Interpreting the Measurement Axiom

Pedro Brunialti Lima de Andrade

August 27, 2021

Abstract

In a typical introductory quantum mechanics course we say that an observable is a self-adjoint operator A on some Hilbert space $\mathcal{H}$, and that the possible measurements for A are its eigenvalues, moreover for a state $\psi = \sum \alpha^i e_i$, where e_i is a normalized eigenvector associated to λ_i, the probability of measuring λ_i is $|\alpha^i|^2$. Note that this is extremely restrictive on the types of possible self-adjoint operators: those that have an orthonormal basis of eingenvectors.

The Measurement Axiom, however, states that for a state $\psi \in \mathcal{H}$ and a Borel set $E \subset \mathbb{R}$, the probability of a measurement of A of a system in the state ψ to lie in E is given by $\mu_\psi^A(E)$, where μ_ψ^A is the measure induced by the unique projected value measure P associated with A. In this text we show how the latter is merely a generalization of the first. In particular, when $\dim V < \infty$ the two are completely equivalent.

1 The Spectral Theorem

Before we talk about the Measurement Axiom, we first recall what exactly the spectral theorem says.

Theorem 1.1 (Spectral Theorem). *Let $A : \mathcal{D}_A \to \mathcal{H}$ be a self-adjoint operator on a Hilbert space $\mathcal{H}$, then there exists a unique projection valued measure*

$$P_A : \mathscr{B}(\mathbb{R}) \longrightarrow \mathcal{L}(\mathcal{H})$$

such that $\int_\mathbb{R} \mathrm{id}_\mathbb{R} dP = A$. Furthermore, the spectrum of A has total measure i.e. $P(\sigma(A)) = \mathrm{id}_\mathcal{H}$.

Now we briefly discuss everything we need to actually understand what this theorem is stating. We need to know two things: what is a projection valued measure and what is the meaning of the integration of a measurable function $f : \mathbb{R} \to \mathbb{C}$ with respect to it.

Definition 1.2. Let $\mathscr{B}(\mathbb{R})$ be the Borel sigma algebra of $\mathbb{R}$ and $\mathcal{L}(\mathcal{H})$ the space of bounded linear operators from $\mathcal{H}$ to itself. A projection valued measure on $\mathcal{H}$ is a function

$$P : \mathscr{B}(\mathbb{R}) \longrightarrow \mathcal{L}(\mathcal{H})$$

such that the following conditions hold:

(i) $P(E) \circ P(E) = P(E)$;

(ii) $P(E)^* = P(E)$;

(iii) $P(\mathbb{R}) = \mathrm{id}_{\mathcal{H}}$;

(iv) for any disjoint union $\bigcup E_n$, $E_n \in \mathscr{B}(\mathbb{R})$ and $\psi \in \mathcal{H}$, $P(\bigcup E_n)\psi = (\sum P(E_n))\psi$.

Conditions (i) and (ii) are equivalent to saying that $P(E)$ is an orthogonal projection, thus the name projection valued measure. We will abbreviate "projection valued measure" by "PVM".

Now we proceed to integration using a PVM. For a measurable function $f : \mathbb{R} \to \mathbb{C}$, its integral with respect to P will be an operator

$$\int_{\mathbb{R}} f dP : \mathcal{D}_f \to \mathcal{H},\ \mathcal{D}_f \subset \mathcal{H}$$

as we want $\int \mathrm{id}_{\mathbb{R}} dP$ to be a self-adjoint operator. As we do for regular integration of functions, we start by looking at the simple functions.

Let $f = \sum_{i=1}^{n} r_i 1_{E_i}$, where 1_{E_i} are characteristic functions of Borel sets and $r_i \in \mathbb{C}$, then we define

$$\int_{\mathbb{R}} f dP := \sum_{i=1}^{n} r_i P(E_i). \tag{1}$$

For our goals, this will be enough, but we may extend this definition to any measurable function, and this is necessary to deal with more general operators, such as the position and momentum operators. Just for the sake of completeness we state the following:

Proposition 1.3. *For any measurable function $f : \mathbb{R} \to \mathbb{C}$ there exists a nice way to define a, possibly unbounded, operator $\int_{\mathbb{R}} f dP : \mathcal{D}_f \to \mathcal{H}$. If f is simple, then its integral is given by Equation 1. Furthermore, if $f(\mathbb{R}) \subset \mathbb{R}$, then $\int_{\mathbb{R}} f dP$ is self-adjoint.*

Remark 1.4. By the above result, the integral of a real valued function is self-adjoint, then in the Spectral Theorem it makes sense that $\int \mathrm{id}_{\mathbb{R}} dP$ is self-adjoint.

Note that for any $\psi \in \mathcal{H}$, we may define a finite measure

$$\mu_\psi : \mathscr{B}(\mathbb{R}) \longrightarrow \mathbb{R}$$

$$\mu_\psi(E) := \langle \psi, P(E)\psi \rangle. \tag{2}$$

Since $P(E)$ is an orthogonal projection, we can infer that $\mu_\psi(\mathbb{R}) = ||\psi||$, in particular when $||\psi|| = 1$, μ_ψ is a probability measure.

Let $f = \sum_{i=1}^{n} r_i 1_{E_i}$ be a simple function, then

$$\langle \psi, \int_{\mathbb{R}} f dP \psi \rangle = \sum r_i \langle \psi, P(E_i)\psi \rangle = \sum r_i \mu_\psi(E_i)$$

but, by definition of regular integrals, $\sum r_i \mu_\psi(E_i) = \int_{\mathbb{R}} f d\mu_\psi$. This holds for any measurable function, that is

$$\langle \psi, \int_{\mathbb{R}} f dP \psi \rangle = \int_{\mathbb{R}} f d\mu_\psi \tag{3}$$

for any ψ in the domain of $\int_{\mathbb{R}} f dP$. This serves as an intuition of why to call it the "integral" of the function with respect to P.

We are now ready to understand the statement of the Spectral Theorem. Let A be a self-adjoint operator, the the theorem tells us that there exists a unique PVM $P_A : \mathscr{B}(\mathbb{R}) \to \mathcal{L}(\mathcal{H})$ such that $\int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} dP_A = A$. By Remark 1.4, we also know that for any PVM P, $\int_P \mathrm{id}_{\mathbb{R}} dP$ is self adjoint, thus the Spectral Theorem gives a bijection

$$\{\text{Self-Adjoint Operators}\} \leftrightsquigarrow \{\text{PVM's}\}.$$

As we shall see below, this is extremely useful for the axiomatization of Quantum Mechanics since it literally tells us how to pair an observable and a state to get a probability measure that predicts the result of measurements of said observable.

Now we interpret what is the meaning of $\sigma(A)$ (the spectrum of A) having total measure in P_A. From $P_A(\sigma(A)) = \mathrm{id}_{\mathbb{R}}$, we may compute $P_A(\sigma(A)^c)$:

$$\psi = P_A(\mathbb{R})\psi = (P(\sigma(A)) + P(\sigma(A)^c))\psi = \mathrm{id}_{\mathcal{H}}\psi + P(\sigma(A)^c) = \psi + P(\sigma(A)^c)\psi$$

so that $P(\sigma(A)^c)\psi = 0$ for every ψ, *i.e.* $P(\sigma(A)^c) = 0$. From this, we conclude that

$$\int_{\sigma(A)} f dP_A := \int_{\mathbb{R}} f \cdot 1_{\sigma(A)} dP_A = \int_{\mathbb{R}} f dP_A \tag{4}$$

since for simple functions components contained in $\sigma(A)^c$ will amount to nothing (see Equation 1). From this we also conclude that

$$\mu_\psi(\sigma(A)^c) = 0$$

so that for any function $f : \mathbb{R} \to \mathbb{C}$

$$\int_{\mathbb{R}} f d\mu_\psi^A = \int_{\sigma(A)} f d\mu_\psi^A.$$

where μ_ψ^A is the measure obtained by P_A as in 2.

2 The Finite Dimension Spectral Theorem

So, the Spectral Theorem mentioned above seems like a radical departure from the good old Spectral Theorem in finite dimension, namely [1]:

Theorem 2.1. *Let V be a finite dimension complex vector space and $T : V \to V$ be a self-adjoint operator and let $\sigma(T) = \{\lambda_1, \dots, \lambda_n\} \subset \mathbb{R}$ be set of eigenvalues of T. Define $\pi_i : V \to V$ to be the orthogonal projection onto* $\mathrm{Ker}(T - \lambda_i)$*. Then*

$$T = \sum_{i=1}^{n} \lambda_i \pi_i.$$

[1]There are several different statements to Spectral Theorem, probably the most common one being that there exists an orthogonal basis of eigenvectors, but they are easily checked to be equivalent.

Okay, so now we must show that this is equivalent to our previous one. First of all, note that in finite dimension, the spectrum of T coincides with its eigenvalues, since a linear map $S : V \to V$ is surjective if and only if it is injective.

Let's define a projection valued measure for T. To do this, we must define a projection $P(E)$ for each Borel set $E \subset \mathscr{B}(\mathbb{R})$. By a discussion in the last section, we know that $P(\mathbb{R} \setminus \sigma(T))$ must be zero if P is a PVM, thus we need only to restrict ourselves to the action of P in $\sigma(T)$. Define for $\{\lambda_i\} \in \sigma(T)$

$$P(\{\lambda_i\}) := \pi_i. \tag{5}$$

This uniquely defines a function $P : \mathscr{B}(\mathbb{R}) \to \mathcal{L}(V)$, such that $P(\mathbb{R} \setminus \sigma(T)) = 0$ and that disjoint unions are taken to the sum of the operators. This function can easily be checked to be a PVM. It remains to show that $\int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} \, dP = T$.

By Equation 4,

$$\int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} \cdot 1_{\sigma(T)} dP = \int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} dP.$$

Observe that $\mathrm{id}_{\mathbb{R}} \cdot 1_{\sigma(T)} = \sum_{i=1}^{n} \lambda_i 1_{\{\lambda_i\}}$, so

$$\int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} \cdot 1_{\sigma(T)} dP = \int_{\mathbb{R}} \sum_{i=1}^{n} \lambda_i 1_{\{\lambda_i\}} dP = \sum_{i=1}^{n} \lambda_i P(\{\lambda_i\}) = \sum_{i=1}^{n} \lambda_i \pi_i = T$$

where in the last equality we are using the finite dimensional Spectral Theorem.

To see that the general Spectral Theorem implies this one we invite the reader to compute $\int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} dP_T$ where P_T is the unique PVM associated with the self-adjoint operator T.

3 The Measurement Axiom in Finite Dimension

The Measurement Axiom states the following:

Axiom 3.1. Let $\mathcal{H}$ be the Hilbert space associated with some physical system, and let A be an observable of said system with PVM P_A. Then, if the system is in the state ψ, the probability of the measure of A to lie is some measurable set $E \in \mathscr{B}(\mathbb{R})$ is given by

$$\mu_\psi^A(E) := \langle \psi, P_A(E)\psi \rangle.$$

Now we show how this is the same as saying that we measure the eigenvalues of A for finite dimensions.

Recall that the probability of measuring anything outside of the spectrum of A is zero for any state, since if $E \subset \mathbb{R} \setminus \sigma(A)$, then $P_A(E) = 0$. In finite dimensions, the spectrum of A is simply its set of eigenvalues, therefore the chance of measuring anything besides them is 0: we measure eigenvectors of A.

Since we are in finite dimension, any state ψ may be written as

$$\psi = \sum \alpha^i e_i, \quad \sum |\alpha^i|^2 = 1$$

where e_i is an orthogonal basis of eigenvectors. Let λ_i be the eigenvalue associated with e_i, and for simplicity assume that all eigenvectors of A are distinct. So, by the discussion of previous section, the PVM associated with A is given by

$$P_A(\{\lambda_i\}) = \pi_i$$

which in this case is the projection onto e_i.

Now let's compute the probability of getting λ_i when measuring A of a system in the state ψ using the Measurement Axiom:

$$\mu_\psi^A(\{\lambda_i\}) = \langle \psi, P_A(\{\lambda_i\})\psi \rangle = \langle \sum \alpha^j e_j, \alpha^i e_i \rangle = |\alpha^i|^2 \langle \psi, \psi \rangle = |\alpha^i|^2.$$

This is exactly what is taught to us in an introductory Quantum Mechanics course.

One important, and quite trivial, corollary is that if a state is an eigenvector, then we will measure its eigenvalue with 100% of certainty ($\alpha^j = e^{i\theta}$).

4 Expected Values

Definition 4.1. In Quantum Mechanics, we define the expected value of an observable A for a system in the state ψ to be

$$\langle A \rangle_\psi := \langle \psi, A\psi \rangle.$$

Here we will show that this is in fact the probabilistic definition of an expected value and give further intuition of why the self-adjoint operator is recovered by $\int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} dP_A$.

First let's see the purely probabilistic definition of the expected value of some measurement. Let $X \subset \mathbb{R}$ be a finite set, and suppose that each x_i has probability p_i of being measured and that we only measure elements in X ($\sum p_i = 1$). An example of this is to take a dice: $X = \{1,2,3,4,5,6\}$ and each of them has probability $1/6$. The expected value of a measurement of X is given by

$$\langle X \rangle := \sum_i x_i p_i.$$

This is simply the weighted average, for instance, for a dice, this number is $3,5 = (1+6)/2$, this means that if keep tossing dices and averaging the sum of the results this value will approach $3,5$.

If A is a self-adjoint operator in a finite dimension vector space, for a state $\psi = \sum \alpha^i e_i$, we have a finite set

$$\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$$

such that each element λ_i has a probability $|\alpha^i|^2$. Thus the expected value of a measurement of X is

$$\sum \lambda^i |\alpha^i|^2 = \langle \psi, A\psi \rangle = \langle A \rangle_\psi.$$

So it does make sense to call $\langle A \rangle_\psi$ the expected value of A.

Continuing with this, note that saying that each element x_i has a probability p_i is the same as defining a probability measure

$$\mu : \mathscr{P}(X) \to [0,1]$$

where $\mathscr{P}(X)$ is the power set of X. Using this, we see that, by definition of the Lebesgue integral,

$$\langle X\rangle = \int_X \mathrm{id}_X d\mu.$$

With this in mind, a straightforward generalization of the expected value for a subset $X \subset \mathbb{R}$ and a probability measure $\mu : \mathcal{A}(X) \to [0, 1]$, where $\mathcal{A}(X)$ is a sigma algebra for A, is

$$\langle X\rangle = \int_X \mathrm{id}_X d\mu.$$

We showed that the Spectral Theorem gives for each observable A and state ψ a measure

$$\mu_\psi^A : \sigma(A) \longrightarrow [0, 1]$$

thus, the expected value of $\sigma(A)$ with said measure is

$$\int_{\sigma(A)} \mathrm{id}_{\sigma(A)} d\mu_\psi^A$$

but by equations 3 and 4 we have, for ψ satisfying certain domain conditions, that

$$\int_{\sigma(A)} \mathrm{id}_{\sigma(A)} d\mu_\psi^A = \langle \psi, \int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} dP_A \psi\rangle$$

but the Spectral Theorem states that $\int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} dP_A = A$, so we conclude that

$$\int_{\sigma(A)} \mathrm{id}_{\sigma(A)} d\mu_\psi^A = \langle \psi, \int_{\mathbb{R}} \mathrm{id}_{\mathbb{R}} dP_A \psi\rangle = \langle \psi, A\psi\rangle = \langle A\rangle_\psi.$$

5 What We Measure

Lastly, I want to talk about what it means for a number to be in the spectrum of some observable. For this section, fix A to be an observable.

As we saw, if λ has an eigenvector ψ, then the chance of measuring λ if the system is in the state ψ is 100%. That is,

$$\mu_\psi^A(\{\lambda\}) = 1 = \langle \psi, P_A(\{\lambda\})\psi\rangle$$

this implies that $\psi \in \operatorname{im} P_A(\{\lambda\})$ and that $P_A(\{\lambda\})$ is the orthogonal projection onto $\ker(A - \lambda)$, in particular $P_A(\{\lambda\}) \neq 0$. This is not at all surprising, given that we are told that "we measure eigenvalues". What is interesting though, is the fact that $P_A(\{\lambda\}) \neq 0$ if and only λ is an eigenvalue! The proof is not difficult, but we omit it here. This discussion can be summarized in the following proposition:

Proposition 5.1. *Let $\lambda \in \mathbb{R}$, then $P_A(\{\lambda\}) \neq 0$ if and only if λ is an eigenvector.*

What about the rest of the spectrum? The following result will tell us how to interpret it:

Proposition 5.2. *Let $\lambda \in \mathbb{R}$, then $\lambda \in \sigma(A)$ if and only if for every $\epsilon > 0$*

$$P_A((\lambda - \epsilon, \lambda + \epsilon)) \neq 0.$$

If $\lambda \in \sigma(A)$ is not an eigenvalue, then there is no state ψ such that whenever we measure a system in ψ we obtain λ since $P_A(\{\lambda\}) = 0$, however for any given $\epsilon > 0$ we can always find a state such that whenever we measure the system in it we get that the outcome lies in $(\lambda - \epsilon, \lambda + \epsilon)$. In fact, let $\epsilon > 0$, then

$$P_A((\lambda - \epsilon, \lambda + \epsilon)) \neq 0,$$

so that there exists $\psi \in \operatorname{im} P_A((\lambda - \epsilon, \lambda + \epsilon))$ with $\|\psi\| = 1$ (a state). Since $P_A((\lambda - \epsilon, \lambda + \epsilon))$ is a projection and ψ is in its image, $P_A((\lambda - \epsilon, \lambda + \epsilon))\psi = \psi$, thus the chance of when measure A in this state to get a value in $(\lambda - \epsilon, \lambda + \epsilon)$ is

$$\mu_\psi^A((\lambda - \epsilon, \lambda + \epsilon)) = \langle \psi, P_A((\lambda - \epsilon, \lambda + \epsilon))\psi \rangle = \langle \psi, \psi \rangle = \|\psi\|^2 = 1$$

as we claimed.

References

[Gri60] David J Griffiths. *Introduction to quantum mechanics*. Pearson International Edition (Pearson Prentice Hall, Upper Saddle River, 2005), 1960.

[Hal13] Brian C Hall. *Quantum theory for mathematicians*, volume 267. Springer, 2013.

[Tes09] Gerald Teschl. *Mathematical methods in quantum mechanics*, volume 99. American Mathematical Society, 2009.