Notes on Quantum Averages and Operators

October 11, 1995

Massachusetts Institute of Technology

Department of Physics

Physics 8.04 October 11, 1995

Quantum kinematics may be carried out entirely with nothing more than the framework laid out in the preceding lecture notes on "Quantum States, Observables and Probability Distributions." In principle, with just the physics in those notes, the student may solve all the problems which will be presented in the rest of this course. However, as complete as the framework we now have is, it is still inconvenient at times in its application as it lacks significant mathematical infrastructure.

Operator theory will provide us with the missing structure. One of the most important uses of operators in quantum theory is in the calculation of averages. We shall thus use this as the context in which to introduce operators into our theory and, in fact, use the calculation of averages as the defining point for our quantum mechanical operators. As no physics will be introduced in this note beyond that of the preceding set of notes, you may think of everything in this note as definitions and derivations based on this physics already presented in previous course material.

1 Notation and Summary of Results

For your reference, we will now present a table summarizing the results presented in these notes. You should know these facts of operator theory well enough to use them freely in algebraic manipulations.

Before we present the table of basic facts of operator theory, we briefly summarize the notation we shall use in this note in the table below.

TABLE I: NOTATIONS INTRODUCED IN THIS NOTE

Observable/ Representation	Pure States	Quantum Amplitudes/ Wave function	Function Operators	Superposition of Generic State
Position, x	$\|x>$	$\psi(x)$	$\hat{x}, \hat{p}, \hat{\mathcal{O}}, \hat{\eta}$	$\|\psi>= \int dx\, \psi(x)\|x>$
Momentum, $p = \hbar k$	$\|\hbar k>$	$\tilde{\psi}(k)$	$\hat{\tilde{x}}, \hat{\tilde{p}}, \hat{\tilde{\mathcal{O}}}, \hat{\tilde{\eta}}$	$\|\psi>= \int dk\, \tilde{\psi}(k)\|\hbar k>$
Generic Observable, $\mathcal{O}$	$\|\mathcal{O}>$	$\overline{\psi}(\mathcal{O})$	$\hat{\overline{x}}, \hat{\overline{p}}, \hat{\overline{\mathcal{O}}}, \hat{\overline{\eta}}$	$\|\psi>= \int d\mathcal{O}\, \overline{\psi}(\mathcal{O})\|\mathcal{O}>$
Alternate Gen. Obs., $\mathcal{R}$	$\|\mathcal{R}>$	$\overline{\overline{\psi}}(\mathcal{R})$	$\hat{\overline{\overline{x}}}, \hat{\overline{\overline{p}}}, \hat{\overline{\overline{\mathcal{O}}}}, \hat{\overline{\overline{\mathcal{R}}}}$	$\|\psi>= \int d\mathcal{R}\, \overline{\overline{\psi}}(\mathcal{R})\|\mathcal{R}>$

We now present the table of facts of operator theory which we discuss in this note.

TABLE II: OPERATOR FACTS DISCUSSED IN THIS NOTE

Fact	Conditions	Section
$< x >= (\psi, x\psi) = (\tilde{\psi}, i\partial_k \tilde{\psi})$	none	2.1
$< p >= (\psi, \frac{\hbar}{i}\partial_x \psi) = (\tilde{\psi}, \hbar k \tilde{\psi})$	none	2.2
$< \mathcal{O} >= (\overline{\overline{\psi}}, \hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\psi}})$	definition of $\hat{\overline{\overline{\mathcal{O}}}}$	4
$< \mathcal{O} >=< \psi \vert \mathcal{O}_{op} \vert \psi >$	definition of $\mathcal{O}_{op}$	4.3
$\hat{\overline{\mathcal{O}}} = \mathcal{O}$	none	4
$(\hat{a} + \hat{b})\psi \equiv \hat{a}\psi + \hat{b}\psi$	definition	3.2.1
$\hat{a} + \hat{b} = \hat{b} + \hat{a}$	none	3.2.1
$\hat{a} + (\hat{b} + \hat{c}) = (\hat{a} + \hat{b}) + \hat{c}$	none	3.2.1
$(\hat{a}\hat{b})\psi \equiv \hat{a}(\hat{b}\psi)$	definition	3.2.2
$(\hat{a}\hat{b}) \neq (\hat{b}\hat{a})$	in general	3.2.2
$(\hat{a}\hat{b})\hat{c} = \hat{a}(\hat{b}\hat{c})$	none	3.2.2
$(\hat{a} + \hat{b})\hat{c} = \hat{a}\hat{c} + \hat{b}\hat{c}$	none	3.2.2
$\hat{c}(\hat{a} + \hat{b}) = \hat{c}\hat{a} + \hat{c}\hat{b}$	$\hat{c}$ linear	4.1
$(\hat{a})^n \equiv \prod_{i=1}^{n} \hat{a}$	definition	3.2.2
$[\hat{a}, \hat{b}] \equiv \hat{a}\hat{b} - \hat{b}\hat{a}$	definition	3.3
$[\hat{x}, \hat{p}] = [\hat{\tilde{x}}, \hat{\tilde{p}}] = i\hbar$	none	3.3
$\hat{a}(f + g) = \hat{a}f + \hat{a}f$	$\hat{a}$ linear (definition)	4.1
$\hat{a} + \hat{b}$ linear	$\hat{a}$ and $\hat{b}$ linear	4.1
$\hat{a}\hat{b}$ linear	$\hat{a}$ and $\hat{b}$ linear	4.1
$(\overline{\overline{\psi}}, \hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\phi}}) = (\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\psi}}, \overline{\overline{\phi}})$	$\hat{\overline{\overline{\mathcal{O}}}}$ Hermitian (definition)	4.2
$\hat{\overline{\overline{\mathcal{O}}}}$ is Hermitian	$\mathcal{O}$ is a physical observable	4.2
$\hat{a}\hat{b}$ Hermitian	$\hat{a}$ and $\hat{b}$ Hermitian *and* $[\hat{a}, \hat{b}] = 0$	4.2
$\mathcal{O}_{op} \vert \mathcal{O}_0 >= \mathcal{O}_0 \vert \mathcal{O}_0 >$	none	4.3

2 Quantum Averages

2.1 Average Position

We shall begin with the calculation of the average observed position of a particle. Suppose that we are given the wave function, or quantum amplitudes, $\psi(x)$ for a system in the position representation and we wish to compute the average position of the particle after many measurements. Within the framework we have established so far (provided that $\psi(x)$ is properly normalized), we would compute simply

$$< x >= \int dx\, \mathcal{P}(x) x = \int dx\, |\psi(x)|^2 x \tag{1}$$

This is a perfectly valid form for computing averages of x; but because it depends so much on the position representation, it is not convenient to work with if we desire a very general framework for computing averages, one which we may use no matter in which representation we are working.

As we have already seen from Parseval's theorem, however, we may write integrals in a *representation-independent* manner by writing them as Hermitian inner products $(\psi, \phi) \equiv \int dx\, \psi^*(x)\phi(x) = (\tilde{\psi}, \tilde{\phi}) \equiv \int dk\, \tilde{\psi}^*(k)\tilde{\phi}(k)$. Accordingly, a much better way to write (1) is as

$$< x >= \int dx\, \psi^*(x)\psi(x) = \int dx\, \psi^*(x) \cdot x\psi(x)$$

$$< x >= (\psi, x\psi). \tag{2}$$

This is clearly the same form as (1). But, now if we had been working in the momentum representation and wanted to compute $< x >$, instead of first doing the Fourier integral $\psi = \mathcal{F}[\tilde{\psi}]$ to find ψ and then using (1), we may now immediately apply Parseval's theorem to (2)

$$< x >= (\tilde{\psi}, \tilde{\phi}) \text{ where } \phi(x) = x\psi(x).$$

Note that here we have the same mathematical structure as (2), $< x >$ is given by the inner product of the wave function ($\tilde{\psi}$) with some other function ($\tilde{\phi}$) which is derived from the wave function through some well-defined procedure. In (2) the "procedure" is simply to multiply by x, the argument of the wave function in the representation. Here, the "procedure" is more complicated and involves three phases. We begin with $\tilde{\psi}$ and 1) Fourier transform it to obtain $\psi = \mathcal{F}[\tilde{\psi}]$. We then 2) multiply by x to get $\phi(x) = x\psi(x) = x\mathcal{F}[\tilde{\psi}]$. Finally, 3) we inverse Fourier transform back to get $\tilde{\phi} = \mathcal{F}^{-1}[\phi] = \mathcal{F}^{-1}\left[x\mathcal{F}[\tilde{\psi}]\right]$. Fortunately, the effect of this entire complex procedure on the function $\tilde{\psi}(k)$ is very simple,

$$\begin{aligned}
\psi(x) &= \int \frac{dk'}{\sqrt{2\pi}} e^{ik'x}\tilde{\psi}(k') \\
\phi(x) = x\psi(x) &= x\int \frac{dk'}{\sqrt{2\pi}} e^{ik'x}\tilde{\psi}(k') \\
\tilde{\phi}(k) = \int \frac{dx}{\sqrt{2\pi}} e^{-ikx}\phi(x) &= \int \frac{dx}{\sqrt{2\pi}} e^{-ikx} x \int \frac{dk'}{\sqrt{2\pi}} e^{ik'x}\tilde{\psi}(k') \\
&= \int \frac{dx}{\sqrt{2\pi}} (i\partial_k e^{-ikx}) \int \frac{dk'}{\sqrt{2\pi}} e^{ik'x}\tilde{\psi}(k') \\
&= i\partial_k \left[\int \frac{dx}{\sqrt{2\pi}} e^{-ikx} \int \frac{dk'}{\sqrt{2\pi}} e^{ik'x}\tilde{\psi}(k')\right] \qquad (*) \\
&= i\partial_k \left[\int dk'\, \tilde{\psi}(k') \int \frac{dx}{2\pi} e^{ix(k'-k)}\right] \\
&= i\partial_k \left[\int dk'\, \tilde{\psi}(k')\delta(k'-k)\right] \\
\tilde{\phi}(k) &= i\partial_k \tilde{\psi}(k).
\end{aligned}$$

In step (*) we used the fact that only the e^{-ikx} term depends on k in order to pull the derivative $i\partial_k$ out of the integral. After writing (*) we could have recognized that the expression inside the square brackets is the expression to transform $\tilde{\psi}(k)$ and then inverse transform the result, which then just returns $\tilde{\psi}(k) = \mathcal{F}^{-1}\left[\mathcal{F}[\tilde{\psi}]\right]$. We proceeded directly, instead, to review for the student the use of the Dirac $\delta-$function.

With ϕ in hand, we now have two equivalent ways of writing $< x >$, one for each representation,

$$\begin{aligned}
< x >= (\psi, \phi) &= (\psi, x\psi) \\
< x >= (\tilde{\psi}, \tilde{\phi}) &= (\tilde{\psi}, i\partial_k \tilde{\psi})
\end{aligned} \qquad (3)$$

In both cases $< x >$ is just given by the Hermitian inner product of the wave function (ψ or $\tilde{\psi}$) with some function derived from the wave function through some mathematical procedure or "operation". As we shall see in section 3, this will lead us to define "function operators." First, however, we will complete our discussion of averages with expressions for averages of momentum, the other primary observable for this course.

2.2 Average Momentum

The average momentum $< p >=< \hbar k >= \int dk\, \tilde{\mathcal{P}}(k)\hbar k = \int dk\, \tilde{\psi}^*(k) \cdot \hbar k \tilde{\psi}(k)$ also clearly fits within the same framework,

$$< p >= (\tilde{\psi}, \tilde{\theta}) \equiv (\tilde{\psi}, \hbar k \tilde{\psi}).$$

Or, in the position representation

$$< p >= (\tilde{\psi}, \tilde{\theta}) = (\psi, \theta),$$

where $\theta(x)$ is just a function derived from $\psi(x)$ through some mathematical operation. In this case, the operation begins with $\psi(x)$ and proceeds as

$$\begin{aligned}
\tilde{\psi}(k) &= \int \frac{dx'}{\sqrt{2\pi}} e^{-ikx'} \psi(x') \\
\tilde{\theta}(k) &= \hbar k \tilde{\psi}(k) = \hbar k \int \frac{dx'}{\sqrt{2\pi}} e^{-ikx'} \psi(x') \\
\theta(x) &= \int \frac{dk}{\sqrt{2\pi}} e^{ikx} \hbar k \int \frac{dx'}{\sqrt{2\pi}} e^{-ikx'} \psi(x') \\
&= \int \frac{dk}{\sqrt{2\pi}} \left(\frac{\hbar}{i}\partial_x e^{ikx}\right) \int \frac{dx'}{\sqrt{2\pi}} e^{-ikx'} \psi(x') \\
&= \frac{\hbar}{i}\partial_x \left[\int \frac{dk}{\sqrt{2\pi}} e^{ikx} \int \frac{dx'}{\sqrt{2\pi}} e^{-ikx'} \psi(x')\right] \\
&= \frac{\hbar}{i}\partial_x \left[\mathcal{F}\left[\mathcal{F}^{-1}[\psi]\right]\right] \\
&= \frac{\hbar}{i}\partial_x \psi(x).
\end{aligned}$$

So, our two equivalent ways (one for each representation) of writing $< p >$ are

$$\begin{aligned}
< p >= (\tilde{\psi}, \tilde{\theta}) &= (\tilde{\psi}, \hbar k \tilde{\psi}) \\
< p >= (\psi, \theta) &= (\psi, \frac{\hbar}{i}\partial_x \psi).
\end{aligned} \quad (4)$$

3 Function Operators and Operator Algebra

In each of the four cases presented in (3) and (4), we find that the average of an observable is computed as the inner product of the wave function with some other function derived from the wave function through some mathematical operation. We will see in section (4) that this structure for the calculation of averages is *completely general* in quantum mechanics.

In this section we will focus on the mathematics of operations performed on functions to produce other functions. We will call such an operation a *function operator*. A function operator is a generalization of our familiar notion of a function. We normally think of a function as a "map" which associates a single real number as output for every real number given to the function as input. However, there is nothing in the idea of a function that requires its input and output to be real numbers. They could be complex numbers or the input and output could be functions themselves. A function operator is then just a special case of this general notion of function. A function operator accepts functions as input, and maps each input function to a single unique function as output. We call these special functions "function" operators to remind us that as input and output, they accept and give functions rather than numbers. Later, in section (4) we shall define *quantum* operators, which map not functions but *quantum states*.

In analogy with the usual notation for the operation of a function f on a real number x

$$y = f(x),$$

we introduce the following notation for the operation of an operator $\hat{\mathcal{O}}$ on a function $\phi(x)$,

$$\phi(x) = \hat{\mathcal{O}}\psi(x).$$

The caret "^" tells us that $\hat{\mathcal{O}}$ is a function operator which acts on the function immediately to its right, $\psi(x)$, giving the result $\phi(x)$.

3.1 Three Simple Operators

There are three simplest forms for function operators which we shall see over and over again in this course, the constant operator, the argument multiplication operator and the differential operator. These are described in detail in the three paragraphs below.

One of the simplest function operators with which we shall deal is just multiplication of the input function by some constant, c. We would then write this operator as $\hat{c} = c$ so that the action of $\hat{c}$ is just

$$\hat{c}\psi(x) \equiv c\psi(x).$$

Note that, with the hat "on," $\hat{c}$ is an operator whereas with the hat "off" c is just a constant number which is multiplying $\psi(x)$. Two very importation constant operators are the zero operator $\hat{0}$, whose operation on any function returns the zero function, $\hat{0}\psi(x) = 0$, and the unit operator $\hat{1}$ whose action on any function is to "do nothing" or just return the input function back as output, $\hat{1}\psi(x) = 1 \cdot \psi(x) = \psi(x)$.

We saw an example of our next more complicated operator in our first expression for $< x >= (\psi, \theta) = (\psi, x\psi)$. Here $\theta(x)$ is derived from $\psi(x)$ by multiplication of $\psi(x)$ by its argument x. For this new function operator, we write $\hat{x} = x$ so that the action of $\hat{x}$ on $\psi(x)$ is just

$$\hat{x}\psi(x) \equiv x\psi(x).$$

We now may rewrite our expression for the average position as $< x >= (\psi, \hat{x}\psi)$. The same type of operation occurs in the calculation of the average momentum *within the momentum representation.* In this case we define $\hat{\tilde{p}} \equiv \hbar k$ so that

$$\hat{\tilde{p}}\tilde{\psi}(k) \equiv \hbar k\tilde{\psi}(k),$$

and $< p >= (\tilde{\psi}, \hat{\tilde{p}}\tilde{\psi})$. Note that we put the tilde "$\tilde{\ }$" on the p below the hat to remind us that this operator is defined for use in the *momentum* representation.

We also had an expression for the average position computed within the *momentum* representation, $< x >= (\tilde{\psi}, \tilde{\theta}) = (\tilde{\psi}, i\partial_k\tilde{\psi})$. Here, $\tilde{\theta}(k) \equiv i\partial_k\tilde{\psi}(k)$ is derived from $\tilde{\psi}(k)$ using the yet more complicated operation of differentiation. We now write $\hat{\tilde{x}}$ as a *differential operator* $\hat{\tilde{x}} = i\partial_k$, so that the action of $\hat{\tilde{x}}$ is just

$$\hat{\tilde{x}}\tilde{\psi}(k) \equiv i\partial_k\tilde{\psi}(k),$$

and the average position computed in the momentum representation is $< x >= (\tilde{\psi}, \hat{\tilde{x}}\tilde{\psi})$. (Again, we put the tilde "$\tilde{\ }$" on the x to remind us that this operator is defined for use in the *momentum* representation.) Finally, to complete the set of average-operator correspondences, for computing the average momentum from within the position representation we define a second differential operator $\hat{p} = \frac{\hbar}{i}\partial_x$ so that

$$\hat{p}\psi(x) \equiv \frac{\hbar}{i}\partial_x\psi(x),$$

and $< p >= (\psi, \hat{p}\psi)$.

3.2 Combining Simple Operators: Operator Algebra

The utility of defining function operators as mathematical entities is that we may then use familiar algebraic rules to manipulate expressions involving our operators and to produce a new, richer set of operators from our known basic operators by combining them into more complex forms. The two basic algebraic operations for combining two operators are through addition of the results of their application (operator addition) or through their application *in sequence* (operator multiplication).

3.2.1 Operator Sum

Given any two operators $\hat{a}$ and $\hat{b}$, we may define a third operator $\hat{c} \equiv \hat{a} + \hat{b}$, the operator sum of $\hat{a}$ and $\hat{b}$, whose operation is defined as applying $\hat{a}$ and $\hat{b}$ separately and then summing the results,

$$\begin{aligned} \hat{c}\psi(x) = (\hat{a} + \hat{b})\psi(x) &\equiv (\hat{a}\psi(x)) + (\hat{b}\psi(x)) \\ (\hat{a} + \hat{b})\psi(x) &= \hat{a}\psi(x) + \hat{b}\psi(x) \end{aligned} \tag{5}$$

Note that, this definition is the same a saying that operator addition exhibits the usual algebraic distributive property when acting on functions to the right. From the definition of the operator sum and the usual associative property of addition, we may show that function operator addition is itself associative,

$$\begin{aligned}\left((\hat{a}+\hat{b})+\hat{c}\right)\psi(x) &= \left(\hat{a}+\hat{b}\right)\psi(x)+\hat{c}\psi(x)\\ &= \left(\hat{a}\psi(x)+\hat{b}\psi(x)\right)+\hat{c}\psi(x)\\ &= \hat{a}\psi(x)+\left(\hat{b}\psi(x)+\hat{c}\psi(x)\right)\\ &= \hat{a}\psi(x)+\left(\hat{b}+\hat{c}\right)\psi(x)\\ &= \left(\hat{a}+(\hat{b}+\hat{c})\right)\psi(x).\end{aligned}$$

Since the result of the operation of $\left((\hat{a}+\hat{b})+\hat{c}\right)$ and $\left(\hat{a}+(\hat{b}+\hat{c})\right)$ is the same on all functions $\psi(x)$, these two operators must be equal,

$$(\hat{a}+\hat{b})+\hat{c}=\hat{a}+(\hat{b}+\hat{c}) \tag{6}$$

Similarly, one may show that operator addition is commutative,

$$\begin{aligned}(\hat{a}+\hat{b})\psi(x) &\equiv \hat{a}\psi(x)+\hat{b}\psi(x)\\ &= \hat{b}\psi(x)+\hat{a}\psi(x)\\ &\equiv (\hat{b}+\hat{a})\psi(x)\\ \Rightarrow \hat{a}+\hat{b} &= \hat{b}+\hat{a} \qquad (7)\end{aligned}$$

Finally, the zero operator mentioned in the previous subsection is the additive identity. (One may easily show following the same procedure as in the proofs above that $\hat{a}+\hat{0}=\hat{a}$.) With an additive identity, subtraction of function operators may be defined in the familiar way.

3.2.2 Operator Product

One could now consider defining multiplication of operators in the same way, with the result of the operation of the operator to be the product of the result of the application of each of the operators separately. However, a much more common combination of two operators is to apply them separately in a sequence of two successive operations. Amazingly, this second definition preserves nearly all of the familiar algebra of multiplication! We will proceed with this definition and define the product $\hat{c}=\hat{a}\hat{b}$ of two operators to be the result of their sequential application,

$$\hat{c}\psi(x)=(\hat{a}\hat{b})\psi(x)\equiv\hat{a}\left(\hat{b}\psi(x)\right). \tag{8}$$

This procedure is sensible since $\hat{b}\psi(x)$ returns a new *function* which may then be fed to $\hat{a}$ as input to produce the final output defined as $\hat{c}\psi(x)$.

Like the familiar multiplication of numbers, the function operator product is also associative, as we see by studying the action of a three-way product on an arbitrary function $\psi(x)$,

$$\begin{aligned}\left(\hat{a}(\hat{b}\hat{c})\right)\psi(x) &\equiv \hat{a}\left((\hat{b}\hat{c})\psi(x)\right)\\ &= \hat{a}\left(\hat{b}(\hat{c}\psi(x))\right)\\ &= (\hat{a}\hat{b})(\hat{c}\psi(x))\\ &= \left((\hat{a}\hat{b})\hat{c}\right)\psi(x)\\ \Rightarrow \hat{a}(\hat{b}\hat{c}) &= (\hat{a}\hat{b})\hat{c}. \qquad (9)\end{aligned}$$

As the operator product is associative, without ambiguity we may write the product $\hat{a}\hat{b}\hat{c}$ without parenthesis. As with normal multiplication, raising an operator to an integer power is just a short-hand for repeated application of the operator,

$$(\hat{a})^n \equiv \prod_{i=1}^{n} \hat{a} \quad = \quad \hat{a}\hat{a}...\hat{a} \ (n \text{ terms}) \tag{10}$$

Note that in our notation there is a subtle difference between $(\hat{a})^n$ and $\hat{a^n}$. The former is the product of n factors of $\hat{a}$ as defined in (10) above, while the later is a single operator, defined physically so that the quantum average $< a^n >= (\psi, \hat{a^n}\psi)$ is given correctly. Although, they will turn out to be equal, the definitions of $(\hat{a})^n$ and $\hat{a^n}$ are entirely different.

There is one key property which our operator product does not share with common multiplication. Defined as the *sequential* application of two operations, there is no reason to expect that the operator product be commutative, that the net affect of two operations be independent of the sequence in which they are applied. Symbolically, we expect in general that $\hat{a}\hat{b} \neq \hat{b}\hat{a}$. This property of operators is closely related to the Heisenberg uncertainty principle. The drive to give this basic principle of quantum physics a direct expression in our mathematical formalism is a key force in our choice of the definition of operator product in terms of sequential application. Because of the prime importance of the non-commutivity of the operator product, we will dedicate the entire next section to the issue of commutivity.

Before discussing non-commutivity, there are two final properties of the operator product which mimic normal multiplication and with which the student should be familiar. First, as with common multiplication, there is an operator multiplicative identity. It is just the constant operator $\hat{1}$ discussed in the previous subsection. It is easily verified, using the procedures above that $\hat{1}\hat{a} = \hat{a}\hat{1} = \hat{a}$. (Note that one must consider both a left and a right identity as the operator product is not commutative.) One may also show that multiplication by $(\hat{-1})$ yields the additive inverse of an operator, $(\hat{a} + ((\hat{-1})\hat{a}) = \hat{0}$.

The final familiar property of the operator product is the distributive property,

$$\begin{aligned} (\hat{a}+\hat{b})\hat{c}\psi(x) &\equiv (\hat{a}+\hat{b})(\hat{c}\psi(x)) \\ &\equiv \hat{a}(\hat{c}\psi(x)) + \hat{b}(\hat{c}\psi(x)) \\ &= (\hat{a}\hat{c})\psi(x) + \hat{b}\hat{c})\psi(x) \\ &= (\hat{a}\hat{c} + \hat{b}\hat{c})\psi(x) \\ \Rightarrow (\hat{a}+\hat{b})\hat{c} &= \hat{a}\hat{c} + \hat{b}\hat{c} \end{aligned} \tag{11}$$

The "commuted" form of this statement, $\hat{a}(\hat{b}+\hat{c}) = \hat{a}\hat{b} + \hat{a}\hat{c}$, does not follow directly. While true for quantum operators, it is not true for general operators and depends on a property known as *linearity* which we shall discuss in detail in section (4.1).

3.3 Commutators

As alluded to above, the non-commutivity of operators is closely related to a very basic principle of quantum physics, the Heisenberg uncertainty principle. Non-commutivity is so central to our theory that we define a special operation on operator, the *commutator* to measure the "degree of non-commutivity" between two operators $\hat{a}$ and $\hat{b}$,

$$[\hat{a}, \hat{b}] \equiv \hat{a}\hat{b} - \hat{b}\hat{a}. \tag{12}$$

If $[\hat{a}, \hat{b}] = \hat{0}$, then $\hat{a}\hat{b} = \hat{b}\hat{a}$, and we say "$\hat{a}$ and $\hat{b}$ commute;" otherwise, $[\hat{a}, \hat{b}]$ gives us a measure of the error we make if we assume $\hat{a}$ and $\hat{b}$ do commute. As an example, consider our operators $\hat{x} \equiv x$ and $\hat{p} \equiv \frac{\hbar}{i}\partial_x$. Then,

$$[\hat{x}, \hat{p}]\psi(x) \quad \equiv \quad (\hat{x}\hat{p} - \hat{p}\hat{x})\psi(x)$$

$$
\begin{aligned}
&= \hat{x}(\hat{p}\psi(x)) - \hat{p}(\hat{x}\psi(x)) \\
&= x(\frac{\hbar}{i}\partial_x\psi(x)) - \frac{\hbar}{i}\partial_x(x\psi(x)) \\
&= \frac{\hbar}{i}x\partial_x\psi(x) - \frac{\hbar}{i}(\partial_x x)\psi(x) - \frac{\hbar}{i}x\partial_x\psi(x) \\
&= i\hbar\psi(x) \\
\Rightarrow [\hat{x},\hat{p}] &= i\hbar \qquad (13)
\end{aligned}
$$

Thus $\hat{x}$ and $\hat{p}$ do not commute and the *extent* of their non-commutivity is measured by $i\hbar$. The extent of this non-commutivity is directly related to the Heisenberg uncertainty principle $\Delta x\,\Delta p \geq \hbar/2$. Note that in the classical limit ($\hbar \to 0$), $[\hat{x},\hat{p}] \to 0$, $\hat{x}$ and $\hat{p}$ do commute. This corresponds directly to the fact that in classical physics there is no limit to how precisely x and p may be defined so that $\Delta x\,\Delta p \to 0$.

As a final exercise, we verify that we get this same result for the commutator when carried out in the momentum representation,

$$
\begin{aligned}
[\hat{\tilde{x}},\hat{\tilde{p}}]\tilde{\psi}(k) &\equiv (\hat{\tilde{x}}\hat{\tilde{p}} - \hat{\tilde{p}}\hat{\tilde{x}})\tilde{\psi}(k) \\
&= i\partial_k(\hbar k\tilde{\psi}(k)) - \hbar k i\partial_k\tilde{\psi}(k) \\
&= i\hbar\tilde{\psi}(k) \\
\Rightarrow [\hat{\tilde{x}},\hat{\tilde{p}}] &= i\hbar \quad (!)
\end{aligned}
$$

4 Quantum Operators

In section (2), we found inner product expressions for the average position $< x >$ and momentum $< p >$ in terms of the quantum amplitudes or "wave functions" in either the position or momentum representation. In all cases we found an expression of the form

$$< \mathcal{O} >= (\overline{\overline{\psi}}, \hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\psi}}), \qquad (14)$$

where $\mathcal{O}$ is the physical observable begin averaged, $\overline{\overline{\psi}}$ are the quantum amplitudes in the representation and $\hat{\overline{\overline{\mathcal{O}}}}$ is a function operator designed to give the correct average when used in the corresponding representation. Here the double over-bar "=" refers to a *generic* representation which could either be position (for which we usually give no over-symbol) or momentum (for which we usually give a tilde " "). With equation (14) and knowledge of the correct operator, we may compute the average of position or momentum in either the position or the momentum representation. The table below summarizes the operators we have found so far:

TABLE III: OPERATORS DERIVED IN THIS NOTE

$< \mathcal{O} >= (\overline{\overline{\psi}}, \hat{\overline{\overline{c\mathcal{O}}}}\,\overline{\overline{\psi}})$	Observable x	Observable p
Position ($\psi(x)$)	$\hat{x} = x$	$\hat{p} = \frac{\hbar}{i}\partial_x$
Momentum ($\tilde{\psi}(k)$)	$\hat{\tilde{x}} = i\partial_k$	$\hat{\tilde{p}} = \hbar k$

As we shall see below, in each and every representation $\overline{\overline{\psi}}(x)$, we may find a (unique) operator $\hat{\overline{\overline{\mathcal{O}}}}$ associate with each and every observable $\mathcal{O}$ of a system so that the form (14) is valid. We call this the *physical operator* associated with the observable, and represent the operator as the symbol for the observable (e.g., x for position, p for momentum, L_z for the z component of angular momentum, H for the energy, or $\mathcal{O}$ for a generic operator, et cetera) with an over-symbol telling the appropriate representation (e.g., no symbol for the position representation, an over-tilde "~" for the momentum representation, an over-bar "-" for the representation associated with pure states of the observable itself and a double over-bar "=" for a generic, unspecified representation) with a final caret or "hat" ("^") on top to remind us that we are dealing with a function operator.

Definition: The operator associated with an observable in a given representation is constructed so that averages of the observable are always given by $<\mathcal{O}>=(\overline{\overline{\psi}},\hat{\overline{\overline{c\mathcal{O}}}}\,\overline{\overline{\psi}})$*, We refer to all operators defined with respect to physical observables in this way as* physical *or* observable *operators.*

All operators in quantum mechanics are either associated directly with physical observables or are constructed from such physical operators using the algebraic rules laid out in section (3). It is thus important to understand the special properties of physical operators. This is the purpose of this section.

To see how the average of any observable in any representation takes the generic form

$$<\mathcal{O}>=(\overline{\overline{\psi}},\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\psi}})$$

we begin by considering the form for the average in the representation of pure states of the associated observable. In this representation, the average clearly has the above form. We will then argue that the above form stays invariant as we move from representation to representation so that the form above is completely general.

To determine the form of operators in their pure state representation, we begin with the principle of superposition and represent the state $|\psi>$ as a combination of pure states of the observable $\mathcal{O}$,

$$|\psi>=\int d\mathcal{O}\,\overline{\psi}(\mathcal{O})|\mathcal{O}>.$$

Here $\overline{\psi}(\mathcal{O})$ are the quantum amplitudes (wave function) for the state in the $\mathcal{O}$ representation so that the probability of measuring the value $\mathcal{O}$ of the observable in an experiment is $\overline{\mathcal{P}}(\mathcal{O})=|\overline{\psi}(\mathcal{O})|^2$. Thus, in the pure representation of any observable,

$$<\mathcal{O}>=\int d\mathcal{O}\,\mathcal{P}(\mathcal{O})\mathcal{O}=\int d\mathcal{O}\,\overline{\psi}^*(\mathcal{O})\cdot\mathcal{O}\overline{\psi}(\mathcal{O})=(\overline{\psi},\mathcal{O}\overline{\psi})\equiv(\overline{\psi},\hat{\overline{\mathcal{O}}}\,\overline{\psi}). \tag{15}$$

We see that in the pure representation we always have that the operator associated with the observable looks like argument multiplication,

$$\hat{\overline{\mathcal{O}}}=\mathcal{O}\quad\text{(in general)}.$$

(We have already seen two examples of this, $\hat{\overline{x}}=x$ and $\hat{\overline{p}}=\hbar k$.)

To transform our expression (15) for general averages to other representations, we write $\overline{\phi}=\mathcal{O}\overline{\psi}$ and apply Parseval's theorem as we did in section (2),

$$<\mathcal{O}>=(\overline{\psi},\overline{\phi})=(\overline{\overline{\psi}},\overline{\overline{\phi}})\equiv(\overline{\overline{\psi}},\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\psi}}),$$

where the operation of $\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\psi}}$ on $\overline{\overline{\psi}}$ proceeds in three well-defined mathematical steps whose application in sequence defines the operator $\hat{\overline{\overline{\mathcal{O}}}}$: 1) transform $\overline{\overline{\psi}}$ to the pure state representation to generate $\overline{\psi}$, 2) multiply $\overline{\psi}(\mathcal{O})$ by $\mathcal{O}$ to make $\overline{\phi}=\mathcal{O}\overline{\psi}(\mathcal{O})$, and 3) transform $\overline{\phi}$ back to the initial representation to get $\overline{\overline{\phi}}\equiv\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\psi}}$. Although the procedure defining $\hat{\overline{\overline{\mathcal{O}}}}$ is complicated and involved, it may be applied to any function $\overline{\overline{\psi}}$ as input and will yield a single $\overline{\overline{\psi}}=\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\psi}}$ as output. This procedure thus qualifies as a valid definition of an operator. As we have seen with position and momentum, although the defining procedure is involved, it often results in a simple final operation on $\overline{\overline{\psi}}$ to yield $\overline{\overline{\phi}}$. In the following sections we shall see further that this definition results in some simple buy very important and general properties for the operators we will encounter in quantum mechanics.

4.1 Linearity

Defined according to the procedure in the discussion above, all of our physical operators have the mathematical property of linearity. This means that the action of a physical operator on a sum of functions is always equal to the sum of its action on each of the operators individually and that the action of an operator on a constant times a function is just that constant times the action of the operator on the function alone

$$\begin{aligned} \hat{a}(f+g) &= \hat{a}f + \hat{a}g \\ \hat{a}(cf) &= c\hat{a}f \end{aligned} \tag{16}$$

Moreover, all operators formed using the algebraic combinations of operator addition and multiplication from linear operators are also linear and thus all of the operators, with a single notable exception, you will encounter in quantum mechanics will be linear. The one exception to this is the "time reversal" operator, which you will learn about in a later course.

From the form of our physical operators, they are all clearly linear in the pure state representation associated with the operator, where the action of the operator is just multiplication by the argument of the function. Once we establish this, we will argue again that as we move to other representations, this statement preserves its mathematical form and so the property of linearity holds for any physical operator in any representation. Now, in the pure state representation, the operator clearly distributes over the sum of the functions,

$$\begin{aligned} \hat{\overline{\mathcal{O}}}(\overline{f}+\overline{g}) &\equiv \mathcal{O}(\overline{f}(\mathcal{O})+\overline{g}(\mathcal{O})) \\ &= \mathcal{O}\overline{f}(\mathcal{O})+\mathcal{O}\overline{g}(\mathcal{O}) \\ \Rightarrow \hat{\overline{\mathcal{O}}}(\overline{f}+\overline{g}) &= \hat{\overline{\mathcal{O}}}\,\overline{f}(\mathcal{O})+\hat{\overline{\mathcal{O}}}\,\overline{g}(\mathcal{O}) \\ \hat{\overline{\mathcal{O}}}(c\overline{f}) &\equiv \mathcal{O}(c\overline{f}) \\ &= c\mathcal{O}\overline{f} \\ \Rightarrow \hat{\overline{\mathcal{O}}}(c\overline{f}) &= c\hat{\overline{\mathcal{O}}}\,\overline{f} \end{aligned} \tag{17}$$

Because of the manner in which we have defined the action of physical observables in representations other than that of their pure states, when we transform the above equation to any other representation, it looks the same. If we take $\overline{h} \equiv \overline{f}+\overline{g}$, by the principle of superposition we will have $\overline{\overline{h}} = \overline{\overline{f}}+\overline{\overline{g}}$ and if $\overline{F} \equiv c\overline{f}$ then $\overline{\overline{F}} = c\overline{\overline{f}}$. Further, by the definition of our operators $\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{h}}$ will be the transform of $\hat{\overline{\mathcal{O}}}\,\overline{h}$, $\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{f}}$ the transform of $\hat{\overline{\mathcal{O}}}\,\overline{f}$ and $\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{g}}$ the transform of $\hat{\overline{\mathcal{O}}}\,\overline{g}$. Thus, the transform of (17) is just the general statement we wish to prove for any physical operator in an arbitrary representation,

$$\begin{aligned} \hat{\overline{\overline{\mathcal{O}}}}(\overline{\overline{f}}+\overline{\overline{g}}) &= \hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{f}}+\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{f}} \qquad \text{Linearity} \\ \hat{\overline{\overline{\mathcal{O}}}}(c\overline{\overline{f}}) &= c\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{f}} \end{aligned} \tag{18}$$

It is easily verified that the operators we have introduced so far are linear.

$$\begin{aligned} \hat{c}(f+g) &\equiv c(f(x)+g(x)) \\ &= cf(x)+cg(x) \\ &= \hat{c}f(x)+\hat{c}g(x) \\ \hat{c}(af) &\equiv c(af(x)) \\ &= acf(x) \\ &= a\hat{c}f(x) \\ \hat{x}(f+g) &\equiv x(f(x)+g(x)) \\ &= xf(x)+xg(x) \\ &= \hat{x}f(x)+\hat{x}g(x) \end{aligned}$$

$$
\begin{aligned}
\hat{x}(af) &\equiv x(af(x)) \\
&= axf(x) \\
&= a\hat{x}f(x) \\
\hat{p}(f+g) &\equiv \frac{\hbar}{i}\partial_x(f(x)+g(x)) \\
&= \frac{\hbar}{i}\partial_x f(x) + \frac{\hbar}{i}\partial_x g(x) \\
&= \hat{p}f(x) + \hat{p}g(x) \\
\hat{p}(af(x)) &\equiv \frac{\hbar}{i}\partial_x(af(x)) \\
&= a\frac{\hbar}{i}\partial_x f(x) \\
&= a\hat{p}f(x)
\end{aligned}
$$

et cetera

A simple example of an operator which is <u>not linear</u> is the operator which add one to any function, $\hat{+}f(x) \equiv f(x) + 1$. We would then have

$$
\begin{aligned}
\hat{+}f + \hat{+}g &\equiv (f(x)+1) + (g(x)+1) \\
&= f(x) + g(x) + 2 \\
&\neq f(x) + g(x) + 1 \equiv \hat{+}(f+g)
\end{aligned}
$$

Because the $\hat{+}$ operator in not linear, we know immediately that it does not correspond to any physical observable and thus may immediately disqualify it as a *physical* operator.

Finally, it is is easily verified that starting with physically observable operators (which we have just demonstrated all to be linear) and forming new combinations by operator addition and multiplication, the result is always another linear operator. This follow directly from induction and the facts that that the sum and operator product of two linear operators is always a third linear operator. These facts are verified directly below.

$$
\begin{aligned}
(\hat{a}+\hat{b})(f+g) &\equiv \hat{a}(f+g) + \hat{b}(f+g) \\
&= \hat{a}f + \hat{a}g + \hat{b}f + \hat{b}g \\
&= \hat{a}f + \hat{b}f + \hat{a}g + \hat{b}g \\
&= (\hat{a}+\hat{b})f + (\hat{a}+\hat{b})g \\
(\hat{a}+\hat{b})(cf) &\equiv \hat{a}(cf) + \hat{b}(cf) \\
&= c\hat{a}f + c\hat{b}f \\
&= c(\hat{a}+\hat{b})f \\
(\hat{a}\hat{b})(f+g) &\equiv \hat{a}(\hat{b}(f+g)) \\
&= \hat{a}(\hat{b}f + \hat{b}g) \\
&= \hat{a}\hat{b}f + \hat{a}\hat{b}g \\
&= (\hat{a}\hat{b})f + (\hat{a}\hat{b})g \\
(\hat{a}\hat{b})(cf) &\equiv \hat{a}(\hat{b}(cf)) \\
&= \hat{a}(c\hat{b}f) \\
&= c\hat{a}\hat{b}f \\
&= c(\hat{a}\hat{b})f
\end{aligned}
$$

Thus we see that all operators which will concern us are linear. With this established, we may now freely exploit a second distributive law, which is companion to (11) and which follows directly from linearity,

$$
\begin{aligned}
\hat{a}(\hat{b}+\hat{c})f &\equiv \hat{a}((\hat{b}+\hat{c})f) \\
&= \hat{a}\left((\hat{b}f)+(\hat{c}f)\right) \\
&= \hat{a}(\hat{b}f)+\hat{a}(\hat{c}f) \quad \text{; from linearity} \\
&= \left((\hat{a}\hat{b})+(\hat{a}\hat{c})\right)f \\
\Rightarrow \hat{a}(\hat{b}+\hat{c}) &= \hat{a}\hat{b}+\hat{a}\hat{c} \quad \text{; provided } \hat{a} \text{ is linear.}
\end{aligned} \tag{19}
$$

4.2 Hermitianness

The second special property of observable operators which follows from our definition that physical operators give correct averages of observables is *Hermitianness.* We say that the operator $\hat{\mathcal{O}}$ has the property of *Hermitianness* or that it is *Hermitian* of we may apply the operator on either side of an inner product and always get the same result,

$$(f, \hat{\mathcal{O}}g) = (\hat{\mathcal{O}}f, g) \quad \text{for all } f \text{ and } g \tag{20}$$

Algebraically, we often wish to make this manipulation and it is good to know that it may be carried out with any observable operator. In this section we will prove that all operators associated with physical observables are Hermitian. Later you will learn that all linear, Hermitian operators are associated with physical observables.

Hermitianness of an operator is also an important concept physically. It provides a quick test of whether an operator is physically observable. For instance, we know that there is no physical observable associated with the product $\hat{x}\hat{p}$ because it is not Hermitian,

$$
\begin{aligned}
(\psi, \hat{x}\hat{p}\phi) - (\hat{x}\hat{p}\psi, \phi) &= (\psi, \hat{x}\hat{p}\phi) - (\hat{p}\psi, \hat{x}\phi) \quad \text{; Hermitianness of } \hat{x} \\
&= (\psi, \hat{x}\hat{p}\phi) - (\psi, \hat{p}\hat{x}\phi) \quad \text{; Hermitianness of } \hat{p} \\
&= (\psi, (\hat{x}\hat{p} - \hat{p}\hat{x})\phi) \quad \text{; linearity of Hermitian inner product} \\
&= (\psi, [\hat{x}, \hat{p}]\phi) \quad (*) \\
&= (\psi, i\hbar\phi) \\
&= i\hbar(\psi, \phi) \quad \text{; linearity of Hermitian inner product} \\
&\neq 0 \quad \text{(in general)} \\
\Rightarrow (\psi, \hat{x}\hat{p}\psi) &\neq (\hat{x}\hat{p}\psi, \psi) \quad \text{(in general)}
\end{aligned}
$$

The fact that the product $\hat{x}\hat{p}$ is not an observable is directly related to the Heisenberg uncertainty principle, which prevents the *simultaneous* measurement of both x and p. Note also that our analysis up to step (*) depended only on the Hermitianness of $\hat{x}$ and $\hat{p}$. If $\hat{x}$ and $\hat{p}$ were commuting operators, then our result would have been zero for all ψ and ϕ implying that the product of two Hermitian operators is Hermitian if and only if they commute. This example demonstrates the important fact that although the operator product of two linear operators is *always* a third linear operator, the product of two Hermitian operators *need not be* Hermitian (unless they commute).

To see how Hermitianness follow from our definition of quantum observable operators, we proceed as we did with linearity. First we see that the statement is clearly true in the pure state representation of the associated observable and then we argue that the statement maintains its form in any representation. First, in the pure state representation,

$$
\begin{aligned}
(\overline{\psi}, \overline{\hat{\mathcal{O}}}\,\overline{\phi}) &\equiv \int d\mathcal{O}\, \overline{\psi}^*(\mathcal{O}) \cdot \overline{\hat{\mathcal{O}}}\,\overline{\phi}(\mathcal{O}) = \int d\mathcal{O}\, \overline{\psi}^*(\mathcal{O}) \cdot \mathcal{O}\overline{\phi}(\mathcal{O}) \\
&= \int d\mathcal{O}\, \left(\mathcal{O}\overline{\psi}(\mathcal{O})\right)^* \overline{\phi}(\mathcal{O}) \quad \text{; } (\mathcal{O}\text{, being physical, is a real number!}) \\
&= \int d\mathcal{O}\, \left(\overline{\hat{\mathcal{O}}}\,\overline{\psi}(\mathcal{O})\right)^* \overline{\phi}(\mathcal{O}) \\
&= (\overline{\hat{\mathcal{O}}}\,\overline{\psi}, \overline{\psi})
\end{aligned}
$$

Now, by Parseval's Theorem, if $\overline{\eta} \equiv \hat{\overline{\mathcal{O}}}\,\overline{\phi}$, then $(\overline{\psi}, \hat{\overline{\mathcal{O}}}\,\overline{\phi}) \equiv (\overline{\psi}, \overline{\eta}) = (\overline{\overline{\psi}}, \overline{\overline{\eta}})$. By our definition of $\hat{\overline{\overline{\mathcal{O}}}}$, $\overline{\overline{\eta}} \equiv \hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\phi}}$, and so $(\overline{\overline{\psi}}, \hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\phi}}) \equiv (\overline{\psi}, \hat{\overline{\mathcal{O}}}\,\overline{\phi})$. Similarly, $(\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\psi}}, \overline{\overline{\phi}}) \equiv (\hat{\overline{\mathcal{O}}}\,\overline{\psi}, \overline{\phi})$, and thus in any representation for a physically observable operator,

$$(\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\phi}}, \overline{\overline{\psi}}) = (\overline{\overline{\phi}}, \hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\psi}}) \tag{21}$$

The observation that physical operators are always Hermitian is a very powerful mathematical fact. While trivial in the pure state representation, Hermitianness is not always apparent in other representations. For instance, while clearly,

$$\begin{aligned}
(\tilde{\psi}, \hat{\tilde{p}}\tilde{\psi}) &= \int dk\, \tilde{\psi}^*(k)\hbar k \tilde{\psi}(k) \\
&= \int dk\, \left(\hbar \tilde{\psi}(k)\right)^* \tilde{\psi}(k) \\
&= (\hat{\tilde{p}}\tilde{\psi}, \tilde{\psi}),
\end{aligned}$$

the same statement is in the position representation is much harder to prove mathematically, requiring the "trick" if integrating by parts at just the right place,

$$\begin{aligned}
(\psi, \hat{p}\psi) &\equiv \int_{-\infty}^{\infty} dx\, \psi^*(x) \left(\frac{\hbar}{i}\partial_x \psi(x)\right) \\
&= \frac{\hbar}{i}\int_{-\infty}^{\infty} dx\, \psi^*(x)\partial_x\psi(x) \\
&= \frac{\hbar}{i}(\psi^*(x)\psi(x)|_{-\infty}^{\infty} - \frac{\hbar}{i}\int_{-\infty}^{\infty} dx\, \psi(x)\partial_x\psi^*(x) \\
&= -\frac{\hbar}{i}\int_{-\infty}^{\infty} dx\, \partial_x\psi(x)\cdot\psi(x) \quad ; \mathcal{P}(\pm\infty) \to 0 \\
&= \int_{-\infty}^{\infty} dx\, \left(\frac{\hbar}{i}\right)^* (\partial_x\psi^*(x))\cdot\psi(x) \\
&= \int_{\infty}^{\infty} dx\, \left(\frac{\hbar}{i}\partial_x\psi(x)\right)^* \psi(x) \\
&= (\hat{p}\psi, \psi).
\end{aligned}$$

4.3 *Quantum* Operators and Definition of Operators from an Eigenvalue Equation

While we have built up a powerful mathematical framework of function operators, one in which the Heisenberg Uncertainty Principle may be seen to follow directly, our framework is still open to the valid criticism that our function operators are very clumsy, requiring an entirely new set of operators for each new representation. In practice, it is inconvenient to keep track of all of the tilde, over-bar and *double* over-bar symbols when $\hat{p}$, $\hat{\tilde{p}}$, $\hat{\overline{p}}$ and $\hat{\overline{\overline{p}}}$ all refer to the same physical observable. While for most of 8.04 we shall find it most convenient to work nearly exclusively in just the position representation, the student should be aware that in abstract discussions there is a more powerful concept known as the *quantum operator* which deals effectively with this criticism. A "quantum," as opposed to "function" operator takes as its input a *quantum* state, rather than a function, and returns as its output another quantum state, as opposed again to a function. Symbolically, we will use *Dirac* notation for this concept and write the action of a *quantum* operator (which we shall always designate with the subscript "op") $\mathcal{O}_{op}$ on a state $|\psi>$ to produce the state $|\phi>$ as output as

$$|\phi> = \mathcal{O}_{op}|\psi> . \tag{22}$$

The power of (22) is that its meaning is entirely clear without reference to any particular choice of representation. To evaluate (22) in practice, we may proceed in three steps. 1) Write $|\psi>$ as a quantum amplitude in some representation, for instance as a superposition in terms of the the pure states of some observable $\mathcal{R}$ (not necessarily the same as $\mathcal{O}$),

$$|\psi>=\int d\mathcal{R}\,\overline{\overline{\psi}}(\mathcal{R})|\mathcal{R}>\,.$$

2) Apply the operator appropriate to the observable $\mathcal{O}$ in this representation to $\overline{\overline{\psi}}(\mathcal{R})$ to produce $\overline{\overline{\phi}}(\mathcal{R})$, and 3) reconstruct the final quantum state from $\overline{\overline{\phi}}(\mathcal{R})$,

$$|\phi>\equiv \mathcal{O}_{op}|\psi>=\int d\mathcal{R}\,\overline{\overline{\phi}}(\mathcal{R})|\mathcal{R}>=\int d\mathcal{R}\,\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\psi}}(\mathcal{R})|\mathcal{R}>\,. \tag{23}$$

We have been careful in the definition of our function operators so that this procedure always results in the same quantum state regardless of the representation $\mathcal{R}$ we choose, and we may thus write (23) without ambiguity.

From our representation-free definition of the quantum inner product and this new representation-free definition of quantum operators, we then have a representation-free form for writing quantum averages,

$$\begin{aligned}<\mathcal{O}>=(\overline{\overline{\psi}},\hat{\overline{\overline{\mathcal{O}}}}\,\overline{\overline{\psi}})\equiv(\overline{\overline{\psi}},\overline{\overline{\phi}})\equiv<\psi|\phi>\equiv<\psi|\cdot\mathcal{O}_{op}|\psi>\\ <\mathcal{O}>\equiv<\psi|\mathcal{O}_{op}|\psi>\end{aligned} \tag{24}$$

A particularly natural choice for the representation in which to carry out (23) is with pure states associated with the observable. This will lead us to a very compact definition of the action of quantum operators known in terms of an eigenvalue equation. We will discuss this special result not only because it is useful for defining operators in abstract discussions but also because it gives us a powerful practical method of recovering the quantum amplitudes (wave functions) of pure states of an observable in arbitrary representations when all we know about the observable in these representations is the form of its operator.

To find the eigenvalue equation we proceed to write (23) for an arbitrary state (described by an arbitrary set of amplitudes $\overline{\psi}(\mathcal{O})$) in the pure state representation associated with $\mathcal{O}$. Using the fact that we know that the operator associated with the observable in that representation is just multiplication by the argument we now have two ways of writing the action of the operator on the state $|\psi>$,

$$\begin{aligned}|\phi> &= \mathcal{O}_{op}|\psi>=\mathcal{O}_{op}\int d\mathcal{O}\,\overline{\psi}(\mathcal{O})|\mathcal{O}>\\ &= \int d\mathcal{O}\,\hat{\overline{\mathcal{O}}}\,\overline{\psi}(\mathcal{O})|\mathcal{O}>=\int d\mathcal{O}\,\overline{\psi}(\mathcal{O})\cdot\mathcal{O}|\mathcal{O}>\end{aligned}$$

From their definition through (23) we see that, like function operators, *quantum operators are linear* (their action on the sum of two states is the sum of their action on each states individually). Noting further that *quantum operators act only on states and not on functions*, we may then move $\mathcal{O}_{op}$ through the integral above to conclude that for arbitrary $\overline{\psi}(\mathcal{O})$,

$$\begin{aligned}0=|\phi>-|\phi> &= \int d\mathcal{O}\overline{\psi}(\mathcal{O})\cdot\mathcal{O}|\mathcal{O}>-\mathcal{O}_{op}\int d\mathcal{O}\,\overline{\psi}(\mathcal{O})|\mathcal{O}>\\ &= \int d\mathcal{O}\,\overline{\psi}(\mathcal{O})\cdot\mathcal{O}|\mathcal{O}>-\int d\mathcal{O}\,\overline{\psi}(\mathcal{O})\cdot\mathcal{O}_{op}|\mathcal{O}>\\ &= \int d\mathcal{O}\,\overline{\psi}(\mathcal{O})\,(\mathcal{O}|\mathcal{O}>-\mathcal{O}_{op}|\mathcal{O}>)\,.\end{aligned}$$

Since this relation must be true for *any* $\overline{\psi}(\mathcal{O})$, $\overline{\psi}(\mathcal{O})=\delta(\mathcal{O}-\mathcal{O}_0)$ for instance, we must have for all $\mathcal{O}_0$,

$$\begin{aligned}0 &= \mathcal{O}_0|\mathcal{O}_0>-\mathcal{O}_{op}|\mathcal{O}_0>\\ \Rightarrow \mathcal{O}_{op}|\mathcal{O}_0> &= \mathcal{O}_0|\mathcal{O}_0>\,.\end{aligned} \tag{25}$$

This gives us a completely general and representation-free definition for the action of a quantum operator. This definition arises ultimately from the physical fact that the square of quantum amplitudes correspond to probabilities and our construction of operators to give correct averages:

Definition: The action of a quantum operator on a pure state with respect to the observable associated with the operator is just to return the pure state of the observable multiplied by the value of the observable in that state.

The form of the equation (25) is very common in mathematics, it is known as an "eigenvalue" (or, sometimes "characteristic") equation of the operator. $\mathcal{O}_0$ is called the eigenvalue (characteristic value) and $|\mathcal{O}_0>$ is the eigenstate (characteristic state).

As mentioned above, we may use the eigenvalue equation to recover the representation of all of the pure states of an observable in any representation from only the form of the operator. As an example, if our operator is the momentum operator and we are working in the position representation, (25) becomes

$$\begin{aligned}
p_{op}|p_0> &= p_0|p_0> \\
\hat{p}\psi_{p_0}(x) &= p_0\psi_{p_0}(x) \\
\frac{\hbar}{i}\partial_x\psi_{p_0}(x) &= p_0\psi_{p_0}(x) \\
\frac{\hbar}{i}\frac{d\psi_{p_0}}{dx} &= p_0\psi_{p_0} \\
\int\frac{\hbar}{i}\frac{d\psi_{p_0}}{\psi_{p_0}} &= \int p_0\,dx \\
\frac{\hbar}{i}\ln\psi_{p_0} &= C + p_0 x \\
\psi_{p_0} &= e^C e^{i\frac{p_0}{\hbar}x} \\
\psi_{p_0}(x) &= C' e^{2\pi i\left(\frac{x}{h/p_0}\right)},
\end{aligned}$$

which apart from the undetermined constant of proportionality, C', is just the de Broglie hypothesis that the pure state of momentum p_0 is a wave of wavelength $\lambda = h/p_0$. We see that the first de Broglie hypothesis is now built *directly* into our operator framework!