, but concluded it was irrelevant to the predictions made by quantum theory:

> "It has been said that we always start with a division of the world into an object, which we are going to study, and the rest of the world, and that this division is to some extent arbitrary. It should indeed not make any difference in the final result if we, e. g., add some part of the measuring device or the whole device to the object and apply the laws of quantum theory to this more complicated object. It can be shown that such an alteration of the theoretical treatment would not alter the predictions concerning a given experiment" [68].

In practical terms, it makes no difference which system is the "true" responsible for the measurement.

While Heisenberg's solution is mathematically correct, this problem was put into more dramatic terms independently by Hugh Everett [117] (1930–1982) and Eugene Wigner [79] (1902–1995). They asked: what if the system E is the scientist performing the observation? Can a human being be in a superposition? As Everett's complete work remained unpublished for more than a decade, this problem became known as *Wigner's friend*.

Wigner presents a situation where a physicist ("Wigner's friend") is performing an experiment that consists of taking note of whether a flash is emitted from a quantum system. In case the flash is emitted, Wigner's friend will record the result corresponding to the eigenstate $|s_1\rangle$; otherwise, his friend records a result corresponding to the eigenstate $|s_2\rangle$. We assume that the system is initially in the superposition state:

$$\frac{1}{2} |s_1\rangle + \frac{\sqrt{3}}{2} |s_2\rangle. \tag{9.2}$$

Suppose that Wigner's friend is in an isolated lab and that, after he has performed the measurement, Wigner approaches him to ask what the result was. The question is whether both the measured quantum system and Wigner's friend will be in a superposition state before Wigner enters the room and asks about the result. If we suppose that $|e_1\rangle$ is the state of Wigner's friend corresponding to recording the result $|s_1\rangle$ and $|e_2\rangle$ is the state corresponding to recording $|s_2\rangle$, then, following Eq. (9.1):

$$\left(\frac{1}{2}|s_1\rangle + \frac{\sqrt{3}}{2}|s_2\rangle\right)|e_0\rangle \longrightarrow \frac{1}{2}|s_1\rangle|e_1\rangle + \frac{\sqrt{3}}{2}|s_2\rangle|e_2\rangle, \tag{9.3}$$

where $|e_0\rangle$ is the state of the friend before any measurement.

Wigner's conclusion was that consciousness privileged the human observer, so that there is no longer a superposition from the moment that his friend reads the result. Wigner does not know what the state is when he enters the room, but it has already been determined [79, 87]—the probabilities of each possible result do not indicate a quantum superposition, but merely Wigner's ignorance.

Of course, formulating the problem in this manner would be unacceptable in orthodox quantum mechanics, where a macroscopic system such as a human being should not be considered in quantum terms. But, should this be possible, Bohm would argue that the superposition disappears much before that:

> "the interaction between the observer and his apparatus is such that statistical fluctuations arising from the quantum nature of the interaction are negligible in comparison with experimental error. It is therefore correct for us to approximate the relation between the investigator and his observing apparatus in terms of the simplified notion that these are two separate and distinct systems interacting only according to the laws of classical physics." [113]

In this way, Bohm proposes a natural manner of breaking von Neumann's chain and separating the observer from the measurement apparatus.

However, this solution was not unanimous and the discussion of whether the measurement apparatus or even the scientist could be described according to quantum mechanics would be important for foundational questions in the future.

9.2 The EPR problem

Albert Einstein (1879–1955), Boris Podolsky (1896–1966), and Nathan Rosen (1909–1995) published in 1935 an article that proposed what would be known as the EPR problem, an attempt to decide whether the current formulation of quantum mechanics could be considered both *correct* and *complete*. The authors defined completeness as the condition that every element of physical reality must have a counterpart in the physical theory. According to them, an *element of physical reality* corresponds to the situation where it is possible to predict the value of some physical quantity with certainty without perturbing the system in any way.

We will analyze the EPR problem from Bohm's simplified re-formulation [113], which provided a clearer analysis in terms of the spin [65, 96, 118]. Assume that we have a single spin-1/2 particle, and let $\{|+\rangle, |-\rangle\}$ be the eigenbasis of the operator $\hat{\sigma}_z$:

$$\hat{\sigma}_z|\pm\rangle = \pm|\pm\rangle \tag{9.4}$$

while $\{|+_x\rangle, |-_x\rangle\}$ is the eigenbasis of $\hat{\sigma}_x$:

$$|\pm\rangle = \frac{|+_x\rangle \pm |-_x\rangle}{\sqrt{2}}. \tag{9.5}$$

If a particle is in one of the eigenstates $|\pm\rangle$, the inherent uncertainty in the measurement of $\hat{\sigma}_z$ will be zero and this will be an element of reality – but the value of $\hat{\sigma}_x$ will be undefined. In the terminology employed by EPR, when the value of the measurement of $\hat{\sigma}_z$ is unknown, then $\hat{\sigma}_x$ has no physical reality.

Now, suppose that two particles are initially in the *singlet* state:

$$\frac{|+-\rangle - |-+\rangle}{\sqrt{2}} = \frac{|+_x-_x\rangle - |-_x+_x\rangle}{\sqrt{2}}. \tag{9.6}$$

This state is invariant under rotations, which means that, no matter what measurement we perform on the first particle, the second particle will be found on the *opposite state* of the same observable. If the first particle is measured in $\hat{\sigma}_z$, the second particle will be in the other eigenstate of $\hat{\sigma}_z$. If the first particle is measured in $\hat{\sigma}_x$, the second particle will be in the other eigenstate of $\hat{\sigma}_x$. Therefore, the state of the second particle is determined by a measurement of the first particle that does not act on it directly. Particles displaying this property are known as an *EPR pair* [119].

Due to this fact, the authors conclude that, despite being *correct*, the quantum description of reality given by wave functions is not *complete* [65]. In their words,

> "This makes the reality of $[\hat{\sigma}_z]$ and $[\hat{\sigma}_x]$ depend upon the process of measurement carried out on the first system, which does not disturb the second system in any way. No reasonable definition of reality could be expected to permit this." [120]

Bohr soon replied [121], emphasizing the complementary character of quantum mechanics. As he had stated multiple times, complementarity resides exactly in the impossibility of considering a system individually, because quantum theory deals with the joint set of the system and the measurement apparatus [65, 122]. For this reason, Bohr concluded that quantum theory is not only correct, but also complete, not in the sense sought by EPR, but in the sense that it allows us to know everything that can be know [23].

Nevertheless, the *"spooky actions at a distance"* [123] between these two particles would remain intriguing scientists, and lead to the formulation of the important concept of *entanglement* by Schrödinger in the following year.

9.3   Schrödinger's cat

Schrödinger formalized the concept of entanglement in two works [124, 125] published soon after the EPR paper [120]. This is a term he coined for the situation where the state of one quantum system depends on the state of another quantum system. In his words:

> "When two systems, of which we know the states by their respective representatives, enter into temporary physical interaction due to known forces between them, and when after a time of mutual influence the systems separate again, then they can no

longer be described in the same way as before, viz. by endowing each of them with a representative of its own. I would not call that *one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought. By the interaction the two representatives (or ψ-functions) have become entangled." [124]

This phenomenon was implied when we discussed von Neumann's measurements and density matrices. Indeed, in von Neumann's model, the measured quantum system S and the measurement apparatus M become *entangled*.

Schrödinger further proposes that "*the [EPR] paradox could be avoided*" if "*the knowledge of the phase relations*" between the different states had "*been entirely lost in consequence of the process of separation*" of the two systems, because "*this would mean that not only the parts, but the whole system, would be in the situation of a mixture, not of a pure state*" [125]. Schrödinger associates the truly quantum features of a system to the interference terms between its superposed states, which disappear after measurement, as we saw in Eq. (8.17).

To illustrate his point in an eloquent manner, he describes the problem that would be known as *Schrödinger's cat*:

"A cat is penned up in a steel chamber, along with the following diabolical device (which must be secured against direct interference by the cat): in a Geiger counter there is a tiny bit of radioactive substance, *so* small, that *perhaps* in the course of one hour one of the atoms decays, but also, with equal probability, perhaps none; if it happens, the counter tube discharges and through a relay releases a hammer which shatters a small flask of hydrocyanic acid. If one has left this entire system to itself for an hour, one would say that the cat still lives *if* meanwhile no atom has decayed. The first atomic decay would have poisoned it. The ψ-function of the entire system would express this by having in it the living and the dead cat (pardon the expression) mixed or smeared out in equal parts." [79, 126]

The controversial point here is to assume that the superposition of alternatives in a microscopic system (the atoms of radioactive substance) would also imply a superposition in the macroscopic system (the cat) [118]. Discussions of this kind would spark interesting new interpretations of quantum mechanics, and eventually experimental proposals to create macroscopic "Schrödinger cat states".

10   Many worlds and decoherence

As mentioned in the previous section, during his doctoral work, Everett became interested in the problem that would be known as Wigner's friend. But, unlike Wigner, who invoked the observer's consciousness as the end point where the system could no longer be in a superposition, Everett assumed that the system, the measurement apparatus, and the observers would all be entangled and embedded in a universal wave function [127–129]. In Everett's words, we are asked to

> "assume the universal validity of the quantum description, by the complete aban-
> donment of Process 1. The general validity of pure wave mechanics, without any
> statistical assertions, is assumed for all physical systems, including observers and
> measuring apparata. Observation processes are to be described completely by the
> state function of the composite system which includes the observer and his object-
> system, and which at all times obeys the wave equation (Process 2)." [117]

The proposal that the observers themselves could be in a superposition state was met with skepticism. Bryce DeWitt (1923–2004) made the following objection after reading the manuscript Everett submitted for publication in 1957:

> "As Everett quite explicitly says: 'With each succeeding observation... the observer
> state ≪branches≫ into a number of different states.' The trajectory of the memory
> configuration of a real physical observer, on the other hand, does not branch. I can
> testify to this from personal introspection, as can you. I simply do not branch." [130]

In response, Everett added a footnote to his paper clarifying that the subjective experience of each observer would correspond to simply seeing a measurement outcome, because each branch would be unable to sense the others:

> "[S]eparate elements of a superposition individually obey the wave equation with
> complete indifference to the presence or absence ('actuality' or not) of any other
> elements. This total lack of effect of one branch on another also implies that no
> observer will ever be aware of any 'splitting' process." [127]

Therefore, in a measurement state such as Eq. (8.4), each possible state of the observer $|m_n\rangle$ will only be aware of a single state of the system, its *relative state* $|s_n\rangle$. It would be as if each of these branches inhabited a different "world" inside our universe. For this reason, this interpretation would be known alternatively as the *relative-state interpretation* and as the *many-worlds interpretation*.

However, Everett's interpretation suffers from the problem of basis ambiguity [131, 132]. We can illustrate it by defining a new orthonormal basis for the measurement apparatus:

$$|m_n\rangle_\pm = \frac{|m_{2n}\rangle \pm |m_{2n+1}\rangle}{\sqrt{2}},$$

in terms of which the measurement state from Eq. (8.4) can be written as:

$$\hat{\Delta}\left|\phi^{(S+M)}(0)\right\rangle = \sum_n \left(\frac{c_{2n}\,|s_{2n}\rangle + c_{2n+1}\,|s_{2n+1}\rangle}{\sqrt{2}}\,|m_n\rangle_+ + \frac{c_{2n}\,|s_{2n}\rangle - c_{2n+1}\,|s_{2n+1}\rangle}{\sqrt{2}}\,|m_n\rangle_- \right).$$

If we can represent the system in this different basis, why is it that, given a fixed experimental set-up, the measured observable is always the one corresponding to the eigenstates $|s_n\rangle$ of the observable $\hat{S}$? Why do we not observe it in the eigenstates $(c_{2n}|s_{2n}\rangle \pm c_{2n+1}|s_{2n+1}\rangle)/\sqrt{2}$ of a different observable? Why does the observer always have the impression of being in a definite state like $|m_{2n}\rangle$ rather than a superposition like $(|m_{2n}\rangle \pm |m_{2n+1}\rangle)/\sqrt{2}$?

This difficulty arises from the fact that the state after the measurement in Eq. (8.4) is still a coherent superposition of the possible measurement results, allowing the "branches" of the universe to interfere with each other. The solution found to this problem lies in performing the same kind of operation that the measurement apparatus exerted on the system in Eq. (8.17), but this time it is an external environment that makes the coherences vanish. Once this environment E becomes entangled with the measurement apparatus, the complete state becomes:

$$\left|\phi^{(S+M+E)}\right\rangle = \sum_n c_n |s_n\rangle |m_n\rangle |e_n\rangle,$$

where we used the same notation as Eq. (9.1). In this case, the joint state of the system S and the measurement apparatus M can no longer be represented by a state vector. Instead, it becomes a diagonal density matrix, where all the coherences have vanished:

$$\hat{\rho}^{(S+M)} = \sum_n |c_n|^2 |s_n\rangle |m_n\rangle \langle m_n| \langle s_n|.$$

In the words of Wojciech Hubert Zurek (born in 1951), who pioneered this field, "*in a certain sense it is the environment of the apparatus which participates in deciding what the apparatus measures*" [133]. Zurek popularized the term *decoherence* [134] for this phenomenon and inserted the fundamental role of the environment in the description, making the decoherence program a question of great importance [134–140] that has already taken up an existence independent of Everett's interpretation of quantum mechanics. Omnès [96, 98, 141–145], for example, uses it together with the consistent histories formalism [13, 146, 147] to arrive at an enhanced version of the orthodox interpretation. And, despite having received also a series of criticisms [148], decoherence plays an increasingly important role in the theory of quantum information, where defeating the undesired effects of the environment remains an important challenge to be overcome in the construction of quantum computers [3, 139].

We will discuss more consequences of the theory of open quantum systems to quantum measurements later in this article. For the moment, we will illustrate our explanation of von Neumann's model by applying it to a specific modern question – weak measurements.

11 Weak measurements

Despite the fact that von Neumann's model describes the measurement as a dynamical process, so far we have not studied its exact evolution in time before the measurement ends. If we return to the interaction Hamiltonian from Eq. (8.6) and remove the condition that the measurement is impulsive, we end up with an incomplete measurement, a situation where the coherences are smaller, but not yet zero. This situation, where $\varepsilon\tau$ is not sufficiently large, is what is called a *weak measurement* [149, 150].

We can write the final state of the system after von Neumann's measurement following the model from Eq. (8.7):

$$\hat{\Delta}\left|\phi^{(S+M)}(0)\right\rangle = \int_{-\infty}^{\infty} dq \, \exp\left\{-\varepsilon\tau\hat{S}^{(S)}\frac{\partial}{\partial q}\right\}\phi_0(q)\left|\phi^{(S)}(0)\right\rangle|q\rangle, \tag{11.1}$$

where we employed the definition of $\left|\phi^{(S)}(0)\right\rangle$ given in Eq. (8.1).

Next, we introduce a *post-selection*. This means that we will only consider the final state of the measurement apparatus that is the relative state corresponding to a certain state $\left|\Phi_{\text{out}}^{(S)}\right\rangle$ of the measured system. Experimentally, this is done by performing a complete, strong measurement after the weak measurement and discarding the run unless the result is $\left|\Phi_{\text{out}}^{(S)}\right\rangle$. In the calculations, the post-selected state of the measurement apparatus, $\left|\Phi_{\text{post}}^{(M)}\right\rangle$, is found by applying $\left\langle\Phi_{\text{out}}^{(S)}\right|$ from the left to Eq. (11.1):

$$\left|\Phi_{\text{post}}^{(M)}\right\rangle \equiv \left\langle\Phi_{\text{out}}^{(S)}\right|\hat{\Delta}\left|\phi^{(S+M)}(0)\right\rangle = \int_{-\infty}^{\infty} dq \, \left\langle\Phi_{\text{out}}^{(S)}\right|\exp\left\{-\varepsilon\tau\hat{S}^{(S)}\frac{\partial}{\partial q}\right\}\left|\phi^{(S)}(0)\right\rangle\phi_0(q)|q\rangle. \tag{11.2}$$

Now we use the fact that the measurement is weak, so that ε is small and we can expand the exponential of the interaction Hamiltonian into a power series:

$$\left\langle\Phi_{\text{out}}^{(S)}\right|\exp\left\{-\varepsilon\tau\hat{S}^{(S)}\frac{\partial}{\partial q}\right\}\left|\phi^{(S)}(0)\right\rangle = \left\langle\Phi_{\text{out}}^{(S)}\right|\phi^{(S)}(0)\rangle\left[\sum_{n=0}^{\infty}\frac{(-\varepsilon\tau)^n}{n!}\,(S^n)_{\text{w}}\frac{\partial^n}{\partial q^n}\right], \tag{11.3}$$

where we defined:

$$(S^n)_{\text{w}} \equiv \frac{\left\langle\Phi_{\text{out}}^{(S)}\right|\left[\hat{S}^{(S)}\right]^n\left|\phi^{(S)}(0)\right\rangle}{\left\langle\Phi_{\text{out}}^{(S)}\right|\phi^{(S)}(0)\rangle}. \tag{11.4}$$

The term with $n = 1$ is known as weak value, and will be represented by S_{w}. In terms of it, the sum in Eq. (11.3) can be written as:

$$\sum_{n=0}^{\infty}\frac{(-\varepsilon\tau)^n}{n!}\,(S^n)_{\text{w}}\frac{\partial^n}{\partial q^n} = \exp\left\{-\varepsilon\tau S_{\text{w}}\frac{\partial}{\partial q}\right\} + \sum_{n=0}^{\infty}(-\varepsilon\tau)^n\frac{[(S^n)_{\text{w}} - S_{\text{w}}^n]}{n!}\frac{\partial^n}{\partial q^n}. \tag{11.5}$$

Using the same reasoning employed in Eq. (8.8), we can conclude that the exponential of the derivative present in Eq. (11.5) displaces the wave function:

$$\exp\left\{-\varepsilon\tau S_{\text{w}}\frac{\partial}{\partial q}\right\}\phi_0(q) = \sum_{n=0}^{\infty}\frac{1}{n!}(-\varepsilon\tau S_{\text{w}})^n\frac{\partial^n}{\partial q^n}\phi_0(q) = \phi_0\left(q - \varepsilon\tau S_{\text{w}}\right). \tag{11.6}$$
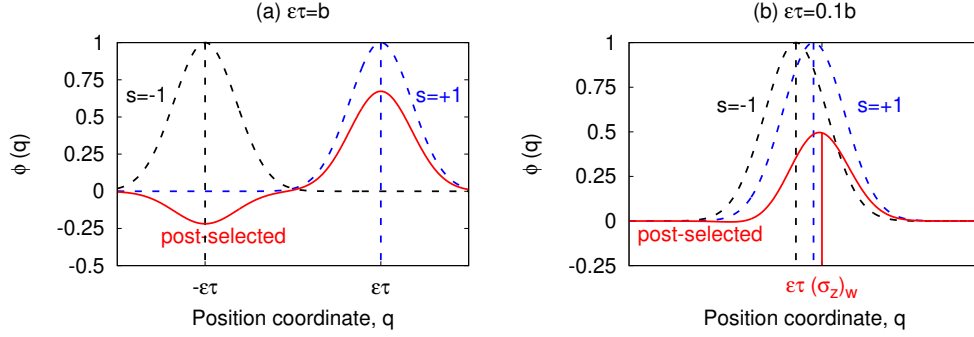
Fig. 11.1: Wave function of an apparatus measuring $\hat{\sigma}_z$ for (a) strong and (b) weak measurements. Dashed lines represent the wave functions corresponding to the two possible results, and the continuous line represents the wave function after the post-selection. The initial state of the apparatus is a Gaussian centered at the origin with compact support $[-b, b]$, and the pre and post-selected states are given by Eq. (11.10) with $\theta = 0.35\pi$. After a weak measurement, the non-normalized post-selected state is approximately a Gaussian centered at $\varepsilon\tau(\sigma_z)_w$, where $(\sigma_z)_w$ is greater than the maximum allowed value of the measurement.

Replacing Eqs. (11.6), (11.5), and (11.3) in Eq. (11.2), we find:

$$\left|\Phi_{\text{post}}^{(M)}\right\rangle = \int_{-\infty}^{\infty} dq \left[\phi_0(q - \varepsilon\tau S_w) + \sum_{n=0}^{\infty} (-\varepsilon\tau)^n \frac{[(S^n)_w - S_w^n]}{n!} \frac{\partial^n}{\partial q^n}\phi_0(q)\right]|q\rangle. \tag{11.7}$$

The terms in the sum on the right-hand side of Eq. (11.7) vanish for $n < 2$. Noticing that, we can define a *weak uncertainty* for $n = 2$, the highest order that does not vanish:

$$(\Delta S)_w \equiv \sqrt{\left|(S^2)_w - S_w^2\right|}. \tag{11.8}$$

As long as ε is sufficiently small so that the highest order term becomes negligible:

$$\left|\frac{1}{2}\varepsilon^2\tau^2 (\Delta S)_w^2 \frac{\partial^2\phi_0}{\partial q^2}\right| \ll 1,$$

we can discard the sum in Eq. (11.7), including the smaller higher-order terms, finding:

$$\left|\Phi_{\text{post}}^{(M)}\right\rangle \approx \int_{-\infty}^{\infty} dq \, \phi_0(q - \varepsilon\tau S_w)|q\rangle.$$

Fig. 11.1 illustrates how this displacement occurs.

It is interesting to notice that, according to Eq. (11.4), the weak value is a complex number. We can make this fact explicit by defining:

$$S_w = S_R + iS_I, \quad \begin{cases} S_R & \equiv \text{Re}\{S_w\} \\ S_I & \equiv \text{Im}\{S_w\} \end{cases}$$

If the wave function $\phi_0(q)$ is real, the imaginary part of the weak value will not affect the probability distribution. However, it will affect the probability distribution of the conjugate momentum $|p\rangle$:

$$\left|\Phi_{\text{post}}^{(M)}\right\rangle \approx \int_{-\infty}^{\infty} dq\, \phi_0(q - \varepsilon\tau S_w) \int_{-\infty}^{\infty} dp\, |p\rangle\langle p|\,q\rangle = \int_{-\infty}^{\infty} dp\, e^{-ip\varepsilon\tau S_R/\hbar} e^{p\varepsilon\tau S_1/\hbar} \tilde{\phi}_0(p)\,|p\rangle,$$

where $\tilde{\phi}_0(p)$ is the Fourier transform of $\phi_0(q)$:

$$\tilde{\phi}_0(p) \equiv \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dq\, e^{-ipq/\hbar} \phi_0(q).$$

Hence, we see that the imaginary part of the weak value contributes as an exponential $e^{p\varepsilon\tau S_1/\hbar}$ multiplied by the wave function of the conjugate variable $\tilde{\phi}_0(p)$.

Weak measurements have been used to derive some unexpected results in quantum mechanics. One of them was presented in an article with the suggestive title "*How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100?*" [149]. The answer involves nothing more than conventional quantum mechanics. Suppose we write the initial (pre-selected) and final (post-selected) states for the quantum system using the notation from Eqs. (9.4) and (9.5):

$$\begin{cases} \left|\phi_0^{(S)}\right\rangle & = a_+\,|+\rangle + a_-\,|-\rangle \\ \left|\Phi_{\text{post}}^{(S)}\right\rangle & = b_+\,|+\rangle + b_-\,|-\rangle \end{cases}, \quad |a_+|^2 + |a_-|^2 = |b_+|^2 + |b_-|^2 = 1. \tag{11.9}$$

In Table 11.1, we see the expectation and weak values for this situation.

Observable	Weak value	Expectation value of $\left|\phi_0^{(S)}\right\rangle$	Expectation value of $\left|\Phi_{\text{post}}^{(S)}\right\rangle$
$\hat{\sigma}_x$	$\dfrac{a_- b_+^* + a_+ b_-^*}{a_+ b_+^* + a_- b_-^*}$	$2\text{Re}\,\{a_+ a_-^*\}$	$2\text{Re}\,\{b_+ b_-^*\}$
$\hat{\sigma}_y$	$i\dfrac{a_+ b_-^* - a_- b_+^*}{a_+ b_+^* + a_- b_-^*}$	$-2\text{Im}\,\{a_+ a_-^*\}$	$-2\text{Im}\,\{b_+ b_-^*\}$
$\hat{\sigma}_z$	$\dfrac{a_+ b_+^* - a_- b_-^*}{a_+ b_+^* + a_- b_-^*}$	$|a_+|^2 - |a_-|^2$	$|b_+|^2 - |b_-|^2$

Tab. 11.1: Comparison between weak values and expectation values.

Suppose we choose the following pre-selected and post-selected states:

$$\left|\phi_0^{(S)}\right\rangle = \frac{(\cos\theta + \sin\theta)\,|+\rangle + (\cos\theta - \sin\theta)\,|-\rangle}{\sqrt{2}}, \quad \left|\Phi_{\text{post}}^{(S)}\right\rangle = \frac{|+\rangle + |-\rangle}{\sqrt{2}}. \tag{11.10}$$

In this case, and according to Table 11.1, the weak values will be the following:

$$(\hat{\sigma}_x)_w = 1, \quad \left(\hat{\sigma}_y\right)_w = i\tan\theta, \quad (\hat{\sigma}_z)_w = \tan\theta. \tag{11.11}$$

This situation allows us to analyze two aspects of the weak values. First, we notice that the weak value of $\hat{\sigma}_y$ is imaginary. Second, if $\theta \to \pi/2$, so that $\left|\phi_0^{(S)}\right\rangle \to (|+\rangle - |-\rangle)/\sqrt{2}$, the weak values will diverge:

$$(\hat{\sigma}_z)_w = -i\left(\hat{\sigma}_y\right)_w \to \infty.$$

In this situation, where the pre and post-selected states tend towards orthogonality, the result of the measurement of the spin component of a spin-1/2 particle can turn out to be much larger than the usual limit [149]. For it to be 100, suffices to choose $\theta \approx 89,427^o$.

We see that weak measurements can be understood in the usual framework of quantum mechanics, as long as we consider von Neumann's measurement model. They are important for foundational arguments [151] and also show parallels with the more practical atomic collision theory [152]. However, despite having the appearance of an intrisically quantum phenomenon, similar results can be obtained also in classical experiments, as long as we generalize the concept of weak value. This was done by Christopher Ferrie and Joshua Combes [153], and relies on realizing that the first weak value from Eq. (11.4) can be written purely in terms of probabilities. Indeed, consider:

$$S_w = \frac{1}{Q} \frac{\left\langle \Phi_{out}^{(S)} \middle| Q \hat{S}^{(S)} \middle| \phi^{(S)}(0) \right\rangle}{\left\langle \Phi_{out}^{(S)} \middle| \phi^{(S)}(0) \right\rangle},$$

where Q is a random variable dependent on the pre and post-selected states. This variable will represent some form of estimate of the state of the system at the intermediary moments between preparation and measurement. This is analogous to the position of the wave function of the measurement apparatus in the quantum prescription. For this reason, we will further assume that Q can only take the discrete values ± 1, just like $\hat{S}^{(S)}$. Using this fact (which implies $Q^{-1} = Q$), and further restricting S_w to real values, so that it is equal to the average of itself and its complex conjugate, we find:

$$S_w = \frac{Q}{2} \frac{\left\langle \Phi_{out}^{(S)} \middle| Q \hat{S}^{(S)} \middle| \phi^{(S)}(0) \right\rangle}{\left\langle \Phi_{out}^{(S)} \middle| \phi^{(S)}(0) \right\rangle} + \frac{Q}{2} \frac{\left\langle \phi^{(S)}(0) \middle| Q \hat{S}^{(S)} \middle| \Phi_{out}^{(S)} \right\rangle}{\left\langle \phi^{(S)}(0) \middle| \Phi_{out}^{(S)} \right\rangle},$$

or, using the fact that Q^2 is always one, we can sum over all Q and divide by 2:

$$S_w = \frac{1}{2} \sum_{Q \in \{-1,1\}} \frac{Q}{2} \frac{\left\langle \Phi_{out}^{(S)} \middle| Q \hat{S}^{(S)} \middle| \phi^{(S)}(0) \right\rangle \left\langle \phi^{(S)}(0) \middle| \Phi_{out}^{(S)} \right\rangle}{\left| \left\langle \Phi_{out}^{(S)} \middle| \phi^{(S)}(0) \right\rangle \right|^2}$$

$$+ \frac{1}{2} \sum_{Q \in \{-1,1\}} \frac{Q}{2} \frac{\left\langle \Phi_{out}^{(S)} \middle| \phi^{(S)}(0) \right\rangle \left\langle \phi^{(S)}(0) \middle| Q \hat{S}^{(S)} \middle| \Phi_{out}^{(S)} \right\rangle}{\left| \left\langle \Phi_{out}^{(S)} \middle| \phi^{(S)}(0) \right\rangle \right|^2}.$$

Assuming again that $\varepsilon\tau \ll 1$, we can approximate $\varepsilon\tau Q \hat{S}^{(S)}/2$ for $e^{\varepsilon\tau Q \hat{S}^{(S)}/2} - 1$, so that:

$$S_w \approx \sum_{Q \in \{-1,1\}} \frac{Q}{\varepsilon\tau} \frac{\left| \left\langle \Phi_{out}^{(S)} \middle| 2^{-1/2} e^{\varepsilon\tau Q \hat{S}^{(S)}/2} \middle| \phi^{(S)}(0) \right\rangle \right|^2}{\left| \left\langle \Phi_{out}^{(S)} \middle| \phi^{(S)}(0) \right\rangle \right|^2},$$

where we have eliminated sums over linear Q. The operators $\hat{K}_Q^{(S)} \equiv 2^{-1/2} e^{\varepsilon\tau Q \hat{S}^{(S)}/2}$, with $Q = \pm 1$, represent two processes of the kind that we will describe in Sec. 13.1 – it is possible to verify that they satisfy the condition from Eq. (13.14) up to first order in $\varepsilon\tau$. These processes do not

significantly change the state of the system, but have the effect of slightly projecting it in the eigenstate of $\hat{S}^{(S)}$ with eigenvalue Q. In this sense, these processes labeled by Q are not very different from a weak measurement.

Using this information, we can see that the term in the denominator is the probability of measuring $\left|\Phi_{\text{out}}^{(S)}\right\rangle$ when the state is in $\left|\phi^{(S)}(0)\right\rangle$, whereas the term in the numerator represents the probability of the same happening after the system undergoes a process labelled by a certain value of Q (as opposed to $-Q$).

This allows us to represent the weak value entirely in terms of probabilites of pre-selecting ϕ, post-selecting Φ, and attributing a certain value Q to an intermediary random variable:

$$S_{\text{w}} \approx \sum_{Q\in\{-1,1\}} \frac{Q}{\varepsilon\tau} \frac{p(Q,\Phi|\phi)}{p(\Phi|\phi)} = \sum_{Q\in\{-1,1\}} \frac{Q}{\varepsilon\tau} p(Q|\Phi,\phi). \tag{11.12}$$

Notice that by this point we have eliminated all references to quantum notation, and these probabilities could be applied to variables associated with a classical process. Ferrie and Combes [153] illustrate this with a simple process of coin toss, where heads and tails are represented by values ±1.

Suppose a regular classical coin is tossed and the results are observed by a reliable observer and by an unreliable one. The reliable observer takes two notes of the result of a single coin toss, first representing the result as ϕ, and later as Φ. Between the two measurements of the result of the coin toss, the unreliable observer makes a measurement of the state of the coin, which is registered as the variable Q. This value, however, can be wrong with a probability $(1 - \varepsilon\tau)/2$. Moreover, the unreliable observer can also flip the coin after the observation, with a probability $1 - \delta/(1 + Q\varepsilon\tau)$. This final flip allows the number in the denominator to be non-zero even when $\phi \neq \Phi$. For example, if $\phi = 1$ and $\Phi = -1$, we have:

$$p(Q = 1, \Phi = -1|\phi = 1) = \frac{1 + \varepsilon\tau}{2}\left(1 - \frac{\delta}{1 + \varepsilon\tau}\right) = \frac{1 + \varepsilon\tau - \delta}{2},$$
$$p(Q = -1, \Phi = -1|\phi = 1) = \frac{1 - \varepsilon\tau}{2}\left(1 - \frac{\delta}{1 - \varepsilon\tau}\right) = \frac{1 - \varepsilon\tau - \delta}{2},$$

so that the probability of finding $\Phi = -1$ and $\phi = 1$, regardless of the value of Q, becomes:

$$p(\Phi = -1|\phi = 1) = p(Q = 1, \Phi = -1|\phi = 1) + p(Q = -1, \Phi = -1|\phi = 1) = (1 - \delta).$$

Replacing these results in Eq. (11.12), we find:

$$S_{\text{w}} \approx \frac{1}{\varepsilon\tau} \frac{p(Q = 1, \Phi = -1|\phi = 1)}{p(\Phi = -1|\phi = 1)} - \frac{1}{\varepsilon\tau} \frac{p(Q = -1, \Phi = -1|\phi = 1)}{p(\Phi = -1|\phi = 1)} = \frac{1}{1 - \delta}.$$

By making $1 - \delta$ small, we can make the weak value much larger than the limits ±1 of the coin toss results.

The fact to note here is that weak values can occur beyond the strict quantum domain, as long as there is some uncertainty in the acquisition of information in the intermediary measurement between pre-selection and post-selection, and this uncertainty can affect the final outcome. We also remark that the generalized definition of weak value given in Eq. (11.12) required a different kind of measurement process than what we had considered so far. We will further discuss these generalized measurements in the next sections.

12 Projective measurements

Wolfgang Pauli (1900–1958) proposed a classification of measurements into those of the *first kind* (or *ideal*) and the *second kind*. The former are those measurements that can be repeated in sequence on the same system without changing the result [65]. All other measurements are of the *second kind* [42, 65].

A measurement that follows von Neumann's model is a measurement of the first kind—as we have showed in the previous sections and discussed in Appendix A, it does not cause the demolition of the system. von Neumann's model also represents a *projective measurement*. To explain what this means, consider a state $\left|\phi^{(S)}\right\rangle$ of the system S that can be decomposed into the orthonormal eigenbasis $|s_n\rangle$ of the observable $\hat{S}^{(S)}$:

$$\left|\phi^{(S)}\right\rangle = \sum_n c_n |s_n\rangle.$$

According to Born's rule, the probability of measuring the eigenvalue s_n is given by $|c_n|^2$. Another way to express this probability is:

$$P_{s_n} = \left|\left\langle\phi^{(S)}\middle| s_n\right\rangle\right|^2 = \left\langle\phi^{(S)}\middle|\hat{\Pi}_n^{(S)}\middle|\phi^{(S)}\right\rangle, \tag{12.1}$$

where we defined the operator $\hat{\Pi}_n^{(S)}$ as:

$$\hat{\Pi}_n^{(S)} \equiv |s_n\rangle\langle s_n|. \tag{12.2}$$

The operator $\hat{\Pi}_n^{(S)}$ is a projector, which means that it is Hermitian and idempotent:

$$\left[\hat{\Pi}_n^{(S)}\right]^2 = \hat{\Pi}_n^{(S)}.$$

Both facts that are easily verifiable from Eq. (12.2).

If we obtained s_n as the result, which will happen with a probability P_n, the state $\left|\phi^{(S)}\right\rangle$ will become $|s_n\rangle$. In terms of the density matrix notation, this can be represented by a convex combination of pure states $|s_n\rangle\langle s_n|$, each of them with a weight P_n:

$$\left|\phi^{(S)}\right\rangle\left\langle\phi^{(S)}\right| \rightarrow \sum_n P_n |s_n\rangle\langle s_n|. \tag{12.3}$$

Now, suppose the initial state is represented by a density matrix $\hat{\rho}^{(S)}$. A density matrix is a positive-definite Hermitian operator with unit trace, and therefore can be decomposed in terms of probabilities p_i and an orthonormal set of state vectors $\left\{\left|\phi_i^{(S)}\right\rangle\right\}$:

$$\hat{\rho}^{(S)} = \sum_i p_i \left|\phi_i^{(S)}\right\rangle\left\langle\phi_i^{(S)}\right|. \tag{12.4}$$

As this represents a mixture where there is a probability p_i of finding a system in the state $\left|\phi_i^{(S)}\right\rangle$, the probability of measuring s_n is:

$$P_n = \sum_i p_i \left|\left\langle\phi_i^{(S)}\middle| s_n\right\rangle\right|^2 = \langle s_n|\left(\sum_i p_i \left|\phi_i^{(S)}\right\rangle\left\langle\phi_i^{(S)}\right|\right)|s_n\rangle = \mathrm{Tr}\left\{\hat{\Pi}_n^{(S)}\hat{\rho}^{(S)}(0)\right\}. \tag{12.5}$$

Using Eqs. (12.3) and (12.5), we can conclude that the final state of the system after Process 1 will be:

$$\sum_n \left(\sum_i p_i \left| \left\langle \phi_i^{(S)} \middle| s_n \right\rangle \right|^2 \right) |s_n\rangle \langle s_n| = \sum_n |s_n\rangle \langle s_n| \left(\sum_i p_i \left| \phi_i^{(S)} \right\rangle \left\langle \phi_i^{(S)} \middle| \right) |s_n\rangle \langle s_n| . \tag{12.6}$$

Recognizing Eq. (12.2) and Eq. (12.4) in Eq. (12.6), we find the final expression for the state after Process 1:

$$\hat{\rho}^{(S)} \to \sum_n \hat{\Pi}_n^{(S)} \hat{\rho}^{(S)} \hat{\Pi}_n^{(S)}. \tag{12.7}$$

Eqs. (12.5) and (12.7) characterize a *projective measurement* [76], or *projection valued measure* [12]. It has this name because the operator multiplied by the density matrix to obtain the probabilities in Eq. (12.5) is a projector. Do notice that any sum of commuting projectors of the kind given in Eq. (12.2) would also be a projector, and therefore characterize a projective measurement.

13 Non-projective measurements

13.1 POVMs

Projection valued measures can be generalized into *positive-operator valued measures* (POVMs). This concept was described mathematically in 1966 by Sterling Berberian (born in 1926) [154] within the context of spectral theory, and applied to a concrete physical case one year later by Josef–Maria Jauch (1914–1974) and Constantin Piron (1932–2012) [155]. It represents measurements that might be neither projective nor of the first kind.

This generalization consists in realizing that it is possible to extract probabilities from Eq. (12.5) even if the operator inside the trace is not a projector. It suffices to have a *positive operator* $\hat{A}_n^{(S)}$ multiplied by the density matrix $\hat{\rho}^{(S)}(0)$ to obtain a result that is positive, like a probability should be:

$$P_n = \text{Tr}\left\{\hat{A}_n^{(S)}\hat{\rho}^{(S)}\right\} \geq 0. \tag{13.1}$$

It is easy to verify the inequality above if we write the density matrix as in Eq. (12.4). The trace in Eq. (13.1) then becomes:

$$\text{Tr}\left\{\hat{\rho}^{(S)}\hat{A}_n^{(S)}\right\} = \sum_i p_i \left\langle \phi_i^{(S)} \middle| \hat{A}_n^{(S)} \middle| \phi_i^{(S)} \right\rangle.$$

From the definition of a positive matrix, $\left\langle \phi_i^{(S)} \middle| \hat{A}_n^{(S)} \middle| \phi_i^{(S)} \right\rangle \geq 0$, and Eq. (13.1) holds.

Moreover, in order for the P_n to be probabilities, their sum must add up to one. For this reason, we have to impose the condition that the sum of the $\hat{A}_n^{(S)}$ is the identity:

$$\sum_n \hat{A}_n^{(S)} = \hat{1}^{(S)}, \tag{13.2}$$

so that

$$\sum_n P_n = \text{Tr}\left\{\sum_n \hat{A}_n^{(S)}\hat{\rho}^{(S)}\right\} = \text{Tr}\left\{\hat{1}^{(S)}\hat{\rho}^{(S)}\right\} = 1.$$

Eq. (13.2) includes the tacit assumption that the eigenvalues of $\hat{A}_n^{(S)}$ are limited [155].

We prove in Appendix C that every Hermitian positive operator $\hat{A}_n^{(S)}$ can be written as:

$$\hat{A}_n^{(S)} = \left[\hat{K}_n^{(S)}\right]^\dagger \hat{K}_n^{(S)}, \tag{13.3}$$

where the operator $\hat{K}_n^{(S)}$ might not necessarily be Hermitian. In terms of it, Eq. (13.1) becomes:

$$P_n = \text{Tr}\left\{\hat{K}_n^{(S)}\hat{\rho}^{(S)}\left[\hat{K}_n^{(S)}\right]^\dagger\right\}. \tag{13.4}$$

The significance of the operators $\hat{K}_n^{(S)}$ will be explained later in this section. For the moment, let us focus on illustrating the meaning of a POVM by providing an example.

13.2   Interpreting the POVMs

The meaning of probabilities like the ones from Eq. (13.4) is explained by Asher Peres (1934–2005) as corresponding to the situation where "*the number of available preparations and that of available outcomes may be different from each other, and also different from the dimensionality of Hilbert space*" [12]. In other words, POVMs allow us to generalize projective measurements into the measurement of states that are not orthogonal, and therefore not mutually exclusive. For example, their projectors might no longer commute.

To illustrate the situation [156], consider a spin-1/2 particle whose state can be expressed in terms of the eigenbasis $|\pm\rangle$ of $\hat{\sigma}_z$ or in terms of the eigenbasis $|\pm_x\rangle$ of $\hat{\sigma}_x$—see Sec. 9.2 for a definition of these bases. The projectors

$$\hat{\Pi}_z^{(S)} = |+\rangle\langle+|, \quad \hat{\Pi}_x^{(S)} = |+_x\rangle\langle+_x|$$

represent apparatuses that detect with 100% probability when a system is in the state $|+\rangle$ or $|+_x\rangle$, respectively:

$$\mathrm{Tr}\left\{\hat{\Pi}_z^{(S)} |+\rangle\langle+|\right\} = 1, \quad \mathrm{Tr}\left\{\hat{\Pi}_x^{(S)} |+_x\rangle\langle+_x|\right\} = 1.$$

Therefore, they represent measurements of the type that von Neumann described, projective and of the first kind, although they measure states that are not orthogonal. For this reason, they cannot be performed at the same time, as their projectors do not commute:

$$\left[\hat{\Pi}_z^{(S)}, \hat{\Pi}_x^{(S)}\right] = \frac{1}{\sqrt{2}} \left\{|+\rangle\langle+_x| - |+_x\rangle\langle+|\right\} \neq \hat{0}.$$

Moreover, it is not possible to form a POVM with $\hat{\Pi}_z^{(S)}$ and $\hat{\Pi}_x^{(S)}$. Every POVM must satisfy Eq. (13.2), and so constructing a POVM with $\hat{\Pi}_z^{(S)}$ and $\hat{\Pi}_x^{(S)}$ would imply finding a positive operator (or sum of positive operators) $\hat{A}^{(S)}$ that satisfies:

$$\hat{A}^{(S)} = \hat{1}^{(S)} - \hat{\Pi}_z^{(S)} - \hat{\Pi}_x^{(S)}. \tag{13.5}$$

However, if we apply $\langle+|$ and $|+\rangle$ to Eq. (13.5), we find:

$$\langle+|\hat{A}^{(S)}|+\rangle = \langle+|\hat{1}^{(S)}|+\rangle - \langle+|\hat{\Pi}_z^{(S)}|+\rangle - \langle+|\hat{\Pi}_x^{(S)}|+\rangle = 1 - 1 - \frac{1}{2} = -\frac{1}{2} < 0.$$

Therefore, the operator $\hat{A}^{(S)}$ cannot be a positive operator.

However, we can still have a POVM that simultaneously detects $|+\rangle$ and $|+_x\rangle$ as long as we abandon ideal detectors and consider instead faulty apparatuses that detect $|+\rangle$ with a probability p_z and $|+_x\rangle$ with a probability p_x:

$$\hat{A}_1^{(S)} = p_z\hat{\Pi}_z^{(S)}, \quad \hat{A}_2^{(S)} = p_x\hat{\Pi}_x^{(S)}.$$

Notice that $\hat{A}_1^{(S)}$ and $\hat{A}_2^{(S)}$ are no longer projectors, just positive operators. Even if the system were initially in the state they were designed to detect, the probability of the detector firing is less than 100%, a fact that shows that this is a measurement of the second kind:

$$\mathrm{Tr}\left\{\hat{A}_1^{(S)} |+\rangle\langle+|\right\} = p_z < 1, \quad \mathrm{Tr}\left\{\hat{A}_2^{(S)} |+_x\rangle\langle+_x|\right\} = p_x < 1.$$

For these to add up to 100%, we will need a third operator $\hat{A}_3^{(S)}$, which corresponds to the case where neither detector fires:

$$\hat{A}_3^{(S)} = \hat{1}^{(S)} - \hat{A}_1^{(S)} - \hat{A}_2^{(S)}. \tag{13.6}$$

The probabilities p_x and p_z have to be chosen appropriately so that $\hat{A}_3^{(S)}$ remains a positive operator. We demonstrate in Appendix D that the choice that maximizes the probability of a correct detection in both devices is $p_x = p_z = 2 - \sqrt{2} \approx 58.6\%$.

Therefore, it is possible to create a POVM that corresponds to two imperfect detectors trying to measure incompatible states simultaneously, as long as we include the possibility of both of them failing. While this illustrates what kind of experimental set-up can be described by a POVM, it does not correlate with a precise evolution of the system after the measurement. In this example, the system under observation can even evade detection, in which case there is no final state to the system.

However, the dynamics of some open systems correspond to probabilities given by POVMs, so that the final state can be expressed in terms of Kraus operators.

13.3   Kraus operators

Karl Kraus (1938–1988) and collaborators [157–160], introduced a formalism to describe the dynamics of an open quantum system [161, 162] where operators just like the ones from Eq. (13.4) are responsible for its evolution. For this reason, the operators $\hat{K}^{(S)}$ are known as Kraus operators.

To derive these operators from an open-systems point of view, let us consider a main quantum system S in which we are interested, and an external system E interacting with it. Before they start to interact, the joint state of both systems is the tensor product of the density matrices of each individual system:

$$\hat{\rho}^{(S+E)}(0) = \hat{\rho}^{(S)}(0)\,\hat{\rho}^{(E)}(0)\,.$$

If these two systems are isolated from any other external interference, their evolution via Process 2 after an interval t will be given in terms of some unitary time evolution operator $\hat{U}^{(S+E)}(t)$ that acts on both S and E. This is proven in Eq. (E.2) from Appendix E:

$$\hat{\rho}^{(S+E)}(t) = \hat{U}^{(S+E)}(t)\,\hat{\rho}^{(S+E)}(0)\left[\hat{U}^{(S+E)}(t)\right]^{\dagger}$$
$$= \hat{U}^{(S+E)}(t)\left[\hat{\rho}^{(S)}(0)\,\hat{\rho}^{(E)}(0)\right]\left[\hat{U}^{(S+E)}(t)\right]^{\dagger}.$$

By taking the partial trace of the total density matrix over the Hilbert space of the system E, we find the time evolution of the reduced density matrix of the system of interest:

$$\hat{\rho}^{(S)}(t) = \mathrm{Tr}_E\left\{\hat{\rho}^{(S+E)}(t)\right\} = \mathrm{Tr}_E\left\{\hat{U}^{(S+E)}(t)\left[\hat{\rho}^{(S)}(0)\,\hat{\rho}^{(E)}(0)\right]\left[\hat{U}^{(S+E)}(t)\right]^{\dagger}\right\}. \tag{13.7}$$

Suppose $\{|s_n\rangle\}$ is an orthonormal basis of the Hilbert space of S while $\{|e_n\rangle\}$ is an orthonormal basis of the Hilbert space of E. We can write the time evolution operator $\hat{U}^{(S+E)}(t)$ as a sum of its elements:

$$\hat{U}^{(S+E)}(t) = \sum_{j,k,r,s} c_{j,k,r,s}(t)\,|s_j\rangle\,|e_k\rangle\,\langle s_r|\,\langle e_s|\,, \tag{13.8}$$

where the coefficients $c_{j,k,r,s}(t)$ are chosen appropriately so that this operator is unitary:

$$\sum_{j,k} c^*_{j,k,m,n}(t) c_{j,k,r,s}(t) = \delta_{m,r}\delta_{n,s}. \tag{13.9}$$

Using Eq. (13.8) in Eq. (13.7) and taking the trace in the basis $\{|e_k\rangle\}$, we find:

$$\hat{\rho}^{(S)}(t) = \sum_k \sum_{j,r,s} \sum_{j',r',s'} c_{j,k,r,s}(t) c^*_{j',k,r',s'}(t) \left(|s_j\rangle \langle s_r| \hat{\rho}^{(S)}(0) |s_{r'}\rangle \langle s_{j'}| \right)$$

$$\times \langle e_s| \hat{\rho}^{(E)}(0) |e_{s'}\rangle. \tag{13.10}$$

The reduced density matrix $\hat{\rho}^{(E)}(0)$ can be diagonalized in its eigenbasis $\left\{ \left| \xi_i^{(E)} \right\rangle \right\}$, just like we did for $\rho^{(S)}(0)$ in Eq. (12.4):

$$\hat{\rho}^{(E)}(0) = \sum_i p_i \left| \xi_i^{(E)} \right\rangle \left\langle \xi_i^{(E)} \right|,$$

where the eigenvalues p_i are positive and add up to one. Then, we can define a new matrix $\sqrt{\hat{\rho}^{(E)}(0)}$, which is diagonal in the same basis, but whose eigenvalues are the square root of the eigenvalues of $\hat{\rho}^{(E)}(0)$:

$$\sqrt{\hat{\rho}^{(E)}(0)} \equiv \sum_i \sqrt{p_i} \left| \xi_i^{(E)} \right\rangle \left\langle \xi_i^{(E)} \right|.$$

Clearly, the original density matrix $\hat{\rho}^{(E)}(0)$ is the product of two of its square roots $\sqrt{\hat{\rho}^{(E)}(0)}$:

$$\hat{\rho}^{(E)}(0) = \sqrt{\hat{\rho}^{(E)}(0)} \sqrt{\hat{\rho}^{(E)}(0)} = \sqrt{\hat{\rho}^{(E)}(0)} \sum_m |e_m\rangle \langle e_m| \sqrt{\hat{\rho}^{(E)}(0)}. \tag{13.11}$$

Replacing Eq. (13.11) in Eq. (13.10), we find:

$$\hat{\rho}^{(S)}(t) = \sum_{k,m} \hat{K}_{k,m}^{(S)}(t) \hat{\rho}^{(S)}(0) \left[\hat{K}_{k,m}^{(S)}(t) \right]^\dagger, \tag{13.12}$$

where we defined the Kraus operators $\hat{K}_{k,m}^{(S)}(t)$ as:

$$\hat{K}_{k,m}^{(S)}(t) \equiv \sum_{j,r,s} c_{j,k,r,s}(t) \langle e_s| \sqrt{\hat{\rho}^{(E)}(0)} |e_m\rangle |s_j\rangle \langle s_r|. \tag{13.13}$$

The probability of the evolution due to each Kraus operator in Eq. (13.12) is given by Eq. (13.4), which we derived for a POVM. Moreover, an operator formed by $\left[\hat{K}_{k,m}^{(S)} \right]^\dagger \hat{K}_{k,m}^{(S)}$ will necessarily be positive. We only have to show that these operators add up to the identity to prove that a set of Kraus operators will always be associated with a positive-operator valued measure:

$$\sum_{k,m} \left[\hat{K}_{k,m}^{(S)}(t) \right]^\dagger \hat{K}_{k,m}^{(S)}(t) = \sum_{k,m} \left[\sum_{j,r,s} c^*_{j,k,r,s}(t) \langle e_m| \sqrt{\hat{\rho}^{(E)}(0)} |e_s\rangle |s_r\rangle \langle s_j| \right]$$

$$\times \left[\sum_{j',r',s'} c_{j',k,r',s'}(t) \langle e_{s'}| \sqrt{\hat{\rho}^{(E)}(0)} |e_m\rangle |s_{j'}\rangle \langle s_{r'}| \right]$$

$$= \sum_k \sum_{j,r,s} \sum_{r',s'} c^*_{j,k,r,s}(t) c_{j,k,r',s'}(t) \langle e_{s'}| \hat{\rho}^{(E)}(0) |e_s\rangle |s_r\rangle \langle s_{r'}|.$$

Using the identity from Eq. (13.9), we conclude that:

$$\sum_{k,m} \left[\hat{K}_{k,m}^{(S)}(t)\right]^{\dagger} \hat{K}_{k,m}^{(S)}(t) = \sum_{r,s} \langle e_s|\hat{\rho}^{(E)}(0)|e_s\rangle |s_r\rangle\langle s_r|.$$

Noticing that $\sum_s \langle e_s|\hat{\rho}^{(E)}(0)|e_s\rangle$ is simply the trace of $\hat{\rho}^{(E)}$, which is 1 for every density matrix, we arrive at the identity we wished to prove:

$$\sum_{k,m} \left[\hat{K}_{k,m}^{(S)}(t)\right]^{\dagger} \hat{K}_{k,m}^{(S)}(t) = \sum_{r} |s_r\rangle\langle s_r| = \hat{1}^{(S)}. \tag{13.14}$$

Hence, the probabilities add up to one, as expected of a POVM.

When there is only one Kraus operator, it becomes equivalent to a unitary time evolution operator $\hat{U}^{(S)}(t)$ and the system evolves according to Eq. (E.2), a sign that this is an isolated quantum system. The presence of more than one Kraus operator indicates that there is more into play – such as the interaction with an external system or measurement apparatus – a sign that we are dealing with an open quantum system.

14 Measurement master equations

In the previous section, we analyzed how POVMs generalized von Neumann's projective measurements and how their dynamics could be described by Kraus operators. In this section, we will see how these dynamics can be calculated at every instant using differential equations. We will employ a statistical tool widely used in the treatment of irreversible processes that involve interacting systems, the *master equation* [162–164]. Classically, an equation of this kind provides the time evolution of the probabilities associated with a given property [164], but in the quantum case it will offer a time evolution of the reduced density operator [162, 163].

14.1 The Lindblad equation

Once again, we will consider the situation where our main system S interacts with an external environment E. Their evolution according to Process 2 will be governed by a total Hamiltonian:

$$\hat{H} = \hat{H}^{(S)} + \hat{H}^{(E)} + \hat{H}^{(S+E)}, \tag{14.1}$$

where the term $\hat{H}^{(S)}$ acts only on the system S, $\hat{H}^{(E)}$ acts only on E, and $\hat{H}^{(S+E)}$ contains the interaction terms that act on both.

As derived in Appendix E, the time evolution of the density matrix of the joint system is given by Eq. (E.1), the Liouville-von Neumann equation:

$$\frac{d}{dt}\hat{\rho}^{(S+E)}(t) = -\frac{i}{\hbar}\left[\hat{H}^{(S)} + \hat{H}^{(E)} + \hat{H}^{(S+E)}, \hat{\rho}^{(S+E)}(t)\right]. \tag{14.2}$$

However, the evolution of the reduced density matrix in which we are interested, $\hat{\rho}^{(S)}(t)$, will be given by the partial trace of Eq. (14.2), which is not always trivial.

To simplify Eq. (14.2), we introduce the interaction-picture Hamiltonian [76]:

$$\hat{H}_I^{(S+E)}(t) \equiv e^{i\left(\hat{H}^{(S)}+\hat{H}^{(E)}\right)t/\hbar}\hat{H}^{(S+E)}e^{-i\left(\hat{H}^{(S)}+\hat{H}^{(E)}\right)t/\hbar}, \tag{14.3}$$

and the interaction-picture density matrix:

$$\hat{\rho}_I^{(S+E)}(t) \equiv e^{i\left(\hat{H}^{(S)}+\hat{H}^{(E)}\right)t/\hbar}\hat{\rho}^{(S+E)}(t)\,e^{-i\left(\hat{H}^{(S)}+\hat{H}^{(E)}\right)t/\hbar}. \tag{14.4}$$

Using Eqs. (14.3) and (14.4), we can re-write Eq. (14.2) in a simpler form:

$$\frac{d}{dt}\hat{\rho}_I^{(S+E)}(t) = -\frac{i}{\hbar}\left[\hat{H}_I^{(S+E)}(t), \hat{\rho}_I^{(S+E)}(t)\right]. \tag{14.5}$$

To find the master equation for the system S alone, we take an iterative approach. We integrate Eq. (14.5) in time, and replace the result in Eq. (14.5) itself. After this, we take the partial trace over the degrees of freedom of the environment, obtaining an equation for the reduced density operator $\hat{\rho}_I^{(S)}(t)$:

$$\frac{d}{dt}\hat{\rho}_I^{(S)}(t) = -\frac{1}{\hbar^2}\int_0^t dt'\,\mathrm{Tr}_E\left\{\left[\hat{H}_I^{(S+E)}(t), \left[\hat{H}_I^{(S+E)}(t'), \hat{\rho}_I^{(S+E)}(t')\right]\right]\right\}.$$

If $t - t'$ is sufficiently small, the difference between $\hat{\rho}_I^{(S+E)}(t)$ and $\hat{\rho}_I^{(S+E)}(t')$ can be ignored. Moreover, we can assume that the system and the environment are almost not entangled, and that the environment has almost not changed since the beginning of the interaction—an assumption known as the Born approximation [76]:

$$\hat{\rho}_I^{(S+E)}(t') \approx \hat{\rho}_I^{(S+E)}(t) \approx \hat{\rho}_I^{(S)}(t)\hat{\rho}_I^{(E)}(0).$$

If we extend the integral to infinity, which is known as the Markov approximation [162], we find the Born–Markov master equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}_I^{(S)}(t) = -\frac{1}{\hbar^2}\int_0^\infty \mathrm{d}t'\,\mathrm{Tr}_E\left\{\left[\hat{H}_I^{(S+E)}(t),\left[\hat{H}_I^{(S+E)}(t'),\hat{\rho}_I^{(S)}(t)\hat{\rho}_I^{(E)}(0)\right]\right]\right\}. \tag{14.6}$$

It is possible [165] to develop Eq. (14.6) into the most general kind of Markovian master equation for a density matrix, the Lindblad equation [166–171]:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}^{(S)}(t) = -\frac{i}{\hbar}\left[\hat{H}^{(S)},\hat{\rho}^{(S)}(t)\right] + \gamma\sum_j\left[\hat{L}_j^{(S)}\hat{\rho}^{(S)}(t)\hat{L}_j^{(S)\dagger} - \frac{1}{2}\left\{\hat{L}_j^{(S)\dagger}\hat{L}_j^{(S)},\hat{\rho}^{(S)}(t)\right\}\right], \tag{14.7}$$

where γ is a constant scalar and the Lindblad operators $\hat{L}_j^{(S)}$ give the non-unitary dynamics for the system S. These operators form the Lindbladian, the last term on the right-hand side of Eq. (14.7). Its presence is the only difference between the Lindblad equation from Eq. (14.7) and the Liouville–von Neumann equation from Eq. (E.1): the Lindbladian is responsible for the effects that the external environment causes on the system.

It is of particular interest to us the situation where the Lindblad operators represent the action of an external measurement apparatus on the system [165, 172, 173]. With this, we can obtain the dynamics of a generalized POVM measurement in the same way that the von Neumann model was capable of providing the dynamics of a projective measurement.

14.2   The Markovian master equation for measurements

We will follow the derivation of the Markovian master equation for continuous measurements that is due to James Cresser and collaborators [174]. Suppose that our quantum system of interest S usually evolves according to Process 2, as prescribed by Eq. (E.2). However, the system has a probability λ per unit of time of undergoing a POVM, in which case it evolves according to the Kraus operators from Eq. (13.12). After a small interval of time Δt, the system will have a probability $1 - \lambda\Delta t$ of having evolved according to Process 2, and a probability $\lambda\Delta t$ of having undergone a POVM:

$$\hat{\rho}^{(S)}(t + \Delta t) = (1 - \lambda\Delta t)\left[e^{-i\hat{H}^{(S)}\Delta t/\hbar}\hat{\rho}^{(S)}(t)e^{i\hat{H}^{(S)}\Delta t/\hbar}\right] + \lambda\Delta t\sum_i\hat{K}_i^{(S)}\hat{\rho}^{(S)}(t)\left[\hat{K}_i^{(S)}\right]^\dagger.$$

Keeping only terms up to the first order in Δt and expanding the exponentials into their Taylor series, we find:

$$\hat{\rho}^{(S)}(t + \Delta t) - \hat{\rho}^{(S)}(t) = -\frac{i}{\hbar}\Delta t\left[\hat{H}^{(S)},\hat{\rho}^{(S)}(t)\right] - \lambda\Delta t\hat{\rho}^{(S)}(t) + \lambda\Delta t\sum_i\hat{K}_i^{(S)}\hat{\rho}^{(S)}(t)\left[\hat{K}_i^{(S)}\right]^\dagger$$
$$+ O\left(\Delta t^2\right).$$

If we divide both sides by Δt and take the limit of $\Delta t \to 0$, we find a Markovian master equation that governs this kind of system:

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}^{(S)}(t) = -\frac{i}{\hbar}\left[\hat{H}^{(S)},\hat{\rho}^{(S)}(t)\right] + \lambda\left\{\sum_i \hat{K}_i^{(S)}\hat{\rho}^{(S)}(t)\left[\hat{K}_i^{(S)}\right]^\dagger - \hat{\rho}^{(S)}(t)\right\}. \tag{14.8}$$

If we choose the Lindblad operators in Eq. (14.7) as being equal to the Kraus operators of some POVM ($\hat{L}_i^{(S)} = \hat{K}_i^{(S)}$) and apply the identity from Eq. (13.14), we obtain Eq. (14.8). Therefore, Eq. (14.8) is a specific case of the Lindblad equation that describes a Markovian continuous measurement [175].

However, it is important to notice that the dynamics of these finite-time measurements do not include the "*reduction and backaction effects*" [176] on the system. After all, this is something that is not present in all interpretations of quantum mechanics. For example, Everett's many-worlds notably lack it.

If one wants to introduce this additional backaction term, it is necessary to employ techniques from Itō calculus [175, 177]. But many interesting phenomena do not require this addition, including some with practical applications. We will illustrate this with the protection of a quantum system in a way that is analogous to the Quantum Zeno Effect [178, 179].

15 Application: Quantum Zeno effect

The Quantum Zeno Effect is named after the Ancient Greek philosopher Zeno of Elea, who proposed a series of paradoxes on the impossibility of movement by dividing the path into an infinite series of small fractions. The analogous phenomenon in quantum mechanics occurs when a quantum system under continuous observation does not evolve in time [96, 180].

This phenomenon was theoretically proposed in 1977 by Baidyanath Misra (born in 1937) and Ennackal Chandy George Sudarshan (1931–2018) [180]. Their original demonstration using projective measurements is given below, followed by another demonstration that employs the master equation for continuous measurements.

15.1 Quantum Zeno Effect via projective measurements

Let us consider that the system S is initially in some pure state represented by a density matrix $\hat{\rho}^{(S)}(0)$, which is also a projector $\hat{\Pi}^{(S)}$:

$$\hat{\rho}^{(S)}(0) = |\psi\rangle\langle\psi| \equiv \hat{\Pi}^{(S)}. \tag{15.1}$$

After a given period of time, the probability $P(t)$ of measuring the system and finding it still in the initial state is given by Eq. (12.5):

$$P(t) = \text{Tr}\left\{\hat{\Pi}^{(S)}\hat{\rho}^{(S)}(t)\right\}. \tag{15.2}$$

If the system is found to be in the state $\hat{\Pi}^{(S)}$, its subsequent evolution will once again follow Process 2, and the probability of measuring it still in the state $\hat{\Pi}^{(S)}$ after another period t will be given by Eq. (15.2). The probability $P_n(nt)$ of the system being found in the state $\hat{\Pi}^{(S)}$ in n measurements equally spaced by periods of time t will be simply the nth power of Eq. (15.2):

$$P_n(nt) = [P(t)]^n. \tag{15.3}$$

Fixing a time $T = nt$, we can take the limit where $n \to \infty$ to find the probability of the system still being in the initial state after a period T of constant projective measurements:

$$P_\infty(T) = \lim_{n\to\infty} P_n(T) = \lim_{n\to\infty} [P(T/n)]^n. \tag{15.4}$$

Using the solution given in Eq. (E.2) to describe the state of $\hat{\rho}^{(S)}$ after a period t of evolution due to Process 2, we can re-write Eq. (15.3) in terms of the time evolution operator $\hat{U}^S(t)$:

$$P_\infty(T) = \lim_{n\to\infty}\left[\text{Tr}\left\{\hat{\Pi}^{(S)}\hat{U}^{(S)}(T/n)\hat{\rho}^{(S)}(0)\hat{U}^{(S)}(-T/n)\right\}\right]^n, \tag{15.5}$$

where we used the fact that:

$$\left[\hat{U}^{(S)}(T/n)\right]^\dagger = \hat{U}^{(S)}(-T/n).$$

Using Eq. (15.1), it is possible to re-write the trace in Eq. (15.5) as an inner product:

$$P_\infty(T) = \lim_{n\to\infty}\left[\langle\psi|\hat{U}^{(S)}(T/n)\hat{\rho}^{(S)}(0)\hat{U}^{(S)}(-T/n)|\psi\rangle\right]^n.$$

Eq. (15.1) also allows us to replace the density matrix with a projector:

$$P_\infty(T) = \lim_{n\to\infty} \left[\langle\psi| \hat{U}^{(S)}(T/n) |\psi\rangle \langle\psi| \hat{U}^{(S)}(-T/n) |\psi\rangle \right]^n . \tag{15.6}$$

We can write Eq. (15.6) in the compact form:

$$P_\infty(T) = F(-T) F(T), \tag{15.7}$$

so that we only have to study the form of the function $F(T)$ to determine the probabilities:

$$F(T) \equiv \lim_{n\to\infty} \left[\langle\psi| \hat{U}^{(S)}(T/n) |\psi\rangle \right]^n . \tag{15.8}$$

Now, notice that, given positive real numbers n_1, n_2, t, the following identity holds:

$$\left[\langle\psi| \hat{U}^{(S)}(t) |\psi\rangle \right]^{n_1} \left[\langle\psi| \hat{U}^{(S)}(t) |\psi\rangle \right]^{n_2} = \left[\langle\psi| \hat{U}^{(S)}(t) |\psi\rangle \right]^{n_1+n_2} . \tag{15.9}$$

Defining $T_1 \equiv n_1 t > 0$ and $T_2 \equiv n_2 t > 0$, so that

$$t = \frac{T_1}{n_1} = \frac{T_2}{n_2} = \frac{T_1 + T_2}{n_1 + n_2},$$

Eq. (15.9) becomes:

$$\left[\langle\psi| \hat{U}^{(S)}(T_1/n_1) |\psi\rangle \right]^{n_1} \left[\langle\psi| \hat{U}^{(S)}(T_2/n_2) |\psi\rangle \right]^{n_2} = \left[\langle\psi| \hat{U}^{(S)} [(T_1 + T_2)/(n_1 + n_2)] |\psi\rangle \right]^{n_1+n_2} .$$

Taking the limits $n_1 \to \infty$ and $n_2 \to \infty$ and using Eq. (15.8), we find:

$$F(T_1) F(T_2) = F(T_1 + T_2), \tag{15.10}$$

for any $T_1, T_2 > 0$. Similarly, notice that, given some positive number $a > 0$:

$$F(aT) = \lim_{n\to\infty} \left[\langle\psi| \hat{U}^{(S)}(aT/n) |\psi\rangle \right]^n . \tag{15.11}$$

Replacing n with $m = n/a$, we find:

$$F(aT) = \lim_{m\to\infty} \left[\langle\psi| \hat{U}^{(S)}(T/m) |\psi\rangle \right]^{am} = [F(T)]^a . \tag{15.12}$$

Combining Eqs. (15.10) and (15.12), we find that, for every $T_1, T_2, a, b \geq 0$:

$$[F(T_1)]^a [F(T_2)]^b = F(aT_1 + bT_2). \tag{15.13}$$

Taking the logarithm on both sides of Eq. (15.13), we have:

$$\log[F(aT_1 + bT_2)] = a\log[F(T_1)] + b\log[F(T_2)],$$

which shows that the logarithm of $F(T)$ is a linear function that can be expressed in terms of two scalar constants A and B:

$$\log F(T) = A + BT.$$

According to Eq. (15.8), $F(0) = 1$, which means we must have $A = 0$. Hence, $F(T)$ becomes:

$$F(T) = e^{\text{Re}\{B\}T} e^{i\text{Im}\{B\}T}, \quad T \geq 0. \tag{15.14}$$

It is possible to repeat this argument exchanging t by $-t$ in Eq. (15.9) and T by $-T$ in Eq. (15.11), so that Eq. (15.13) is derived for $T_1, T_2 \leq 0$. In this way, we arrive at an expression that is analogous to Eq. (15.14), but in terms of a different scalar C:

$$F(T) = e^{\text{Re}\{C\}T} e^{i\text{Im}\{C\}T}, \quad T \leq 0. \tag{15.15}$$

However, from the definition given in Eq. (15.8), we can see that

$$F(-T) = \lim_{n \to \infty} \left[\langle \psi | \hat{U}^{(S)} (-T/n) | \psi \rangle \right]^n$$
$$= \lim_{n \to \infty} \left\{ \langle \psi | \left[\hat{U}^{(S)} (T/n) \right]^\dagger | \psi \rangle \right\}^n$$
$$= F^*(T),$$

which is only possible for every $T > 0$ if

$$\text{Re}\{B\} = -\text{Re}\{C\},$$
$$\text{Im}\{B\} = \text{Im}\{C\}.$$

Misra and Sudarshan employed arguments from group theory to prove that B is a purely imaginary number. [180] Here we will do something simpler [181] and write the time evolution operator explicitly in terms of the Hamiltonian $\hat{H}^{(S)}$:

$$\hat{U}^{(S)}(T/n) = \exp\left\{ -\frac{i}{\hbar} \frac{T}{n} \hat{H}^{(S)} \right\} = \sum_{m=0}^{\infty} \frac{1}{m!} \left(-\frac{i}{\hbar} \frac{T}{n} \right)^m \left[\hat{H}^{(S)} \right]^m. \tag{15.16}$$

Replacing Eq. (15.16) in Eq. (15.8), we find:

$$F(T) = \lim_{n \to \infty} \left[\sum_{m=0}^{\infty} \frac{1}{m!} \left(-\frac{i}{\hbar} \frac{T}{n} \right)^m \langle \psi | \left[\hat{H}^{(S)} \right]^m | \psi \rangle \right]^n.$$

Terms with $m > 2$ will vanish in the limit when $n \to \infty$, because they are multiplied by n^{-2} or lower, which vanish even when multiplied by n in the binomial. The only terms left are $m = 0$ and $m = 1$:

$$F(T) = \lim_{n \to \infty} \left[1 + \left(-\frac{i}{\hbar} \frac{T}{n} \right) \langle \psi | \hat{H}^{(S)} | \psi \rangle \right]^n = \exp\left\{ -\frac{i}{\hbar} T \langle \psi | \hat{H}^{(S)} | \psi \rangle \right\}. \tag{15.17}$$

Therefore,

$$B = C = -\frac{i}{\hbar} \langle \psi | \hat{H}^{(S)} | \psi \rangle, \tag{15.18}$$

which is a purely imaginary number, because $\langle \psi | \hat{H}^{(S)} | \psi \rangle$ is the expectation value of a Hermitian operator, and therefore real.

Then, we can replace Eqs. (15.14) and (15.15) in Eq. (15.7) and use the result from Eq. (15.18) to conclude that:

$$P_\infty(T) = e^{\text{Re}\{B\}T} e^{\text{Re}\{C\}T} = e^{2\text{Re}\{-i\langle\psi|\hat{H}^{(S)}|\psi\rangle/\hbar\}T} = e^0 = 1. \tag{15.19}$$

Hence, the probability of finding the system in the initial state becomes 1 as the projective observations become more closely spaced. This is the Quantum Zeno Effect: a system under constant observation of its initial state will never evolve to a different state.

15.2 Quantum Zeno Effect via continuous measurements

Misra and Sudarshan discussed the implications of the existence of such continuous measurements in their original article:

> "Is it a curious but innocent mathematical result or does it have something to say about the foundation of quantum theory? Does it, for example, urge us to have a principle in the formulation of quantum theory that forbids the continuous observation of an observable that is not a constant of motion?" [180]

They conclude that quantum mechanics may lack the proper mathematical methods to describe this type of measurement and to calculate the probability that a system will change its state within a period of time:

> "The quantum Zeno's paradox shows that the seemingly natural approach to this problem discussed in the preceding sections leads to bizarre and physically unacceptable answers. We thus lack a trustworthy quantum-theoretic algorithm for computing such probabilities. Until such a trustworthy algorithm is developed the completeness of quantum theory must remain in doubt." [180]

Since then, however, experimental evidence seems to be demonstrating that the Quantum Zeno Effect occurs in nature and that the quantum theory is correct [182] – although not without some controversy [183].

In terms of formalism, the development of master equations such as Eq. (14.8) have allowed the dynamical description of a system under constant observation, including means of obtaining the Quantum Zeno Effect [179]. As a simple example of this, suppose we have a spin-1/2 system initially in the state $|+\rangle$, but subjected to a Hamiltonian that flips this state:

$$\hat{H}^{(S)} = \hbar\omega_0 \left(|+\rangle\langle-| + |-\rangle\langle+|\right) = \hbar\omega_0\hat{\sigma}_x, \tag{15.20}$$

where ω_0 is a frequency. Concomitantly, we observe it with an apparatus that continuously detects whether the system has changed its initial state, which corresponds to the following Kraus operators:

$$\hat{K}_1^{(S)} = |+\rangle\langle+|, \quad \hat{K}_2^{(S)} = |-\rangle\langle-|. \tag{15.21}$$

It is easy to see that they add up to the identity, as demanded by Eq. (13.14).

Replacing Eqs. (15.20) and (15.21) in Eq. (14.8), we find:

$$\frac{d}{dt}\hat{\rho}^{(S)}(t) = -i\omega_0\left[\hat{\sigma}_x, \hat{\rho}^{(S)}(t)\right] + \frac{1}{2}\lambda\left\{\hat{\sigma}_z\hat{\rho}^{(S)}(t)\hat{\sigma}_z - \hat{\rho}^{(S)}(t)\right\}, \tag{15.22}$$

where we used the following identity to simplify the Lindbladian term in Eq. (15.22):

$$\hat{\sigma}_z\hat{\rho}^{(S)}(t)\hat{\sigma}_z - \hat{\rho}^{(S)}(t) = (|+\rangle\langle+| - |-\rangle\langle-|)\,\hat{\rho}^{(S)}(t)\,(|+\rangle\langle+| - |-\rangle\langle-|) - \hat{\rho}^{(S)}(t)$$

$$= 2\,|+\rangle\langle+|\hat{\rho}^{(S)}(t)|+\rangle\langle+| + 2\,|-\rangle\langle-|\hat{\rho}^{(S)}(t)|-\rangle\langle-|$$

$$- 2\,|+\rangle\langle+|\hat{\rho}^{(S)}(t)|+\rangle\langle+| - 2\,|-\rangle\langle-|\hat{\rho}^{(S)}(t)|-\rangle\langle-|$$

$$- 2\,|-\rangle\langle-|\hat{\rho}^{(S)}(t)|+\rangle\langle+| - 2\,|+\rangle\langle+|\hat{\rho}^{(S)}(t)|-\rangle\langle-|$$

$$= 2\,|+\rangle\langle+|\hat{\rho}^{(S)}(t)|+\rangle\langle+| + 2\,|-\rangle\langle-|\hat{\rho}^{(S)}(t)|-\rangle\langle-|$$

$$- 2\,(|+\rangle\langle+| + |-\rangle\langle-|)\,\hat{\rho}^{(S)}(t)\,(|+\rangle\langle+| + |-\rangle\langle-|)$$

$$= 2\sum_i \hat{K}_i^{(S)}\hat{\rho}^{(S)}(t)\left[\hat{K}_i^{(S)}\right]^\dagger - 2\hat{\rho}^{(S)}(t).$$

The probability of the system being detected in the same state as the initial state by a projective measurement at the instant t is found by multiplying the density matrix by $|+\rangle\langle+|$ and taking the trace:

$$P_\infty(t) = \mathrm{Tr}\left\{\hat{\rho}^{(S)}(t)|+\rangle\langle+|\right\} = \langle+|\hat{\rho}^{(S)}(t)|+\rangle.$$

Multiplying Eq. (15.22) by $\langle\pm|$ and $|\pm\rangle$, we find the differential equations for the components of the density matrix:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle+|\hat{\rho}^{(S)}(t)|+\rangle = -i\omega_0\langle-|\hat{\rho}^{(S)}(t)|+\rangle + i\omega_0\langle+|\hat{\rho}^{(S)}(t)|-\rangle,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle+|\hat{\rho}^{(S)}(t)|-\rangle = -i\omega_0\langle-|\hat{\rho}^{(S)}(t)|-\rangle + i\omega_0\langle+|\hat{\rho}^{(S)}(t)|+\rangle - \lambda\langle+|\hat{\rho}^{(S)}(t)|-\rangle,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle-|\hat{\rho}^{(S)}(t)|+\rangle = -i\omega_0\langle+|\hat{\rho}^{(S)}(t)|+\rangle + i\omega_0\langle-|\hat{\rho}^{(S)}(t)|-\rangle - \lambda\langle-|\hat{\rho}^{(S)}(t)|+\rangle,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle-|\hat{\rho}^{(S)}(t)|-\rangle = -i\omega_0\langle+|\hat{\rho}^{(S)}(t)|-\rangle + i\omega_0\langle-|\hat{\rho}^{(S)}(t)|+\rangle.$$

In matrix form, this system of four linear equations can be written as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \langle+|\hat{\rho}^{(S)}(t)|+\rangle \\ \langle+|\hat{\rho}^{(S)}(t)|-\rangle \\ \langle-|\hat{\rho}^{(S)}(t)|+\rangle \\ \langle-|\hat{\rho}^{(S)}(t)|-\rangle \end{pmatrix} = \begin{pmatrix} 0 & i\omega_0 & -i\omega_0 & 0 \\ i\omega_0 & -\lambda & 0 & -i\omega_0 \\ -i\omega_0 & 0 & -\lambda & i\omega_0 \\ 0 & -i\omega_0 & i\omega_0 & 0 \end{pmatrix}\begin{pmatrix} \langle+|\hat{\rho}^{(S)}(t)|+\rangle \\ \langle+|\hat{\rho}^{(S)}(t)|-\rangle \\ \langle-|\hat{\rho}^{(S)}(t)|+\rangle \\ \langle-|\hat{\rho}^{(S)}(t)|-\rangle \end{pmatrix},$$

whose solution for the initial state $|+\rangle\langle+|$ is the matrix exponential:

$$\begin{pmatrix} \langle+|\hat{\rho}^{(S)}(t)|+\rangle \\ \langle+|\hat{\rho}^{(S)}(t)|-\rangle \\ \langle-|\hat{\rho}^{(S)}(t)|+\rangle \\ \langle-|\hat{\rho}^{(S)}(t)|-\rangle \end{pmatrix} = \exp\left\{\begin{pmatrix} 0 & i\omega_0 & -i\omega_0 & 0 \\ i\omega_0 & -\lambda & 0 & -i\omega_0 \\ -i\omega_0 & 0 & -\lambda & i\omega_0 \\ 0 & -i\omega_0 & i\omega_0 & 0 \end{pmatrix}t\right\}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{15.23}$$

The exponential of the matrix in Eq. (15.23) is not easy to be calculated analytically, but its numerical simulation showing the Quantum Zeno Effect can be seen in Fig. 15.1. With this, we see that the formalism of continuous measurements allows us to quantify the Quantum Zeno Effect at each stage of the system's evolution with less hassle than the very formal result obtained before for projective measurements.
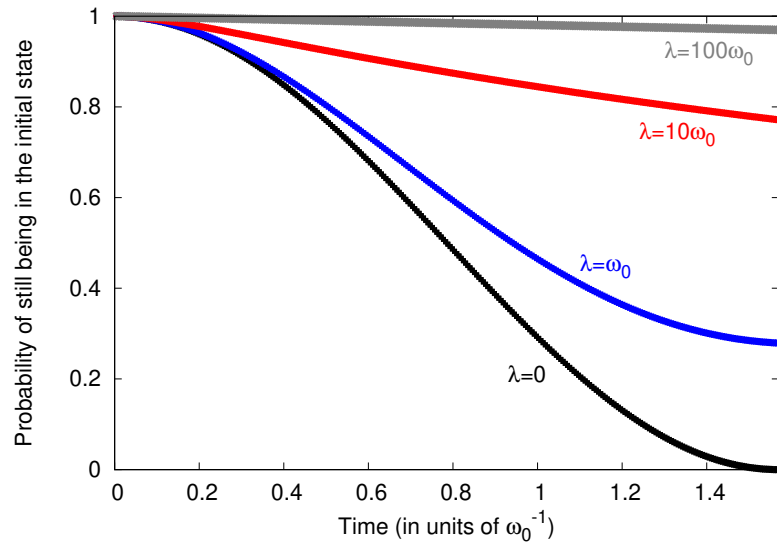
Fig. 15.1: Probability $P_\infty(t)$ of keeping a system under continuous measurement in its initial state, calculated by solving numerically Eq. (15.23). The unperturbed system ($\lambda = 0$) transitions to a different state, but as the intensity λ of the interaction with the measurement apparatus increases, the system almost does not leave the initial state, analogously to the Quantum Zeno Effect for discrete projective measurements.

16   Summary

The emergence of quantum mechanics in the end of the 19th century and beginning of the 20th century resulted in the quantization of magnitudes that used to be continuously measured. After a few years, this discontinuity was recognized as typical of atomic processes in general. Meanwhile, new experiments and theoretical advances by Einstein and de Broglie revealed that atomic systems also displayed both wave and particle aspects. This inspired Schrödinger's introduction of the wave function, which would be interpreted according to the rule of Born-Dirac, which turned measurements into probabilistic processes. If in classical physics we could talk about an isolated system, in the orthodox interpretation of quantum mechanics it is necessary to consider the measurement apparatus and its interaction with the observed system. According to the uncertainty principle, this imposes intrinsic limits to the precision of what can be measured.

These developments were concerned with what one *cannot* do with a quantum measurement, and constitute the bulk of the problem of measurement that persists to this day. The main difficulty is that the measurement in quantum mechanics necessarily requires the interaction between a quantum system and a classical apparatus, in a process that neither theory can fully describe. But if the apparatus is considered quantum as well, it is possible to understand its dynamics and what *can* be done with quantum measurements.

This is the step taken by von Neumann with his influential model of measurement, which has had lasting influence. von Neumann boldly treated the measurement apparatus as a quantum system, and described the system-measurer interaction as unitary, thus constructing a theory for the transmission of information from the atomic world to our classical macroscopic instruments.

Treating the apparatus as a quantum object sparked a fruitful discussion about where the quantum world ends. While Wigner believed that the answer lied in the consciousness of the observer, which would make it a privileged system, Bohm attributed classical features to the intensity of the interaction between the apparatus and the observer, which would make the interference terms vanish. Everett went further, treating the whole universe as a wave function, where worlds containing all the possible outcomes would co-exist without interacting with each other. These different views are summarized in Table 16.1.

A discussion of the finer points of Everett's interpretation led to the decoherence program, now used independently of the many-worlds interpretations to describe errors in quantum computers, for example. This is not the only practical contribution of von Neumann's model: the Quantum Phase Estimation algorithm, for example, proposes using an analogous interaction between two discrete systems to determine the energy of a quantum state encoded as qubits [184]. Moreover, reducing the intensity of the interaction during von Neumann's measurement led to weak measurements, which are under intense research due to the surprising effects that they cause, such as obtaining results that lie much beyond the expected limits.

This is a field still open to investigations, as exemplified by the several books [185–189], reports [190, 191] and collected works [79, 192] on the subject. We have not intended to write a complete treatise on quantum measurements, but rather to provide a broad introduction to the subject centered on the main paradigm introduced by von Neumann, discussing both the orthodoxy and some attempts to break with it. Other approaches that, for concision, were not included here include quantum state diffusion [193]; dynamical reduction models [190] like the original proposal by Giancarlo Ghirardi, Alberto Rimini and Tulio Weber (GRW) [194] and related developments [195, 196]; superselection rules [197–201]; and consistent histories [13,

Interpretation	Observed system (S)	Measurement apparatus (M)	Human observer (E)
Orthodox	Superposition	Incoherent mixture	Incoherent mixture
Bohm's formulation of von Neumann's model	Superposition	Superposition when interacting with S, incoherent mixture when interacting with E	Incoherent mixture
Wigner	Superposition	Superposition	Consciousness destroys superposition
Everett	Superposition	Superposition	Superposition

Tab. 16.1: Which parts of von Neumann's chain can be in superposition states and which parts cannot be more than incoherent mixtures, according to the interpretations mentioned in this article.

146, 147]. We hope that, with the tools we have provided, the readers will feel prepared to make an informed choice about where to continue pursuing their studies.

Appendices

A   Quantum non-demolition measurements

A quantum non-demolition (QND) measurement is constructed so that, if its result is an eigenvalue s_n of the observable $\hat{S}^{(S)}$, then any subsequent measurement of this observable will yield the same result. This means that the experimental apparatus will not insert more uncertainty into the system beyond the uncertainty inherent to quantum mechanics. Examples of QND measurements can be found in quantum optics [202–204] and in electromechanical systems [205].

An example of a measurement that does not satisfy this is the position of a free particle [206]. There is no restriction in principle about how precise the position can be determined, but this will create some uncertainty Δp in the associated momentum. A new measurement of the position performed at a time t after the first one will have an uncertainty corresponding to how much the particle may have moved during this interval, $\Delta x = t\,(\Delta p)\,/m$, where m is the mass of the free particle. Hence, a new measurement of the position will not necessarily yield the same value as the original measurement. On the other hand, a measurement of the momentum will not be affected by the uncertainty in the position, and can therefore be turned into a QND measurement.

More generally, suppose that we want to measure a system originally in a certain eigenstate $|s_n\rangle$ of $\hat{S}^{(S)}$, while the measuring apparatus is in the original state $|m_0\rangle$. If the measurement is a QND, the state after the interaction $\hat{\Delta}$ should still be an eigenstate with the same eigenvalue s_n:

$$\hat{S}^{(S)}\left[\hat{\Delta}\left(|s_n\rangle\,|m_0\rangle\right)\right] = s_n\left[\hat{\Delta}\left(|s_n\rangle\,|m_0\rangle\right)\right]. \tag{A.1}$$

But as $\hat{S}^{(S)}|s_n\rangle = s_n|s_n\rangle$, Eq. (A.1) can be re-written as:

$$\hat{S}^{(S)}\hat{\Delta}\left(|s_n\rangle\,|m_0\rangle\right) = \hat{\Delta}\left(\hat{S}^{(S)}|s_n\rangle\,|m_0\rangle\right). \tag{A.2}$$

In order for Eq. (A.2) to be true for any s_n, the following identity must be satisfied:

$$\left[\hat{S}^{(S)}, \hat{\Delta}\right]|m_0\rangle = 0. \tag{A.3}$$

Eq. (A.3) is the necessary and sufficient condition for a measurement to be considered a QND measurement [207]. A sufficient condition that satisfies Eq. (A.3) is if $\hat{S}^{(S)}$ and $\hat{\Delta}$ commute:

$$\left[\hat{S}^{(S)}, \hat{\Delta}\right] = \hat{0}. \tag{A.4}$$

Furthermore, as $\hat{\Delta} = e^{-i\hat{H}\tau/\hbar}$, Eq. (A.4) will be satisfied if the observable commutes with the Hamiltonian, which is another sufficient, but not necessary condition:

$$\left[\hat{S}^{(S)}, \hat{H}\right] = \hat{0}. \tag{A.5}$$

Eq. (A.5) is equivalent to saying that $\left\langle \hat{S} \right\rangle$ is a constant of motion, which means that the eigenvalue s_n will not change in time.

Returning to the example given in the beginning of this section, we can see that, for a free particle, the Hamiltonian of the system is:

$$\hat{H}^{(S)} = \frac{\hat{P}^2}{2m}.$$

Unless the interaction Hamiltonian $\hat{H}^{(S+M)}$ is chosen very carefully, the total Hamiltonian $\hat{H} = \hat{H}^{(S)} + \hat{H}^{(M)} + \hat{H}^{(S+M)}$ will not commute with the position operator $\hat{Q}$, because the term $\hat{H}^{(S)}$ does not commute with it. This means that the condition from Eq. (A.5) will not be satisfied. On the other hand, operator $\hat{P}$ commutes with this $\hat{H}^{(S)}$, and will also commute with the interaction term as long as $\hat{H}^{(S+M)}$ takes a form similar to Eq. (8.6). Therefore, a measurement of the position causes the demolition of the system, while a measurement of the momentum can be a QND measurement.

B   von Neumann's derivation of the density matrix formalism

Here we reproduce the procedure employed in von Neumann's textbook [91] using modern notation. Write the observable $\hat{Q}^{(S)}$ of the system S in matrix form, using the eigenbasis $\{|s_n\rangle\}$:

$$\hat{Q}^{(S)} = \sum_{m,n} Q_{m,n} |s_m\rangle \langle s_n|.$$

Now, divide the elements of the matrix into diagonal and off-diagonal:

$$\hat{Q}^{(S)} = \sum_m Q_{m,m} |s_m\rangle \langle s_m| + \sum_{m>n} \{Q_{m,n} |s_m\rangle \langle s_n| + Q_{n,m} |s_n\rangle \langle s_m|\}. \tag{B.1}$$

Next, define three new sets of observables:

$$\hat{A}_m^{(S)} \equiv |s_m\rangle \langle s_m|,$$
$$\hat{B}_{m,n}^{(S)} \equiv |s_m\rangle \langle s_n| + |s_n\rangle \langle s_m|,$$
$$\hat{C}_{m,n}^{(S)} \equiv i |s_m\rangle \langle s_n| - i |s_n\rangle \langle s_m|,$$

and replace them in Eq. (B.1), obtaining the following decomposition of $\hat{Q}^{(S)}$:

$$\hat{Q}^{(S)} = \sum_m Q_{m,m} \hat{A}_m^{(S)} + \frac{1}{2} \sum_{m>n} \left\{ (Q_{m,n} + Q_{n,m}) \hat{B}_{m,n}^{(S)} - i (Q_{m,n} - Q_{n,m}) \hat{C}_{m,n}^{(S)} \right\}. \tag{B.2}$$

Using Eq. (B.2), the expectation value of any observable $\hat{Q}^{(S)}$ can be calculated from the expectation values of the auxiliary observables $\hat{A}_m^{(S)}$, $\hat{B}_{m,n}^{(S)}$, and $\hat{C}_{m,n}^{(S)}$. Thus, von Neumann defines the density operator $\hat{\rho}^{(S)}$ as the Hermitian matrix formed by these expectation values:

$$\hat{\rho}^{(S)} \equiv \sum_m \left\langle \hat{A}_m^{(S)} \right\rangle \hat{A}_m^{(S)} + \frac{1}{2} \sum_{m>n} \left\{ \left\langle \hat{B}_{m,n}^{(S)} \right\rangle \hat{B}_{m,n}^{(S)} - \left\langle \hat{C}_{m,n}^{(S)} \right\rangle \hat{C}_{m,n}^{(S)} \right\},$$

so that the expectation value of any observable $\hat{Q}^{(S)}$ will be given by the trace of it multiplied by the density matrix, just as in Eq. (8.13):

$$\begin{aligned}
\left\langle \hat{Q}^{(S)} \right\rangle &= \sum_m Q_{m,m} \left\langle \hat{A}_m^{(S)} \right\rangle + \frac{1}{2} \sum_{m>n} \left\{ (Q_{m,n} + Q_{n,m}) \left\langle \hat{B}_{m,n}^{(S)} \right\rangle \right. \\
&\quad \left. -i (Q_{m,n} - Q_{n,m}) \left\langle \hat{C}_{m,n}^{(S)} \right\rangle \right\} \\
&= \sum_m \text{Tr}_S \left\{ \hat{Q}^{(S)} \hat{A}_m^{(S)} \right\} \left\langle \hat{A}_m^{(S)} \right\rangle + \frac{1}{2} \sum_{m>n} \left\{ \text{Tr}_S \left\{ \hat{Q}^{(S)} \hat{B}_{m,n}^{(S)} \right\} \left\langle \hat{B}_{m,n}^{(S)} \right\rangle \right. \\
&\quad \left. -\text{Tr}_S \left\{ \hat{Q}^{(S)} \hat{C}_{m,n}^{(S)} \right\} \left\langle \hat{C}_{m,n}^{(S)} \right\rangle \right\} \\
&= \text{Tr}_S \left\{ \hat{\rho}^{(S)} \hat{Q}^{(S)} \right\}.
\end{aligned}$$

C   Splitting a positive operator

Suppose we want to split a positive Hermitian operator $\hat{A}^{(S)}$ into the product of two operators $\hat{B}^{(S)}$ and $\hat{C}^{(S)}$:

$$\hat{A}^{(S)} = \hat{B}^{(S)}\hat{C}^{(S)}. \tag{C.1}$$

This is clearly feasible if we choose $\hat{B}^{(S)} = \hat{A}^{(S)}$ and $\hat{C}^{(S)} = \hat{1}^{(S)}$ (or vice-versa), but we are seeking non-trivial solutions.

The fact that $\hat{A}^{(S)}$ is Hermitian and positive means that it can be diagonalized in its eigenbasis $|a_i\rangle$, leaving its positive eigenvalues a_i in the diagonal:

$$\hat{A}^{(S)} = \sum_i a_i |a_i\rangle \langle a_i|. \tag{C.2}$$

We can write $\hat{B}^{(S)}$ and $\hat{C}^{(S)}$ in this same basis, although they are not necessarily diagonal:

$$\hat{B}^{(S)} = \sum_{i,j} b_{i,j} |a_i\rangle \langle a_j|, \quad \hat{C}^{(S)} = \sum_{i,j} c_{i,j} |a_i\rangle \langle a_j|. \tag{C.3}$$

Replacing Eqs. (C.2) and (C.3) in Eq. (C.1), we find:

$$\sum_i a_i |a_i\rangle \langle a_i| = \sum_{i,j} \sum_k b_{i,k} c^*_{k,j} |a_i\rangle \langle a_j|.$$

This is true only if the coefficients that multiply $|a_i\rangle \langle a_j|$ on both sides are equal:

$$\delta_{i,j} a_i = \sum_k b_{i,k} c^*_{k,j}. \tag{C.4}$$

The trivial solution to this is $b_{i,k} = \delta_{i,k} a_i$ and $c_{k,j} = \delta_{k,j}$, which corresponds to $\hat{B}^{(S)} = \hat{A}^{(S)}$ and $\hat{C}^{(S)} = \hat{1}^{(S)}$. However, as $a_i > 0$, it is far more interesting to choose

$$b_{i,k} = \delta_{i,k} e^{i\phi_i} \sqrt{a_i}, \quad c_{k,j} = \delta_{k,j} e^{-i\phi_i} \sqrt{a_i}, \tag{C.5}$$

where ϕ_i is a phase. Eq. (C.5) clearly satisfies Eq. (C.4), but it also means that $\hat{B}^{(S)} = \left[\hat{C}^{(S)}\right]^\dagger$. Therefore, every positive Hermitian operator $\hat{A}^{(S)}$ can be written as the product:

$$\hat{A}^{(S)} = \left[\hat{C}^{(S)}\right]^\dagger \hat{C}^{(S)}.$$

D   Deriving the POVM

We want to make sure that the $\hat{A}_3^{(S)}$ defined in Eq. (13.6) is a positive operator. This means that, for every state vector

$$|\psi\rangle = \cos\theta\,|+\rangle + \sin\theta\,|-\rangle,$$

the expectation value of $\hat{A}_3^{(S)}$ must be positive:

$$\langle\psi|\,\hat{A}_3^{(S)}\,|\psi\rangle = \langle\psi|\,\psi\rangle - \langle\psi|\,\hat{A}_1^{(S)}\,|\psi\rangle - \langle\psi|\,\hat{A}_2^{(S)}\,|\psi\rangle = 1 - p_z\,|\cos\theta|^2 - p_x\frac{1}{2}\,|\cos\theta + \sin\theta|^2 \geq 0.$$

Assuming that both detectors are equally important to our experiment, we choose $p_x = p_z$. Then, a little algebra turns

$$1 - p_z\,|\cos\theta|^2 - p_z\frac{1}{2}\,|\cos\theta + \sin\theta|^2 \geq 0$$

into the inequality

$$p_z f(\theta) \leq 1, \tag{D.1}$$

where we defined the real function

$$f(\theta) \equiv \cos^2\theta + \frac{1}{2}\,(\cos\theta + \sin\theta)^2 = 1 + \frac{1}{2}\cos(2\theta) + \frac{1}{2}\sin(2\theta).$$

The value of p_z must be chosen so that Eq. (D.1) is satisfied for every state vector, and hence for every value of θ. This is true as long as it is satisfied for the maximum value of $f(\theta)$.

The extremes of the function $f(\theta)$ occur when its derivative is zero:

$$\frac{\mathrm{d}}{\mathrm{d}\theta}f(\theta) = -\sin(2\theta) + \cos(2\theta) = 0,$$

which is to say that

$$\sin 2\theta = \cos 2\theta,$$

meaning that $2\theta = \pi/4$. A second derivative shows that this is indeed a maximum of the function:

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2}f(\theta)\bigg|_{2\theta=\pi/4} = -2\,\cos(2\theta)|_{2\theta=\pi/4} - 2\,\sin(2\theta)|_{2\theta=\pi/4} = -2\frac{1}{\sqrt{2}} - 2\frac{1}{\sqrt{2}} < 0.$$

Hence, the largest value of p_z that we can choose, the one that will maximize the precision of our devices, is:

$$p_z = \frac{1}{f(2\theta = \pi/4)} = \frac{1}{1 + 1/\sqrt{2}} = \frac{\sqrt{2}}{\sqrt{2} + 1} = 2 - \sqrt{2}.$$

E   The Liouville–von Neumann equation

Here we will derive the evolution of a density matrix via Process 2. Any density matrix can be expressed as a convex combination of pure states:

$$\hat{\rho}(t) = \sum_i p_i \, |\Psi_i(t)\rangle \, \langle\Psi_i(t)| \,,$$

where the p_i are probabilities and the $|\Psi_i\rangle$ are state vectors.

To find the equation of motion for Process 2, it suffices to take the time derivative of a density matrix:

$$\frac{d\hat{\rho}(t)}{dt} = \sum_i p_i \frac{d\,|\Psi_i(t)\rangle}{dt} \, \langle\Psi_i(t)| + \sum_i p_i \, |\Psi_i(t)\rangle \, \frac{d\,\langle\Psi_i(t)|}{dt} \,.$$

Replacing the Schrödinger equation:

$$i\hbar \frac{d\,|\Psi_i(t)\rangle}{dt} = \hat{H}\,|\Psi_i(t)\rangle$$

and its Hermitian conjugate:

$$-i\hbar \frac{d\,\langle\Psi_i(t)|}{dt} = \langle\Psi_i(t)|\,\hat{H},$$

we find the Liouville–von Neumann equation [76, 91] in terms of the Hamiltonian $\hat{H}$:

$$\frac{d\hat{\rho}(t)}{dt} = -\frac{i}{\hbar}\left[\hat{H}, \hat{\rho}(t)\right]. \tag{E.1}$$

The solution for Eq. (E.1) is given by the time evolution operator $\hat{U}(t)$,

$$\hat{\rho}(t) = \hat{U}(t)\hat{\rho}(0)\left[\hat{U}(t)\right]^{\dagger}, \tag{E.2}$$

which is defined as the operator that satisfies

$$\frac{d\hat{U}(t)}{dt} = -\frac{i}{\hbar}\hat{H}\hat{U}(t), \tag{E.3}$$

and which becomes the following when the Hamiltonian $\hat{H}$ is time-independent:

$$\hat{U}(t) = \exp\left\{-\frac{i}{\hbar}\hat{H}t\right\}.$$

If we take the time derivative of Eq. (E.2) and replace Eq. (E.3) in it, we obtain the Liouville-von Neumann equation as expressed in Eq. (E.1), thus proving that Eq. (E.2) is indeed the solution.

References

[1] P.-S. de Laplace. *A philosophical essay on probabilities*. John Wiley & Sons, New York, 1902.

[2] A. Pais. *Subtle is the Lord... – The science and the life of Albert Einstein*. Oxford University Press, Oxford, 1982.

[3] P. W. Shor. Scheme for reducing decoherence in quantum computer memory. *Physical Review A*, 52, 1995.

[4] H. J. Briegel and R. Raussendorf. Persistent entanglement in arrays of interacting particles. *Physical Review Letters*, 86:910–913, 2001.

[5] H. J. Briegel, D. E. Browner, W. Dür, R. Raussendorf, and M. V. D. Nest. Measurement-based quantum computation. *Nature Physics*, 5:19–26, 2009.

[6] R. Eisberg and R. Resnick. *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles*. John Wiley & Sons, New York, 1985.

[7] A. Messiah. *Quantum mechanics, vol. 1*. North-Holland Publishing Company, Amsterdam, 1961.

[8] L. I. Schiff. *Quantum mechanics*. McGraw-Hill Book Company, New York, 1968.

[9] R. Eisberg and R. Resnick. *Fundamentals of Modern Physics*. John Wiley & Sons, New York, 1961.

[10] E. Merzbacher. *Quantum mechanics*. John Wiley & Sons Inc, New York, 1998.

[11] L. E. Ballentine. *Quantum mechanics: A modern development*. World Scientific Publishing, Singapore, 1988.

[12] A. Peres. *Quantum theory: Concepts and methods*. Kluwer Academic Publishers, New York, 2002.

[13] R. B. Griffiths. *Consistent Quantum Theory*. Cambridge University Press, Cambridge, 2002.

[14] E. M. Purcell. *Electricity and magnetism – Berkeley Physics Course vol. 2*. McGraw–Hill Book Company, New York, 1985.

[15] F. W. Sears and G. L. Salinger. *Thermodynamics, kinetic theory, and statistical thermodynamics*. Addison-Wesley Publishing Company, Reading, 1982.

[16] A. P. French and E. F. Taylor. *Special relativity*. W. W. Norton and Company Inc, New York, 1968.

[17] M. Bucher. Rise and premature fall of the old quantum theory, 2008. arXiv:0802.1366.

[18] A. Hermann. *The genesis of quantum theory 1899–1913*. MIT Press, Cambridge, 1971.

[19] M. Jammer. *The conceptual development of quantum mechanics*. McGraw–Hill Book Company, New York, 1966.

[20] B. Feldens, P. M. C. Dias, and W. M. S. Santos. E assim se fez o quantum... / Let there be quantum... *Revista Brasileira de Ensino de Física*, 22:2602, 2010.

[21] S.-I. Tomonaga. *Quantum mechanics vol. 1*. North-Holland Publishing Company, Amsterdam, 1968.

[22] L. A. Castro, C. A. Brasil, and R. d. J. Napolitano. Elliptical orbits in the phase-space quantization. *Revista Brasileira de Ensino de Física*, 38:e3318, 2016.

[23] A. d'Abro. *The rise of the new physics*. Dover Publications Inc, New York, 1951.

[24] A. Duncan and M. Janssen. The trouble with the orbits: The Stark effect in the old and the new quantum theory. *Studies in History and Philosophy of Science*, 48:68–83, 2014.

[25] A. Duncan and M. Janssen. The Stark effect in the Bohr–Sommerfeld theory and in Schrödinger's wave mechanics, 2014. arXiv:1404.5341.

[26] F. A. G. Parente, A. C. F. dos Santos, and A. C. Tort. Os 100 anos do átomo de Bohr / One hundred years of Bohr's atom. *Revista Brasileira de Ensino de Física*, 35:4301, 2013.

[27] N. Studart. A invenção do conceito de quantum de energia segundo Planck. *Revista Brasileira de Ensino de Física*, 22:523–535, 2000.

[28] M. Planck. Über eine verbesserung der wienschen spektralgleichung. *Verhandlungen der Deutschen Physicalishen Gesellschaft*, 2:202–204, 1900.

[29] M. Planck. Ueber das gesetz der energieverteilung im normalspectrum. *Annalen der Physik*, 309:553–563, 1901.

[30] D. ter Haar. *The old quantum theory*. Pergamon Press Ltd, Oxford, 1967.

[31] T. S. Kuhn. *Black-body theory and the quantum discontinuity, 1894–1912*. University of Chicago Press, Chicago, 1987.

[32] A. Beck and P. Havas. *The collected papers of Albert Einstein vol. 2 - The Swiss Years: Writings, 1900-1909*. Princeton University Press, Princeton, 1989.

[33] A. Arons and M. B. Peppard. Einstein's proposal of the photon concept – a translation of the Annalen der Physik paper of 1905. *American Journal of Physics*, 33:367–374, 1965.

[34] J. J. Thomson. The magnetic properties of systems of corpuscles describing circular orbits. *Philosophical Magazine Series 6*, 6:673–693, 1903.

[35] J. J. Thomson. On the structure of the atom: An investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle; with application of the results to the theory of atomic structure. *Philosophical Magazine Series 6*, 7:237–265, 1904.

[36] E. Rutherford. The structure of the atom. *Philosophical Magazine and Journal of Science*, 27:488–498, 1914.

[37] H. Nagaoka. Motion of particles in an ideal atom illustrating the line and band spectra and the phenomena of radioactivity. *Bulletin of the Mathematical and Physical Society of Tokyo*, 2:140–141, 1904.

[38] H. Nagaoka. On a dynamical system illustrating the spectrum lines and the phenomena of radio-activity. *Nature*, 69:392–393, 1904.

[39] H. Nagaoka. Kinetics of a system of particles illustrating the line and band spectrum and the phenomena of radio-activity. *Philosophical Magazine*, 7:445–455, 1904.

[40] W. Heisenberg. *The physical principles of the quantum theory*. Dover, Mineola, 1949.

[41] F. Laloë. Do we really understand quantum mechanics? Strange correlations, paradoxes, and theorems. *American Journal of Physics*, 69:655–701, 2001.

[42] B. d'Espagnat. *Conceptual foundations of quantum mechanics*. Perseus Books, Reading, 1999.

[43] R. L. Liboff. The correspondence principle revisited. *Physics Today*, 37:50–55, 1984.

[44] A. J. Makowski. A brief survey of various formulations of the correspondence principle. *European Journal of Physics*, 27:1133–1139, 2006.

[45] G. Q. Hassoun and D. H. Kobe. Synthesis of the Planck and Bohr formulations of the correspondence principle. *American Journal of Physics*, 57:658–662, 1969.

[46] B. Bhattacharyya. Looking back into Bohr's atom. *European Journal of Physics*, 27:497–500, 2006.

[47] N. Bohr. On the constitution of atoms and molecules. *Philosophical Magazine*, 26:1–25, 1913.

[48] K. Pelogia and C. A. Brasil. Analysis of the Jun Ishiwara's 'The universal meaning of the quantum of action'. *European Physical Journal H*, 43:507–521, 2017.

[49] M. Eckert. How Sommerfeld extended Bohr's model of the atom (1913–1916). *European Physical Journal H*, 39:141–156, 2014.

[50] A. Sommerfeld. On the theory of the Balmer series. *European Physical Journal H*, 39:157–177, 2014.

[51] A. Sommerfeld. The fine structure of hydrogen and hydrogen-like lines. *European Physical Journal H*, 39:179–204, 2014.

[52] A. Sommerfeld. Zur quantentheorie der spektrallinien. *Annalen der Physik*, 51:1–94, 1916.

[53] A. Sommerfeld. Zur quantentheorie der spektrallinien. *Annalen der Physik*, 51:125–167, 1916.

[54] W. Wilson. The quantum-theory of radiation and line spectra. *Philosophical Magazine*, 29:795–802, 1915.

[55] W. Wilson. The quantum of action. *Philosophical Magazine*, 31:156–162, 1916.

[56] J. Ishiwara. Die universelle bedeutung des wirkungsquantums. *Proceedings of Tokyo Mathematico– Physical Society*, 8:106–116, 1915.

[57] J. Ishiwara. The universal meaning of the quantum of action. *European Physical Journal H*, 43:523–536, 2017.

[58] B. L. van der Waerden. *Sources of quantum mechanics*. Dover Publications Inc, Mineola, 2007.

[59] M. Planck. The nobel prize in physics 1918. http://www.nobelprize.org/nobel_prizes/physics /laureates/1918/planck-lecture. Accessed on: September 17, 2015.

[60] R. A. Martins and P. S. Rosa. *História da teoria quântica – a dualidade onda partícula, de Einstein a de Broglie*. Editora Livraria da Física, São Paulo, 2014.

[61] L. de Broglie. *An introduction to the study of wave mechanics*. Methuen & Co. Ltd, Essex, 1930.

[62] L. de Broglie. Recherches sur la théorie des quanta. *Annales de Physique*, 10:22–128, 1925.

[63] L. de Broglie. *On the theory of quanta*. Fondation Louis de Broglie, Paris, 2004.

[64] E. MacKinnon. De Broglie's thesis: a critical retrospective. *American Journal of Physics*, 44:1047–1055, 1976.

[65] M. Jammer. *The philosophy of quantum mechanics: the interpretations of quantum mechanics in historical perspective*. Wiley, New York, 1974.

[66] A. Einstein. Quantentheorie des einatomigen idealen Gases. Zweite Abhandlung. *Sitzungsberichte der Preußischen Akademie der Wissenschaften (Berlin). Physikalisch-mathematische Klasse*, pages 3–14, 1925.

[67] N. Bohr, H. K. Kramers, and J. C. Slater. The quantum theory of radiation. *Philosophical Magazine*, 47:785–802, 1924.

[68] W. Heisenberg. *Physics and philosophy – The revolution in modern science*. Harper Perennial, New York, 2007.

[69] C. Joas and C. Lehner. The classical roots of wave mechanics: Schrödinger's transformations of the optical-mechanical analogy. *Studies in History and Philosophy of Modern Physics*, 40:338–351, 2009.

[70] R. Köberle. Sobre a gênese da mecânica ondulatória. *Revista Brasileira de Física*, 9:243–273, 1979.

[71] L. Wessels. Schrödinger's route to wave mechanics. *Studies in History and Philosophy of Science*, 10:311–340, 1977.

[72] A. L. Fetter and J. D. Walecka. *Theoretical mechanics of particles and continua*. Dover, Mineola, 2003.

[73] E. Schrödinger. Quantisierung als eigenwertproblem. *Annalen der Physik*, 384:489–527, 1926.

[74] E. Schrödinger. Quantisierung als eigenwertproblem. *Annalen der Physik*, 385:361–376, 1926.

[75] E. Schrödinger. *Collected papers on wave mechanics*. Chelsea Pub. Co., New York, 1978.

[76] C. C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum mechanics*. Wiley, New York, 1977.

[77] G. S. Im. Experimental constraints on formal quantum mechanics: The emergence of Born's quantum theory of collision processes in Göttingen, 1924–1927. *Archive for History of Exact Sciences*, 50:73–101, 1995.

[78] M. Born. Zur quantenmechanik der stoßvorgänge. *Zeitschrift für Physik*, 37:863–867, 1926.

[79] J. A. Wheeler and W. H. Zurek. *Quantum theory and measurement*. Princeton University Press, Princeton, 1983.

[80] M. Born. Quantenmechanik der stoßvorgänge. *Zeitschrift für Physik*, 38:803–827, 1926.

[81] G. Ludwig. *Wave mechanics*. Pergamon Press, Oxford, 1968.

[82] E. Butkov. *Mathematical physics*. Addison–Wesley, Reading, 1968.

[83] P. Morse and H. Feshbach. *Methods of theoretical physics*. McGraw–Hill, New York, 1953.

[84] K. Gottfried. P. A. M. Dirac and the discovery of quantum mechanics. *American Journal of Physics*, 79:261–266, 2011.

[85] P. A. M. Dirac. On the theory of quantum mechanics. *Proceedings of the Royal Society of London A*, 112:661–677, 1926.

[86] M. Bunge. Survey of the interpretations of quantum mechanics. *American Journal of Physics*, 24:272–286, 1956.

[87] J. B. Hartle. Quantum mechanics of individual systems. *American Journal of Physics*, 36:704–712, 1968.

[88] L. E. Ballentine. The statistical interpretation of quantum mechanics. *Reviews of Modern Physics*, 42:358–380, 1970.

[89] N. R. Hanson. Copenhagen interpretation of quantum theory. *American Journal of Physics*, 27:1–15, 1959.

[90] H. P. Stapp. The Copenhagen interpretation. *American Journal of Physics*, 40:1098–1116, 1972.

[91] J. von Neumann. *Mathematical foundations of quantum mechanics*. Princeton University Press, Princeton, 1953.

[92] D. Howard. Who invented the 'Copenhagen interpretation'? a study in mythology. *Philosophy of Science*, 71:669–682, 2004.

[93] W. Pauli, L. Rosenfeld, and V. Weisskopf. *Niels Bohr and the development of physics – Essays dedicated to Niels Bohr on the occasion of his seventieth birthday*. Pergamon Press, London, 1955.

[94] N. Bohr. The quantum postulate and the recent development of atomic theory. *Nature*, 121:580–590, 1928.

[95] J. Roldán-Charria. Indivisibility, complementarity and ontology: A Bohrian interpretation of quantum mechanics. *Foundations of Physics*, 44:1336–1356, 2014.

[96] R. Omnès. *The interpretation of quantum mechanics*. Princeton University Press, Princeton, 1994.

[97] N. Bohr. *Atomic physics and human knowledge*. John Wiley & Sons, New York, 1958.

[98] R. Omnès. *Quantum philosophy*. Princeton University Press, Princeton, 1994.

[99] J. R. Henderson. Classes of Copenhagen interpretations: Mechanisms of collapse as typologically determinative. *Studies in History and Philosophy of Modern Physics*, 41:1–8, 2010.

[100] W. Heisenberg. *Physics and beyond – encounters and conversations*. Harper & Row, New York, 1971.

[101] S. Seth. Crafting the quantum: Arnold Sommerfeld and the older quantum theory. *Studies in History and Philosophy of Science*, 39:335–349, 2008.

[102] S. Seth. Zeideutigkeit about Zweideutigkeit: Sommerfeld, Pauli and the methodological origins of quantum mechanics. *Studies in History and Philosophy of Science*, 40:303–315, 2009.

[103] H. Reichenbach. *Philosophic foundations of quantum mechanics*. University of California Press, Berkeley, 1948.

[104] W. Heisenberg. Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. *Zeitschrift für Physik*, 43:172–198, 1927.

[105] R. W. Ditchburn. The uncertainty principle in quantum mechanics. *Proceedings of the Royal Irish Academy*, 39:73–81, 1930.

[106] H. P. Robertson. The uncertainty principle. *Physical Review*, 34:163–164, 1929.

[107] H. P. Robertson. Proceedings of the american physical society, minutes of the new york meeting, Feb. 21 and 22 – A general formulation of the uncertainty principle and its classical interpretation. *Physical Review*, 35:667A, 1930.

[108] H. P. Robertson. An indeterminacy relation for several observables and its classical interpretation. *Physical Review*, 46:794–801, 1934.

[109] A. Einstein. Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt. *Annalen der Physik*, 17:132, 1905.

[110] N. Bohr. On the constitution of atoms and molecules. *Philosophical Magazine*, 26:476–502, 1913.

[111] N. Bohr. On the constitution of atoms and molecules. *Philosophical Magazine*, 26:857–875, 1913.

[112] M. Planck. *Scientific autobiography and other papers*. Williams & Norgate Ltd, London, 1950.

[113] D. Bohm. *Quantum theory*. Dover Publications Inc, New York, 1989.

[114] F. D. Peat. *Infinity potential – The life and times of David Bohm*. Basic Books, 1996.

[115] L. D. Landau. *Collected Papers of L. D. Landau*. Pergamon Press Ltd, Oxford, 1965.

[116] U. Fano. Description of states in quantum mechanics by density matrix and operator techniques. *Reviews of Modern Physics*, 29:74–93, 1957.

[117] B. S. DeWitt and R. N. Graham. *The many-worlds interpretation of quantum mechanics: a fundamental exposition by Hugh Everett III*. Princeton University Press, Princeton, 1973.

[118] J. M. Jauch. *Foundations of quantum mechanics*. Addison–Wesley Publishing Company, Reading, 1968.

[119] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000.

[120] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47:777–780, 1935.

[121] N. Bohr. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 48:696–702, 1935.

[122] J. Mehra. *The quantum principle: its interpretation and epistemology*. D. Reidel Publishing Company, Dordrecht, 1974.

[123] M. Born. *The Born–Einstein Letters*. Macmillan, London, 1971.

[124] E. Schrödinger. Discussion of probability relations between separated systems. *Proceedings of the Cambridge Philosophical Society*, 31:555–563, 1935.

[125] E. Schrödinger. Probability relations between separated systems. *Proceedings of the Cambridge Philosophical Society*, 32:446–452, 1935.

[126] J. D. Trimmer. The present situation in quantum mechanics: A translation of schrödinger's 'cat paradox' paper. *Proceedings of the American Philosophical Society*, 124:323–338, 1980.

[127] H. Everett III. 'Relative state' formulation of quantum mechanics. *Reviews of Modern Physics*, 29:454–492, 1957.

[128] J. A. Wheeler. Assessment of Everett's 'relative state' formulation of quantum theory. *Reviews of Modern Physics*, 29:463–465, 1957.

[129] F. Freitas and O. Freire Jr. A formulação dos 'estados relativos' da teoria quântica. *Revista Brasileira de Ensino de Física*, 30:1–15, 2008.

[130] B. S. DeWitt. Bryce S. DeWitt letter to John Wheeler, Wheeler's copy. http://ucispace.lib.uci.edu/handle/10575/1145.

[131] C. A. Brasil and L. A. Castro. Understanding the pointer states. *European Journal of Physics*, 36:065024, 2014.

[132] O. Lombardi and L. Vanni. Medición cuántica y decoherencia: ¿qué medimos cuando medimos? *Scientiae Studia*, 8:273–291, 2010.

[133] W. H. Zurek. Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse? *Physical Review D*, 24:1516–1525, 1981.

[134] W. H. Zurek. Decoherence and the transition from quantum to classical. *Physics Today*, 44:36, 1991.

[135] S. Fortin and L. Vanni. Quantum decoherence: a logical perspective. *Foundations of Physics*, 44:1258–1268, 2014.

[136] R. Omnès. Results and problems in decoherence theory. *Brazilian Journal of Physics*, 35:207–210, 2005.

[137] M. Schlosshauer. Decoherence, the measurement problem, and interpretations of quantum mechanics. *Reviews of Modern Physics*, 76:1267–1305, 2004.

[138] M. Schlosshauer. *Decoherence and the quantum-to-classical transition*. Springer, Berlin, 2007.

[139] W. H. Zurek. Decoherence, einselection, and the quantum origins of the classical. *Reviews of Modern Physics*, 75:715–775, 2003.

[140] W. H. Zurek. Decoherence and the transition from quantum to classical – revisited. *Los Alamos Science*, 27:2–25, 2002.

[141] R. Omnès. Logical reformulation of quantum mechanics. i. Foundations. *Journal of Statistical Physics*, 53:893–932, 1988.

[142] R. Omnès. Logical reformulation of quantum mechanics. ii. Interferences and the Einstein–Podolsky–Rosen experiment. *Journal of Statistical Physics*, 53:933–955, 1988.

[143] R. Omnès. Logical reformulation of quantum mechanics. iii. Classical limit and irreversibility. *Journal of Statistical Physics*, 53:957–975, 1988.

[144] R. Omnès. Logical reformulation of quantum mechanics. iv. Projectors in semiclassical physics. *Journal of Statistical Physics*, 57:357–382, 1989.

[145] R. Omnès. Consistent interpretations of quantum mechanics. *Reviews of Modern Physics*, 64:329–382, 1992.

[146] R. B. Griffiths. Consistent histories and the interpretation of quantum mechanics. *Journal of Statistical Physics*, 36:212–272, 1984.

[147] F. Dowker and A. Kent. On the consistent histories approach to quantum mechanics. *Journal of Statistical Physics*, 82:1575–1646, 1996.

[148] R. Kastner. 'Einselection' of point observables: The new H-theorem?, 2014. arXiv:1406.4126.

[149] Y. Aharonov, D. Z. Albert, and L. Vaidman. How the result of a measurement of a component of the spin of a spin-$\frac{1}{2}$ particle can turn out to be 100. *Physical Review Letters*, 60:1351–1354, 1988.

[150] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan. The sense in which a 'weak measurement' of a spin-1/2 particle's spin component yields a value 100. *Physical Review D*, 40:2112–2117, 1989.

[151] A. Danan, D. Farfurnik, S. Bar-Ad, and L. Vaidman. Asking photons where they have been. *Physical Review Letters*, 111:240402, 2013.

[152] L. A. Castro, C. A. Brasil, and R. d. J. Napolitano. Weak values in collision theory. *Annals of Physics*, 392:272–286, 2018.

[153] C. Ferrie and J. Combes. How the result of a single coin toss can turn out to be 100 heads. *Physical Review Letters*, 113:120404, 2014.

[154] S. K. Berberian. *Notes on Spectral Theory*. D. Van Nostrand Company, Inc., Princeton, 1966.

[155] J. M. Jauch and C. Piron. Generalized localizability. *Helvetica Physica Acta*, 40:559, 1967.

[156] H. E. Brandt. Positive operator valued measure in quantum information processing. *American Journal of Physics*, 67:434–439, 1999.

[157] K.-E. Hellwig and K. Kraus. Pure operations and measurements. *Communications in Mathematical Physics*, 11:214–220, 1969.

[158] K.-E. Hellwig and K. Kraus. Operations and measurements ii. *Communications in Mathematical Physics*, 16:142–147, 1969.

[159] K. Kraus. *States, effects and operations – Fundamental notions of quantum theory*. Springer–Verlag, Berlin, 1983.

[160] K. Kraus. General state changes in quantum theory. *Annals of Physics*, 64:311–335, 1971.

[161] J. Maziero. A representação de Kraus para a dinâmica de sistemas quânticos abertos – The Kraus representation for the dynamics of open quantum systems. *Revista Brasileira de Ensino de Física*, 38:e2307, 2016.

[162] H.-P. Breuer and F. Petruccione. *The theory of open quantum systems*. Oxford University Press, Oxford, 2002.

[163] A. Caldeira. *An introduction to macroscopic quantum phenomena and quantum dissipation*. Cambridge University Press, Cambridge, 2014.

[164] F. Reif. *Fundamentals of statistical and thermal physics*. Waveland Press Inc, Long Grove, 1965.

[165] C. A. Brasil, F. F. Fanchini, and R. d. J. Napolitano. A simple derivation of the Lindblad equation. *Revista Brasileira de Ensino de Física*, 35:1303, 2013.

[166] G. Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48:119–130, 1976.

[167] S. L. Adler. Derivation of the Lindblad generator structure by use of the Itō stochastic calculus. *Physics Letters A*, 265:58–61, 2000.

[168] S. L. Adler. Corrigendum to: Derivation of the Lindblad generator structure by use of the Itō stochastic calculus. *Physics Letters A*, 267:212, 2000.

[169] A. Peres. Classical interventions in quantum systems. i. The measuring process. *Physical Review A*, 61:022116, 2000.

[170] S. Weinberg. What happens in a measurement? *Physical Review A*, 93:032124, 2016.

[171] J. Distler and S. Paban. von Neumann's formula, measurements and the Lindblad equation, 2017. arXiv:1702.01724v1.

[172] C. A. Brasil and R. d. J. Napolitano. The master equation for the reduced open-system dynamics, including a Lindbladian description of the finite-duration measurement. *European Physical Journal Plus*, 126:91, 2011.

[173] C. A. Brasil, L. A. Castro, and R. d. J. Napolitano. Protecting a quantum state from environmental noise by an incompatible finite-time measurement. *Physical Review A*, 84:022112, 2011.

[174] J. D. Cresser, S. M. Barnett, J. Jeffers, and D. T. Pegg. Measurement master equation. *Optics Communications*, 264:352–361, 2006.

[175] K. Jacobs and D. A. Steck. A straightforward introduction to continuous quantum measurement. *Contemporary Physics*, 47:279–303, 2006.

[176] D. N. Klyshko. Reduction of the wave function: an operational approach. *Physics Letters A*, 243:179–186, 1998.

[177] H. W. Wiseman and G. Milburn. *Quantum measurement and control*. Cambridge University Press, Cambridge, 2009.

[178] C. A. Brasil, L. A. Castro, and R. d. J. Napolitano. How much time does a measurement take? *Foundations of Physics*, 43:642–655, 2015.

[179] C. A. Brasil, L. A. Castro, and R. d. J. Napolitano. Efficient finite-time measurements under thermal regimes. *European Physical Journal Plus*, 129:206, 2014.

[180] B. Misra and E. C. G. Sudarshan. The Zeno's paradox in quantum theory. *Journal of Mathematical Physics*, 18:756, 1977.

[181] P. Facchi and S. Pascazio. Quantum Zeno phenomena: Pulsed versus continuous measurement. *Fortschritte der Physik*, 49:941, 2001.

[182] W. M. Itano, D. J. Heinzen, J. J. Bollinger, and D. J. Wineland. Quantum Zeno effect. *Physical Review A*, 41:2295, 1990.

[183] A. G. Kofman and G. Kurizki. Acceleration of quantum decay processes by frequent observations. *Nature*, 405:546–550, 2000.

[184] S. McArdle, S. Endo, A. Aspuru-Guzik, S. Benjamin, and X. Yuan. Quantum computational chemistry, 2019. arXiv:quant-ph/1808.10402v2.

[185] V. B. Braginsky and F. Y. Khalili. *Quantum measurement*. Cambridge University Press, Cambridge, 1992.

[186] P. Busch, P. J. Lahti, and P. Mittelstaedt. *The quantum theory of measurement*. Springer–Verlag, Berlin, 1996.

[187] K. Jacobs. *Quantum measurement theory and its applications*. Cambridge University Press, Cambridge, 2014.

[188] M. B. Mensky. *Continuous quantum measurements and path integrals*. IOP Publishing, London, 1993.

[189] M. B. Mensky. *Quantum measurements and decoherence – models and phenomenology*. Springer-Science+Business, Dordrecht, 2000.

[190] A. Bassi and G. C. Ghirardi. Dynamical reduction models. *Physics Reports*, 379:257–426, 2003.

[191] M. Namiki and S. Pascazio. Quantum theory of measurement based on the many-hilbert-space approach. *Physics Reports*, 232:201–411, 1993.

[192] P. Tombesi and D. F. Walls. *Quantum measurements in optics – Proceedings of a NATO Advanced Research Workshop on Quantum Measurements in Optics*. Springer Science+Business, New York, 1992.

[193] I. Percival. *Quantum state diffusion*. Cambridge University Press, Cambridge, 1998.

[194] G. Ghirardi, R. Rimini, and T. Weber. Unified dynamics for microscopic and macroscopic systems. 34:470–491, 1986.

[195] G. Ghirardi, P. Pearle, and A. Rimini. Markov processes in Hilbert space and continuous spontaneous localization of systems of identical particles. *Physical Review A*, 42:78–89, 1990.

[196] P. Pearle. Combining stochastic dynamical state-vector reduction with spontaneous localization. *Physical Review A*, 39:2277–2289, 1989.

[197] G. C. Wick, A. S. Wightman, and E. P. Wigner. The intrinsic parity of elementary particles. *Physical Review*, 88:101–105, 1952.

[198] G. C. Wick, A. S. Wightman, and E. P. Wigner. Superselection rule for charge. *Physical Review D*, 1:3267–3267, 1970.

[199] J. Bub. How to solve the measurement problem of quantum mechanics. *Foundations of Physics*, 18:701–722, 1988.

[200] N. P. Landsman. Observation and superselection in quantum mechanics, 1994. arXiv:hep-th/9411173.

[201] D. Giulini. Superselection rules, 2007. arXiv:0710.1516.

[202] G. J. Milburn and D. F. Walls. Quantum nondemolition measurements via quadratic coupling. *Physical Review A*, 28:2065–2070, 1983.

[203] D. F. Walls, M. J. Collet, and G. J. Milburn. Analysis of a quantum measurement. *Physical Review D*, 32:3208, 1985.

[204] P. Grangier, J. A. Levenson, and J.-P. Poizat. Quantum non-demolition measurements in optics. *Nature*, 396:537–542, 1998.

[205] O. P. Sá Neto, M. C. de Oliveira, and G. J. Milburn. Temperature measurement and phonon number statistics of a nanoelectromechanical resonator. *New Journal of Physics*, 17:093010, 2015.

[206] C. S. Unnikrishnan. Quantum non-demolition measurements: concepts, theory and practice. *Current Science*, 109:2052, 2015.

[207] V. B. Braginsky and F. Y. Khalili. Quantum nondemolition measurements. *Reviews of Modern Physics*, 68, 1996.

Leonardo Andreta de Castro studied physics at the São Carlos Institute of Physics (IFSC) of the University of São Paulo (USP), in Brazil, where he obtained his bachelor's degree, his master's degree, and his PhD. He has been researching quantum information theory since his undergraduate research project, and has also published papers on the theory of open quantum systems and foundations of quantum mechanics. He currently works as a postdoctoral associate at Duke University, developing research on the interface between quantum control and quantum error correction theories.

Olímpio Pereira de Sá Neto has a degree in Physics from the Federal University of Piauí (UFPI) (2007), a Master's (2009) and a PhD (2013) in Physics from the Instituto de Física 'Gleb Wataghin' (IFGW) of the State University of Campinas (UNICAMP). He is currently a professor at the State University of Piauí (UE-SPI), having worked in research involving theoretical quantum optics.

Carlos Alexandre Brasil is a theoretical physicist with a bachelor's degree (2005), a master's degree (2012), and doctoral degree (2012) awarded by IFSC/USP. He was a postdoctoral researcher at the IFGW/UNICAMP (2014) and at IFSC/USP (2018). He is currently an associate professor at the Universidade Tecnológica Federal do Paraná (UTFPR), in the Departamento Acadêmico de Ciências da Natureza (DACIN), Campus de Cornélio Procópio (CP). His interests include classical and quantum information theory, history and foundations of quantum mechanics, theory of quantum measurements and Bohm's causal formulation.