Spring 5-20-2016

Elements of the Mathematical Formulation of Quantum Mechanics

Keunjae Go
Washington University in Saint Louis

Elements of the Mathematical Formulation of Quantum States

Keunjae (Harry) Go

Advisors: Renato Feres and Xiang Tang

Washington University in St. Louis

April 1, 2016

Abstract

In this paper, we will explore some of the basic elements of the mathematical formulation of quantum mechanics. In the first section, I will list the motivations for introducing a probability model that is quite different from that of the classical probability theory, but still shares quite a few significant commonalities. Later in the paper, I will discuss the quantum probability theory in detail, while paying a brief attention to some of the axioms (by Birkhoff and von Neumann) that illustrate both the commonalities and differences between classical mechanics and quantum mechanics. This paper will end with a presentation of two theorems that form the core of quantum mechanics.

1 Motivations

1.1 Overview

Given a physical system, we are usually interested in certain physical quantities that we can measure, called **observables**. Some of the examples of observables are mass, linear or angular momentum, and electrical charge. Observables are mathematically equivalent to random variables in classical probability theory.

In quantum mechanics, we are often interested in a system involving one or more particles, such as atoms and photons. In this case, the observables behave according to their probability distributions rather than possessing definite values. The **quantum state** of the system describes the values of the observables at a given time. Therefore, being able to represent the quantum state as a mathematical object is critical to the mathematical formulation of quantum mechanics.

As it turns out, the quantum probability theory bears a lot of resemblances to the classical probability theory. In quantum probability theory, the probability space is replaced by the Hilbert space, which allows us to perform lots of useful operations with the quantum states, including the inner product.

One of the important differences between the classical mechanics and the quantum mechanics arises from one of the postulates of the quantum theory which states that the observables do not commute in general. In general, evaluating a random variable corresponds to performing an experiment to determine the value of an observable. Thus, the commutative property would imply that the order in which the observables are measured does not affect the measurements. However, as it has been verified by countless experiments, this is not the case in quantum mechanics. We will see this later in this paper, as illustrated by the well-known Heisenberg Uncertainty Principle.

Thus, in order to study a quantum physical system, we need the framework of noncommutative probability theory.

For the purpose of this paper, we will consider a finite, n-level system.

1.2 Why Hilbert Space?

1.2.1 Time-dependent Schrödinger Equation

Recall that the time-dependent Schrödinger Equation is given as

$$\hbar i \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial x^2} + V(x)\psi$$

where ψ is a function of both x (position) and t (time).

We note that this equation is the core of the quantum mechanical description of a moving particle.

In regards to the time-dependent Schrödinger equation, Max Born came up with two hypotheses:

1. Each solution ψ satisfying certain boundary conditions is called a state of the particle, and the values of the physical quantities associated are determined by the states.

2. Each state ψ associates the position of the particle with a probability measure M^ψ on the real numbers such that: if the particle is in state ψ at time t, then the probability that a measurement for the position of the particle at time t will take a value in real interval [a, b] is given by

$$M_t^\psi[a,b] = \int_a^b |\psi(x,t)|^2 dx$$

In particular, since M^ψ is a probability measure, we know that the above integral takes a value in [0, 1]. This implies that the integral above, when evaluated instead from $-\infty$ to ∞, yields 1.

The above integral also implies that ψ is a square-integrable function. Thus, we naturally think of $L^2(R)$, the space of all real-valued square-integrable functions. This space is, in fact, a Hilbert space.

Among physicists, $\psi(x,t)$ is also known as a **wavefunction**.

Assume for simplicity that $\psi(x,t)$ can be brokwn down as a product of two functions, as $\psi(x,t) = f(t)\gamma(x)$

If we substitute this into the Schrdinger Equation and manipulate it in a smart way, we get $H\gamma = E\gamma$, where H is an operator on Hilbert space $L^2(R)$ and E is an eigenvalue associated with γ.

1.2.2 Quantum States

Recall that a quantum state describes the given quantum system. The first thing to note here is that a quantum state can either be a pure state or a mixture of pure states. Before we perform a measurement on the given quantum system, we only have the knowledge of a probability distribution for the actual state of the system, and only after the measurement does the system collapse into one particular pure state. A pure state can be represented by a **state vector** in a Hilbert space over complex numbers. The state vector is always normalised to 1. A mixed state is, in fact, given by the distribution of pure states. It is described by a **density matrix**. An n-level system has n pure states, which are, mathematically, state vectors with unit norm in an n-dimensional Hilbert space over complex numbers, denoted H_n. We call this the **phase space** of the system. A **state** is simply a point in the phase space. In fact, the state vectors form an orthonormal basis for H_n. Physicists use the ket notation to describe these orthonormal state vectors, originally introduced by Paul Dirac. Using the

ket notation, we can write any state of the given quantum system as

$$\sum_{k=1}^{n} \alpha_k |s_k\rangle = \alpha_1|s_1\rangle + \alpha_2|s_2\rangle + ... + \alpha_n|s_n\rangle$$

where α_k are complex numbers satisfying $\sum_{k=1}^{n} |\alpha_k|^2 = 1$. We call these α_k the **amplitudes**

Example: The spin of an electron have two possible values. Each of the two possible spins represents a pure state, whereas by a mixed state we can think of a system of one electron that has a probability of having either of the spins. Formally, $s = \sum_{k=1}^{n} p_k s_k$ satisfying $\sum_{k=1}^{n} p_k = 1$, where the s_k's are the pure states and p_k is the probability that the system would collapse into the state s_k upon measurement.

2 Classical Probability Theory

When discussing quantum mechanics, it's useful to first discuss the probability theory, since everything that happens in the quantum world is probabilistic.

As we have seen earlier, in quantum mechanics Recall that, in classical probability theory, we can describe the outcomes of any real-world process using a probability space. Any probility space consists of a sample space Ω, a σ-algebra F, and a probability function P, denoted as (Ω, F,μ) where:

A sample space Ω is the set of all possible outcomes for the process.

A σ-algebra F is a set containing subsets of Ω, called events.

A probability measure μ assigns to an event a real number $\in$ [0, 1]

Also recall that a σ-algebra F has following properties:

1. Both ϕ and Ω are in F
2. If A $\in$ F, $A^c = \Omega - A \in$ F
3. If A_1, A_2, A_3, ... are in F, then so is $A = A_1 \cup A_2 \cup A_3 \cup ...$

Recall that μ is a real-valued function defined by μ: F$\rightarrow$[0, 1] having the following properties:

1. $\mu(\phi) = 0$
2. $\mu(\Omega) = 1$
3. $\mu(A) \geq 0$ for all A $\in$ F

3 Quantum Probability Theory

3.1 Hilbert Space

When we switch from classical to quantum probability, it is useful to consider a Hilbert space as our sample space because of some of the interesting properties of Hilbert spaces. First, recall the following definitions:

Definition: V is called an **inner product space** if it has the following properties:

1. V is a vector space over complex numbers
2. The $<,>$ operator associates a complex number with every pair of vectors in V satisfying for all x, y, z $\in$ V and $\lambda \in C$:
   (a) $\langle x, y\rangle = \langle y, x\rangle^*$ where * denotes complex conjugate
   (b) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ and $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$
   (c) $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$
   (d) $\langle x, x\rangle \in R$ and $\langle x, x\rangle \geq 0$, with equality if and only if x is the zero vector in V

Definition: An inner product space V is said to be **complete** if and only if every Cauchy sequence in V converges in norm to some vector in V

The properties of inner product space are particularly useful in quantum mechanics, because we can utilise them to define the density matrix as follows.

Consider an orthonormal basis of a Hilbert space H, which we denote as $u_1, ..., u_n$, and real numbers $p_1, ..., p_n \in [0, 1]$ then the density matrix is given by $\sum_{i=1}^{n} p_i |u_i\rangle\langle u_i|$ where $\langle u_i|$ is called the **bra vector**, which is dual to the ket vector $|u_i\rangle$. In fact, $\langle u_i|$ is simply the complex conjugate transpose of $|u_i\rangle$. The bra and ket vectors have a property that their product equals to 1.

The density matrix serves to represent the quantum system in a mixed state.

Recall that, earlier in this paper, we saw that any state of the given quantum system can be written as

$$\sum_{k=1}^{n} \alpha_k |s_k\rangle = \alpha_1|s_1\rangle + \alpha_2|s_2\rangle + ... + \alpha_n|s_n\rangle$$

Now, instead, we consider an nxn matrix, which we also call a **density operator**. The density operator ρ has the following properties:

1. Projector: $\rho^2 = \rho$
2. Hermiticity: $\rho^\dagger = \rho$

3. Normalisation: $Tr\rho = 1$

4. Positivity: $\rho \geq 0$

Where the trace operator is defined by $Tr\rho = \sum_n \langle n|\rho|n\rangle$

One of the advantages of using this operator is demonstrated in the ease with which we can calculate the expected value of an observable in a state represented by the density matrix. This is illustrated in the following theorem.

Theorem 1 *The expectated value of an observable A in a state represented by a density matrix ρ is given by*

$$\langle A\rangle_\rho = Tr(\rho A) \tag{1}$$

Proof: $Tr(\rho A) = Tr(|\psi\rangle\langle\psi|A) = \sum_n \langle n|\psi\rangle\langle\psi|A|n\rangle = \sum_n \langle\psi|A|n\rangle\langle n|\psi\rangle = \langle\psi|A|\psi\rangle = \langle A\rangle$

3.2 Mathematical definition of some important terms

Now we define a few more important terms.

1. A **phase state** is as defined earlier. In quantum mechanics, this is also a space of the wavefunctions, and is a Hilbert space.

2. An **observable** is a self-adjoint (i.e.) operator on Hilbert space

3. A **state** is equivalent to the probability measure

4. An **event** is an orthogonal projection operation P; according to Birkhoff and von Neumann, the quantum event structure is the algebra of closed subspaces of a Hilbert space.

5. A pair of **compatible events** are events that can occur in the same experiment

4 Axioms

There are some important axioms based on the set theory that are shared in common by the classical mechanics and quantum mechanics, as first laid out by Garrett Birkhoff and John von Neumann. Let x, y, z, p be events, then:

1. $x \leq x$

2. $If x \leq y$ and $y \leq z$ then $x \leq z$

3. $If x \leq y$ and $y \leq x$ then $x = y$

4. $0 \leq x \leq 1$

5. $x \cap y \leq x$, and $x \cap y \leq y$, and if $z \leq x$ and $z \leq y$ then $z \leq x \cap y$
6. $x \leq x \cup y$, and $y \leq x \cup y$, and if $x \leq z$ and $y \leq z$ then $x \cup y \leq z$
7. $(x^{\perp})^{\perp} = x$
8. $x \cap x^{\perp} = 0$ and $x \cup x^{\perp} = 1$
9. $x \leq y$ implies $y^{\perp} \leq x$
10. $If x \leq y$ then $y = x \cup (y \cap x^{\perp})$

All these axioms are true both in classical and quantum mechanics, except the last one does not apply in quantum mechanics. In classical mechanics, this axiom is pretty easily shown by using the distributive property, but this also implies that the distributive property need not hold in quantum mechanics, related to the noncommutative nature of quantum observables.

Here is another list of axioms shared by the clasical and quantum mechanics, with the exception of the last one. These axioms are more specific to the physical context.

1. Atomism: if $x < y$, then there is an atom p such that $p \leq y$ and $p > x$

   Remark: an **atom** is a nonzero $p \in L$ such that $x \leq p$ implies x = 0 or x = p
2. Covering property: for all atoms p and all elements x, if $x \cap p = 0$ then $x \leq y \leq x \cup p$ implies $y = x$ or $y = x \cup p$
3. Completeness: if $S \subset L$ then $\cup_{a \in S(a)}$ and $\cap_{a \in S(a)}$ exist
4. Irreducibility: If z satisfies $x = (x \cap z) \cup (x \cap z^{\perp})$ for all $x \in L$, then z = 0 or z = 1

The last axiom listed here is actually specific to the quantum mechanics, because the distributive property does not generally hold. In the classical mechanics, we don't have the restriction on z because the equality would hold for all x and z.

Assume that x and z are atoms. The last axiom is equivalent to the statement that x and z are incompatible.

5 Important Results of the Mathematical Formulation of Quantum Mechanics

Now I will present some of the very important theorems that result from this mathematical formulation discussed thus far.

Definition: two observables A and B are said to be **commutative** if AB = BA

Definition: the **commutator** of A and B is a mapping $[A, B] = AB - BA$

Also recall that the variance of A is the expected value of observable $(A-\mu)^2$, where μ is the mean. Note that, if A is an observable (i.e. self-adjoint), then $(A - \mu)^2$ is also an observable, since it's also self-adjoint.

Theorem 2 (Uncertainty Principle) *If A, B are self-adjoint operators on H and μ is a pure state (rank 1 projection), then*

$$Var_\mu(A)Var_\mu(B) \geq \frac{1}{4}|Tr(\mu([A, B]))|^2 \tag{2}$$

Note: the expected value of A in state x is given as $\mu = E_x(A) = \langle x|Ax\rangle$, so we can write the right hand side of the inequality in the theorem as follows:

$$\frac{1}{4}|Tr(\mu([A, B]))|^2 = \frac{1}{4}|\langle x|[A, B]x\rangle|^2$$

A special example of this theorem is Heisenberg's Uncertainty Principle, where A and B correspond to the position and the momentum. often let A denote the position and B denote the momentum to conclude that there exists a lower bound on the product of the measurement errors of the position and the momentum. In other words, we cannot measure both the position and the momentum to the level of accuracy we want.

Heisenberg's Uncertainty Principle is an important example that illustrates two observables that are noncommutative. Since they're noncommutative, their commutator [A, B] = AB - BA is nonzero, and we have a positive lower bound for the product of their variances.

Theorem 3 (Gleason) *For each μ on a Hilbert space of dimension ≥ 3, and for any self-adjoint projection (i.e. an event) P, there exists an operator T with tr(T) = 1 such that $\mu(P) = tr(TP)$*

The implication of this theorem is that all observables for the system are associated with Hermitian operators on $L^2(R)$.

This theorem is helpful for us who study quantum mechanics, because this theorem allows us to assign these events to probabilites, simply put.

Gleason's theorem also offers a proof of Born's rule, which allows us to calculate the probability that a measurement on a quantum system will yield a given result.

6 References

1. Birkhoff, G. and von Neumann, J. The logic of quantum mechanics. Annals of Mathematics 37 , 823 (1936)

2. Pitowsky, I. Quantum mechanics as a theory of probability. (2008)

3. Cohen, D. An Introduction to Hilbert Space and Quantum Logic. New York: Springer-Verlag (1989)

4. Hirvensalo, M. Quantum Computing. Heidelberg: Springer-Verlag (2004)