Lecture notes (these are from ny earlier version of the course - we may follow these at a slightly different order, but they should still be relevant!) Physics 3220, Steve Pollock.

Basic Principles of Quantum Mechanics

The first part of Griffith's Ch 3 is in many ways a *review* of what we've been talking about, just stated a little formally, introducing some new notation and a few new twists. We will keep coming come back to it -

First, a quick review of ordinary *vectors*. Think carefully about each of these - nothing here is unfamiliar (though the notation may feel a little abstract), but if you get comfortable with each idea for regular vectors, you'll find it much easier to generalize them to more abstract vectors!

- Vectors live in N-dimensional space. (You're used to 3-D!)
- We have (or can choose) *basis* vectors**:** $\hat{e}_i$ (N of them in N-dim space.) (Example in an "older Phys 1110 notation" of these would be the old familiar unit vectors: $\hat{\imath},\hat{\jmath},\hat{k}$
- They are *orthonormal***:** $\hat{e}_i \cdot \hat{e}_j = \delta_{ij}$ (This is the *scalar*, or *inner*, or *dot* product.)
- They are *complete***:**
  This means *any* vector $\vec{\mathrm{v}} = \sum_i \mathrm{v}_i \hat{e}_i$ is a unique linear combo of basis vectors.
- The basis set *spans* ordinary space. This is like *completeness,* but backwards - every linear combination of basis vectors is again an N-Dim vector, and all such linear combos generate all possible vectors in the space.
- We can choose a specific *representation* of **v**, namely $\{\mathrm{v}_1,\mathrm{v}_2,\mathrm{v}_3,\ldots \mathrm{v}_n\}$, but it is *not* unique, it depends on the choice of basis. (e.g. polar vs. rectangular, and even which particular rotation of the rectangular axes.)
- Each number v_i is the *projection* of **v** in the $\hat{e}_i$ direction, and can be obtained by the formula $\boxed{\mathrm{v}_i = \vec{\mathrm{v}} \cdot \hat{e}_i}$.
  (This involves the same scalar product, again, as we used above in the statement of *orthonormality.*) You can *prove* the last formula by using orthogonality and completeness.
- Addition, or multiplication by a scalar (number), keeps you in the same N-dim "vector space". (adding or scaling vectors gives another vector.)

We can make *very* powerful analogies to *all* of the above in the world of square integrable functions: Note the one to one correspondence between each of the following statements about functions, with the preceding ones about vectors.

- Square integrable functions live in *Hilbert* space. (Never mind what this means for now!)

- We have (or can choose) *basis* functions**:** $u_n(x)$ (*Infinitely* many of them.)

(This infinity might be *countable* (discrete), or it might be uncountable, in which case you can't use integers as labels, but need real numbers.)
We have already met some examples of both such types of u_n's, as eigenfunctions of operators.

- They are *orthonormal***:** $\int dx\, u_n(x)\, u_m(x) = \delta_{n\,m}$. This is apparently our new way of writing the *inner* product. (If the labeling is continuous, the right side will be a Dirac delta function!)

- They are *complete***:** *Any* function $\psi(x) = \sum_n c_n u_n(x)$ is a unique linear combo of the basis vectors. (If the labeling is continuous, then $\psi(x) = \int dE\; c(E)\, u_E(x)$)

- The basis *spans* Hilbert space. This is similar to *completeness,* but backwards - any linear combo of basis functions is itself another square integrable function, and all such linear combos generate all possible functions in the space.

- We can choose a specific *representation* of ψ, e.g. $\{c_1, c_2, c_3, \ldots c_n, \ldots c_\infty\}$ if the labeling is discrete. (Otherwise, we need a function, like e.g. $\varphi(p)$)
  But, it is *not* unique, it depends on the choice of basis. (More on this soon!!)

- The number c_n is the *projection* of ψ in the u_n direction, and
  $c_n = \int u_n^*(x)\psi(x)dx$.
  (This involves the same scalar product, again, as we used above in the statement of *orthonormality*)
  You can *prove* the last formula by using orthogonality and completeness.

- Addition, or multiplication by a scalar (number), keeps you in the same Hilbert space. (adding or scaling functions gives another function.)

As I said, each of the above is a *direct analogy* to a familiar statement about regular vectors on the previous page. It's worth your while to go through them one at a time, side by side, so you can get a sense for how our ideas about vectors in 3-D and our ideas about wave functions in quantum mechanics are closely aligned (even though you might think a function and a little arrow are awfully different things. They're more alike than you might think!)

The analogies don't even end here!

- *Operators* transform vectors (functions): $\hat{O}\vec{V}_1 = \vec{V}_2$, or $\hat{O}\psi_1(x) = \psi_2(x)$
- *Linear* operators are defined by: $\hat{O}(a_1\vec{V}_1 + a_2\vec{V}_2) = a_1\hat{O}\vec{V}_1 + a_2\hat{O}\vec{V}_2$
- *Linear independence says:* **A** is linearly independent of a *set* $\{\vec{\alpha},\vec{\beta},\vec{\gamma},\ldots\}$ if you cannot possibly write **A** as a linear combination of the vectors in that set, i.e $\vec{A} \neq c_1\vec{\alpha} + c_2\vec{\beta} + c_3\vec{\gamma} + \ldots$ for *any* set of constants, c_1, c_2, c_3,...
  (Same for functions!)

E.g, $\hat{z} \neq a\hat{i} + b\hat{j}$ for *any* "a" or "b", or e.g. $e^x \neq a + bx^2$ for *any* "a" or "b".

This analogy is powerful, because it allows us to think of wave functions more abstractly, and to make use of lots of math we know and understand regarding vectors.

ψ (x) is, after all, just a collection of numbers, in the same way that **v**=(v_x, v_y, v_z) is a collection of numbers. Think how often in physics it is better to work with the symbol, or concept, of **v**: independent of any particular *representation*, or basis choice. (Choosing polar or rectangular or rotated coordinates will change the *numbers* v_x, v_y, v_z, but it will *not* change the vector, **v**!)

We will often use the notation $|\mathbf{v}\rangle$, instead of **v**, for the abstract vector.
Similarly, we will work with the abstract wave function, and call it $|\psi\rangle$.
We sometimes call this a "ket". This is our first example of **Dirac Notation**.

We may choose to write any abstract vector in terms of its components in some particular basis:

$$\vec{\mathbf{v}} = |\mathbf{v}\rangle = \sum_i \mathrm{v}_i \hat{e}_i = \sum_i \mathrm{v}_i \left|\hat{e}_i\right\rangle = \sum_i \mathrm{v}_i |i\rangle$$

The first summation notation is the normal old one, the second is where I begin to think of my basis vectors as *themselves* abstract vectors (!!), and the third is just an obvious briefer notation (a shorthand notation for the basis vectors.)
The components of any vector, you will recall, are found by

$$\mathrm{v}_i = \hat{e}_i \cdot \vec{\mathrm{v}} \equiv \left\langle \hat{e}_i \middle| \vec{\mathrm{v}} \right\rangle$$

(This last is just a *definition* for an alternative *notation* for the dot product. Instead of writing "dot" I write a "bracket". Stare at it, we will need to get used to this!)

In complete analogy, we will write

$$|\psi\rangle = \sum_n \mathrm{c}_n \left|u_n\right\rangle = \sum_n \mathrm{c}_n |n\rangle$$

(The final notation is again just an obvious briefer notation, |n> is just a shorthand notation for the basis function $|u_n\rangle$.)

Again, I am thinking of my basis states as *themselves* being abstract kets!

Any function, even basis functions, can be thought of as a ket, just like any vector, even basis vectors, are themselves just vectors.

The "ket" $|\psi\rangle$ represents the *state* of a particle. But it doesn't have to be given in any particular representation, it's more general, more abstract. It's analogous to talking about some vector **v**: I don't have to tell you its components, I might describe it in some more general way, [perhaps it solves some particular equation....] If I need the components, I can pick some convenient coordinate system (rectangular, or polar, or whatever) and *then* figure them out.

E.g., we have already learned that $\psi(x)$ and $\varphi(p)$ (the Fourier transform) refer to the exact same particle in the exact same state. They are fully equivalent, carry identical and complete information about the quantum state. It's just that one is specifically in position space, the other is in momentum space. They will turn out to be two different *representations* of the one and only true abstract state $|\psi\rangle$.

It is a little unfortunate that we usually use the same Greek letter "Psi" for the name of the space wave function and *also* for the name of the abstract state. There's <u>*nothing*</u> special about the space wave function, it is really no better than $\varphi(p)$. We just had to pick some name for the abstract state, and $|\psi\rangle$ seemed better than $|Fred\rangle$. (Using kets has the same potential for simplifying our life as does using vectors without specifying a coordinate system ahead of time.)

More Dirac Notation:

As we have already said, any state can be expanded in basis states:

$$|\psi\rangle = \sum_n c_n |u_n\rangle = \sum_n c_n |n\rangle$$

We have also seen the formula for the coefficients in this expansion:

$$c_n = \int u_n^*(x)\psi(x)dx \equiv \langle u_n|\psi\rangle \equiv \langle n|\psi\rangle$$

This is our scalar product for functions, and as I've mentioned earlier, we often use the "bracket" notation instead of the integral. It's a definition, the bracket notation is a shorthand for the integral. (The final expression is just a lazy shorthand for the first, when I get tired of writing u_n, and just label it by "n" alone)

By the way, the right half of a "bracket" symbol looks like this, $|\ \rangle$, that's why we called it a "ket". And, I kid you not, the name for the left half, which we write like this, $\langle\ |$ is called a "bra".

There is already lots to observe about our new Dirac notation:

$$c_n = \langle u_n | \psi \rangle = \left\langle u_n \middle| \sum_m c_m u_m \right\rangle = \sum_m c_m \langle u_n | u_m \rangle$$

(The first is just the usual old formula, c_n is the inner product of Psi with u_n.
The next statement is just obtained by expanding Psi as a linear combo of basis vectors. The final statement is just an observation that the summation comes out of the inner product [*linearity!*])

But $\langle u_n | u_m \rangle = \int u_n^*(x)\, u_m(x) dx = \delta_{nm}$: this is our statement of orthonormality.
What we had a few lines above now looks like

$$c_n = \sum_m c_m \langle u_n | u_m \rangle = \sum_m c_m \delta_{nm},$$

which is manifestly correct. Don't let these formulas slip by, work through them. There's nothing very difficult here, it's largely NOTATION, but you want to get used to it. Try writing the above statements out for yourself, convince yourself there's *less* here than meets the eye!

$\langle \varphi | \psi \rangle = \int \varphi^*(x)\, \psi(x)\ dx$ is in general a *complex, constant number.* (And, VERY important, don't ever forget the star, complex conjugate, on the "bra"'d function when you do the integral!)

$$\langle \varphi | \psi \rangle^* = \int \left(\varphi^*(x)\, \psi(x) \right)^* dx = \int \varphi(x) \psi^*(x)\ \ dx = \int \psi^*(x)\, \varphi(x)\, dx = \langle \psi | \varphi \rangle$$

Starring a bracket just flips the two entries.
You do *not*, in general, get back what you started with (unless it was real to begin with.)

$\langle \varphi | a\psi \rangle = a \langle \varphi | \psi \rangle$ (constants, even complex ones, slide right out of kets.)
But, $\langle a\varphi | \psi \rangle = a^* \langle \varphi | \psi \rangle$ (note the complex conjugate! Convince yourself of this.)

I keep writing $\langle \varphi | \psi \rangle = \int \varphi^*(x)\, \psi(x)\ dx$, but I should stop, because that is *representation dependent*! It's like writing $\vec{a} \cdot \vec{b} = a_x b_x + a_y b_y + a_z b_z$. The equation is correct, but it depends on my specific choice of Cartesian coordinates. (in polar coordinates, the *formula* for the dot product looks different, but it's still the same result.)

For instance, it is equally true that $\langle \varphi | \psi \rangle = \int \tilde{\varphi}^*(p)\, \tilde{\psi}(p)\ dp$, where the tilde on top means "Fourier transform of"... (We'll prove this later.)

You should try to think of $\langle \varphi | \psi \rangle$ as an abstract dot product. (If that's hard, its not *wrong* to think of it as that specific integral, it's just a little limiting! It's like being "stuck" with always and only thinking of a vector as its components in a particular Cartesian coordinate system!)

The "bra", $\langle\varphi|$ is a slightly odder beast than the "ket", $|\psi\rangle$.
The latter is just like a vector.
But the former is *not* a vector. In fact, it's more like an operator, except its not that *either*, because operators operate on vectors to give back another *vector*, while bras act on vectors to give back a *scalar* (a number).

It's more like a function of vectors. If I want to think of it concretely, I usually think of it like this:
$\langle\varphi|\ldots = \int dx\varphi(x)\ldots$
and you generally fill in the blanks with some ket.
(It's formally called a *dual*, if you really want to know! And of course, this notation is back in my "x-space way of thinking", so you can see it's not wrong, but also not fully general)

In the world of ordinary (complex) 3-D vectors, where
$|\mathbf{b}\rangle = \sum_i b_i \hat{e}_i = b_1\hat{e}_1 + b_2\hat{e}_2 + b_3\hat{e}_3$ (where the b's might be complex), you might be used to many equivalent notations, e.g.

$$|\mathbf{b}\rangle \text{ or } \vec{\mathbf{b}} \text{ or } \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

In this "new notational world" we're building, the notation for the dual would be

$$\langle\mathbf{a}| = \begin{pmatrix} a_1^* & a_2^* & a_3^* \end{pmatrix}$$

With this notation, the dot product is very simple:

$$\langle\mathbf{a}|\mathbf{b}\rangle = \begin{pmatrix} a_1^* & a_2^* & a_3^* \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \sum_i a_i^* b_i$$

You're not used to those stars, but then, you're probably not used to complex vectors!
The *reason* we put them there is so that the *norm* of any vector, namely the dot product of the vector with itself, is guaranteed to be positive:

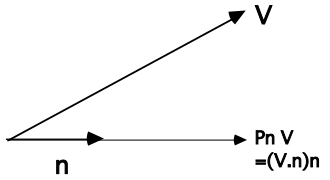
$$\langle\mathbf{a}|\mathbf{a}\rangle = \sum_i a_i^* a_i = \sum_i |a_i|^2 \geq 0.$$

The norm is technically the square root of this:

$$\text{Norm}(\mathbf{a}) = \text{Length}(\mathbf{a}) = \sqrt{\langle\mathbf{a}|\mathbf{a}\rangle} = \sqrt{\sum_i |a_i|^2},$$

For ordinary vectors you're used to, the star is irrelevant, and you just get the usual dot product and the usual norm.

A rather abstract and formal looking relation (but *very* useful), can now be introduced. First, think in ordinary vector space, and consider a *projection operator*, $\hat{P}_{\hat{n}}$, which operates on a vector, and returns the projection of the vector in the **n** direction (which is again a vector.)



I claim that I can write the projection operator as

$$\hat{P}_{\hat{n}} = \left|\hat{e}_n\right\rangle\left\langle\hat{e}_n\right|$$

We need to think about this - *Why* should this be?

Well, just operate **P** on a vector, and see what happens: (Think carefully about this one...)

$\hat{P}_{\hat{n}}|\mathbf{v}\rangle = \left|\hat{e}_n\right\rangle\left\langle\hat{e}_n\middle|\mathbf{v}\right\rangle = \hat{n}(\mathbf{v}\cdot\hat{n})$, precisely what the projection operator is supposed to do!

It returns a vector in the **n** direction, with length $\mathbf{v}\cdot\hat{n}$. Look at the notation, think about vectors, convince yourself this is all ok.

In Hilbert space, this projection operator is written $\hat{P}_n = \left|u_n\right\rangle\left\langle u_n\right|$, it projects into the "$u_n$" direction. (Yikes!) This means

$$\hat{P}_n|\psi\rangle = \left|u_n\right\rangle\left\langle u_n\middle|\psi\right\rangle = c_n\left|u_n\right\rangle$$

In words: projecting in the "n direction" gives back a "vector in the n direction" (here, the ket u_n) with a magnitude c_n.

Now, I claim that $\boxed{\sum_n \hat{P}_n = \sum_n \left|u_n\right\rangle\left\langle u_n\right| = \mathbf{1}}$

(That **1** is the unit *operator*, since projections are themselves operators.)

This mathematical statement is formally called "completeness", it says that if you add up projections in all possible directions, you always get back what you started with.

We'll use it OFTEN!

It's rather abstract, and admittedly a bit hard to understand at first! Here's a rough proof:

Take any vector $|\psi\rangle = \sum_n c_n\left|u_n\right\rangle = \sum_i \left\langle u_n\middle|\psi\right\rangle\left|u_n\right\rangle$

The above familiar looking statement is equivalent to

$$\mathbf{1}|\psi\rangle = \sum_n \left\langle u_n\middle|\psi\right\rangle \left|u_n\right\rangle = \sum_n \left|u_n\right\rangle \left\langle u_n \middle|\psi\right\rangle$$

(The bracket of u_n with psi is just a number; it can be freely moved around)

But since this relation is true for any psi, we have an operator identity: $\mathbf{1} = \sum_n \left|u_n\right\rangle\left\langle u_n\right|$.

You might want to check with ordinary vectors, to see how the identity looks:

$$\sum_i \left|\hat{e}_i\right\rangle\left\langle\hat{e}_i\right| = |\hat{x}\rangle\langle\hat{x}| + |\hat{y}\rangle\langle\hat{y}| = \begin{pmatrix}1\\0\end{pmatrix}(1\ 0) + \begin{pmatrix}0\\1\end{pmatrix}(0\ 1) = \begin{pmatrix}1&0\\0&0\end{pmatrix} + \begin{pmatrix}0&0\\0&1\end{pmatrix} = \begin{pmatrix}1&0\\0&1\end{pmatrix} = \mathbf{1}$$

Note that putting the bra and ket in the *other* order gives

$\left\langle u_n\middle|u_n\right\rangle = 1$, a number not an operator. (It's a totally different beast!)

Next, we add *operators* to the Dirac notation game we're playing:

Suppose we have an operator **O** that does something like this:
$\hat{O}\psi(x) = \beta(x)$, or abstractly $\hat{O}|\psi\rangle = |\beta\rangle$
Then, we will define a new ket, which we call $\left|\hat{O}\psi\right\rangle$.
This is simply the *name,* or *label*, of a new ket. It is *defined* by $\left|\hat{O}\psi\right\rangle \equiv \hat{O}|\psi\rangle$
We thus have $\left|\hat{O}\psi\right\rangle = |\beta\rangle$

These are just two different *labels* for the *same* ket, so naturally
$\langle\beta| = \left\langle\hat{O}\psi\right|$

Putting these together with any other bra, $\langle\varphi|$ gives
$\langle\varphi|\beta\rangle = \left\langle\varphi\middle|\hat{O}\psi\right\rangle = \langle\varphi|\hat{O}|\psi\rangle$
(The last two expressions are interchangeable notation, the 2nd vertical bar is basically there to guide the eye.)

This all looks awfully abstract, but we know what it means when we look at it in our more familiar "x-representation":

$$\boxed{\langle\varphi|\hat{O}|\psi\rangle = \left\langle\varphi\middle|\hat{O}\psi\right\rangle = \int \varphi^*(x)\beta(x)dx = \int \varphi^*(x)\hat{O}\psi(x)dx}$$

For reasons we'll see later, this expression is usually referred to as a *matrix element of operator* ***O***.

In the expression $\left\langle\varphi\middle|\hat{O}\psi\right\rangle$ or equivalently $\langle\varphi|\hat{O}|\psi\rangle$ the operator **O** operates on psi (i.e. to its right) as usual. It does not operate on phi, which is to its left.

However, I can *define* a new operator, related to **O** but possibly quite different, cleverly designed so that if *it* operates on phi, we somehow get back the same number as when **O** operates on psi. (Remember, the matrix element we're looking at is just some complex number.)

This new operator is called the *adjoint* of **O**, or the *Hermitian adjoint*, or sometimes the *Hermitian conjugate*, or sometimes the *dagger of* **O**, and is denoted $\mathbf{O}^+$. (That little "plus sign" should really look like a cross or "dagger symbol", †, but my Word formula processor doesn't have that symbol!)
All the above words are stated mathematically as follows, we *define* $\mathbf{O}^+$ by:

$\left\langle\mathbf{O}^+\varphi\middle|\psi\right\rangle \equiv \left\langle\varphi\middle|\hat{\mathbf{O}}\psi\right\rangle$ for *all* functions phi and psi (!!!)

(I suppose it is not at all obvious that such a dagger operator even exists, but it really does for all operators **O** of interest to us. We will not prove it - it's getting into some pretty hairy linear algebra theory.)

If you star (complex conjugate) the expression above defining $\mathbf{O}^+$, and remember from earlier that starring a bracket just reverses it:

$\left\langle \hat{\mathbf{O}}\psi \middle| \varphi \right\rangle \equiv \left\langle \psi \middle| \mathbf{O}^+\varphi \right\rangle$ (which is again true for *all* functions phi and psi.)

If you stare at this, you should be able to convince yourself that it means

$\mathbf{O}^{++} = \mathbf{O}$. (Perhaps a little reminiscent of c**=c for NUMBERS?)

We can also write the definition of the dagger in integral form:

$\int \left(\mathbf{O}^+\varphi\right)^* \psi(x)dx \equiv \int \varphi^*(x)\hat{O}\psi(x)\,dx$ for *all* functions phi and psi.

Hermitian operators are operators whose expectation values are always real, like the operators for physical observables we've used all term. There is a very different definition of Hermitian operators, a more basic one, and we will see the connection between the two definitions later (we'll prove that they're really equivalent).

Our new definition will be that an operator, **O**, is Hermitian iff (if and only if) $\mathbf{O}^+ = \mathbf{O}$.
Using our definitions above, this says that if **O** is Hermitian, then

$\boxed{\left\langle \hat{\mathbf{O}}\varphi \middle| \psi \right\rangle \equiv \langle \varphi | \mathbf{O}\psi \rangle}$ for *all* functions phi and psi.

Hermitian operators are very special, and as we've already stated, they play a key role in quantum mechanics. (We'll see more reasons shortly)

Example 1:

O=c (That means our operator is the "multiply by a complex constant "c" operator.)
Is this operator Hermitian? Let's test it out:

$\langle \varphi | c\psi \rangle = c\langle \varphi | \psi \rangle$

$\langle c\varphi | \psi \rangle = c^*\langle \varphi | \psi \rangle$

These two expression are *not* equal, unless c is real.
So, the operator "3" is Hermitian, but the operator "i" is *not*.

It seems like a Hermitian operator might mean "real operator" in some way. This is sort of true: remember our old definition, **O** is Hermitian means **<O>** is real.
But you have to be a bit careful, the presence of "i" can be misleading:

Example 2:

$\hat{O} \equiv \frac{d}{dx}$, the derivative operator. It sure looks "real". Is it Hermitian?

$$\left\langle \varphi \middle| \hat{\mathbf{O}}\psi \right\rangle = \int \varphi^*(x)\frac{d\psi(x)}{dx}\,dx = \varphi^*\psi\Big|_{-\infty}^{\infty} - \int \left(\frac{d\varphi^*}{dx}\right)\psi\,dx = -\int \left(\frac{d\varphi^*}{dx}\right)\psi\,dx$$

$$\left\langle \hat{\mathbf{O}}\varphi \middle| \psi \right\rangle = \int \left(\frac{d\varphi^*}{dx}\right)\psi\,dx$$

The signs disagree, the two lines are *not* the same. This operator is most certainly *not* Hermitian!

Example 3:

$\hat{O} \equiv \frac{\hbar}{i}\frac{d}{dx}$, the momentum operator. Is it Hermitian?

$$\left\langle \varphi \middle| \hat{\mathbf{O}}\psi \right\rangle = \int \varphi^*(x)\frac{\hbar}{i}\frac{d\psi(x)}{dx}\,dx = \frac{\hbar}{i}\left(\varphi^*\psi\Big|_{-\infty}^{\infty} - \int\left(\frac{d\varphi^*}{dx}\right)\psi\,dx \right)$$

$$\left\langle \hat{\mathbf{O}}\varphi \middle| \psi \right\rangle = \int\left(\frac{\hbar}{i}\frac{d\varphi}{dx}\right)^*\psi\,dx = \frac{\hbar}{-i}\int\left(\frac{d\varphi^*}{dx}\right)\psi\,dx$$

The explicit presence of an "i" has fixed up the minus sign. The two lines above agree, for any psi and phi (as long as they vanish off at infinity, which is true for any physically valid wave functions) and thus this operator is Hermitian.

So momentum is Hermitian, and its expectation value in ANY state is always real. (You could have seen this from the momentum space representation, $< p >= \int \varphi^*(p) p \varphi(p)\,dp$, manifestly real!)

Example 4:

$\hat{O} \equiv V(x)$. (This is the "multiply by a [real] function V(x)" operator.)

If V(x) is real, this is just like case 1, the operator is Hermitian.

Earlier this semester, even before the introduction of Dirac notation, we used to write that the expectation value of **O** in a state psi is

$$\left\langle \hat{O} \right\rangle \equiv \int \psi^*(x)\hat{O}\psi(x)\,dx$$

Now we see that the proper Dirac notation for this expression is

$\langle \psi | \hat{O} | \psi \rangle$, and the above is a pretty obvious shorthand notation!

Brief mathematical digression- *Matrices*

Go back to N-dimensional vector spaces, and define the operator **O** by $\hat{O}|\alpha\rangle = |\beta\rangle$ We can understand what **O** does in *general*, as long as we know what it does only to our *basis* vectors!

$$\hat{O}\left|\hat{e}_1\right\rangle = \text{some vector} \equiv O_{11}\left|\hat{e}_1\right\rangle + O_{21}\left|\hat{e}_2\right\rangle + \ \ldots\ + O_{n1}\left|\hat{e}_n\right\rangle$$

$$\hat{O}\left|\hat{e}_2\right\rangle \qquad\qquad \equiv O_{12}\left|\hat{e}_1\right\rangle + O_{22}\left|\hat{e}_2\right\rangle + \ \ldots\ + O_{n2}\left|\hat{e}_n\right\rangle$$

$$\vdots$$

$$\hat{O}\left|\hat{e}_j\right\rangle \equiv \sum_{i=1}^{n} O_{ij}\left|\hat{e}_i\right\rangle$$

(Note the slightly screwy ordering of indices above - it's a definition!)

As always with formal equations like this, stare at it for a moment, just get used to this notation.

Given the above, how do I know what **O** does to $|\alpha\rangle$?
Just expand $|\alpha\rangle$ in some basis, and you see

$$\hat{O}|\alpha\rangle = \hat{O}\sum_j a_j \left|\hat{e}_j\right\rangle \equiv \sum_j a_j \hat{O}\left|\hat{e}_j\right\rangle = \sum_j \sum_{i=1}^{n} a_j O_{ij} \left|\hat{e}_i\right\rangle$$

By inspection, we find

$|\beta\rangle \equiv \hat{O}|\alpha\rangle = \sum_{ij} O_{ij} a_j \left|\hat{e}_i\right\rangle$, or in component form,

$$b_i = \sum_j O_{ij} a_j$$

Now the ordering of subscripts is making some sense, this looks like a matrix multiplication:

$$|\beta\rangle \equiv \hat{O}|\alpha\rangle = \begin{pmatrix} O_{11} & O_{12} & \cdots \\ O_{21} & \ddots & \\ \vdots & & \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix}$$

The N^2 elements in that matrix *fully characterize* the operator! If you know those N^2 *numbers*, you can figure out what **O** does to *any* vector! So an operator can also be described by a set of numbers (typically written in matrix form, and know as "matrix elements"... Kind of cool!
The particular numbers are, unfortunately, representation dependent. If you change your choice of basis vectors, all the elements of that matrix can change.

This means that *operators can be expressed explicitly as matrices*, although the specific numerical entries do depend on the choice of basis vectors.

Once you've got this, you can work out products of matrices quite easily, e.g.

$$\hat{O}_1\hat{O}_2 \Rightarrow (O_1O_2)_{ij} = \sum_l (O_1)_{il}(O_2)_{lj}$$

If you have a product of operators, just multiply the matrices!

And now, some definitions of matrix terminology, basically some linear algebra review:

***Transpose*:**

$$\tilde{O}_{ij} \equiv O_{ji}, \qquad \text{i.e.} \quad \tilde{O} = \begin{pmatrix} O_{11} & O_{21} & \cdots \\ O_{12} & O_{22} & \\ \vdots & & \ddots \end{pmatrix}$$

You can transpose a vector too,

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} \quad \Rightarrow \quad \tilde{\mathbf{a}} = \begin{pmatrix} a_1 & a_2 & \cdots \end{pmatrix}$$

Conjugate:

$$O^*_{ij} \equiv \left(O_{ij}\right)^*, \qquad \text{i.e.} \quad O^* = \begin{pmatrix} O^*_{11} & O^*_{12} & \cdots \\ O^*_{21} & O^*_{22} & \\ \vdots & & \ddots \end{pmatrix}$$

Adjoint: (or, Hermitian Adjoint)

$$O^+ = \tilde{O}^*, \qquad \text{i.e.} \quad O^+ = \begin{pmatrix} O^*_{11} & O^*_{21} & \cdots \\ O^*_{12} & O^*_{22} & \\ \vdots & & \ddots \end{pmatrix}$$

Which means $O^{++} = O$ (daggering twice gets back to where you started, like complex conjugating twice does for a number)
You can adjoint a vector too,

$$\begin{pmatrix} a_1 \\ \vdots \end{pmatrix}^+ = \begin{pmatrix} a^*_1 & \cdots \end{pmatrix}$$

Note that our Dirac notation for vectors now looks like:

$$\langle\alpha|\beta\rangle = a^*_1 b_1 + a^*_2 b_2 + \ldots = \begin{pmatrix} a_1 \\ \vdots \end{pmatrix}^+ \begin{pmatrix} b_1 \\ \vdots \end{pmatrix} = \alpha^+ \beta$$

i.e. $\langle\alpha| = |\alpha\rangle^+$

Which tells us in a more formal way what the "bra" is - it's the adjoint of the ket.

Inverse:

$$\hat{O}^{-1}: \quad \hat{O}^{-1}\hat{O} = \hat{O}\hat{O}^{-1} = \mathbf{1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \ddots \end{pmatrix}$$

Unitary:

A matrix (or operator) **O** is *unitary* iff (if and only if) $\hat{O}^+ = \hat{O}^{-1}$

Hermitian:

A matrix (or operator) **O** is *Hermitian* iff $\hat{O}^+ = \hat{O}$
(This is really the *same* definition as I gave above for vector operators, can you see why?
I'll talk about this on the next page, think about it yourself a bit first!)

Products of operators: If you write out the indices, you should not have *too* hard a time proving:

$\widetilde{(ST)} = \tilde{T}\tilde{S}$, and

$(\mathrm{ST})^+ = (\mathrm{T})^+(\mathrm{S})^+$
(See below for an alternative proof of the latter relation**)

When we were dealing with operators, I had a rather different looking *definition* of the adjoint (also called the Hermitian conjugate). My old definition, written now for vectors, looked like:
$\left\langle \mathbf{O}^+\alpha \middle| \beta \right\rangle \equiv \left\langle \alpha \middle| \hat{\mathbf{O}}\beta \right\rangle$ for all vectors α,β.

My *new* definition is $O^+ = \tilde{O}^*$.
Are these two definitions equivalent? Yes!
Using *only* the new matrix definitions we can *derive* the old definition:
$\left\langle \alpha \middle| \hat{\mathbf{O}}\beta \right\rangle = \alpha^+\hat{\mathbf{O}}\,\beta = (\mathbf{O}^+\alpha)^+\,\beta$ (because $\mathbf{O}^{++} = \mathbf{O}$)

$= \left\langle \mathbf{O}^+\alpha \middle| \beta \right\rangle$

** Similarly, here's a proof that $(\hat{\mathrm{A}}\hat{\mathrm{B}})^+ = \hat{\mathrm{B}}^+\hat{\mathrm{A}}^+$ (using my older "function" defs:)
$\left\langle \alpha \middle| (\mathbf{AB})^+\beta \right\rangle = \left\langle \mathbf{AB}\alpha \middle| \beta \right\rangle = \left\langle \mathbf{A}(\mathbf{B}\alpha) \middle| \beta \right\rangle = \left\langle (\mathbf{B}\alpha) \middle| \mathbf{A}^+\beta \right\rangle$

$= \left\langle \alpha \middle| \mathbf{B}^+\mathbf{A}^+\beta \right\rangle$

This is no matter what alpha and beta are, so it must be that $(\mathrm{AB})^+ = \mathrm{B}^+\mathrm{A}^+$

A Ridiculous Example:
$\big(\langle\varphi|\mathbf{AB}|\psi\rangle\lambda|x\rangle\langle y| \big)^+ = |y\rangle\,\langle \mathrm{x}|\lambda^*\langle\psi|\mathbf{B}^+\mathbf{A}^+|\varphi\rangle$
Just stare, think about the ordering and stars, this is the "game" of daggering!

We've been defining operators in vector space, but as we've seen, we can continue with our analogy and now think about operators as *matrices* in Hilbert space. (They may, however, be infinite dimensional matrices, that's the price we pay!)

So, e.g. we can continue to use notations like
$\langle\psi| = |\psi\rangle^+$
even though psi isn't really an ordinary column vector anymore!

One immediate consequence of all the above is that we can often make quick proofs of important things, like e.g., we already know that momentum (**p**) was Hermitian. How about $\mathbf{p}^2$?

$(\hat{p}\,\hat{p})^+ = \hat{p}^+\hat{p}^+ = \hat{p}\,\hat{p}$

This statement proves that p^2 is Hermitian!
(Stare at it, and convince yourself that you understand this quick and simple little proof.)

Since we just showed p^2 is Hermitian, and in an earlier example we showed that V(x) is Hermitian, it means that the Hamiltonian $\hat{H} = \hat{p}^2/2m + V(x)$ is also Hermitian.

What's the deal here? Why do we care whether an operator is Hermitian?
It's because of a few *theorems:*

1) The eigenvalues of Hermitian operators are always real.

2) The expectation values of Hermitian operators are always real.

3) The eigenvectors of Hermitian operators span the Hilbert space.

4) The eigenvectors of Hermitian operators belonging to distinct eigenvalues are orthogonal.

In quantum mechanics, these characteristics are essential if you want to associate *measurements* with *operators*. Operators must be Hermitian so that observables are real.
And, you must be able to expand in the eigenfunctions (the expansion coefficients, as we have seen, give you probabilities!)

We proceed to prove these (I'll skip number 3, which will probably become fairly obvious, and the proof really doesn't *teach* us anything especially useful...)

Proof of theorem 1): *The eigenvalues of Hermitian operators are real.*
Let **O** be Hermitian, and write the usual eigenvalue eqn:
$\hat{O}u_n(x) = o_n u_n(x)$ with o_n some *number*, the eigenvalue of **O**.
In Dirac notation, we would write this equation as
$\hat{O}\left|u_n\right\rangle = o_n\left|u_n\right\rangle$
Now hit the left hand side with $\left\langle u_n\right|$, giving
$\left\langle u_n\middle|\hat{O}\middle|u_n\right\rangle = o_n\left\langle u_n\middle|u_n\right\rangle = o_n$
But, since **O** is Hermitian, then we also find
$\left\langle u_n\middle|\hat{O}\middle|u_n\right\rangle \equiv \left\langle \hat{O}^+u_n\middle|u_n\right\rangle = \left\langle \hat{O}u_n\middle|u_n\right\rangle$ (because $\hat{O}^+ = \hat{O}$)

$= \left\langle o_n u_n\middle|u_n\right\rangle = o_n^*\left\langle u_n\middle|u_n\right\rangle = o_n^*$

Combining with the previous expression, we see $\boxed{o_n^* = o_n}$
Which proves that each and every eigenvalue is real. (Don't glaze over - go through the steps!)

Proof of theorem 2): *Expectation values of Hermitian operators are real.*

Let **O** be Hermitian, and let ψ be any state at all, then

$$\langle\psi|\hat{O}|\psi\rangle = \int \psi^*(x)\hat{O}\psi(x)\,dx = \int\left(\hat{O}^+\psi\right)^*\psi\,dx = \int\left(\hat{O}\psi\right)^*\psi\,dx$$

$$= \int \psi(x)\,\left(\hat{O}\psi(x)\right)^*\,dx = \left(\int \psi^*(x)\;\hat{O}\psi(x)\,dx\right)^* = \left(\langle\psi|\hat{O}|\psi\rangle\right)^*$$

Comparing the beginning and end of that expression, we see that $\langle\psi|\hat{O}|\psi\rangle$ must always be real, no matter what.

I can give you an alternative proof of theorem 2, which is a little cooler because it is "coordinate free", I never need to invoke functions in x-space:

$$\langle\psi|\hat{O}|\psi\rangle = \sum_n \langle\psi|\hat{O}\left|u_n\right\rangle\left\langle u_n\right|\psi\rangle \quad \text{(this is "completeness"**!)}$$

$$= \sum_n \left\langle\psi\right|o_n u_n\rangle\left\langle u_n\right|\psi\rangle \;=\; \sum_n o_n\left\langle\psi\right|u_n\rangle\left\langle u_n\right|\psi\rangle \;=\; \sum_n o_n\left\langle\psi\right|u_n\rangle\left\langle\psi\right|u_n\rangle^*$$

$$= \sum_n o_n\left|\left\langle\psi\right|u_n\rangle\right|^2 \;=\; \text{real, manifestly.}$$

It's real because theorem 1 proved the o_n's are all real, and the absolute value squared of any bracket is always real.

(** This proof shows one of the nice uses of "completeness" - you can always stick $\boxed{\mathbf{1} = \sum_n \left|u_n\right\rangle\left\langle u_n\right|}$ into the middle of just about any expression.)

Proof of theorem 3) *E-vectors of Hermitian operators span the Hilbert space.* As I mentioned, we won't prove this one. (In the next theorem, though, we will prove it for a special case.)

Proof of theorem 4) *The eigenvectors of Hermitian operators belonging to distinct eigenvalues are orthogonal.*

Suppose I have two different eigenvectors of **O**, labeled 1 and 2,

$\hat{O}\left|u_1\right\rangle = o_1\left|u_1\right\rangle$

$\hat{O}\left|u_2\right\rangle = o_2\left|u_2\right\rangle$

and *suppose further that o_1 differs from o_2.* (The eigen-values are *distinct.*)
Hit the 1st eqn. on the left with the bra $\langle u_2|$, and hit the 2nd eqn with $\langle u_1|$:

$\left\langle u_2\left|\hat{O}\right|u_1\right\rangle = o_1\left\langle u_2\middle|u_1\right\rangle$

$\left\langle u_1\left|\hat{O}\right|u_2\right\rangle = o_2\left\langle u_1\middle|u_2\right\rangle$

Now, star the second equation. What happens to the LHS?

$\left\langle u_1\left|\hat{O}\right|u_2\right\rangle^* = \left\langle \hat{O}u_2\middle|u_1\right\rangle = \left\langle u_2\left|\hat{O}^+\right|u_1\right\rangle = \left\langle u_2\left|\hat{O}\right|u_1\right\rangle$ (because $\hat{O}$ is Hermitian).

Meanwhile, starring the RHS of the second equation gives

$o_2^*\left\langle u_2\middle|u_1\right\rangle = o_2\left\langle u_2\middle|u_1\right\rangle$ (Theorem 1).

Putting together the last two lines, we just showed that

$\left\langle u_2\left|\hat{O}\right|u_1\right\rangle = o_2\left\langle u_2\middle|u_1\right\rangle$

Above, we began with two equations, the first of which was

$\left\langle u_2\left|\hat{O}\right|u_1\right\rangle = o_1\left\langle u_2\middle|u_1\right\rangle$

Subtracting these last two equations from each other, and recalling that o_1 and o_2 are by assumption different, the only way this can make sense is if

$\left\langle u_2\middle|u_1\right\rangle = 0$, the two states are orthogonal. (As desired!)

In N-dimensional vector space, if all eigenvalues of **O** are distinct, there will be N of them [since **O** is an NxN matrix], and they're all orthogonal by the last theorem, so they must span the space (Theorem 3 is proven in this case).

In Hilbert space, and/or if some eigenvalues are repeated, the proof is trickier, but still true.

Our *formalism* is mostly complete now. We'll be making lots of use of it, working out many examples where Dirac notation (and/or Hermiticity, the "matrix picture", completeness, etc.) will help us solve problems quickly and efficiently.

At this point, we're ready to try to pull everything we've learned so far this semester together, and summarize it all in some formal way.

The general structure of quantum:

The postulates of Quantum Mechanics... (The numbering, and even ordering, of these is *not* especially universal, and may differ from what we did in class)

(1) The state of a particle is completely represented by a normalized vector in Hilbert space, which we call $|\psi\rangle$.

(2) All physical observables, Q, are associated with Hermitian operators $\hat{\mathbf{Q}}$, and the expectation value of Q in some state $|\psi\rangle$ is $\langle\psi|\hat{\mathbf{Q}}|\psi\rangle$.

(3) A measurement of Q on a particle in state $|\psi\rangle$ is certain to return a particular value, λ, iff ("if and only if") $\hat{\mathbf{Q}}|\psi\rangle = \lambda|\psi\rangle$
(i.e. if and only if $|\psi\rangle$ is already an eigenvector of $\hat{\mathbf{Q}}$, with eigenvalue λ)

(3a) If you measure Q (in any state $|\psi\rangle$), you are certain to obtain *one* of the eigenvalues of $\hat{\mathbf{Q}}$.
The probability of measuring some eigenvalue λ is given by $\left|\left\langle u_\lambda \middle| \psi \right\rangle\right|^2$,
(where $\left|u_\lambda\right\rangle$ is defined to be the eigenvector of $\hat{\mathbf{Q}}$, with eigenvalue λ.)

(The above has to be modified slightly if there are degenerate eigenvectors.
It is *always* possible to arrange the set of $\left|u_\lambda\right\rangle$'s form an orthonormal basis, no matter what, but when we discuss degeneracy we'll see that it isn't always trivial. We'll talk about degeneracy shortly.)

(3b) After a measurement gives you the value λ, the system will collapse into the state $\left|u_\lambda\right\rangle$.
(Again, modulo some subtleties if eigenvalues are degenerate)

(4) The time evolution of the state $|\psi\rangle$ is given by Schrodinger's equation:

$$i\hbar\frac{\partial|\psi\rangle}{\partial t} = \hat{H}|\psi\rangle$$

These statements are a modification of Griffith's collection of the postulates of quantum mechanics. They are perhaps a little redundant (The statements 3a and 3b are closely related to 3, so I didn't give them separate numbers), and they are all consistent with one another. I claim that the above is pretty close to all you need to "derive" everything else in quantum mechanics...)

Read them over again, think about them, try to "translate" them back into the way you've been thinking throughout the semester. This is the formal basis for quantum mechanics!

I claim that statements 2 and 3b are also closely related, as follows:

$$\langle\psi|\hat{\mathbf{Q}}|\psi\rangle = \sum_{\lambda}\langle\psi|\hat{\mathbf{Q}}\big|u_\lambda\big\rangle\big\langle u_\lambda\big|\psi\big\rangle$$

(This is *completeness* again, just math, inserting a projection operator sum like we've done before!)

$$= \sum_{\lambda}\big\langle\psi\big|\lambda u_\lambda\big\rangle\big\langle u_\lambda\big|\psi\big\rangle = \sum_{\lambda}\lambda\big|\big\langle u_\lambda\big|\psi\big\rangle\big|^2$$

But postulate 3b says that $\big|\big\langle u_\lambda\big|\psi\big\rangle\big|^2$ is the probability of measuring λ,

so what we have written is equivalent to

$$\langle\psi|\hat{\mathbf{Q}}|\psi\rangle = \sum_{\lambda}\lambda\ \text{Prob}(\lambda),$$

which is precisely what you would think the expectation value of an observable should be.

START DIGRESSION: *Some rather advanced material about representations:*

[The following is a pretty intense discussion of how you pull *explicit functions* out of the abstract kets, which is in the end what we usually want.]

We're very used to thinking in the x-representation, i.e.

I usually think of $|\psi\rangle$ as being equivalent to $\psi(x)$.

They are closely related, but $|\psi\rangle$ is more general (it's like a vector **V**, while $\psi(x)$ is like the specific components (Vx, Vy, Vz) in some particular frame.)

Just as we can extract the components of a vector, we can also extract $\psi(x)$ from $|\psi\rangle$. Here's how:

Consider the **x** operator, and it's eigenfunctions, which we'll call $|x\rangle$.

To be specific, $\hat{\mathbf{x}}\big|x_0\big\rangle = x_0\big|x_0\big\rangle$

I normally use completeness to write $\sum_n \big|u_n\big\rangle\big\langle u_n\big| = \mathbf{1}$, but now, with **x** having continuous (not discrete) eigenfunctions, I should write completeness as $\int dx_0\ \big|x_0\big\rangle\big\langle x_0\big| = \mathbf{1}$. (Seem reasonable? Just replace the sum with an integral, since the eigenvalues of **x** are continuous, not discrete.)

Now Look at $\langle\psi|\psi\rangle = \int dx_0 \psi^*(x_0)\psi(x_0)$. But, I can *also* use completeness here, to say

$\langle\psi|\psi\rangle = \int dx_0 \big\langle\psi\big|x_0\big\rangle\big\langle x_0\big|\psi\big\rangle = \int dx_0 \big\langle x_0\big|\psi\big\rangle^*\big\langle x_0\big|\psi\big\rangle$ (Are you getting used to the "completeness" math trick yet?) Comparing the two expressions, apparently

$\psi(x_0) = \big\langle x_0\big|\psi\big\rangle$, or $\psi(x) = \langle x|\psi\rangle$

Stare at that last expression: $|\psi\rangle$ is the general Q.M. state vector.

$\psi(x)$ is the probability amplitude fn for finding a particle at a location x

By postulate 3a, with **Q**=**x**, this is completely consistent:

Probability according to 3b is $\big|\langle x|\psi\rangle\big|^2 = |\psi(x)|^2$

Completeness says that $|\psi\rangle = \int dx_0\ \big|x_0\big\rangle\big\langle x_0\ \big|\psi\big\rangle = \int dx_0 \big|x_0\big\rangle\psi(x_0)$.

Of course I could equally well have chosen **P** as my operator instead of **x**, with eigenfunctions u_p:
$\hat{\mathbf{p}}|p\rangle = p|p\rangle$, in which case completeness would tell me
$|\psi\rangle = \int dp \ \ |p\rangle\langle p| \ \psi\rangle$
and I would know from Postulate 3a that
$|\langle p|\psi\rangle|^2 = |\varphi(p)|^2$ = probability of finding momentum p.

This is telling me that
$\psi(x) = \langle x|\psi\rangle$: "What does the state $|\psi\rangle$ look like as a function of x?"
$\varphi(p) = \langle p|\psi\rangle$: "What does the state $|\psi\rangle$ look like as a function of p?"

It is interesting to ask what, e.g., $\langle x|p\rangle$ is.
(What does the eigenfunction of momentum look like, as a function of x?)

Recall that $\varphi(p) = \int dx \frac{1}{\sqrt{2\pi\hbar}} e^{-ipx/\hbar}\psi(x)$, so I have
$\langle p|\psi\rangle = \int dx \langle p|x\rangle\langle x|\psi\rangle = \int dx \langle x|p\rangle^* \psi(x)$,
comparing these tells me $\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{+ipx/\hbar}$
(We've seen this before! These are plane waves, the eigenfunctions of **p**, written in x space)
Similarly, $\langle x|x_0\rangle = \delta(x_0 - x)$, the eigenfunctions of **x**, i.e. $|x_0\rangle$, written in x space.

I can also write the above in the following equivalent pair of expressions:
$\langle x'|\psi\rangle = \int dx \langle x'|x\rangle\langle x|\psi\rangle$
$\psi(x') = \int dx \, \delta(x'-x)\psi(x)$

But we're getting awfully abstract, and I think will stop with "pushing on Dirac notation" so hard at this point!

-END DIGRESSION

Degeneracy and simultaneous observables:

Suppose we have Hermitian operators **A** and **B**, each of which of course has its own complete set of eigenfns. Let's *define* the e-vectors and values of **A**:

$\hat{A}\left|u_a\right\rangle = a\left|u_a\right\rangle$

S*uppose further* that each and every eigenfunction $\left|u_a\right\rangle$ is *also* an eigenfunction of **B**.

This might be possible, or it might not!
(We've seen that, e.g., the eigenfunctions of **H** *are* also all eigenfunctions of Parity, **P**, as long as V(x) is symmetric).

Anyway, *if this supposition is o.k.*, then we have

$\hat{B}\left|u_a\right\rangle = b\left|u_a\right\rangle$.

But now observe the following:

$\hat{A}\hat{B}\left|u_a\right\rangle = \hat{A}b\left|u_a\right\rangle = b\hat{A}\left|u_a\right\rangle = ba\left|u_a\right\rangle$

$\hat{B}\hat{A}\left|u_a\right\rangle = \hat{B}a\left|u_a\right\rangle = a\hat{B}\left|u_a\right\rangle = ab\left|u_a\right\rangle$

$\Rightarrow (\hat{A}\hat{B} - \hat{B}\hat{A})\left|u_a\right\rangle = 0$

This wouldn't be all that interesting except remember that I'm supposing that it is true for each and every eigenfunction $\left|u_a\right\rangle$! Consider what the commutator **[A,B]** does when acting on any old function, $\psi(x)$.

A is Hermitian => the eigenfunctions of **A** span the space=>
we can expand $\psi(x)$ in eigenfunctions of **A** =>

$$[\hat{A},\hat{B}]\psi(x) = (\hat{A}\hat{B} - \hat{B}\hat{A})\psi(x) = (\hat{A}\hat{B} - \hat{B}\hat{A})\sum_a c_a\left|u_a\right\rangle$$

$$= \sum_a c_a(\hat{A}\hat{B} - \hat{B}\hat{A})\left|u_a\right\rangle \;\; = 0$$

which means that **[A,B]**=0.

We say that **A** and **B** are *compatible* in this case. We can simultaneously know the eigenvalue of *both* of the operators, or we say **A** and **B** have *simultaneous eigenvectors*. (Remember, that was our starting assumption at the top - that every eigenfunction of **A** was also an eigenfunction of **B.**)

We have only shown that *if* **A** and **B** have simultaneous eigenvalues, then they must commute.
Does it work the other way?
It *does*, although there are again some subtleties involved if there are degeneracies.
To see it work the other way, *suppose* **[A,B]**=0. Then:

$$\hat{A}\hat{B}\left|u_a\right\rangle = \hat{B}\hat{A}\left|u_a\right\rangle = \hat{B}a\left|u_a\right\rangle = a\hat{B}\left|u_a\right\rangle$$

$$\Rightarrow \hat{A}\left(\hat{B}\left|u_a\right\rangle\right) = a\left(\hat{B}\left|u_a\right\rangle\right)$$

Stare at that last line. It is of the form **A** ("something") = a ("something"),
and implies that the "something" is an eigenvector of **A**, with eigenvalue a.
Now, *if the eigenfunctions of A are non-degenerate*, then each eigenvalue has by definition one and only one corresponding eigenfunction, implying:

$\hat{B}\left|u_a\right\rangle \propto \left|u_a\right\rangle$, i.e. $\hat{B}\left|u_a\right\rangle = b\left|u_a\right\rangle$.

Thus, if **A** and **B** commute, then indeed the operators will have simultaneous eigenvectors, but only if **A** has a "non-degenerate" spectrum.

What if **A** does have some degeneracy? Let's suppose for simplicity that it is just 2-fold, i.e.

$$\hat{A}u_a^{(1)}(x) = a\,u_a^{(1)}(x)$$

$$\hat{A}u_a^{(2)}(x) = a\,u_a^{(2)}(x)$$

Knowing (as I showed above) that $\hat{A}\left(\hat{B}\left|u_a^{(i)}\right\rangle\right) = a\left(\hat{B}\left|u_a^{(i)}\right\rangle\right)$ merely tells me that $\hat{B}\left|u_a^{(i)}\right\rangle$ must be some linear combination of $\left|u_a^{(1)}\right\rangle$ and $\left|u_a^{(2)}\right\rangle$.

The *Gram-Schmidt* procedure is what you use here two find 2 orthogonal combinations of $\left|u_a^{(1)}\right\rangle$ and $\left|u_a^{(2)}\right\rangle$ which are themselves eigenfunctions of **B.**
(They might be degenerate in **B**, or might not be.)

Bottom line: If [**A,B**]=0, then (even if **A** has degeneracies), we can still always generate an orthogonal basis of vectors which are *all* simultaneously eigenfunctions of **B** as well.

If we are unlucky, and some degeneracy in **B** still exists, there must be a 3rd operator, **C**, which satisfies **[A,C]=[B,C]=0** and now we can find an orthonormal basis which are simultaneously e-fns of **A,B**, and **C**, and (hopefully) with distinct eigenvalues for operator **C**. (If not, you may need a **D,** and **E**, etc...) At some point, you will have a set of operators, **A, B, C, ... M**, which all commute with each other, and which have a common set of eigenfunctions, and each and every such eigenfunction will have a distinct set of eigenvalues.
This set of operators (and eigenvalues) completely characterizes any wave function:

$$\hat{A}u_{abc...m}(x) = a\,u_{abc...m}(x)$$

$$\hat{B}u_{abc...m}(x) = b\,u_{abc...m}(x)$$

etc. , and the set of numbers a,b,... m characterize the state fully.

What if [**A,B**] is *not* zero? Then, we cannot simultaneously know the eigenvalues of **A** and **B**, in general. Let's quantify:

Pick a state, $|\psi_0\rangle$, *any* state. Define the uncertainty of an operator, **A**, in that particular state in the usual way we have all semester:

$$(\Delta A)^2 \equiv \langle A^2\rangle - \langle A\rangle^2 = \langle A^2\rangle - 2\langle A\rangle^2 + \langle A\rangle^2 = \langle A^2 - 2A\langle A\rangle + \langle A\rangle^2\rangle = \langle (A - \langle A\rangle)^2\rangle,$$

The *Schwartz inequality* is a completely general theorem, which says

$\boxed{\langle\psi|\psi\rangle\langle\varphi|\varphi\rangle \geq |\langle\psi|\varphi\rangle|^2}$ for any phi and psi. (Kind of looks like (1)(1) >=$\cos^2\theta$, doesn't it?)

Next, we *define a particular* psi and phi (which, like anything, will have to obey the above Schwartz inequality):

$$\boxed{\begin{aligned}|\psi\rangle &\equiv (\hat{A} - \langle A\rangle)|\psi_0\rangle, \\ |\varphi\rangle &\equiv (\hat{B} - \langle B\rangle)|\psi_0\rangle,\end{aligned}}$$

Observe:

$$\boxed{\langle\psi|\psi\rangle \equiv \langle\psi_0|(\hat{A} - \langle A\rangle)(\hat{A} - \langle A\rangle)|\psi_0\rangle \equiv \Delta A^2.}$$

and, similarly,

$$\boxed{\langle\varphi|\varphi\rangle \equiv \Delta B^2.}$$

So the Schwartz inequality for this particular case reads

$$\begin{aligned}(\Delta A)^2(\Delta B)^2 &\geq \left|\langle\psi_0|(A - \langle A\rangle)(B - \langle B\rangle)|\psi_0\rangle\right|^2 \\ &= \left|\langle\psi_0|(AB - \langle A\rangle B - \langle B\rangle A + \langle A\rangle\langle B\rangle)|\psi_0\rangle\right|^2 \\ &= \left|\langle\psi_0|(AB - \langle A\rangle\langle B\rangle)|\psi_0\rangle\right|^2\end{aligned}$$

But we know expectation values of Hermitian operators are always real. And we can also write that the square of *any* complex number z obeys the following inequality:

$$|z|^2 = (\mathrm{Re}\, z)^2 + (\mathrm{Im} z)^2 \geq \mathrm{Im}(z)^2 = \left[\frac{1}{2i}(z - z^*)\right]^2$$

Making use of this in our last expression we find

$$\begin{aligned}(\Delta A)^2(\Delta B)^2 &\geq \left[\frac{1}{2i}\left(\langle\psi_0|\hat{A}\hat{B}|\psi_0\rangle - \langle\psi_0|\hat{A}\hat{B}|\psi_0\rangle^*\right)\right]^2 \\ &= \left[\frac{1}{2i}\left(\langle\psi_0|\hat{A}\hat{B}|\psi_0\rangle - \langle\psi_0|\hat{B}\hat{A}|\psi_0\rangle\right)\right]^2 = \left[\frac{1}{2i}\langle\psi_0|[\hat{A},\hat{B}]|\psi_0\rangle\right]^2 \\ &= \frac{1}{4}\langle i[\hat{A},\hat{B}]\rangle^2\end{aligned}$$

If [**A,B**]=0, then we just showed basically nothing at all, we have the trivial identity $(\Delta A)^2(\Delta B)^2 \geq 0$ (Which implies that you could know **A** and **B** both, simultaneously)

But if [**A,B**] is NOT zero, then you get something interesting!

Consider for example **x** and **p**, where we have $[x,p] = i\hbar$,

this means $(\Delta x)^2(\Delta p)^2 \geq \frac{1}{4}\langle i\ i\hbar\rangle^2 = \hbar^2/4$

i.e., $\Delta x \Delta p \geq \hbar/2$

Apparently, x and **p are *incompatible observables***. (Incompatible observables do not share a complete set of common eigenvectors)

We have just derived the Heisenberg Uncertainty principle purely by using operator properties. There was no discussion of waves, or wavefunctions, or Fourier transforms, or anything! It depends only on the observed properties, and not on the particular object which you are looking at! It is a fundamental statement about the operators x and p!

Final topic: Time dependence

Consider the time evolution of the expectation value of any operator, **O**:

$$\frac{d}{dt}\langle\hat{O}\rangle = \frac{d}{dt}\langle\psi|\hat{O}|\psi\rangle = \left\langle\frac{\partial\psi}{\partial t}\middle|\hat{O}|\psi\right\rangle + \langle\psi|\frac{\partial\hat{O}}{\partial t}|\psi\rangle + \left\langle\psi|\hat{O}\middle|\frac{\partial\psi}{\partial t}\right\rangle.$$

The Schrod equation tells us what the time evolution of wave fns is:

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi$$

Putting this into the first equation gives

$$\frac{d}{dt}\langle\hat{O}\rangle = \frac{1}{-i\hbar}\left\langle\hat{H}\psi\middle|\hat{O}|\psi\right\rangle + \left\langle\frac{\partial\hat{O}}{\partial t}\right\rangle + \frac{1}{i\hbar}\left\langle\psi|\hat{O}\middle|\hat{H}\psi\right\rangle$$

But H is Hermitian => $\left\langle\hat{H}\psi\middle|\hat{O}|\psi\right\rangle = \langle\psi|\hat{H}\hat{O}|\psi\rangle$

We thus have

$$\boxed{\frac{d}{dt}\langle\hat{O}\rangle = \frac{i}{\hbar}\langle\psi|\hat{H}\hat{O} - \hat{O}\hat{H}|\psi\rangle + \left\langle\frac{\partial\hat{O}}{\partial t}\right\rangle = \frac{i}{\hbar}\langle\psi|[\hat{H},\hat{O}]|\psi\rangle + \left\langle\frac{\partial\hat{O}}{\partial t}\right\rangle}$$

Most operators don't have any *explicit* time dependence, so that last term generally vanishes! (E.g. if **O** is x, or p, or p^2/2m, or V(x), or Parity, or anything like this, the last term vanishes.) It's only if you explicitly see a "t" in the operator that you must include it, like e.g. if you had a potential which looked like V= $A\sin\omega t$, or something...

Example 1:

$$\frac{d}{dt}\langle x\rangle = \frac{i}{\hbar}\langle\psi|[\frac{p^2}{2m} + V(x), x]|\psi\rangle$$

[V(x),x]=0, and it's not hard to work out $[p^2, x] = \frac{2\hbar}{i}p$ which means

$\boxed{\frac{d}{dt}\langle x\rangle = \langle p\rangle/m}$ An old friend! We've seen and used this many times this term.

Example 2:

$$\frac{d}{dt}\langle p\rangle = \frac{i}{\hbar}\langle\psi|[\frac{p^2}{2m} + V(x), p]|\psi\rangle$$

This time, $[p^2,p]=0$, and we need to work out [V(x), p]:

$$[V(x),p]f(x) = V(x)\frac{\hbar}{i}\frac{\partial f}{dx} - \frac{\hbar}{i}\frac{\partial}{dx}(V(x)f(x)) = -\frac{\hbar}{i}\frac{\partial V(x)}{dx}f(x)$$

$$\Rightarrow [V(x),p] = -\frac{\hbar}{i}V'(x)$$

In this case,

$$\boxed{\frac{d}{dt}\langle p\rangle = -\langle\psi|\frac{dV(x)}{dx}|\psi\rangle}$$

which is *Ehrenfest*'s theorem. Proving it directly using Schrodinger's equation (integrating by parts repeatedly) is a big pain, and took over a full page of algebra in my solutions to an old homework problem from a very early set. With this technique, it's an easy 2 liner!

Example 3:

$$\frac{d}{dt}\langle\hat{H}\rangle = \frac{i}{\hbar}\langle\psi|[\hat{H},\hat{H}]|\psi\rangle + \left\langle\frac{\partial\hat{H}}{\partial t}\right\rangle$$

Clearly, any operator commutes with itself. So, unless the Hamiltonian has an explicit time dependence (e.g. an explicitly time dependent potential), the expectation value of energy is conserved, as you would certainly hope! Nice....