November 2009

The Dirac Approach to Quantum Theory

Paul Renteln

Department of Physics
California State University
5500 University Parkway
San Bernardino, CA 92407

Table of Contents

1. Introduction

1.1 These notes provide a brief description of the mathematical framework that underlies the Dirac approach to quantum theory. [1] Students who are interested in more of the details can do no better than to consult the master himself. [2]

2. Abstract Vector Spaces

2.1 One of the postulates of quantum theory, based on the experimental result that both matter and light exhibit interference, is that the space of possible quantum states of a system form a kind of vector space. It therefore seems appropriate to begin with a

Definition. A **vector space** V is a set of elements, called *vectors*, that is closed under addition and multiplication by scalars, *i.e.,* for all $\phi, \psi \in V$ and all $a, b \in \mathbb{C}$ we have

$$a\phi + b\psi \in V. \tag{2.1}$$

(We call the expression in (2.1) a **linear combination** of ϕ and ψ.)

2.2 Example. The vector space with which we are perhaps most familiar is the usual two dimensional cartesian space $\mathbb{R}^2$. A vector in $\mathbb{R}^2$ is labeled by a pair of numbers (x, y) and we define scalar multiplication and vector addition as follows:

$$a(x, y) := (ax, ay) \tag{2.2}$$

and

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2). \tag{2.3}$$

2.3 Example. The space of $C[a, b]$ of continuous functions from the interval $[a, b]$ to the real numbers $\mathbb{R}$ is a vector space. A vector in $C[a, b]$ is a function f, and scalar multiplication and vector addition are defined as follows:

$$(af)(x) := af(x) \tag{2.4}$$

[1] Parts of the following treatment are inspired by that of L.E. Ballentine, *Quantum Mechanics* (Prentice-Hall, Englewood Cliffs, NJ, 1990).

[2] P.A.M. Dirac, *The Principles of Quantum Mechanics*, 4th ed. (Clarendon, Oxford, 1958).

and

$$(f+g)(x) := f(x) + g(x). \tag{2.5}$$

That is, the function obtained by summing the two functions f and g is simply defined as the function whose value at x is the sum of the values of f and g at x. (We say the functions add **pointwise**.)

2.4 We need some way of characterizing sets of vectors. The first such characterization is provided by the notion of 'linear independence':

DEFINITION. A set of vectors $\{\phi_n\} \subset V$ is said to be **linearly independent** if

$$\sum_n c_n \phi_n = 0 \qquad \text{implies} \qquad c_n = 0 \quad \forall n. \tag{2.6}$$

Essentially this means that no member of a set of linearly independent vectors may be expressed as a linear combination of the others.

2.5 EXAMPLE. The vectors $(1,1)$ and $(2,1)$ in $\mathbb{R}^2$ are linearly independent.

2.6 EXAMPLE. The vectors $(1,1)$ and $(2,2)$ in $\mathbb{R}^2$ are linearly dependent.

DEFINITION. A maximal set of linearly independent vectors is called a **basis** for the vector space. The cardinality of a basis is called the **dimension** of the space.

2.7 EXAMPLE. $\mathbb{R}^2$ is two dimensional, because I can find two linearly independent vectors (*e.g.,* $(1,0)$ and $(0,1)$), but I cannot find three. The vectors $(1,0)$ and $(0,1)$ form a basis for the space $\mathbb{R}^2$, as do, for example, $(1,0)$ and $(1,4)$. In general, there are an infinite number of bases for a vector space.

2.8 EXAMPLE. The space $C[-L,L]$ of continuous functions on the interval $[-L,L]$ is infinite dimensional. A basis is given by the functions

$$f(x) = \left\{ \begin{array}{c} \cos(n\pi x/L) \\ \sin(n\pi x/L) \end{array} \right\}.$$

In other words, any function in $C[-L,L]$ may be expressed as an infinite linear combination of these basis functions (provided we use something called 'convergence in the mean'). We call this a **Fourier series**. The dimension of $C[-L,L]$ is therefore countably infinite (because the basis vectors can be put in one-to-one correspondence with the integers). There are other infinite dimensional vector spaces that have uncountably infinite dimension, as we shall see later.

3. Inner Products

3.1 Although it is certainly true that the space of states in quantum theory forms a vector space, this is not enough. We need some way of computing **probabilities** for measurements carried out on these states. The main tool allowing us to do this is something called an 'inner product'. We have the following

Definition. An **inner product** (or scalar product, or dot product) on a linear vector space associates a complex number with every pair of vectors. That is, it is a map

$$(\, , \,) : V \times V \longrightarrow \mathbb{C}$$

satisfying the following three axioms:

$$\begin{array}{lll} (i) & (\phi, \psi) = (\psi, \phi)^* & \text{(Hermiticity)} \\ (ii) & (\phi, c_1\psi_1 + c_2\psi_2) = c_1(\phi, \psi_1) + c_2(\phi, \psi_2) & \text{(linearity on second entry)} \\ (iii) & (\phi, \phi) \geq 0, = 0 \Leftrightarrow \phi = 0 & \text{(positive definiteness)} \end{array}$$

Remark. A vector space V equipped with an inner product is called an **inner product space**.

Remark. It follows immediately from the definition that the inner product is **antilinear** in the first entry:

$$(c_1\phi_1 + c_2\phi_2, \psi) = c_1^*(\phi_1, \psi) + c_2^*(\phi_2, \psi). \tag{3.1}$$

3.2 Example. In the vector space $\mathbb{C}^n$ an inner product may be defined as follows. If $\psi = (a_1, a_2, ..., a_n)$ and $\phi = (b_1, b_2, ..., b_n)$ then

$$(\psi, \phi) = a_1^* b_1 + a_2^* b_2 + ... + a_n^* b_n \tag{3.2}$$

is an inner product, *i.e.,* it satisfies all the axioms. We recognize this as the ordinary dot product of two vectors if we restrict ourselves to the subspace $\mathbb{R}^n$.

3.3 Example. In the vector space of square integrable functions [3] we may define an inner product given by

$$(f, g) = \int_{-\infty}^{\infty} f^* g \, dx. \tag{3.3}$$

[3] A square integrable function is one for which $\int |f|^2 \, dx$ exists.

Definition. Two vectors ϕ and ψ for which $(\phi, \psi) = 0$ are said to be **orthogonal**.

Definition. The **norm** or **length** of a vector ϕ is defined to be

$$\|\phi\| := (\phi, \phi)^{1/2}. \tag{3.4}$$

Remark. The norm is well-defined, as the axioms imply (ϕ, ϕ) is always real.

4. The Dual Space

4.1 Next, we turn to a very important construction in the theory of linear vector spaces.

Definition. A **linear functional** F on V is a map

$$F : V \longrightarrow \mathbb{C}$$

satisfying

$$F(a\phi + b\psi) = aF(\phi) + bF(\psi) \tag{4.1}$$

for all $a, b \in \mathbb{C}$ and $\phi, \psi \in V$.

Definition. The **dual space** $\widetilde{V}$ of V is the set of all linear functionals on V.

The dual space $\widetilde{V}$ is a linear vector space in its own right. To see this, we must define how to multiply a functional by a number and how to add functionals. This is done the natural way (*cf.* Example **2.3**):

$$(aF)(\phi) := aF(\phi) \tag{4.2}$$

and

$$(F_1 + F_2)(\phi) := F_1(\phi) + F_2(\phi). \tag{4.3}$$

4.2 We have now seen that, given a vector space V, we may construct a new vector space $\widetilde{V}$, called the dual space, consisting of the set of all linear functionals on V. According to a very important theorem, due to the mathematician F. Riesz, the dual space $\widetilde{V}$ is intimately related to the original vector space V:

4.3 Theorem (Riesz). *There is a canonical isomorphism between V and $\widetilde{V}$:*

$$f \leftrightarrow F,$$

where $f \in V$ and $F \in \widetilde{V}$, given by

$$F(\phi) = (f, \phi) \tag{4.4}$$

for all $\phi \in V$.

Remark. The notion of 'isomorphism' is explained below. Right now you can just take it to mean that the spaces V and $\widetilde{V}$ may, for most purposes, be treated as one and the same. The theorem says that there is a one-to-one correspondence between vectors f in V and linear functionals F in $\widetilde{V}$. The Riesz theorem works for all finite dimensional and some infinite dimensional vector spaces. Note that the space V must be an inner product space, because the inner product is central to the statement of the theorem.

Proof. Given $f \in V$ it is clear that

$$(f, \cdot)$$

is a linear functional on V. (Just check the definitions of linear functional and the axioms of the inner product.) So we need only show that, given an arbitrary linear functional F we can construct a unique vector f satisfying (4.4).

Let $\{\phi_n\}$ be an orthonormal basis for V (*i.e.,* $\{\phi_n\}$ is a basis for V and $(\phi_n, \phi_m) = \delta_{nm}$ where δ_{nm} is the Kronecker delta; we can always find such a basis *via* Gram-Schmidt orthonormalization). Let

$$\psi = \sum a_n \phi_n$$

be an arbitrary vector in V. Then

$$F(\psi) = \sum a_n F(\phi_n)$$

(this follows by linearity of F). Define

$$f := \sum [F(\phi_n)]^* \phi_n.$$

This vector does the trick:

$$\begin{aligned} (f, \psi) &= (\sum [F(\phi_n)]^* \phi_n, \psi) \\ &= \sum F(\phi_n)(\phi_n, \psi) \\ &= \sum a_n F(\phi_n) \\ &= F(\psi). \end{aligned}$$

Next, suppose $(g,\psi) = (f,\psi)$ for all ψ. Then we have $(g-f,\psi) = 0$. Let $\psi = g-f$. By the nondegeneracy of the inner product, we must therefore have $g-f = 0$. So f is unique. ∎

5. Dirac Bra-Ket Notation

5.1 According to the postulates of quantum theory, the state space of a physical system is modeled by a certain kind of vector space called a **Hilbert space**, after the great mathematician David Hilbert. In finite dimensions, a Hilbert space is just an inner product space, and we already know what this means. [4] In Dirac's notation, vectors in the Hilbert space are denoted by the symbol $|\phi\rangle$ and are called **ket-vectors**. Vectors in the dual of the Hilbert space (linear functionals on the Hilbert space) are denoted by the symbol $\langle F|$ and are called **bra-vectors**.

5.2 Now we know the secret of bras—they are just linear functionals on the Hilbert space (or, equivalently, vectors in the dual of the Hilbert space)! But there is more. We know what linear functionals do for a living: they eat vectors and spit out numbers. We denote this process by means of the **Dirac bra(c)ket**:

$$F(\phi) =: \langle F|\phi\rangle. \tag{5.1}$$

So now we know the secret of the Dirac bracket—it simply represents the natural map from linear functionals and vectors to the complex numbers.

5.3 But we may go one step farther. According to the Riesz theorem, there is a one-to-one correspondence between vectors on a vector space and linear functionals on that space. This means there is a natural one-to-one correspondence between kets and bras. Going back to the proof of the Riesz theorem, we constructed a vector f that corresponded to the linear functional F. So, by a slight abuse of notation, we may denote the linear functional F by the bra vector $\langle f|$. In this way, we have established that there is a one-to-one correspondence between bras and kets given by

$$\langle f| \longleftrightarrow |f\rangle. \tag{5.2}$$

[4] In infinite dimensions a Hilbert space is a *complete* inner product space, where 'complete' means essentially that every infinite linear combination of the basis elements converges to a vector in the space.

In his book, Dirac postulated this correspondence, but you can now see that it follows directly from the Riesz theorem. [5]

Furthermore, with this mathematical device, it follows (*cf.* (4.4) and (5.1)) that

$$\boxed{\langle f|\phi\rangle = (f, \phi)} \tag{5.3}$$

In other words, the Dirac bracket is simply the inner product. [6]

5·4 What is the physical significance of the Dirac bracket? We have the following axioms (inspired by wave mechanics)

- $\langle\phi|\psi\rangle$ is the **probability amplitude** to find the system in the state $|\phi\rangle$ when it is prepared in the state $|\psi\rangle$.
- $|\langle\phi|\psi\rangle|^2$ is the **probability** to find the system in the state $|\phi\rangle$ when it is prepared in the state $|\psi\rangle$.

5·5 There is only one loose end to clear up before we go on. We would like to understand the correspondence (5.2) a little more fully. In particular, we said that, to every ket vector there is a bra vector, which is simply obtained by turning the ket around. But what bra corresponds to the ket $c|\phi\rangle$ where c is some complex number? Well, strictly speaking, $c|\phi\rangle$ is not a ket—it is a constant times a ket, which we have not yet defined. But kets are vectors, and the product of a scalar and a vector is again a vector. A moment's thought should convince you that the only possible definition is

$$c|\phi\rangle = |c\phi\rangle. \tag{5.4}$$

The object on the right is the ket vector representing the vector $c\phi$.

It follows that the bra corresponding to the ket $c|\phi\rangle = |c\phi\rangle$ is simply $\langle c\phi|$. But what is this? Well, from (5.3) we can write

$$\langle c\phi|\psi\rangle = (c\phi, \psi) = c^*(\phi, \psi) = c^*\langle\phi|\psi\rangle, \tag{5.5}$$

[5] Technically, this is only true in finite dimensions.

[6] You might wonder why we bother with all this stuff about bras when we could have simply defined the bracket to be the inner product. Well, the answer is because bras turn out to be very useful for defining things like projection operators later on.

so we may conclude that

$$\langle c\phi| = c^*\langle\phi|, \tag{5.6}$$

whereupon we observe that the correspondence (5.2) is **antilinear**. That is

$$c_1|\phi_1\rangle + c_2|\phi_2\rangle \leftrightarrow c_1^*\langle\phi_1| + c_2^*\langle\phi_2|. \tag{5.7}$$

Once again we see that this correspondence is not arbitrary, but follows from the properties of the inner product. [7]

6. Linear Transformations

6.1 In the real world, things change. How are these changes represented in quantum theory? The answer is: by linear operators on the Hilbert space.

Definition. Let V and W be vector spaces. A map $T : V \to W$ is called a **linear transformation** provided

$$T(av_1 + bv_2) = aTv_1 + bTv_2, \tag{6.1}$$

for all $a, b \in \mathbb{C}$ and $v_1, v_2 \in V$.

Remark. Linear transformations are also called **natural maps** in this context, because they preserve the linear structure of the two spaces. They are also called **homomorphisms**, from the Greek '$\eta o\mu o\sigma$' meaning similar or akin and '$\mu o\rho\phi\epsilon$' meaning shape or form. Two spaces are **homomorphic** if there is a homomorphism between them. In that case they are indeed similar in form.

Definition. Let V and W be vector spaces. A map $T : V \to W$ is called a **linear isomorphism** if T is bijective (one-to-one and onto) and both T and its inverse are homomorphisms. V and W are said to be **isomorphic** if there is an isomorphism between them.

[7] Of course, this begs the question of why the inner product must be Hermitian. The reason has to do with the requirement that probability amplitudes be complex, in order to give rise to quantum mechanical interference.

Remark. If V and W are isomorphic, there are essentially the same (the Greek word '$\iota\sigma o\sigma$' means 'equal'). Now we understand the meaning of the word 'isomorphism' in the context of the Riesz theorem. [8]

Definition. A homomorphism $T : V \to V$ from a vector space to itself is called an **endomorphism**.

Definition. A bijective endomorphism $T : V \to V$ is called an **automorphism**.

Remark. An endomorphism is not necessarily one-to-one or onto, which means it could map V to part of itself in a funny way. An automorphism is just about the most innocuous sort of map around—it basically permutes the vectors in the space in a linear way.

6.2 With all this terminology out of the way, we may move back in the direction of physics. For a physicist, an endomorphism is called a **linear operator**, so henceforth we shall mostly use this terminology. Written out in Dirac notation, a linear operator $\widehat{A} : V \to V$ acts on a ket vector to give another ket vector:

$$\widehat{A}|\phi\rangle = |\psi\rangle, \tag{6.2}$$

and does so in a linear way:

$$\widehat{A}(c_1|\phi_1\rangle + c_2|\phi_2\rangle) = c_1\widehat{A}|\phi_1\rangle + c_2\widehat{A}|\phi_2\rangle. \tag{6.3}$$

6.3 It will be important for us to have a working criterion of when two operators are equal: two operators $\widehat{A}$ and $\widehat{B}$ are **equal** if

$$\widehat{A}|\psi\rangle = \widehat{B}|\psi\rangle \tag{6.4}$$

for all states $|\psi\rangle$.

6.4 What is the physical significance of linear operators? Well, according to one of the axioms of quantum theory, physical observables are represented by special kinds of linear operators. We will see below precisely what sorts of operators these are. In the mean time, we must introduce yet another important construction.

[8] The other word there, 'canonical', simply means that the isomorphism is in some sense natural or "God-given". That is, there is basically only one choice for the map.

7. Left Actions of Operators

7.1 So far, we have defined the action of operators to the *right* on ket vectors. We can also define the action of an operator to the *left* on bra vectors. How? Well, what we are really asking for is, how can we define the action of a linear operator on a linear functional? The answer is to use what mathematicians call the *natural action*, namely **pullback**. The operator $\widehat{A}$ acts on the functional F to give another functional $\widehat{A}^\star F$, called the pullback of F by $\widehat{A}$. [9]

7.2 We define $\widehat{A}^\star F$ by saying what it does to ket vectors. Thinking like a mathematician for a moment, we reason as follows. We want $\widehat{A}^\star F$ to be a functional. That is, it must eat ket vectors and spit out numbers. Clearly we cannot simply define $\widehat{A}^\star F$ to act first with F on ket vectors, because the result would then be a number, and $\widehat{A}$ acting on a number c would give an operator $c\widehat{A}$, not a functional. Instead, we let things act in the most natural way: first act with $\widehat{A}$ on the ket vector, then operate on this with the linear functional F. That is

$$(\widehat{A}^\star F)|\phi\rangle := F(A|\phi\rangle). \tag{7.1}$$

This makes it clear that the pullback of the functional F by the operator $\widehat{A}$ is simply the composition of F and $\widehat{A}$. The whole idea is best illustrated by a picture:

$$\begin{tikzcd}
V \arrow[r, "\widehat{A}"] \arrow[rr, bend left, "\widehat{A}^\star F = F \circ \widehat{A}"] & V \arrow[r, "F"] & \mathbb{C}
\end{tikzcd}$$

Figure 1. Pullback of a linear functional by a linear operator.

In this picture we see that $\widehat{A}$, being a linear operator on V, is a map from V to another copy of itself, and F is a linear functional taking elements of V to the complex numbers. The functional F is "pulled back" from the second copy of V to first copy of V to give a new functional, which is just the composition.

7.3 All of this becomes almost transparent in the Dirac notation. [10] The functional F is

[9] Note the difference between '$\star$', meaning pullback, and '$*$' meaning 'complex conjugate'.
[10] Indeed, this is one of the virtues of Dirac notation, but is also the reason why the Dirac notation is so slippery. It is so powerful, one sometimes forgets what one is doing.

represented by the bra $\langle f|$, so the right hand side of (7.1) is written

$$F(\widehat{A}|\phi\rangle) = \langle f|\widehat{A}|\phi\rangle. \tag{7.2}$$

Comparing (7.2) to the left side of (7.1) we see that, in Dirac notation, pullback of a linear functional is represented by the *left* action of $\widehat{A}$ on the corresponding bra:

$$\widehat{A}^\star F \leftrightarrow \langle f|\widehat{A}. \tag{7.3}$$

With this notation, then, we have

$$(\langle f|\widehat{A})|\phi\rangle = \langle f|(\widehat{A}|\phi\rangle) = \langle f|\widehat{A}|\phi\rangle. \tag{7.4}$$

That is, in the last expression, $\widehat{A}$ may be thought of as acting either to the left on $\langle f|$ or to the right on $|\phi\rangle$.

8. The Adjoint

8.1 Now, we just finished showing that there is a natural left action of operators on bras. If you look back at the ideas presented in the previous section, you will see that pullback is a linear map from the dual space to itself. This means that

$$\langle\phi|\widehat{A} = \langle\chi| \tag{8.1}$$

for some bra $\langle\chi|$. Now, according to the Riesz theorem, there is a one-to-one correspondence between bras and kets. This means that the left action of $\widehat{A}$ above defines a unique mapping that associates $|\phi\rangle$ to $|\chi\rangle$. This map is called the **adjoint** of $\widehat{A}$ and is denoted $\widehat{A}^\dagger$:

$$\widehat{A}^\dagger|\phi\rangle = |\chi\rangle. \tag{8.2}$$

8.2 One way to get a handle on the adjoint map is to see what it does in the old notation, then translate everything back into Dirac notation. Denote the linear functional $\langle\phi|$ by F, and the linear functional $\langle\chi|$ by G. Then what (8.1) is saying is that

$$\widehat{A}^\star F = G. \tag{8.3}$$

It follows from our definitions that the vector corresponding to the functional F (respectively, G) under the Riesz theorem is ϕ (respectively, χ). That is

$$F(\cdot) = (\phi, \cdot) \tag{8.4}$$

and

$$G(\cdot) = (\chi, \cdot), \tag{8.5}$$

where (,) is the inner product. It follows from the definition of pullback and the adjoint and Equation (8.2) in the old notation that

$$(\widehat{A}^\dagger\phi, \psi) = (\chi, \psi) = G(\psi) = \widehat{A}^\star F(\psi) = F(\widehat{A}\psi) = (\phi, \widehat{A}\psi), \tag{8.6}$$

whereupon we see that another definition of the adjoint is

$$(\widehat{A}^\dagger\phi, \psi) = (\phi, \widehat{A}\psi). \tag{8.7}$$

8.3 Translating (8.7) back into Dirac notation is a little awkward, because we do not really want to write $\langle \widehat{A}^\dagger\phi|$—that is, we want to keep our operators and kets separate. We may circumvent this notational crisis by the following trick. We observe from (8.7) that

$$(\phi, \widehat{A}\psi) = (\psi, \widehat{A}^\dagger\phi)^*. \tag{8.8}$$

Recalling that the Dirac bracket is the same thing as the inner product (*cf.* (5.3)) we may write (8.8) in Dirac notation as

$$\boxed{\langle\phi|\widehat{A}|\psi\rangle = \langle\psi|\widehat{A}^\dagger|\phi\rangle^*}. \tag{8.9}$$

If we wish, we could take this to be the *definition* of the adjoint. [11]

8.4 As a consequence of our definitions, the following operator identities hold:

$$\begin{aligned}(c\widehat{A})^\dagger &= c^*\widehat{A}^\dagger \\ (\widehat{A}+\widehat{B})^\dagger &= \widehat{A}^\dagger + \widehat{B}^\dagger \\ (\widehat{A}\widehat{B})^\dagger &= \widehat{B}^\dagger\widehat{A}^\dagger \\ (\widehat{A}^\dagger)^\dagger &= \widehat{A}.\end{aligned} \tag{8.10}$$

For example, the first identity may be proved as follows:

$$\begin{aligned}\langle\psi|(c\widehat{A})^\dagger|\phi\rangle &= \langle\phi|c\widehat{A}|\psi\rangle^* \\ &= [c\langle\phi|\widehat{A}|\psi\rangle]^* \\ &= c^*\langle\phi|\widehat{A}|\psi\rangle^* \\ &= c^*\langle\psi|\widehat{A}^\dagger|\phi\rangle \\ &= \langle\psi|c^*\widehat{A}^\dagger|\phi\rangle,\end{aligned}$$

where we twice used the fact that bras are linear functionals. As the two sides are equal for arbitrary states $\langle\psi|$ and $|\phi\rangle$ the two operators must be equal. [12]

[11] We could have done so initially, and spared ourselves the intervening mathematical exposition. But the motivation for its introduction would have been obscure.

[12] Looking back at the definition (6.4) of operator equality, we see we actually have a bit

9. The Outer Product

9.1 We have spent a lot of time with the object we called the 'inner product'. We may also define the 'outer product', which is equally important:

Definition. The **outer product** of a ket $|\psi\rangle$ and a bra $\langle\phi|$ is the operator

$$|\psi\rangle\langle\phi|.$$

Remark. This is an operator, because it acts on kets and returns kets:

$$(|\psi\rangle\langle\phi|)|\lambda\rangle = |\psi\rangle\langle\phi|\lambda\rangle. \tag{9.1}$$

The latter ket is the number $\langle\phi|\lambda\rangle$ times the ket $|\psi\rangle$. It is easy to see that the operator so defined is linear, because the bra is a linear functional.

Remark. We prove below that

$$(|\psi\rangle\langle\phi|)^\dagger = |\phi\rangle\langle\psi|. \tag{9.2}$$

Combining this with the third identity in the list (8.10) we *could* write

$$(|\psi\rangle)^\dagger = \langle\psi|, \tag{9.3}$$

but this is an abuse of notation, as the adjoint really only applies to *operators* and not *kets*. We shall therefore avoid it. [13] The proof of (9.2) is very slick:

$$\begin{aligned} \langle\xi|(|\psi\rangle\langle\phi|)^\dagger|\eta\rangle &= [\langle\eta|(|\psi\rangle\langle\phi|)|\xi\rangle]^* \\ &= [\langle\eta|\psi\rangle\langle\phi|\xi\rangle]^* \\ &= \langle\xi|\phi\rangle\langle\psi|\eta\rangle \\ &= \langle\xi|(|\phi\rangle\langle\psi|)|\eta\rangle. \qquad \blacksquare \end{aligned}$$

---

more to prove. We must show that $\langle\psi|\widehat{A}|\phi\rangle = \langle\psi|\widehat{B}|\phi\rangle$ implies that $\widehat{A} = \widehat{B}$. I claim this follows from the fact that $\langle\psi|$ is arbitrary. We may see this as follows. Let $\{|\phi_n\rangle\}$ be an orthonormal basis for the Hilbert space. Let $|\psi\rangle = \sum a_n|\phi_n\rangle$ be an arbitrary ket vector. Then $\langle\phi_n|\psi\rangle = a_n$. Now if $|\chi\rangle = \sum b_n|\phi_n\rangle$ is another arbitrary state, then $|\psi\rangle = |\chi\rangle$ if and only if $a_n = b_n$ for all n. But this may be written $\langle\phi_n|\psi\rangle = \langle\phi_n|\chi\rangle$ for all n. So, if $\langle\phi|\psi\rangle = \langle\phi|\chi\rangle$ for all $\langle\phi|$, then $|\psi\rangle = |\chi\rangle$. Hence $\langle\psi|\widehat{A}|\phi\rangle = \langle\psi|\widehat{B}|\phi\rangle$ for arbitrary $\langle\psi|$ implies that $\widehat{A}|\phi\rangle = \widehat{B}|\phi\rangle$ for all $|\phi\rangle$. By (6.4), then, we are truly done.

[13] This notation is convenient, though, when considering representations of kets. If an operator $\widehat{A}$ is represented by a matrix A in a given basis, then $\widehat{A}^\dagger$ is represented by A^{*T}, the conjugate transpose matrix. Representing kets as column vectors we find that the corresponding bra is naturally represented as the conjugate transpose vector.

10. Self-Adjoint or Hermitian Operators

10.1 The most important kinds of operators in quantum theory are those that are self-adjoint. We have the following

Definition. If $\widehat{A}^\dagger = \widehat{A}$ the operator $\widehat{A}$ is said to be **self-adjoint** or **Hermitian**.

Remark. Hermitian operators (spelled 'Hermitean' if you grew up on the other side of the Atlantic) are named after the famous French mathematician Charles Hermite. In some cases (pathological infinite dimensional cases) the two terms 'self-adjoint' and 'Hermitian' mean different things. Fortunately for us, we may safely ignore those cases.

10.2 From (8.9) we see that, if $\widehat{A}$ is Hermitian, then

$$\langle\phi|\widehat{A}|\psi\rangle = \langle\psi|\widehat{A}|\phi\rangle^* \tag{10.1}$$

holds for all states $|\phi\rangle$ and $|\psi\rangle$. Conversely, if (10.1) holds, then it follows (by the same arguments employed in Footnote 11) that $\widehat{A}^\dagger = \widehat{A}$. One can also show that an apparently weaker condition implies Hermiticity:

10.3 Theorem. *If $\langle\psi|\widehat{A}|\psi\rangle = \langle\psi|\widehat{A}|\psi\rangle^*$ for all states $|\psi\rangle$ then $\widehat{A}$ is Hermitian.*

Proof. Exercise.

10.4 Hermitian operators are central to quantum theory for the simple reason that they represent physical observables. That is, to every physical observable there is a Hermitian operator, and only Hermitian operators are physical (though not all of them are). Shortly, we will see the reason why.

11. Projection Operators

11.1 Projection operators are of crucial importance to modern quantum theory as they represent the result of measurements. So we introduce yet another

Definition. An operator $\widehat{A}$ satisfying $(\widehat{A})^2 = \widehat{A}$ is called **idempotent**.

Definition. An idempotent, Hermitian operator is called a **projection operator**.

11.2 Example. The operator $|\phi\rangle\langle\phi|$ is a projection operator, provided $|\phi\rangle$ is normalized to unity. The proof is easy. First, it is Hermitian, as $(|\phi\rangle\langle\phi|)^\dagger = |\phi\rangle\langle\phi|$, according to (9.2). Second, it is idempotent

$$(|\phi\rangle\langle\phi|)^2 = (|\phi\rangle\langle\phi|)(|\phi\rangle\langle\phi|) = |\phi\rangle\langle\phi|\phi\rangle\langle\phi| = |\phi\rangle\langle\phi|,$$

because $\langle\phi|\phi\rangle = 1$. ▮

12. Eigenvectors and Eigenvalues

12.1 You should already be familiar with this concept from elementary quantum theory. When solving the Schrödinger equation in the time independent case to find the wavefunction, you are solving an eigenvalue equation in which the wavefunction is the eigenfunction and the energy is the eigenvalue. The same idea holds in this more general setting.

Definition. If

$$\widehat{A}|\phi\rangle = a|\phi\rangle \tag{12.1}$$

for some $a \in \mathbb{C}$ and some $|\phi\rangle \in V$ we call $|\phi\rangle$ an **eigenvector** or **eigenket** of the operator $\widehat{A}$ and a its **eigenvalue**.

Remark. The physical significance of this is the following. In the quantum theory, dynamical observables A, such as the energy, momentum, position, or spin of a particle, are represented by Hermitian operators $\widehat{A}$. If the system is in an eigenstate of the operator $\widehat{A}$ with eigenvalue a, then we may say that the value of the observable A for that state is a with probability 1. That is, eigenstates correspond to systems with definite values for the corresponding observable.

Definition. The totality of eigenvalues of an operator $\widehat{A}$ is called its **spectrum**.

Remark. The significance of the spectrum is that it gives you the set of possible values of a dynamical observable. If the spectrum is discrete, then the dynamical observable is quantized (takes only discrete values). Such is the case, for example, with the energy levels of the hydrogen atom. On the other hand, the energy states of a particle scattering off a potential step form a continuum, so the Hamiltonian (which represents the energy of the system) has a continuus spectrum in that case.

We are now in a position to appreciate the importance of Hermitian operators. This comes about because of the following two theorems:

12.2 THEOREM. *The eigenvalues of a Hermitian operator are real.*

REMARK. As the eigenvalues of an operator representing a dynamical observable are the results of measurements on the system, they had better be real. It is for this reason that such operators are required to be Hermitian.

Proof. Let $\widehat{A}|\phi\rangle = a|\phi\rangle$. Then

$$\begin{aligned} \langle\phi|\widehat{A}|\phi\rangle &= \langle\phi|\widehat{A}|\phi\rangle^* && \text{(Equation (10.1))} \\ \Rightarrow \langle\phi|a|\phi\rangle &= \langle\phi|a|\phi\rangle^* && \text{(by hypothesis)} \\ \Rightarrow a\langle\phi|\phi\rangle &= a^*\langle\phi|\phi\rangle^* && \text{(linearity of bras).} \end{aligned}$$

But $\langle\phi|\phi\rangle$ is real, so $a = a^*$. ∎

Next we have a theorem about the eigenvectors of a Hermitian operator.

12.3 THEOREM. *The eigenvectors corresponding to distinct eigenvalues of a Hermitian operator are orthogonal.*

Proof. Let $\widehat{A}|\phi_1\rangle = a_1|\phi_1\rangle$ and $\widehat{A}|\phi_2\rangle = a_2|\phi_2\rangle$. Then

$$\begin{aligned} 0 &= \langle\phi_1|\widehat{A}|\phi_2\rangle - \langle\phi_2|\widehat{A}|\phi_1\rangle^* && \text{(Hermiticity of } \widehat{A}\text{)} \\ &= a_2\langle\phi_1|\phi_2\rangle - a_1^*\langle\phi_2|\phi_1\rangle^* && \text{(linearity of bras)} \\ &= (a_2 - a_1)\langle\phi_1|\phi_2\rangle && \text{(property of inner product),} \end{aligned}$$

where we used the fact that a_1 is real by Theorem 12.2. By hypothesis, $a_1 \neq a_2$. Hence $\langle\phi_1|\phi_2\rangle = 0$. ∎

12.4 Theorem 12.3 says nothing about eigenvectors corresponding to *degenerate* eigenvalues. That is, we often find situations in which some subset of the eigenvectors of an operator have the same eigenvalue. In that case we say the eigenvalue is **degenerate**. Theorem 12.3 does not apply in that case. However, one can show, using something called

Gram-Schmidt orthonormalization, that the eigenvectors corresponding to a degenerate eigenvalue may be chosen to be orthogonal as well. By normalizing all of them, we have the following result:

12.5 Corollary. *The entire set of eigenvectors of a Hermitian operator may be chosen so as to form an orthonormal set.*

13. Completeness

Definition. The orthonormal set of eigenvectors of a Hermitian operator is said to be **complete** if it forms a basis for the Hilbert space (in other words, if the set spans the space).

Remark. According to the **spectral theorem** of linear algebra, the set of eigenvectors of a Hermitian operator always forms a complete set if the Hilbert space is finite dimensional (and note, the dimension of the Hilbert space is determined by the spectrum of all the operators acting in the space, because the possible results of measurements each correspond to at least one distinct orthogonal direction). In infinite dimensions, the problem is much harder; indeed, in some cases the spectral theorem fails. We shall nonetheless assume it to be true! The assumption of completeness is crucial, for it essentially makes calculations possible.

13.1 Henceforth, we assume that $\{|\phi_n\rangle\}$ is a complete set of orthonormal eigenkets of some Hermitian operator. In that case, any vector may be expanded in terms of these basis eigenkets:

$$|v\rangle = \sum_n v_n|\phi_n\rangle. \tag{13.1}$$

Using the orthonormality properties of the eigenkets, it is easy to see that

$$\langle\phi_k|v\rangle = \sum_n v_n\langle\phi_k|\phi_n\rangle = v_k. \tag{13.2}$$

So we may write

$$\begin{aligned} |v\rangle &= \sum_n v_n|\phi_n\rangle \\ &= \sum_n |\phi_n\rangle(\langle\phi_n|v\rangle) \\ &= \left(\sum_n |\phi_n\rangle\langle\phi_n|\right)|v\rangle. \end{aligned}$$

Comparing the left and right hand sides of this expression we see that, as $|v\rangle$ is arbitrary, we may identify the expression in parentheses with the identity operator:

$$\boxed{\sum_n |\phi_n\rangle\langle\phi_n| = \mathbf{1}}. \tag{13.3}$$

The relation (13.3) is called a **completness relation** or a **resolution of unity**; it expresses the completeness of the basis of eigenkets. The operators $P_n = |\phi_n\rangle\langle\phi_n|$ are projection operators. P_n acts on $|v\rangle$ by projecting the vector onto the $|\phi_n\rangle$ direction.

13.2 Suppose we prepare a system in the state $|\psi\rangle$. We may write this state as a superposition of states $|\psi\rangle = \sum_n P_n|\psi\rangle$, where each summand $P_n|\psi\rangle$ gives the component of $|\psi\rangle$ in the $|\phi_n\rangle$ direction. If we now perform a measurement of the observable A, we will get one and only one of the eigenvalues a_n. Furthermore, if we immediately make a second measurement of A, we will find a_n with probability one. We conclude that, in the act of measuring the state $|\psi\rangle$, we have caused it to be projected into one of the eigenstates $|\phi_n\rangle$. That is, before the first measurement $|\psi\rangle$ exists as a superposition of states, but just after the measurement the system is forced into the state $|\phi_n\rangle$. This discontinuous jump from one ket to another during the act of measurement is called the **collapse of the state vector**. Because it seems unnatural, many people have objected to this part of quantum theory. Unfortunately, after years of attempting to exorcise quantum theory of the demon of state vector collapse, there is still no viable alternative. The basic rule is: a measurement of a dynamical observable causes a state vector to be projected (*via* a projection operator) into one of the possible eigenkets of the Hermitian operator representing that dynamical observable.

14. Spectral Decomposition of a Hermitian Operator

14.1 There are times when we would like to be able to speak about things like the square root of an operator. But it is not clear what such an expression would mean. We can give meaning to it by means of something called the spectral decomposition of the operator, and the key is to use the completeness relation. In particular, we have the following

14.2 Lemma. *Let $\{|\phi_n\rangle\}$ denote a complete set of orthonormal eigenkets of a Hermitian operator $\widehat{A}$, with $\widehat{A}|\phi_n\rangle = a_n|\phi_n\rangle$. Then the operator $\widehat{A}$ may be written as follows:*

$$\widehat{A} = \sum_n a_n|\phi_n\rangle\langle\phi_n|. \tag{14.1}$$

Proof. Exercise.

Remark. This way of writing the operator $\widehat{A}$ in terms of its eigenvalues and eigenkets is called the **spectral decomposition** of the operator $\widehat{A}$. It expresses the operator as a superposition of projection operators, weighted by the possible eigenvalues of the operator $\widehat{A}$.

14.3 We are now in a position to define what we mean by the **function of an operator** $f(\widehat{A})$. It is

$$f(\widehat{A}) := \sum_n f(a_n)|\phi_n\rangle\langle\phi_n|. \tag{14.2}$$

14.4 Example. The square root of an operator $\widehat{A}$ is the operator

$$\widehat{A}^{1/2} = \sum_n \sqrt{a_n}\ |\phi_n\rangle\langle\phi_n|. \tag{14.3}$$

15. Expectations Values and Uncertainties

15.1 We may now justify the statement made in Section **12.1** that the eigenvectors of a Hermitian operator represent states with definite values for the corresponding observable. To do so, we introduce the idea of expectation values and uncertainties of an operator in a particular state.

15.2 Let A be a dynamical observable represented by the Hermitian operator $\widehat{A}$, and suppose the system is in the state represented by $|\psi\rangle$. Then the **average** or **expectation value** $\langle A\rangle_{|\psi\rangle}$ of the dynamical observable A in the state $|\psi\rangle$ is given by

$$\boxed{\langle A\rangle = \langle\widehat{A}\rangle = \langle\psi|\widehat{A}|\psi\rangle}. \tag{15.1}$$

For notational ease, we have dropped the dependence on $|\psi\rangle$ because the state is clear from context. Also, we have assumed that the state vector $|\psi\rangle$ is normalized. [14]

[14] If not, we must write

$$\langle A\rangle = \frac{\langle\psi|\widehat{A}|\psi\rangle}{\langle\psi|\psi\rangle}$$

for the expectation value. This is equivalent to replacing the state $|\psi\rangle$ by $|\psi\rangle/\langle\psi|\psi\rangle^{1/2}$.

15.3 To see that this really is the average value, suppose that

$$|\psi\rangle = \sum_n b_n|\phi_n\rangle, \tag{15.2}$$

where $\{|\phi_n\rangle\}$ is a complete set of orthonormal eigenkets for $\widehat{A}$. We know that

$$|\langle\phi_n|\psi\rangle|^2 = |b_n|^2 \tag{15.3}$$

represents the probability to find the system in the state $|\phi_n\rangle$ given that it was prepared in the state $|\psi\rangle$. From (15.1), (15.2), and (15.3) we compute

$$\begin{aligned}\langle A\rangle &= \langle\psi|\widehat{A}|\psi\rangle \\ &= \sum_{nm} b_n b_m^* \langle\phi_m|\widehat{A}|\phi_n\rangle \\ &= \sum_n |b_n|^2 a_n,\end{aligned} \tag{15.4}$$

which is clearly the average value of A in the state $|\psi\rangle$.

15.4 By definition, the **uncertainty** in the value of A in the state $|\psi\rangle$ is simply the standard deviation of the probability distribution of the values of A in that state:

$$\boxed{\Delta A := [\langle\widehat{A}^2\rangle - \langle\widehat{A}\rangle^2]^{1/2}}. \tag{15.5}$$

Now, if $|\psi\rangle$ happens to be one of the possible eigenstates $|\phi_n\rangle$ of the operator $\widehat{A}$ representing the dynamical observable A, we have

$$\langle\widehat{A}\rangle = \langle\psi|\widehat{A}|\psi\rangle = \langle\phi_n|\widehat{A}|\phi_n\rangle = a_n \tag{15.6}$$

and

$$\langle\widehat{A}^2\rangle = \langle\psi|\widehat{A}^2|\psi\rangle = \langle\phi_n|\widehat{A}^2|\phi_n\rangle = a_n^2, \tag{15.7}$$

so $\Delta A = 0$. This shows that, if the system is in an eigenstate of the operator $\widehat{A}$ corresponding to the dynamical observable A, then the uncertainty of A in that state is zero. That is, eigenstates of an operator correspond to states with a definite value for the corresponding dynamical observable.

16. Operators With Continuous Spectra

16.1 In everything we have done up to now, our operators have had a discrete spectrum. That is, the dimension of the Hilbert space has been either finite or countably infinite. What happens if the dimension of the Hilbert space is uncountable? When would this ever arise? Well, surely we can measure the x position of a particle. This means there must be a Hermitian operator $\widehat{x}$ representing this dynamical observable x on the Hilbert space. We call $\widehat{x}$ the **position operator**. The position of a particle may take a continuum of possible values. As the spectrum of an operator is the set of possible values that operator can have, it is clear that the spectrum of the position operator must be the continuum of possible x positions of the particle.

16.2 We label the eigenkets of the position operator $\widehat{x}$ by their corresponding eigenvalues: $|x\rangle$. So we have

$$\widehat{x}|x\rangle = x|x\rangle. \tag{16.1}$$

In this formula $\widehat{x}$ is the position operator, $|x\rangle$ is the position eigenket, and x is the eigenvalue (position) to which $|x\rangle$ corresponds. The meaning of (16.1) may perhaps be made a little clearer if we apply the position operator to the ket $|x'\rangle$ corresponding to a particle at the position x'. In that case we find

$$\widehat{x}|x'\rangle = x'|x'\rangle. \tag{16.2}$$

16.3 The resolution of unity now becomes

$$\int |x\rangle\langle x|\, dx = \mathbf{1}. \tag{16.3}$$

and the spectral decomposition of the position operator reads

$$\widehat{x} = \int x|x\rangle\langle x|\, dx. \tag{16.4}$$

16.4 Consider an arbitrary vector $|\psi\rangle$ in the Hilbert space. By using the resolution of unity (16.3) we may expand $|\psi\rangle$ in terms of the basis of position eigenkets:

$$\begin{aligned} |\psi\rangle &= \left(\int |x\rangle\langle x|\, dx\right)|\psi\rangle \\ &= \int |x\rangle\langle x|\psi\rangle\, dx. \end{aligned} \tag{16.5}$$

What is the significance of the object $\langle x|\psi\rangle$? Well, physically we can guess what is must be. We know that the bracket $\langle x|\psi\rangle$ represents the probability amplitude to find the particle in the state $|x\rangle$ given that it was prepared in the state $|\psi\rangle$. That is, it represents the probability amplitude to find the particle at the point x. But this is simply the **wavefunction** $\psi(x)$ with which we are already familiar! In other words

$$\boxed{\langle x|\psi\rangle = \psi(x)}. \tag{16.6}$$

16.5 This is consistent, because, as we now show, $\int |\psi|^2\, dx = 1$. To begin, consider the case in which the particle is located at the position x', so its state is $|\psi\rangle = |x'\rangle$. Then from (16.5) we have

$$|x'\rangle = \int |x\rangle\langle x|x'\rangle\, dx. \tag{16.7}$$

This expresses the position eigenket $|x'\rangle$ as a superposition of the position eigenkets $|x\rangle$ with coefficients $\langle x|x'\rangle$. The only way this could be true is if

$$\boxed{\langle x|x'\rangle = \delta(x - x')}. \tag{16.8}$$

That is, the bracket $\langle x|x'\rangle$ must equal the Dirac delta function. This is sometimes called the **generalized orthonormality** condition. It is the analogue of $\langle\phi_n|\phi_m\rangle = \delta_{nm}$ in finite dimensions.

16.6 Using another resolution of unity we may write

$$\langle\psi| = \int \langle\psi|x'\rangle\langle x'|\, dx', \tag{16.9}$$

so combining (16.5), (16.6), (16.9) and the generalized orthonormality condition, we have

$$\begin{aligned}\langle\psi|\psi\rangle &= \int \langle\psi|x'\rangle\langle x'|x\rangle\langle x|\psi\rangle\, dx\, dx' \\ &= \int |\langle x|\psi\rangle|^2\, dx \\ &= \int |\psi(x)|^2\, dx. \end{aligned} \tag{16.10}$$

But we know that $\langle\psi|\psi\rangle = 1$ because (i) $|\langle\psi|\psi\rangle|^2$ is the probability of measuring the particle in the state $|\psi\rangle$ given that it was prepared in the state $|\psi\rangle$, and it therefore equals one, and (ii) $\langle\psi|\psi\rangle$ must be real and positive (by virtue of the properties of the inner product).

16.7 As a final illustration, recall from our study of wave mechanics that the position operator acts on wavefunctions by simple multiplication. Let us use all the formalism we have developed to prove this. We have

$$
\begin{aligned}
\hat{x}\psi(x) &:= (\widehat{x}\psi)(x) && \text{(definition)} \\
&= \langle x|\widehat{x}|\psi\rangle && \text{(Equation (16.6) generalized)} \\
&= \langle x| \int x' |x'\rangle\langle x'|\, dx' \ |\psi\rangle && \text{(spectral decomposition)} \\
&= \int x' \langle x|x'\rangle\langle x'|\psi\rangle\, dx' && \text{(linearity)} \\
&= x\langle x|\psi\rangle && \text{(generalized orthonormality)} \\
&= x\psi(x) && \text{(Equation (16.6) again).}
\end{aligned}
$$

∎

Looking back at this computation we can see why, whenever possible, it is better to stick to wavefunctions when doing quantum theory. But in order to truly understand the meaning of quantum theory, Dirac's notation is indispensable.

17. Change of Basis

17.1 Equations (16.5) and (16.6) allow us to view the wavefunction $\psi(x)$ as the **components** of the ket vector $|\psi\rangle$ in the **position basis**. But this is not the only possible basis we could use. For example, we could use the **momentum basis** $|p\rangle$ consisting of eigenkets of the momentum operator $\widehat{p}$:

$$\widehat{p}|p\rangle = p|p\rangle. \tag{17.1}$$

These satisfy the generalized orthonormality condition

$$\langle p|p'\rangle = \delta(p - p'). \tag{17.2}$$

17.2 We know the momentum operator $\widehat{p}$ is a Hermitian operator (or, we demand that it be one), so its eigenkets are assumed to span the Hilbert space. This means we may expand the ket $|\psi\rangle$ in terms of the momentum basis. We define the components of $|\psi\rangle$ in the momentum basis to be

$$\phi(p) := \langle p|\psi\rangle. \tag{17.3}$$

I claim this is just what we previously called the **momentum space wavefunction**. We may see this as follows.

17.3 First we find the "change of basis matrix" $\langle x|p\rangle$. Following (16.6) we may write

$$\widehat{A}\psi(x) := [\widehat{A}\psi](x) = \langle x|\widehat{A}|\psi\rangle. \tag{17.4}$$

Now, acting on wavefunctions we argued previously that the momentum operator is given by

$$\widehat{p} = -i\hbar\frac{\partial}{\partial x}. \tag{17.5}$$

We may think of $\langle x|p\rangle$ as a function of either x or p; here we suppose it to be a function of x so that, using (17.1) and (17.4) we get

$$-i\hbar\frac{\partial}{\partial x}\langle x|p\rangle = \langle x|\widehat{p}|p\rangle = p\langle x|p\rangle. \tag{17.6}$$

This is a differential equation which may be easily solved:

$$\langle x|p\rangle = c(p)e^{ipx/\hbar}, \tag{17.7}$$

where $c(p)$ is some normalization factor that could depend on p. We compute it as follows:

$$\begin{aligned}
\delta(p-p') &= \langle p|p'\rangle && \text{(generalized orthonormality)}\\
&= \int dx\,\langle p|x\rangle\langle x|p'\rangle && \text{(resolution of unity)}\\
&= c^*(p)c(p')\int dx\, e^{i(p'-p)x/\hbar} && \text{(Equation (17.7))}\\
&= 2\pi\hbar c^*(p)c(p')\delta(p-p') && \text{(delta function identity)},
\end{aligned}$$

whereupon we may conclude that (up to an irrelevant phase)

$$c(p) = \frac{1}{\sqrt{2\pi\hbar}}, \tag{17.8}$$

so

$$\langle x|p\rangle = (2\pi\hbar)^{-1/2}e^{ipx/\hbar}. \tag{17.9}$$

Finally then, we get

$$\begin{aligned}
\phi(p) &= \langle p|\psi\rangle\\
&= \int dx\,\langle p|x\rangle\langle x|\psi\rangle\\
&= (2\pi\hbar)^{-1/2}\int dx\, e^{-ipx/\hbar}\,\psi(x).
\end{aligned} \tag{17.10}$$

This proves that $\phi(p)$ really is the momentum space wavefunction that we encountered in wave mechanics, because it is the Fourier transform of the position space wavefunction. We finally see that Fourier transformation is nothing more or less than an infinite dimensional change of basis!