Mathematics of Quantum Theory

Iván Mauricio Burbano Aldana

January 11, 2018

1 Axioms of Quantum Theory

Let us start by studying the axioms of Quantum Mechanics which resemble the most undergraduate approaches to the theory. Understanding these axioms will be an excuse to introduce various mathematical concepts which are often overlooked in elementary discussions. We follow for this [1].

Axiom 1. *We can assign to every physical system a separable Hilbert space.*

We assume the reader is aware of what an inner product space is. In particular every inner product space V with inner product $\langle \cdot, \cdot \rangle$ is a normed space with norm

$$
\begin{aligned}
\| \cdot \| : V &\to [0, \infty) \\
v &\mapsto \|v\| := \sqrt{\langle v, v \rangle}.
\end{aligned}
\tag{1}
$$

Norms are useful because they determine the topology of the space. How does this happen? Although it is not necessary (for now) to delve into point-set topology, the topology of a normed space is given by its metric structure. A metric space is a set X equipped with a metric

$$
d : X \times X \to [0, \infty),
\tag{2}
$$

thats is, a function which satisfies that for all $x, y, z \in X$:

1. $d(x, y) = 0$ if and only if $x = y$;

2. (symmetric) $d(x, y) = d(y, x)$;

3. (triangle inequality) $d(x, z) \leq d(x, y) + d(y, z)$.

In particular, a normed space V has a natural metric

$$
\begin{aligned}
V \times V &\to [0, \infty) \\
(u, v) &\mapsto \|u - v\|.
\end{aligned}
\tag{3}
$$

In a metric space we may define what a limit of a sequence is. Consider a sequence (x_n) on a metric space X. We say that $x_n \to x$ or

$$
\lim_{n \to \infty} x_n = x
\tag{4}
$$

if for every $\epsilon \in (0, \infty)$ there exists an $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}^{>N} := \{m \in \mathbb{N} | m > n\}$ we have that $d(x, x_n) < \epsilon$. A closely related but different notion is that of a Cauchy sequence. We say that a sequence (x_n) in a metric space is Cauchy if for all $\epsilon \in (0, \infty)$ there exists an $N \in \mathbb{N}$ such that for all $n, m \in \mathbb{N}^{>N}$ we have that $d(x_n, x_m) < \epsilon$. Notice that we have avoided giving an "intuitive" explanation to both the definition of limit and the definition of Cauchy sequence. This is done to avoid misjudgments due to lack of rigor. Indeed, both definitions look very similar and one would expect that a sequence converges if and only if it is Cauchy. Indeed, it is true that every convergent sequence is Cauchy. Let us prove it to get a feeling of how analysis proofs usually go. Suppose (x_n) converges to x. Let $\epsilon \in (0, \infty)$. Then there exists an $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}^{>N}$ we have that $d(x, x_n) < \epsilon/2$. Then, for all $n, m \in \mathbb{N}^{>N}$ we have by the triangle inequality $d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \epsilon/2 + \epsilon/2 = \epsilon$. On the other hand, it turns out that not every Cauchy sequence converges. For example the sequence (x_n) with

$$x_n = \left(1 + \frac{1}{n}\right)^n \tag{5}$$

in the metric space $\mathbb{Q}$ (notice that $\mathbb{Q}$ is a vector space over itself and has a natural metric induced by its norm, the absolute value). This sequence is Cauchy but its limit if we considered the sequence in $\mathbb{R}$ is e which is not in $\mathbb{Q}$. A metric space in which every Cauchy sequence converges is called complete. As it turns out, every metric space X can be completed to a complete space $\overline{X}$ much like $\mathbb{Q}$ is completed by $\mathbb{R}$.

Now we are ready to define a Hilbert space. A Hilbert space is a complex inner product space which is complete. We are left with the problem of separability. Like completeness, separability is a topological property. Still, we don't have to go as far as to point-set topology to understand it. Take a subset A of a metric space X. We say x is a limit point of A if for all $\epsilon \in (0, \infty)$ the ball $B(x, \epsilon) := \{y \in X | d(x, y) < \epsilon\}$ of radius ϵ around x intersects A. If a set A contains all of its limit points it is said to be closed. On the other hand, we can take the closure $\overline{A}$ of an arbitrary subset $A \subseteq X$ by taking the union of A with all of its limit points. A metric space X is said to be separable if it contains a countable dense subset, that is, a countable subset D whose closure is $\overline{D} = X$. The axiom of separability is important for the theory since a lot of the Hilbert spaces appearing in quantum mechanics are infinite dimensional. The axiom of separability assures that the space is no too big, in the sense that there exists a countable complete orthonormal set, that is, a countable set $\{\psi_n | n \in \mathbb{N}\}$ such that for all $n, m \in \mathbb{N}$ we have that $\langle \psi_n, \psi_m \rangle = \delta_{nm}$ and the only vector which orthogonal to ψ_n for all $n \in \mathbb{N}$ is 0. The existence of such a set is important because then for every vector ψ we have the Fourier series

$$\psi = \sum_{n=0}^{\infty} \langle \psi_n, \psi \rangle \psi_n. \tag{6}$$

Notice that this doesn't make $\{\psi_n | n \in \mathbb{N}\}$ a basis since we are allowing for

2

infinite linear combinations[2].

Example 1. *The square integrable functions (up to differences on a set of measure 0) $L^2(\mathbb{R})$ on $\mathbb{R}$ form a Hilbert space. Mind that this is only true if we consider integrability with respect to the Lebesgue measure[1], not the usual Riemann-Stieltjes integral. This fact is often overlooked since both integrals coincide on continuous functions. Nevertheless, if we attempt to prove that $L^2(\mathbb{R})$ is indeed complete we will soon find out that Lebesgue's theory is necessary. One can proof that the stationary states of the harmonic oscillator $|n\rangle$ (the Hermite polynomials) are a countable orthonormal complete set in $L^2(\mathbb{R})$. This space is usually assigned to a particle moving in one dimension.*

As a rule of thumb, the Hilbert space $L^2(\Omega)$ is assigned to a particle moving in a measurable subset $\Omega \subseteq \mathbb{R}^n$ in which only the translational degrees of freedom are of importance, $L^2(S^2)$ is assigned to a rotating particle (say a stationary molecule), and $\mathbb{C}^{2j+1}$ is assigned to a spin-j system. If a system can be decomposed into two subsystems described by the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ we describe the whole system by the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$[2]. For example, a spin-1/2 particle moving in $\mathbb{R}^3$ would be described by the Hilbert space $L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$.

Axiom 2. *The observables of a physical system described by a Hilbert space $\mathcal{H}$ are described by densely defined self-adjoint operators on $\mathcal{H}$.*

In elementary linear algebra an operator is a linear transformation from a vector space into itself. Nevertheless, in quantum mechanics we need to allow for operators to be defined only on a subset of the Hilbert space. For example, thinking about momentum, there are square integrable functions which are not differentiable. Therefore operators will in general be linear mappings $T : \mathcal{D}_T \subseteq \mathcal{H} \to \mathcal{H}$ and to define operators we must also define their domain. This is particularly pressing when talking about self-adjointness. In elementary linear algebra an operator T is said to be self-adjoint if for all vectors v and u we have $\langle u, Tv \rangle = \langle Tu, v \rangle$. We might attempt the same definition in our setting but restricting ourselves to the vectors in the domain of the operators. Operators $T : \mathcal{D}_T \subseteq \mathcal{H} \to \mathcal{H}$ such that for all $\psi, \phi \in \mathcal{D}_T$ we have $\langle \psi, T\phi \rangle = \langle T\psi, \phi \rangle$ are called symmetric. Often in elementary treatments this notion is mixed with that of self-adjointness. We will now define the correct notion. For a densely defined operator $T : \mathcal{D}_T \subseteq \mathcal{H} \to \mathcal{H}$ we define $\mathcal{D}_{T^*} \subseteq \mathcal{H}$ to be the set of vectors $\psi \in \mathcal{H}$ such that for all $\phi \in \mathcal{D}_T$ there exists and $\eta \in \mathcal{H}$ such that $\langle \psi, T\phi \rangle = \langle \eta, \phi \rangle$. We define $T^*\psi = \eta$ for all $\psi \in \mathcal{D}_{T^*}$. Notice that T^* is well defined because we assume that the domain of T is dense. We say an operator is self-adjoint

---

[1]Measure theory is of crucial importance to the theory of quantum mechanics. An excellent account of the subject may be found in [2]. We will from now on use the notion of measurable set without defining it. In simple terms a measurable set in the real numbers is one in which the notion of size makes sense. For most purposes it is enough to know that open and closed sets are measurable in $\mathbb{R}$.

[2]The tensor product of Hilbert spaces is slightly different from that of vector spaces since we need to define a suitable inner product and to ensure that the resulting structure is complete[1].

if it is equal (including domain-wise) to its adjoint. Therefore the matter of self-adjointness becomes highly non-trivial since we need to define the domain of the operator carefully.

Axiom 3. *The state of a physical system described by a Hilbert space $\mathcal{H}$ is a positive self-adjoint trace-class normalized operator $\rho : \mathcal{H} \to \mathcal{H}$.*

First note that states will be defined on all of the Hilbert space. An operator $\rho : \mathcal{H} \to \mathcal{H}$ is said to be positive if for all $\psi \in \mathcal{H}$ we have that $\langle \psi, \rho\psi \rangle \geq 0$. It is said to be trace-class on a separable Hilbert space if for a countable complete orthonormal set $\{\psi_n \in \mathcal{H} | n \in \mathbb{N}\}$ the series

$$\sum_{n=0}^{\infty} \langle \psi_n, \rho\psi_n \rangle \tag{7}$$

converges. If ρ is trace-class, we define its trace $\mathrm{tr}(\rho)$ by equation (7). We say it is normalized if $\mathrm{tr}(\rho) = 1$. An important example of these operators are the pure states. These are the orthogonal projections onto the subspace spanned by $\psi \in \mathcal{H}$

$$\rho_\psi : \mathcal{H} \to \mathcal{H}$$
$$\phi \mapsto \frac{\langle \psi, \phi \rangle}{\langle \psi, \psi \rangle}. \tag{8}$$

Axiom 4. *The probability that the measurement of an observable A on a system in the state ρ yields a result in a measurable set $E \subseteq \mathbb{R}$ is given by*

$$\mathrm{tr}(P_A(E)\rho) \tag{9}$$

where P_A is the spectral decomposition of A.

We have now reached a point where the mathematics involved become highly non-trivial. We will just limit ourselves to saying that P_A takes measurable sets E into projections $P_A(E)$ in a way similar to how measures take measurable sets to real numbers. P_A is usually called a projection-valued measure. As in measure theory, we can define an integral with respect to P_A and it will satisfy

$$A = \int_{\mathbb{R}} \lambda dP_A(\lambda). \tag{10}$$

The existence of such a P_A is guaranteed by the spectral theorem for self-adjoint operators. More information may be found in [1]. Note that even though the mathematics involved are quite difficult, the spectral theorem is an absolute must in every discussion of quantum mechanics. Without it we cannot even describe axiom 4 which answers the most fundamental question in quantum mechanics linking the mathematical structure to experiment.

4

Example 2. *Take a particle moving in one dimension. This is described by the Hilbert space $L^2(\mathbb{R})$. The position operator is*

$$x : \mathcal{D}_x \to L^2(\mathbb{R}) \tag{11}$$

where $x\psi(y) = y\psi(y)$. As it turns out, the position operator has the spectral decomposition $P_x(E)(\psi) = \chi_E \psi$ and therefore the probability that the particle is found in the measurable set E when it is in a pure state ρ_ψ with $\|\psi\| = 1$ is

$$\mathrm{tr}(P_x(E)\rho_\psi) = \langle \psi, P_x(E)\psi \rangle = \int_{\mathbb{R}} \psi^* \chi_E \psi \, dx = \int_E \|\psi\| dx. \tag{12}$$

In here χ_E is the characteristic function of the set E.

Example 3. *Consider a spin-1/2 system. It is described by the Hilbert space $\mathbb{C}^2$. The spin in the third direction is described by the operator $S_3 = \frac{\hbar}{2}\sigma_3$ where in the canonical basis*

$$[\sigma_3]_{can} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{13}$$

The spectral decomposition of S_3 is the projection $P_{S_3}(E)$ onto the span of $(1,0)$ if E contains $\hbar/2$ but not $-\hbar/2$, onto the span of $(0,1)$ if E contains $-\hbar/2$ but not $\hbar/2$, $\mathrm{id}_{\mathbb{C}^2}$ if E contains both, and 0 if it contains none. Therefore the probability that a particle has spin in the third direction with value on the set E on a pure state ρ_ψ with $\|\psi\| = 1$ is

$$\begin{aligned}
\mathrm{tr}(P_{S_3}(E)\rho_\psi) &= \langle \psi, P_{S_3}(E)\psi \rangle \\
&= \begin{cases}
\langle \psi, \langle (1,0), \psi \rangle (1,0) \rangle = \|\langle (1,0), \psi \rangle\|^2 & \frac{\hbar}{2} \in E \subset \mathbb{R} \setminus \left\{ -\frac{\hbar}{2} \right\} \\
\langle \psi, \langle (0,1), \psi \rangle (0,1) \rangle = \|\langle (0,1), \psi \rangle\|^2 & -\frac{\hbar}{2} \in E \subset \mathbb{R} \setminus \left\{ \frac{\hbar}{2} \right\} \\
\langle \psi, \psi \rangle = 1 & \frac{\hbar}{2}, -\frac{\hbar}{2} \in E \\
\langle \psi, 0 \rangle = 0 & \frac{\hbar}{2}, -\frac{\hbar}{2} \notin E.
\end{cases}
\end{aligned} \tag{14}$$

References

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

[2] E. Hewitt and K. Stromberg, *Real and Abstract Analysis.* New York: Springer-Verlag, 1975.