The Importance of Mathematics in Quantum Mechanics

Iván Mauricio Burbano Aldana
Universidad de los Andes

January 23, 2018

Abstract

Quantum mechanics is one of the most important theories in every physicist's toolbox. Nevertheless, in many situations the mathematical underpinnings of the theory are often disregarded as non important by many instructors. To show why this is not the case, we will briefly discuss arguably the most important question in quantum mechanics: "what is the probability that the measurement of the observable A on a system in the state ρ will yield a value in the measurable set $E \subseteq \mathbb{R}$?". We will see that it cannot be treated without knowledge of the mathematical structure of the theory. In the way we will be forced to consider at depth Hilbert spaces, measure theory, bounded and unbounded operators, spectrum, states, and the spectral theorem.

1 Introduction

It is not the purpose of this document to discuss the importance of mathematics in physics, to explain the mathematics involved in physical theories, or to discuss the properties of a good physical explanation. Instead, through the discussion of questions commonly found in the minds of physics students, we will find ourselves in the need of presenting (at least by name) various concepts in modern mathematics not usually found in physics curricula. As a disclaimer, we do not intend to develop a mathematical discussion. Instead, we intend to expose some of the ways these objects appear in the theory, hoping that the reader will gain a deeper appreciation for these, even if the do not want to take a look at the suggested references.

The discussion will be carried out assuming knowledge of quantum mechanics at a level of [1].

2 The Question in Quantum Mechanics

One of the most striking departures of quantum mechanics from classical theories is that the following question does not make sense in general: "what is

the value of the observable A?" Instead the question is replaced by: "what is the probability that a measurement of the observable A will yield a value in the subset $E \subseteq \mathbb{R}$?" Of course, the answer to the latter amounts to establishing what mathematicians call a probability measure[1] on the real numbers. This is a function μ which to every measurable subset $E \subseteq \mathbb{R}$ assigns a probability $\mu(E) \in [0, 1]$. Establishing what measurable means, that is, to which sets can we assign a probabilities, is also part of this task. Any discussion of quantum mechanics should provide an answer to this question. Therefore, a clear and complete understanding of quantum mechanics requires that we give a prescription to obtain probability measures on the real line from an observable and a state.

3   Preliminary Examples

Before we discuss the general procedure, let us examine the case of the position operator on a particle moving in one dimension in a pure state defined by a normalized wave function ψ. The probability measure is then given by

$$\mu(E) = \int_E \|\psi(x)\|^2 dx. \tag{1}$$

This means that the probability of finding the particle on the set E is given by (1). Let us comment this result. If one wants (1) to define a honest probability measure, we cannot use the notion of Riemann integration. We may inspire this through our physical intuition. In this system we know that (unless we have pathological examples such as "dirac functions" which are in any case not physical) the probability of finding the particle in a set with only one point $\{x\}$ is 0. We therefore conclude that the probability of finding a particle in $\mathbb{Q}$, the set of rational points, is 0 since the rational numbers can be listed as $q_1, q_2, \dots$ and therefore the probability should be the sum of the probabilities corresponding to $\{q_1\}, \{q_2\}, \dots$. However, attempting to integrate over the rational numbers is ill defined according to Riemann's theory. Indeed, if one tried to integrate (1) with $E = \mathbb{Q}$ and

$$\psi(x) = \chi_{[0,1]}(x) = \begin{cases} 1 & x \in [0, 1] \\ 0 & x \notin [0, 1] \end{cases} \tag{2}$$

the upper Riemann sum would yield 1 whilst the lower 0 for every partition of the real numbers. The theory we must consider is called Lebesgue integration. Moreover, Lebesgue integration is also necessary to define the notion of a state in this system. Indeed, the square integrable functions form a Hilbert space with the usual operations only under this integral. It is also important to note that the space of square integrable functions does not distinguish between two functions which differ on a set of size 0, that is a measurable set $E \subseteq \mathbb{R}$ where

---

[1]All of the new mathematical objects and notions that will appear in this section can be studied more throughly with the excellent book [2].

$\int_E 1 dx = 0$. Therefore a state as pathological as

$$\psi(x) = \chi_{[0,1]\cap\mathbb{Q}}(x) = \begin{cases} 1 & x \in [0,1] \cap \mathbb{Q} \\ 0 & x \notin [0,1] \cap \mathbb{Q} \end{cases} \tag{3}$$

is equivalent to $\psi(x) = 0$ in quantum mechanics.

It is also instructive to consider the case of a stationary spin-1/2 particle. In here, the spin along the third direction operator S_3 has a matrix representation in the canonical basis of $\mathbb{C}^2$

$$\frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{4}$$

In here the probability measure is given for a pure normalized state described by the vector $\psi = (\psi_1, \psi_2) \in \mathbb{C}^2$ by

$$\mu(E) = \begin{cases} 0 & \frac{\hbar}{2}, -\frac{\hbar}{2} \notin E \\ \|\psi_1\|^2 & \frac{\hbar}{2} \in E, \ -\frac{\hbar}{2} \notin E \\ \|\psi_2\|^2 & -\frac{\hbar}{2} \in E, \ \frac{\hbar}{2} \notin E \\ 1 & \frac{\hbar}{2}, -\frac{\hbar}{2} \in E. \end{cases} \tag{5}$$

Therefore the only possible values for the measurement are $\pm\frac{\hbar}{2}$, that is, the eigenvalues of S_3. This can be seen from the fact that $\mu(\{\frac{\hbar}{2}\}) = \|\psi_1\|^2$, $\mu(\{-\frac{\hbar}{2}\}) = \|\psi_2\|^2$, and $1 = \|\psi\| = \|\psi_1\|^2 + \|\psi_2\|^2$.

Trying to connect both examples leads to the wrong statement that the only possible values that the measurement of an observable could yield are its eigenvalues. This is wrong because the position operator x which acts on wave functions through the prescription $x\psi(y) = y\psi(y)$ has no eigenvalues in the non trivial square integrable functions. Indeed, if $\psi \neq 0$ was an eigenvector with eigenvalue x_0 then for all y we have $x_0\psi(y) = y\psi(y)$. If $\psi(y) \neq 0$ for some $y \neq x_0$ then we have $y = x_0$, which is a contradiction. Therefore $\psi(y) = 0$ for all $y \neq x_0$. But then ψ differs from 0 only on the set $\{x_0\}$ of measure 0. Therefore, with respect to square integrable functions $\psi = 0$ is trivial, which is also a contradiction. To make a correct statement about the connection between both of our examples we must introduce the notion of spectra which generalizes the concept of eigenvalues. In the finite dimensional case, such as the spin-1/2 system, both notions coincide. In general we will have that the probability measure induced by an operator and a state on the real numbers will only give positive probabilities for sets which intersect the spectra of the operator.

4   States

We must also discuss the notion of a state before we delve into the prescription. Thus far we have discussed pure states, which in our first example were the normalized elements of the set of square integrable functions $L^2(\mathbb{R})$ and in the second the normalized elements of $\mathbb{C}^2$. These are examples of Hilbert spaces.

Although inner product spaces are a common subject in Linear Algebra discussions and are part of the usual physics curriculum, Hilbert spaces have a very special quality which distinguishes them from the rest of inner product spaces. This is the topological notion of completeness. The fact that they are complete allows us to treat infinite dimensional cases such as $L^2(\mathbb{R})$ with generalizations of well known theorems from the finite dimensional environment. Now, pure states are not the most general notion of state we can have. In general a state should be an object which contains the information of the system we need to answer the question asked at the beginning of this section. Of course, our question can be rearranged as "what is the expectation value of the observable corresponding to the measuring the observable A and obtaining a value in the measurable set $E \subseteq \mathbb{R}$?". The observable we are now considering is very special because it has only two answers, either you measure a value in E or you do not. We may associate to those outcomes 1 and 0. The only observables with the spectrum $\{0, 1\}$ are the orthogonal projections. This are very especial because they can be defined on all of the Hilbert space $\mathcal{H}$ (the existence of square integrable functions which are not differentiable shows that this is not possible for the momentum operator) and are bounded, that is, the least upper bound

$$\sup\{\|A\psi\| \text{ for all } \psi \in \mathcal{H} \text{ such that } \|\psi\| = 1\} \tag{6}$$

is finite. For bounded operators A we call (6) the operator norm $\|A\|$. Therefore states should be objects which assign to every projection their expectation value. As it turns out, for every assignation of expectation values to bounded operators there exists a bounded, self-adjoint, non-negative operator ρ of unit trace, called density operators, such that the expectation value of the bounded operator A is $\mathrm{tr}(A\rho)$. Mind that the notion of self-adjointness and trace are used without care here. Self-adjointness in the case of unbounded operators requires more careful definition than the one of Linear Algebra because they may not be defined on all of the Hilbert space. On the other hand, defining the trace of a bounded operator requires that the Hilbert space is separable, another topological notion. Forgetting this warnings for the moment, we may conclude that states can be identified with density operators ρ. The case of a pure state $\psi \in \mathcal{H}$ may be identified with the density operator ρ_ψ corresponding to the orthogonal projection onto the subspace spanned by ψ. More explicitly

$$\rho_\psi(\phi) = \langle \psi, \phi \rangle \psi. \tag{7}$$

A state which is not pure is called a mixed state.

5 The Answer: The Spectral Theorem

We are finally ready to answer our question. Remember that we want to obtain a probability measure on the real numbers from an observable A and a state ρ on a separable Hilbert space $\mathcal{H}$. For this we need a generalization of the spectral theorem from the case of finite dimensional inner product spaces. Instead of

decomposing and operator with respect to the projections onto its eigenspaces as a sum, we will now decompose them as an integral with respect to a projection valued measure

$$A = \int_{\mathbb{R}} \lambda dP_A(\lambda). \tag{8}$$

Understanding the actual meaning of this formula is well above the scope of this text. The important thing to note is that for every self-adjoint operator A we can construct a function P_A which to every measurable set $E \subseteq \mathbb{R}$ assigns an orthogonal projection $P_A(E)$. The answer to the question "what is the probability that the measurement of the observable A on a system in the state ρ will yield a value in the measurable set $E \subseteq \mathbb{R}$?" is then simply $\mu_\rho^A(E) := \text{tr}(P_A(E)\rho)$. Incidentally, we may very well answer the question "what is the expected value of a bounded observable A on a system in the state ρ?" by $\text{tr}(A\rho)$. This of course familiar in the case of pure state ρ_ψ since $\text{tr}(A\rho_\psi) = \langle \psi, A\psi \rangle$. Although this seems like a way to avoid the spectral theorem, we would still need an axiom which states that the observable corresponding to measuring an observable A and obtaining a value in the measurable set $E \subseteq \mathbb{R}$ is $P_A(E)$. The construction of such a function P_A won't be discussed. Instead, we will take our previous examples and explicitly show the function P_A.

In the case of the particle moving in one dimension, we have $P_x(E)\psi = \chi_E \psi$ where the characteristic function χ_E of the measurable set E is defined by

$$\chi_E(y) = \begin{cases} 1 & y \in E \\ 0 & y \notin E. \end{cases} \tag{9}$$

Note that for a pure state ρ_ψ we have

$$\mu_{\rho_\psi}^x(E) = \text{tr}(P_x(E)\rho_\psi) = \langle \psi, P_x(E)\psi \rangle = \langle \psi, \chi_E \psi \rangle$$
$$= \int_{\mathbb{R}} \psi^*(y)\chi_E(y)\psi(y)dy = \int_E \|\psi(y)\|^2 dy \tag{10}$$

agreeing with (1). On the other hand, for the case of the spin-1/2 system, we have $\psi = (\psi_1, \psi_2) \in \mathbb{C}^2$

$$P_{S_3}(E)\psi = \begin{cases} 0 & \frac{\hbar}{2}, -\frac{\hbar}{2} \notin E \\ \langle(1,0),\psi\rangle(1,0) = (\psi_1, 0) & \frac{\hbar}{2} \in E, \ -\frac{\hbar}{2} \notin E \\ \langle(0,1),\psi\rangle(0,1) = (0,\psi_2) & -\frac{\hbar}{2} \in E, \ \frac{\hbar}{2} \notin E \\ \psi & \frac{\hbar}{2}, -\frac{\hbar}{2} \in E. \end{cases} \tag{11}$$

One can then check that for a pure state ρ_ψ

$$\mu_{\rho_\psi}^{S_3}(E) = \text{tr}(P_{S_3}(E)\rho_\psi) = \langle \psi, P_{S_3}(E)\psi \rangle$$
$$= \begin{cases} \langle \psi, 0 \rangle = 0 & \frac{\hbar}{2}, -\frac{\hbar}{2} \notin E \\ \langle \psi, (\psi_1, 0) \rangle = \|\psi_1\|^2 & \frac{\hbar}{2} \in E, \ -\frac{\hbar}{2} \notin E \\ \langle \psi, (0, \psi_2) \rangle = \|\psi_2\|^2 & -\frac{\hbar}{2} \in E, \ \frac{\hbar}{2} \notin E \\ \langle \psi, \psi \rangle = \|\psi\|^2 = 1 & \frac{\hbar}{2}, -\frac{\hbar}{2} \in E \end{cases} \tag{12}$$

agreeing with (5). In this last case (8) reduces to the usual spectral theorem

$$S_3 = \int_{\mathbb{R}} \lambda dP_{S_3}(\lambda) = \frac{\hbar}{2}\left(P_{S_3}\left(\left\{\frac{\hbar}{2}\right\}\right) - P_{S_3}\left(\left\{-\frac{\hbar}{2}\right\}\right)\right) \tag{13}$$

Note that the spectral theorem has other uses in quantum mechanics. A favorite is that of defining functions of operators. It is common to observe definitions such as

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} \tag{14}$$

for an observable A. This is okay to do as long we stick to bounded operators. This is because for this definition to make sense, we need a way to evaluate the convergence of such a series. In the case of bounded operators, the operator norm provides a criteria for convergence in exactly the same manner that the absolute value allows us to understand convergence in the real numbers. However, in the case of unbounded operators, this is no longer the case. Still, in the case of unbounded operators, we may define such an expression by[2]

$$f(A) = \int_{\mathbb{R}} f(\lambda) dP_A(\lambda) \tag{15}$$

for a measurable function $f : \mathbb{R} \to \mathbb{C}$. Therefore the spectral theorem allows us to define such operators as the time evolution e^{iHt}.

References

[1] N. Zettili, *Quantum Mechanics: concepts and applications.* John Wiley & Sons, Inc., 2 ed., 2009.

[2] B. C. Hall, *Quantum Theory for Mathematicians.* Springer, 2013.

[3] P. A. M. Dirac, *Principles of quantum mechanics.* Oxford University Press, 4 ed., 1958.

---

[2]Compare with the definition given by Dirac in [3].