AN INTRODUCTION TO HILBERT SPACES AND THE HEISENBERG UNCERTAINTY PRINCIPLE

YU XIAO

ABSTRACT. In this paper, we will focus on one particular application of Hilbert spaces in quantum mechanics. We start with simple definitions of certain kinds of vector spaces that lead up to Hilbert spaces and follow by defining the Hermitian operator. Relevant, basic postulates of quantum mechanics will be given. We prove the Heisenberg Uncertatinty Principle and finish with a brief account of Hilbert spaces' importance and wider applications.

Contents

1. Hilbert Spaces

Why are Hilbert spaces so important? Essentially, a Hilbert space is a space that allows geometry. Exact analogs of the Pythagorean Theorem and Parallelogram Law hold in Hilbert spaces. To satisfy those requirements, a Hilbert space naturally must have many restrictions. To define Hilbert spaces, we shall start with an inner product space as it is simply a vector space with the additional structure of inner product.

Definition 1.1. (**Inner Product**) Let E be a complex vector space. A mapping $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{C}$ is called an *inner product* in E if for any $x, y, z \in E$ and $\alpha, \beta \in \mathbb{C}$ the following conditions are all satisfied:

(1) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ (the bar denotes the complex conjugate);
(2) $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$;
(3) $\langle x, x \rangle \geq 0$, and $\langle x, x \rangle = 0$ implies $x = 0$.

A space with an inner product is called an *inner product space*. It is also called a *pre-Hilbert space*. Its alternate name clearly signifies its importance in defining Hilbert spaces.

Date: August 18, 2017.

Definition 1.2. (**Norm**) A function $x \to ||x||$ from a vector space E into $\mathbb{R}$ is a norm if it satisfies the following conditions:

(1) $||x|| = 0$ if and only if $x = 0$;
(2) $||\lambda x|| = |\lambda|||x||$ for every $x \in E$ and $\lambda \in \mathbb{C}$;
(3) $||x + y|| \leq ||x|| + ||y||$ for every $x, y \in E$.

A vector space with a *norm* is called a normed space. It turns out that every inner product space is also a normed space with the norm defined by

$$||x|| = \sqrt{\langle x, x \rangle}$$

The norm above $||x|| = \sqrt{\langle x, x \rangle}$ is well defined because condition (3) of Definition 1.1 clearly implies that $||x|| = 0$ if and only if $x = 0$. Moreover,

$$||\lambda x|| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\lambda \bar{\lambda} \langle x, x \rangle} = |\lambda|||x||.$$

Therefore, we only need to prove for condition (3) of Definition 1.2, which we will begin by proving the Schwarz's Inequality Theorem.

Theorem 1.3. ***(Schwarz's Inequality)*** *For any two elements x and y of an inner product space, we have*

$$|\langle x, y \rangle| \leq ||x||||y|| \tag{1.4}$$

The equality $|\langle x, y \rangle| = ||x||||y||$ holds if and only if x and y are linearly dependent.

Proof. If $y = 0$, then the theorem holds since both sides equal zero.
If $y \neq 0$, then

$$0 \leq \langle x + \alpha y, x + \alpha y \rangle = \langle x, x \rangle + \bar{\alpha} \langle x, y \rangle + \alpha \langle y, x \rangle + |\alpha|^2 \langle y, y \rangle \tag{1.5}$$

Now substitute $\alpha = -\langle x, y \rangle / \langle y, y \rangle$, then we have

$$0 \leq \langle x, x \rangle \langle y, y \rangle - |\langle x, y \rangle|^2$$

Clearly, $|\langle x, y \rangle| \leq ||x||||y||$, giving us the theorem of Schwarz's Inequality.
If x and y are linearly dependent, then $x = \alpha y$ for some $\alpha \in \mathbb{C}$. Therefore,

$$|\langle x, y \rangle| = |\langle x, \alpha x \rangle| = |\bar{\alpha}| \langle x, x \rangle = |\alpha|||x||||x|| = ||x||||\alpha x|| = ||x||||y||$$

Now, let x and y be such vectors that $\langle x, y \rangle = ||x||||y||$.
Equivalently, $\langle x, y \rangle \langle y, x \rangle = \langle x, x \rangle \langle y, y \rangle$.
We have

$$\begin{aligned} \Big\langle \big\langle y, y \big\rangle x - \big\langle x, y \big\rangle y, \big\langle y, y \big\rangle x - \big\langle x, y \big\rangle y \Big\rangle = &\big\langle y, y \big\rangle^2 \big\langle x, x \big\rangle - \big\langle y, y \big\rangle \big\langle y, x \big\rangle \big\langle x, y \big\rangle \\ &- \big\langle x, y \big\rangle \big\langle y, y \big\rangle \big\langle y, x \big\rangle + \big\langle x, y \big\rangle \big\langle y, x \big\rangle \big\langle y, y \big\rangle = 0 \end{aligned}$$

So $\big\langle y, y \big\rangle x - \big\langle x, y \big\rangle y = 0$, which means that x and y are linearly dependent, thus completing the proof. □

Using Theorem 1.3 proven above, we can obtain condition (3) of Definition 1.2 as a corollary.

Corollary 1.6. ***(Triangle Inequality)*** *For any two elements x and y of an inner product space,*

$$||x + y|| \leq ||x|| + ||y|| \tag{1.7}$$

Proof. Taking $\alpha = 1$ in Equation 1.5 and applying Schwarz's Inequality, we have

$$\begin{aligned} ||x+y||^2 &= \langle x+y, x+y\rangle = \langle x,x\rangle + 2\Re\langle x,y\rangle + \langle y,y\rangle \\ &\le \langle x,x\rangle + 2|\langle x,y\rangle| + \langle y,y\rangle \\ &\le ||x||^2 + 2||x||||y|| + ||y||^2 \\ &= (||x|| + ||y||)^2 \end{aligned}$$

Therefore,

$$||x+y|| \le ||x|| + ||y||$$

□

So a inner product space is also a normed space. Obtaining an inner product space is the important first step in the process of trying to define Hilbert spaces, partly because the Parallelogram Law already holds in an inner product space.

Theorem 1.8. ***(Parallelogram Law)*** *For any two elements x and y of an inner product space we have*

$$||x+y||^2 + ||x-y||^2 = 2(||x||^2 + ||y||^2) \tag{1.9}$$

Proof. We have

$$||x+y||^2 = \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle,$$

hence

$$||x+y||^2 = ||x||^2 + \langle x,y\rangle + \langle y,x\rangle + ||y||^2. \tag{1.10}$$

Now replace y by $-y$ in Equation 1.10, to obtain

$$||x-y||^2 = ||x||^2 - \langle x,y\rangle - \langle y,x\rangle + ||y||^2 \tag{1.11}$$

Adding Equations 1.10 and 1.11 together, we have the parallelogram law. □

Having seen the usefulness of inner product spaces, we shall also give a proposition about the continuity of the inner product.

Proposition 1.12. ***(Continuity of Inner Product)*** *If (x_n) and (y_n) are sequences of elements of the inner product space E such that $||x_n - x|| \to 0$ and $||y_n - y|| \to 0$ when $n \to \infty$ with $x, y \in E$, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$.*

Proof. Applying Schwarz's Inequality, we have

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x,y\rangle| &= |\langle x_n, y_n - y\rangle + \langle x_n - x, y\rangle| \\ &\le |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\le ||x_n||||y_n - y|| + ||x_n - x||||y||. \end{aligned}$$

Since $||x_n - x|| \to 0$ and $||y_n - y|| \to 0$ when $n \to \infty$, $|\langle x_n, y_n\rangle - \langle x,y\rangle| \to 0$.

So, $\langle x_n, y_n\rangle \to \langle x,y\rangle$ when $n \to \infty$. □

Of course, the continuity of inner products can also be extended to norms. Now that we have proven that an inner product space is also a normed space and given a proposition of the inner product, we can move on to the next step of defining Hilbert spaces: defining the completeness of a space.

Definition 1.13. (**Cauchy Sequence**) A sequence of vectors (x_n) in a normed space is called a *Cauchy* sequence if for every $\epsilon > 0$ there exists a number M such that $||x_m - x_n|| \le \epsilon$ for all $m, n > M$.

Definition 1.14. (**Banach Space**) A normed space E is called *complete* if every Cauchy sequence in E converges to an element of E. A complete normed space is called a *Banach space*.

Definition 1.15. (**Hilbert Space**) A complete inner product space is called a Hilbert space.

As the standard formulations of quantum mechanics require the Hilbert spaces used to be separable, we will here give the further definition of separable Hilbert spaces as well. A Hilbert space is separable if and only if it admits a countable orthonormal basis.

Definition 1.16. (**Orthogonal Vectors**) Two vectors x and y in an inner product space are called *orthogonal* if $\langle x, y\rangle$=0, denoted by $x \perp y$.

Definition 1.17. (**Countable Set**) A *countable set* is a set that has the same cardinality (number of elements) as some subbset of the set of all natural numbers.

Definition 1.18. (**Orthonormal Basis**) Let E be an inner product space. A basis S of E with $||x|| = 1$ for all $x \in S$ and $x \perp y$ for any two distinct elements of S is called an *orthonormal basis*.

Definition 1.19. (**Separable Space**) A Hilbert space is called *separable* if it admits a countable orthonormal basis.

Finite dimensional Hilbert spaces are separable, and orthogonal bases are used often when dealing with finite dimensional Hilbert spaces. It is only natural and sensible that we wish to still have them when working with infinite dimensions.

2. Hermitian Operator

Linear maps are functions between two vector spaces that preserve the operations of vector addition and scalar multiplication. An operator (or transformation) is generally a mapping that maps elements of a space to produce elements of the same space. Linear operators on a normed vector space are widely used to represent physical quantities, hence they are very important in applied mathematics and mathematical physics. The most important operators include differential, integral, and matrix operators. For now, we are interested in the bounded linear operators on Hilbert space, particularly adjoint and self-adjoint (Hermitian) operators.

Let us give some basic definitions that will be of use in the following definitions of adjoint and self-adjoint Hilbert operators and the relevant proofs.

Definition 2.1. (**Bounded Operator**) An operator A is called bounded if there is a number K such that $||Ax|| \leq K||x||$ for every x in the domain of A.

Definition 2.2. (**Linear Functional**) A linear functional is a linear map from a vector space to its field of scalars. For vector space V over field $\mathbb{C}$, a linear functional is a function $f : V \to \mathbb{C}$ that satisfies the following:

(1) $f(v + w) = f(v) + f(w)$ for all $v, w \in V$;
(2) $f(\alpha v) = \alpha f(v)$ for all $v \in V$ and $\alpha \in \mathbb{C}$

And based on the above definitions, we can prove our first simple and quite straightforward lemma, the continuity of bounded linear functionals.

Lemma 2.3. *Any bounded linear functional is continuous.*

Proof. For any bounded linear functional L on a Hilbert space H, there exists a number K such that $||L(x)|| \leq K||x||$ for any $x \in H$. Let $x \to 0$, then $||L(x)|| \leq K||x|| \to 0$. So L is continuous at 0. Let x be an arbitrary element of H and (x_n) be a sequence convergent to x. Then $(x_n - x + 0)$ converges to 0 and so we have $||L(x_n) - L(x)|| = ||L(x_n - x + 0) - L(0)|| \to 0$.

Thus, any bounded linear functional is continuous. □

We shall put aside the above lemma for now and come back for it later. As the theorems below utilize the ideas of orthogonal complement and convex sets, we of course need to list those definitions first and then prove the related theorems.

Definition 2.4. (**Orthogonal Complement**) Let S be a non-empty subset of a Hilbert space H. An element $x \in H$ is said to be orthogonal to S, denoted by $x \perp S$, if $\langle x, y\rangle = 0$ for every $y \in S$. The set of all elements of H orthogonal to S, denoted by $S^\perp$, is called the *orthogonal complement* of S.

Theorem 2.5. *For any subset S of a Hilbert space H, the set $S^\perp$ is a closed subspace of H.*

Proof. If $\alpha, \beta \in \mathbb{C}$ and $x, y \in S^\perp$, then for every $z \in S$,

$$\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle = 0.$$

Thus, $S^\perp$ is a subspace of H. We next need to prove that $S^\perp$ is closed.

Let $(x_n) \in S^\perp$ and $x_n \to x$ for some $x \in H$. From the continuity of the inner product (Proposition 1.12), we have for every $y \in S$,

$$\langle x, y\rangle = \langle \lim_{n\to\infty} x_n, y\rangle = \lim_{n\to\infty} \langle x_n, y\rangle = 0.$$

This shows that $x \in S^\perp$, and thus $S^\perp$ is closed. □

Definition 2.6. (**Convex Sets**) A set U in a vector space is called *convex* if for any $x, y \in U$ and $\alpha \in (0, 1)$, we have $\alpha x + (1 - \alpha)y \in U$.

Theorem 2.7. ***(Closest Point Property)*** *Let S be a closed convex subset of a Hilbert space H. For every point $x \in H$ there exists a unique point $y \in S$ such that*

$$||x - y|| = \inf_{z\in S} ||x - z||. \tag{2.8}$$

Proof. Let (y_n) be a sequence in S such that $\lim_{n\to\infty} ||x - y_n|| = \inf_{z\in S} ||x - z||$. Denote $d = \inf_{z\in S} ||x - z||$. For all $m, n \in \mathbb{N}$, since $\frac{1}{2}(y_m + y_n) \in S$, we have

$$||x - \frac{y_m + y_n}{2}|| \geq d.$$

By Parallelogram Law (Theorem 1.8), we obtain

$$\begin{aligned} ||y_m - y_n||^2 &= ||y_m - y_n||^2 + 4||x - \frac{y_m + y_n}{2}||^2 - 4||x - \frac{y_m + y_n}{2}||^2 \\ &= ||(x - y_m) - (x - y_n)||^2 + ||(x - y_m) + (x - y_n)||^2 - 4||x - \frac{y_m + y_n}{2}||^2 \\ &= 2(||x - y_m||^2 + ||x - y_n||^2) - 4||x - \frac{y_m + y_n}{2}||^2. \end{aligned}$$

Since $2(||x - y_m||^2 + ||x - y_n||^2) \to 4d^2$ as $m, n \to \infty$, and $||x - \frac{y_m+y_n}{2}||^2 \geq d^2$, we have $||y_m - y_n||^2 \to 0$. So (y_n) is a Cauchy sequence. Since H is complete and S

is closed, the limit $\lim_{n\to\infty} y_n = y$ exists and $y \in S$. From Proposition 1.12 the continuity of inner product and by extension norm, we obtain

$$||x - y|| = ||x - \lim_{n\to\infty} y_n|| = \lim_{n\to\infty} ||x - y_n|| = d.$$

So there exists a point y that satisfies Equation 2.8. All that remains is to prove its uniqueness. Suppose then that there is another point $y_1 \in S$ that satisfies the said equation. Then since $\frac{y+y_1}{2} \in S$, we have $||y - y_1||^2 = 4d^2 - 4||x - \frac{y_1+y_2}{2}||^2 \leq 0$, which can only happen if $y = y_1$. □

Note that a vector subspace is clearly a convex set by definition. So any theorem that applies to convex sets also apply to vector subspaces. Therefore, Theorem 2.7 can be used for the proofs of Theorem 2.9 and Theorem 2.12.

And now we can prove the Theorem of Orthogonal Projection.

Theorem 2.9. ***(Orthogonal Projection)*** *If S is a closed subspace of a Hilbert space H, then every element $x \in H$ has a unique decomposition in the form $x = y+z$ with $y \in S$ and $z \in S^\perp$. (Should S be a one-dimensional subspace and y_0 be a unit vector in S, then $y = \langle x, y_0\rangle y_0$.)*

Proof. If $x \in S$, then the obvious decomposition is $x = x + 0$. Suppose then that $x \notin S$. Let y be the unique point in S satisfying $||x - y|| = \inf_{w\in S} ||x - w||$, as in Theorem 2.7. We shall show that $x = y + (x - y)$ is the desired decomposition. If $w \in S$ and $\lambda \in \mathbb{C}$, then $y + \lambda w \in S$ and

$$||x - y||^2 \leq ||x - y - \lambda w||^2 = ||x - y||^2 - 2\Re\lambda\langle w, x - y\rangle + |\lambda|^2||w||^2.$$

Hence,

$$-2\Re\lambda\langle w, x - y\rangle + |\lambda|^2||w||^2 \geq 0.$$

If $\lambda > 0$, then dividing by λ and letting $\lambda \to 0$ gives

$$\Re\langle w, x - y\rangle \leq 0. \tag{2.10}$$

Similarly, replacing λ by $-i\lambda(\lambda > 0)$, dividing by λ, and letting $\lambda \to 0$ gives

$$\Im\langle w, x - y\rangle \leq 0. \tag{2.11}$$

Since $y \in S$ also implies $-y \in S$, the inequalities 2.10 and 2.11 also hold with $-w$ instead of w. Therefore $\langle w, x - y\rangle = 0$ for every $w \in S$, which means $x - y \in S^\perp$.

For uniqueness, note that if $x = y_1 + z_1$ for some $y_1 \in S$ and $z_1 \in S^\perp$, then $y - y_1 \in S$ and $z - z_1 \in S^\perp$. Since $y - y_1 = z_1 - z$, we must have $y - y_1 = z_1 - z = 0$.

Furthermore, if S is a one-dimensional subspace and y_0 is a unit vector in S, then intuitive geometric understanding of the orthogonal decomposition gives us $y = \langle x, y_0\rangle y_0$. The more rigorous proof is as follows.

If $y = \langle x, y_0\rangle y_0$ does not satisfy the equation $||x - y|| = \inf_{w\in S} ||x - w||$, then there exists some $\alpha \in C$ such that $||x - y|| > ||x - \alpha y_0||$. So for any $x \in H$, we obtain

$$\begin{aligned} 0 &< ||x - y||^2 - ||x - \alpha y_0||^2 \\ &= (||y_0||^2 - 2)|\langle x, y_0\rangle|^2 + 2\Re(\alpha\langle x, y_0\rangle) - |\alpha|^2||y_0||^2 \\ &\leq (||y_0||^2 - 2)|\langle x, y_0\rangle|^2 + 2|\alpha|||\langle x, y_0\rangle|| - |\alpha|^2||y_0||^2 \end{aligned}$$

So the determinant $\Delta = 4|\alpha|^2 - 4(||y_0||^2 - 2)(-|\alpha|^2||y_0||^2) = (2|\alpha|||y_0||^2 - 2|\alpha|)^2 < 0$. However, this is impossible. So the assumption is false and $y = \langle x, y_0\rangle y_0$ must satisfy the equation $||x - y|| = \inf_{w\in S} ||x - w||$. So in a one-dimensional subspace

S with a unit vector y_0, $y = \langle x, y_0\rangle$, $x = y + (x - y)$ gives the unique decomposition of x for any $x \in H$. □

We shall also give another quite straightforward theorem as a result of the just-proved Orthogonal Projection Theorem.

Theorem 2.12. *If S is a closed subspace of a Hilbert space H, then $S^{\perp\perp} = S$.*

Proof. If $x \in S$, then for every $z \in S^{\perp}$ we have $\langle x, z\rangle = 0$, which means $x \in S^{\perp\perp}$. Thus, $S \subset S^{\perp\perp}$. To prove $S^{\perp\perp} \subset S$ consider $x \in S^{\perp\perp}$. Since S is closed, by Theorem 2.9, $x = y + z$ for some $y \in S$ and $z \in S^{\perp}$. Since $S \subset S^{\perp\perp}$, we have $y \in S^{\perp\perp}$ and thus $z = x - y \in S^{\perp\perp}$ because $S^{\perp\perp}$ is a vector subspace. But $z \in S^{\perp}$, so we must have $z = 0$, which means $x = y \in S$. This shows that $S^{\perp\perp} \subset S$. So we have $S^{\perp\perp} = S$. □

The above Theorem of Orthogonal Projection (Theorem 2.9) is one important stepping stone for the proof of the Riesz Representation Theorem (Theorem 2.15). But we also need another stepping stone that has to do with null spaces.

Definition 2.13. (**Null Space**) Let E_1 and E_2 be vector spaces and L be a mapping from E_1 to E_2. The *null space* of L, denoted by $\mathcal{N}(L)$, is the set of all vectors $x \in \mathcal{D}(L)$ such that $L(x) = 0$. ($\mathcal{D}(L)$ denotes the domain of L.)

Since for $\alpha, \beta \in \mathbb{C}$, and $x, y \in \mathcal{N}(L)$, $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = 0$, the null space of L is a vector subspace.

As we now continue on to proving another lemma about bounded linear functionals, the previously put aside Lemma 2.3 will serve its use.

Lemma 2.14. *Let f be a bounded linear functional on an inner product space E. Then* $\dim \mathcal{N}(f)^{\perp} \leq 1$.

Proof. If $f = 0$, then $\mathcal{N}(f) = E$. So $\mathcal{N}(f)^{\perp} = \mathbf{0}$ and $\dim \mathcal{N}(f)^{\perp} = 0 \leq 1$.

When f is not zero, f is continuous by Lemma 2.3. Since f is continuous, for all $x \in E\backslash\mathcal{N}(L)$, there exists $\epsilon > 0$ such that $B(x, \epsilon) \subseteq E\backslash\mathcal{N}(L)$. So $E\backslash\mathcal{N}(L)$ is open, implying that $\mathcal{N}(f)$ is a closed subspace of E. Suppose that $\mathcal{N}(f)^{\perp} = \mathbf{0}$, then $\mathcal{N}(f)^{\perp\perp} = E$. By Theorem 2.12, $\mathcal{N}(f) = \mathcal{N}(f)^{\perp\perp} = E$, which is impossible. Therefore $\mathcal{N}(f)^{\perp}$ contains more than the zero vector. Let $x_1, x_2 \in \mathcal{N}(f)^{\perp}$ be non-zero vectors. Therefore, $x_1, x_2 \notin \mathcal{N}(f)$. By Definition 2.13, $f(x_1) \neq 0$ and $f(x_2) \neq 0$. So there exists a scalar $a \neq 0$ such that $f(x_1) + af(x_2) = 0$ or $f(x_1 + ax_2) = 0$. Thus $x_1 + ax_2 \in \mathcal{N}(f)$. However, since $x_1, x_2 \in \mathcal{N}(f)^{\perp}$, we must have $x_1 + ax_2 \in \mathcal{N}(f)^{\perp}$. This is only possible if $x_1 + ax_2 = 0$, which implies that x_1 and x_2 are linearly dependent since $a \neq 0$. So $\dim \mathcal{N}(f)^{\perp} = 1$.

Therefore, $\dim \mathcal{N}(f)^{\perp} \leq 1$. □

And now we finally have everything needed to prove the following Riesz Representation Theorem.

Theorem 2.15. *(**Riesz Representation Theorem**) Let f be a bounded linear functional on Hilbert Space H. There exists a unique $x_0 \in H$ such that $f(x) = \langle x, x_0\rangle$ for all $x \in H$.*

Proof. If $f(x) = 0$ for all $x \in H$. Then $x_0 = 0$ has the desired properties and the theorem obviously holds.

Assume then that f is a non-zero functional. By Lemma 2.14, $\dim \mathcal{N}(f)^\perp = 1$. Let z_0 be a unit vector in $\mathcal{N}(f)^\perp$. Then for every $x \in H$, we have

$$x = x - \langle x, z_0 \rangle z_0 + \langle x, z_0 \rangle z_0.$$

By Theorem 2.9, we have $\langle x, z_0 \rangle z_0 \in \mathcal{N}(f)^\perp$, hence $x - \langle x, z_0 \rangle z_0 \in \mathcal{N}(f)$, which means that

$$f(x - \langle x, z_0 \rangle z_0) = 0.$$

As a result,

$$f(x) = f(\langle x, z_0 \rangle z_0) = \langle x, z_0 \rangle f(z_0) = \langle x, \overline{f(z_0)} z_0 \rangle.$$

Therefore, if we put

$$x_0 = \overline{f(z_0)} z_0,$$

then $f(x) = \langle x, x_0 \rangle$ for all $x \in H$.

Suppose now that there is another point x_1 such that $f(x) = \langle x, x_1 \rangle$ for all $x \in H$. Then $\langle x, x_0 - x_1 \rangle = 0$ for all $x \in H$. Take $x = x_0 - x_1$, we have $\langle x_0 - x_1, x_0 - x_1 \rangle = 0$, which is only possible when $x_0 - x_1 = 0$. Therefore, such a x_0 must be unique. □

Let A be a bounded operator on a Hilbert space H such that $||Ax|| \leq K||x||$ for all $x \in H$. For every fixed $x_0 \in H$, the functional f defined on H by $f(x) = \langle Ax, x_0 \rangle$ is a bounded linear functional on H, since applying Schwarz's Inequality gives us $||f(x)|| \leq ||Ax|| ||x_0|| \leq K||x_0|| ||x||$ with $K||x_0||$ fixed for every fixed $x_0 \in H$. By the Riesz Representation Theorem (Theorem 2.15) above, there exists unique $y_0 \in H$ such that $f(x) = \langle x, y_0 \rangle$ for all $x \in H$. Thus, $\langle Ax, x_0 \rangle = \langle x, y_0 \rangle$ for all $x \in H$. If we denote by A^* the operator which to every $x_0 \in H$ assigns that unique y_0, then we have $\langle Ax, y \rangle = \langle x, A^* y \rangle$ for all $x, y \in H$.

Definition 2.16. (**Adjoint Operator**) Let A be a bounded operator on a Hilbert space H. The operation $A^* : H \to H$ defined by $\langle Ax, y \rangle = \langle x, A^* y \rangle$ for all $x, y \in E$ is called the *adjoint operator* of A.

Definition 2.17. (**Hermitian/Self-Adjoint Operator**) If $A = A^*$, that is $\langle Ax, y \rangle = \langle x, Ay \rangle$ for all $x, y \in H$, then A is called self-adjoint, or *Hermitian*.

Let us also define eigenvalues and eigenvectors in terms of operators.

Definition 2.18. (**Eigenvalue**) Let A be an operator on a complex vector space E. A complex number λ is called an *eigenvalue* of A if there is a non-zero vector $u \in E$ such that

$$Au = \lambda u \tag{2.19}$$

Every vector u satisfying Equation 2.19 is called an *eigenvector* of A corresponding to the eigenvalue λ. If E is a function space, then those eigenvectors are often called eigenfunctions.

3. The Five Postulates of Quantum Mechanics

Having defined all the necessary math involved, we now turn to the physics side of things. In general, the mechanics description of a physical system requires three things: (i) variables or observables, (ii) states, and (iii) equations of motion. In classical mechanics, observables refer to experimentally measurable quantities, such as position, momentum, and energy. State, which describe values of the observables at given times, is uniquely determined by appropriate physics laws and the initial state of the system. And equations of motion determine how the values of the observables change in time.

But classical mechanics breaks down at the level of atoms and molecules. Historically, the first indication of a breakdown occurs in the phenomenon of black body radiation, which essentially deals with electromagnetic radiation in a container in equilibrium with its surroundings, that is to say the thermodynamics of the exchange of energy between radiation and matter. Classical mechanics holds that energy can be transferred continuously. However, under such assumptions, a hydrogen atom would become very unstable and radiate energy over a continuous range of frequencies, which contradicts the experiments done and phenomenon observed. Failure of classical mechanics led to the rise of quantum mechanics.

In 1900, Max Planck first postulated that energy is absorbed and emitted in discrete quantities. The radiation of frequency ν can only exchange energy with matter in units of $h\nu$, where h is the *Planck constant* of numerical value

$$h = 2\pi\hbar = 6.625 \times 10^{-27}\ erg\ sec = 4.14 \times 10^{-21}\ MeV\ sec \tag{3.1}$$

and $\hbar$ is called the *universal constant.*

In this section, we will present some basic principles of quantum mechanics as postulates. No attempt will be made to justify or derive these postulates.

While classical mechanics identifies the *state* of a physical system with the values of certain *observables* of the system, quantum mechanics makes a clear distinction between states and observables. We begin with the first postulate concerning the *state* of a quantum system.

Postulate 3.2. (**State Vector**) The collection of all possible states of a given system in quantum mechanics corresponds to a separable Hilbert space over the complex number field. A state of the system is represented by a non-zero vector in the space, and every non-zero scalar multiple of a state vector represents the same state. Conversely, every non-zero vector in the Hilbert space and its non-zero scalar multiples represent the same physical state of the system. Usually, a state vector is denoted by $|\psi\rangle$.

For any two state vectors $|\phi\rangle$ and $|\psi\rangle$, the complex number $\langle\phi|\psi\rangle$ represents their inner product in space. This postulate also asserts that all physical properties of a given system are unchanged if multiplied by a non-zero scalar. We can remove this arbitrariness by imposing the normalizing condition $\langle\psi|\psi\rangle = 1$.

Postulate 3.3. (**Observable Operators and Their Values**)

(1) To every physical observable in quantum mechanics, there corresponds in the Hilbert space a linear Hermitian operator $\hat{A}$, which has an orthonormal basis of orthogonal eigenvectors ψ_n with the corresponding eigenvalues λ_n such that

$$\hat{A}\psi_n = \lambda_n\psi_n, n = 1, 2, 3, \cdots \tag{3.4}$$

Conversely, to each such operator in the Hilbert space there corresponds some physical observable.

(2) The only possible values of a physical observable are the various eigenvalues.

According to this postulate, a quantum observable is represented by a linear Hermitian operator on an infinite dimensional separable Hilbert space. Such operators are called *observable operators* and usually satisfy the rules of operators in Hilbert space. It should be pointed out that quantum observables are operators in Hilbert space, but an observable in classical mechanics is just a real function on the set of all possible system states in Euclidean space.

The most important difference between the rules of two observable operators and the rules for classical observables is that two classical observables, A and B always commute ($AB = BA$), but two observable operators $\hat{A}$ and $\hat{B}$ generally do not commute. The *commutator* of these operators is defined by

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{3.5}$$

In the case of position observable $\hat{x}$ and the corresponding momentum observable $\hat{p}$ ($\hat{p} = -i\hbar\,\partial/\partial x$), the commutator becomes the canonical commutation relation:

$$[\hat{x}, \hat{p}] = i\hbar.$$

The canonical commutation relation refers to the fundamental relation between canonical conjugate quantities, quantities related by definition such that one is the Fourier transform of the other. Other examples include time and frequency, Doppler and range, etc. The basic canonical commutation relation above $[\hat{x}, \hat{p}] = i\hbar$ follows immediately from

$$[\hat{x}, \hat{p}]\psi = (\hat{x}\hat{p} - \hat{p}\hat{x})\psi = -i\hbar\hat{x}\frac{\partial\psi}{\partial x} + i\hbar\hat{x}\frac{\partial\psi}{\partial x} + i\hbar\psi = i\hbar\psi.$$

Postulate 3.6. (**Correspondence Principle**) A quantum observable operator corresponding to a dynamical variable is obtained by replacing the canonical variable in classical mechanics by the corresponding quantum mechanical operator.

This postulate gives a method of formulating quantum mechanical operators from classical dynamical variables which are made of canonical conjugate variables. In general, any observable in classical mechanics is some well-behaved function of position and momentum $f(x, p)$ and is represented in quantum mechanics by the operator $f(\hat{x}, \hat{p})$.

However, we should also note that some observables in quantum mechanics have no analogues in classical mechanics, such as spin and isospin. This postulate also does not provide how to specify the quantum observables corresponding to the basic canonical variables. Luckily, for that we have Postulate 3.7.

Postulate 3.7. (**Quantization**) Every pair of canonically conjugate observable operators satisfies the following Heisenberg commutation relations:

$$[\hat{q}_m, \hat{q}_n] = 0 = [\hat{p}_m, \hat{p}_n], \tag{3.8}$$

$$[\hat{q}_m, \hat{p}_n] = i\hbar\hat{\delta}_{mn}, \tag{3.9}$$

where $\hat{q}_m$ is the observable operator corresponding to the generalized coordinates, and $\hat{p}_m$ is the momentum operator corresponding to the generalized momentum.

Postulate 3.10. (**Outcome of Quantum Measurement**) If an observable operator $\hat{A}$ has eigenbasis ψ_n with the corresponding eigenvalues λ_n, then the probability that the measurement will give the eigenvalue λ_n of $\hat{A}$ of the system in the normalized state $\psi(x)$ is

$$P(\lambda_n) = |\langle\psi_n|\psi\rangle|^2 \tag{3.11}$$

4. Heisenberg Uncertainty Principle

We have listed all the postulates needed for a sound, if limited quantum theory. According to Postulate 3.5, the only real values that can be predicted for an observable are the eigenvalues of the corresponding operator of a given physical system. Which one of the eigenvalues will be obtained depends on the particular form of a state vector of the system at the time of measurement. But the operations of measurement generally affect the system. Therefore it is impossible to predict a single measurement with absolute certainty. Naturally, we deal with this problem by considering a large number of measurements under the same conditions, and use the average of measurements obtained.

Definition 4.1. (**Expectation Value**) The expectation value $\langle \hat{A} \rangle$ of an observable operator $\hat{A}$ in the state $\psi(x)$ of a physical system is defined by

$$\langle \hat{A} \rangle = \frac{\langle \psi | \hat{A}\psi \rangle}{\langle \psi | \psi \rangle}. \tag{4.2}$$

If the state ψ is normalized with $\langle \psi | \psi \rangle = 1$, then the expectation value is

$$\langle \hat{A} \rangle = \langle \psi | \hat{A}\psi \rangle \tag{4.3}$$

Definition 4.4. (**Root-mean-square Deviation**) The root-mean-square deviation $(\nabla \hat{A})$ is defined by the square root of the expectation value of $\left\langle (\hat{A} - \left\langle \hat{A} \right\rangle)^2 \right\rangle$ in the state ψ in which $\langle \hat{A} \rangle$ is computed.

The root-mean-square deviation measures the dispersion around $\langle \hat{A} \rangle$.

If $\hat{A}$ and $\hat{B}$ are Hermitian observables, the uncertainties $\Delta\hat{A}$ and $\Delta\hat{B}$ in the measurement of $\hat{A}$ and $\hat{B}$ in the given state $\psi(x)$ (equivalent here to $\nabla\hat{A}$ and $\nabla\hat{B}$) are given by

$$\begin{aligned}
\Delta\hat{A} &= (\left\langle (\hat{A} - \left\langle \hat{A} \right\rangle)^2 \right\rangle)^{1/2} \\
&= \left\langle \psi | (\hat{A} - \left\langle \hat{A} \right\rangle)^2 \psi \right\rangle^{1/2} \\
&= \left\langle (\hat{A} - \left\langle \hat{A} \right\rangle)\psi | (\hat{A} - \left\langle \hat{A} \right\rangle)\psi \right\rangle^{1/2} \\
&= ||(\hat{A} - \left\langle \hat{A} \right\rangle)\psi|| = ||\psi_1||
\end{aligned}$$

Where $\psi_1 = (\hat{A} - \left\langle \hat{A} \right\rangle)\psi$.

Similarly, we write $\Delta\hat{B} = ||\psi_2||$, where $\psi_2 = (\hat{B} - \left\langle \hat{B} \right\rangle)\psi$

Since the state vector ψ is an element of a Hilbert space, there exists a correlation between the uncertainties. This correlation is well known as the *Generalized Uncertainty Principle.*

Theorem 4.5. ***(The Uncertainty Principle)*** *If $\hat{A}$ and $\hat{B}$ are Hermitian operators, then for any state vector*

$$\Delta\hat{A}\Delta\hat{B} \geq \frac{1}{2}\left| \frac{1}{i}\left\langle [\hat{A}, \hat{B}] \right\rangle \right| \tag{4.6}$$

Proof. We have

$$\Delta\hat{A} = ||\psi_1||, \psi_1 = (\hat{A} - \left\langle \hat{A} \right\rangle)\psi,$$

$$\Delta\hat{B} = ||\psi_2||, \psi_2 = (\hat{B} - \left\langle \hat{B} \right\rangle)\psi.$$

According to Schwarz's Inequality, for any two vectors ψ_1 an ψ_2 we have

$$\begin{aligned}||\psi_1||||\psi_2|| &\geq |\langle\psi_1|\psi_2\rangle| \\ &\geq |Im\langle\psi_1|\psi_2\rangle| \\ &\geq \left|\frac{1}{2i}\{\langle\psi_1|\psi_2\rangle - \overline{\langle\psi_1|\psi_2\rangle}\}\right| \\ &\geq \left|\frac{1}{2i}\{\langle\psi_1|\psi_2\rangle - \langle\psi_2|\psi_1\rangle\}\right|.\end{aligned}$$

Evidently, since $\hat{A}$ and $\hat{B}$ are Hermitian operators,

$$\begin{aligned}\Delta\hat{A}\Delta\hat{B} &\geq \left|\frac{1}{2i}\{\langle(\hat{A}-\langle\hat{A}\rangle)\psi|(\hat{B}-\langle\hat{B}\rangle)\psi\rangle - \langle(\hat{B}-\langle\hat{B}\rangle)\psi|(\hat{A}-\langle\hat{A}\rangle)\psi\rangle\}\right| \\ &\geq \left|\frac{1}{2i}\langle\langle\psi|(\hat{A}\hat{B}-\hat{B}\hat{A})\psi\rangle\rangle\right| \\ &\geq \frac{1}{2}\left|\frac{1}{i}\langle[\hat{A},\hat{B}]\rangle\right|\end{aligned}$$

□

Corollary 4.7. *For conjugate operators q_j and p_k, velocity vector v_j and mass m, the following holds:*

$$\Delta q_j \Delta p_k \geq \frac{\hbar}{2}\delta_{jk}, \tag{4.8}$$

$$\Delta x_j \Delta p_j \geq \frac{\hbar}{2}, \tag{4.9}$$

$$\Delta x_j \Delta v_j \geq \frac{\hbar}{2m}. \tag{4.10}$$

Proof. Substitute $\hat{A} = q_j$ and $\hat{B} = p_k$ in (4.6), we obtain (4.8). In Cartesian coordinates, (4.8) reduces to (4.9). Finally, with $p_j = v_j m$, the product of the velocity vector v_j and mass m, (4.9) becomes (4.10). □

We should point out that (4.9) is the famous uncertainty relation between the position and momentum first discovered by Heisenberg. According to the Heisenberg uncertainty relation, the position and momentum of a particle cannot be determined exactly and simultaneously, and such an experiment is impossible to design.

The Uncertainty Principle is inherent in the properties of all wave-like systems, and arises in quantum mechanics due to the matter wave nature of all quantum objects, stating a fundamental property of quantum systems. Quantum mechanics is essentially built upon this principle, and hence the Heisenberg Uncertainty Principle is perhaps one of the most fundamental results in quantum theory.

Additionally, since $\hbar$ is very small, $\frac{\hbar}{2m}$ is extremely small for any macroscopic system, and so (4.10) tells us why the Uncertainty Principle is of no importance in classical mechanics.

5. Importance and Wider Applications

The mathematical concept of Hilbert space generalizes the notion of Euclidean space. In addition to quantum mechanics, it is incredibly useful in the theories of partial differential equations, Fourier analysis, and ergodic theory. David Hilbert first formulated and studied the Hilbert space. But it was John von Neumann who coined the term Hilbert space. Von Neumann later leaned on much of Hilbert's work when proving the equivalence of Werner Heisenberg's matrix mechanics and Erwin Schrodinger's wave equations.

Quantum mechanics usually only use separable Hilbert spaces. Meanwhile, the Wightman axioms in an attempt to give a mathematically rigorous formulation of standard quantum field theory (the theoretical framework for constructing quantum mechanical models of subatomic particles in particle physics and quasi-particles in condensed matter physics) also require separable Hilbert spaces. However, we should acknowledge that non-separable Hilbert spaces are also important in quantum field theory, for example as the natural state space of a boson. This is because in quantum field theory, systems possess an infinite number of degrees of freedom and any infinite Hilbert tensor product is non-separable.

Coming back to quantum mechanics, states of physical quantum systems can be represented by vectors in separable Hilbert space, observables by Hermitian operators, symmetries by unitary operators, and measurements by orthogonal projections. David Hilbert brought much of mathematical rigor into physics, studied and formulated many of the important definitions and mathematical foundations, which is perhaps even more influential than his actual achievements in physics though these cannot be discounted either.

Acknowledgments. Many thanks to Professor Peter May for giving me this opportunity to participate in REU 2017. I thank my mentors Karen Butt and Dylan Quintana for their advice and help with this paper. I would also like to thank Mao Tian-Tan for his recommendations on quantum mechanics related readings.

References

[1] Lokenath Debnath and Piotr Mikusinski. Introduction to Hilbert Spaces with Applications. Academic Press. 1999. ISBN 0-12-208436-5.

[2] John.S.Townsend. A Modern Approach to Quantum Mechanics. University Science Books. 2000. ISBN 978-1891389139.

[3] Ramamurti Shankar. Principles of Quantum Mechanics. Plenum Press. 2000. ISBN 978-1891389139.