Physics 221A  
Fall 1996  
Notes 2  
The Postulates of Quantum Mechanics

In these notes we present the postulates of quantum mechanics, which allow one to connect experimental results with the mathematical formalism described in Notes 1. Actually, we are not ready to state the postulates in their complete and final form, since that requires the use of the density operator, which we discuss in Notes 3. Therefore our first version of the postulates will involve some undefined terminology, and will be incomplete. Nevertheless, even in their incomplete form, the postulates explain a good deal about the mathematical formalism of quantum mechanics and its relation to physical reality.

In their incomplete form, the postulates of quantum mechanics are the following:

1. Every physical system is associated with a Hilbert space $\mathcal{E}$. We call the vectors of this space \textit{kets}.

2. Every pure state of a physical system is associated with a definite ray in $\mathcal{E}$. We postpone for a while the definition of a pure state, but a ray was defined in Eq. (1.9). In practice we often represent the state of a system by some nonzero ket lying in the ray in question.

3. Every measurement process which can be carried out on the system corresponds to a complete Hermitian operator $A$.

4. The possible results of the measurement are the eigenvalues of $A$, either the discrete eigenvalues $a_1, a_2, \ldots$ or the continuous ones $a(\nu)$.

5. In the discrete case, the probability of measuring $A = a_n$ is

$$\text{Prob}(A = a_n) = \frac{\langle \psi | P_n | \psi \rangle}{\langle \psi | \psi \rangle}, \quad (2.1)$$

where $P_n$ is the projection operator onto the eigenspace $\mathcal{E}_n$ corresponding to eigenvalue $a_n$, as indicated by Eq. (1.98), and where $|\psi\rangle$ is any nonzero ket in the ray representing the state of the system. In the continuous case, the probability of measuring $A$ to lie in some interval $I = [a_0, a_1]$ of the continuous spectrum is

$$\text{Prob}(a_0 \leq A \leq a_1) = \frac{\langle \psi | P_I | \psi \rangle}{\langle \psi | \psi \rangle}, \quad (2.2)$$

where $P_I$ is the projection operator corresponding to interval $I$, as in Eq. (1.102).

6. After a measurement with discrete outcome $A = a_n$, the system is represented by the ket $P_n |\psi\rangle$, where $|\psi\rangle$ represents the system before the measurement. Note that in this
case, the system is in an eigenstate of $A$ with eigenvalue $a_n$ after the measurement. In the continuous case, with outcome $a_0 \leq A \leq a_1$, the system is represented by $P_I|\psi\rangle$ after the measurement, where again $P_I$ is as in Eq. (1.102).

As postulate 5 shows, the predictions made by quantum mechanics are of a statistical nature. The probabilities referred to in that postulate are determined experimentally by repeating the same measurement on a large number of identically prepared systems. The fact that identically prepared systems can yield different results on successive measurements is a point which we will discuss after we give some examples.

Postulate 6 is called the *collapse* postulate, and it is the one which is most troubling from a philosophical standpoint. We will have more to say about it later, but for now it is more important to see how these postulates actually work in practice. For example, we will have to answer questions such as the following: How do we know what Hilbert space is associated with a given physical system? What operator corresponds to definite measuring processes? Which ray corresponds to a definite (pure) state of system? To answer such questions, we will analyze the Stern-Gerlach experiment, pretending that we know nothing except the experimental results and the postulates listed above. In particular, we will pretend that we know nothing about wave functions, Pauli matrices, etc.

Suppose we are working with a beam of silver atoms, as in the original Stern-Gerlach experiment. From a modern perspective, we know that silver atoms possess a magnetic moment because of their single unpaired valence electron, so that the measured value of any component of magnetic moment is $\pm \mu_0$, where $\mu_0 = e\hbar/2mc$ is a Bohr magneton. (In fact, the original experiment of Stern and Gerlach in 1921 provided a measurement of $\mu_0$, with a result in good agreement with Bohr’s value.) We also know that the magnetic moment operator is proportional to the spin operator, $\mathbf{\mu} = (e/mc)\mathbf{S}$, where any component of spin takes on the values $\pm \hbar/2$. In the following discussion, however, we will play dumb and ignore all of this, and instead we will work solely with the experimental results. For the same reason, we will speak in terms of measurements of magnetic moment, not spin.

![Fig. 2.1. A beam of silver atoms is subjected to a measurement of $\mu_x$, after which the atoms with $\mu_x = +\mu_0$ are passed to a second magnet which measures $\mu_z$.](image)
Consider the tandem Stern-Gerlach apparatus illustrated in Fig. 2.1, in which we first measure the $x$-component of the magnetic moment $\mu$, and then pass the beam with $\mu_x = +\mu_0$ to a second magnet which measures $\mu_z$. The experimental result is that the two beams emerging from the second magnet, with $\mu_z = \pm \mu_0$, each carry 50% of the atoms which entered that magnet.

Since a measurement of $\mu_x$ gives rise to two possible outcomes, we must assume according to postulates 3 and 4 that the operator $\mu_x$ has two possible eigenvalues, $\pm \mu_0$. (The same is true for $\mu_y$ and $\mu_z$.) Therefore the Hilbert space $\mathcal{E}$ must be at least 2-dimensional, since the eigenspaces corresponding to $\mu_x = \pm \mu_0$ are orthogonal. The easiest way to proceed with the following argument is simply to assume that the eigenvalues $\pm \mu_0$ are nondegenerate, and then to come back later to the question of degeneracies. Under this assumption, the eigenspaces are 1-dimensional, the ket space $\mathcal{E}$ has exactly two dimensions, and it is spanned by the eigenkets of $\mu_x$ with eigenvalues $\pm \mu_0$. It is also spanned by eigenkets of $\mu_y$ or $\mu_z$ with eigenvalues $\pm \mu_0$, and therefore the pair of eigenkets of any one of these operators must be expressible as linear combinations of the eigenkets of any other operator.

From this point we follow Sakurai closely. Let us denote some normalized eigenvector of $\mu_x$ with eigenvalue $+\mu_0$ by $|\mu_x+\rangle$; we say “some” because any other vector differing from this one by an overall phase would work just as well at this stage of the argument. Similarly, let us choose normalized eigenvectors $|\mu_x-\rangle$, $|\mu_y\pm\rangle$, and $|\mu_z\pm\rangle = |\pm\rangle$ (we will henceforth omit the $\mu_z$ specification in the eigenkets of this operator). Because eigenkets corresponding to distinct eigenvalues are orthogonal, we have

$$\langle \mu_x+ | \mu_x- \rangle = \langle \mu_y+ | \mu_y- \rangle = \langle + | - \rangle = 0. \quad (2.3)$$

In accordance with postulate 2, the state of the atomic beam at various stages in the apparatus are described by some state vector (as we will discuss later, the spin system is in a pure state anywhere after the measurement of $\mu_x$ in the first magnet). Also, according to postulate 6, after the measurement of $\mu_x$ with result $+\mu_0$, the state of the system is specified by $|\mu_x+\rangle$, as indicated by the ket attached to the upper beam in Fig. 2.1. Similarly, the two emerging beams from the second magnet, which measures $\mu_z$, are represented by the state vectors $|\mu_z\pm\rangle$. We represent the state $|\mu_x+\rangle$ entering the second magnet as a linear combination of the eigenkets $|\pm\rangle$,

$$|\mu_x+\rangle = c_+ |+\rangle + c_- | - \rangle, \quad (2.4)$$

where $c_\pm$ are the expansion coefficients. Since $|+\rangle$ and $| - \rangle$ are orthogonal, this implies

$$c_\pm = \langle \pm | \mu_x+ \rangle. \quad (2.5)$$
Then, according to postulate 5, we have

$$\text{Prob}(\mu_z = +\mu_0) = \langle \mu_x + | P_+ | \mu_x + \rangle,$$  \hspace{1cm} (2.6)

where

$$P_+ = |+\rangle\langle +|$$  \hspace{1cm} (2.7)

is the projection operator onto the $+\mu_0$ eigenspace of $\mu_z$. The denominator in (2.2) is unity because $|\mu_x + \rangle$ is normalized. Therefore Eq. (2.6) becomes

$$\text{Prob}(\mu_z = +\mu_0) = |\langle + | \mu_x + \rangle|^2 = |c_+|^2 = \frac{1}{2},$$  \hspace{1cm} (2.8)

where we use Eq. (2.4) and the experimentally measured probability. This implies

$$c_+ = \frac{1}{\sqrt{2}} e^{i\alpha_1},$$  \hspace{1cm} (2.9)

where $e^{i\alpha_1}$ is an unknown phase factor. But since $\text{Prob}(\mu_z = -\mu_0)$ is also 1/2, a similar argument gives

$$c_- = \frac{1}{\sqrt{2}} e^{i\beta_1},$$  \hspace{1cm} (2.10)

where $e^{i\beta_1}$ is another unknown phase. Thus, we have

$$|\mu_x + \rangle = \frac{1}{\sqrt{2}}(e^{i\alpha_1} |+\rangle + e^{i\beta_1} |-\rangle).$$  \hspace{1cm} (2.11)

---

**Fig. 2.2.** Same as Fig. 2.1, except the beam with $\mu_x = -\mu_0$ is fed into the second magnet.

This can be simplified by changing the phase of the ket $|\mu_x + \rangle$ so as to absorb the phase $e^{i\alpha_1}$ on the right hand side. After we have done this we have established a phase convention for this ket and we cannot change its phase again. This gives

$$|\mu_x + \rangle = \frac{1}{\sqrt{2}}(|+\rangle + e^{i\beta_1} |-\rangle).$$  \hspace{1cm} (2.12)
By a similar analysis, if we feed the beam emerging from the first magnet with $\mu_x = -\mu_0$ into the second magnet, as illustrated in Fig. 2.2, and analyze the probabilities exactly as we have done here, we obtain

$$|\mu_x^-\rangle = \frac{1}{\sqrt{2}}(|+\rangle + e^{i\gamma_1}|-\rangle), \quad (2.13)$$

where $e^{i\gamma_1}$ is another phase. Now we have fixed the phases of both kets $|\mu_x\pm\rangle$.

But the phases $e^{i\beta_1}$ and $e^{i\gamma_1}$ are related, for by orthogonality we have

$$\langle \mu_x^+ | \mu_x^- \rangle = \frac{1}{2} [1 + e^{i(-\beta_1 + \gamma_1)}] = 0, \quad (2.14)$$

or $e^{i\gamma_1} = -e^{i\beta_1}$. Thus, Eqs. (2.12) and (2.13) become

$$|\mu_x\pm\rangle = \frac{1}{\sqrt{2}}(|+\rangle \pm e^{i\beta_1}|-\rangle). \quad (2.15)$$

Similarly, suppose we measure $\mu_y$ instead of $\mu_x$ in the first magnet in Figs. 2.1 and 2.2. Then the two beams emerging from the second magnet are still split with equal probabilities, and an analysis just like that presented above gives

$$|\mu_y\pm\rangle = \frac{1}{\sqrt{2}}(|+\rangle \pm e^{i\beta_2}|-\rangle), \quad (2.16)$$

where $e^{i\beta_2}$ is another phase. Now the phases of kets $|\mu_y\pm\rangle$ have been fixed.

[At this point it must be said that we have entered the realm of gedankenexperiments, because only the components of the magnetic moment transverse to the beam are easy to measure in a Stern-Gerlach apparatus. Thus, if the $x$- and $z$-directions are transverse to the beam, then the $y$-direction is along the beam, and we cannot measure $\mu_y$ easily. Nevertheless, we will proceed as if the 50% probabilities quoted are actual experimental results. Certainly if we believe that the distribution of magnetic moments of the atoms emerging from the oven is isotropic, then we would expect these probabilities.]

We can now write down the operators representing the components of $\mu$ in the $|\pm\rangle$ basis. For example, according to Eq. (1.101) we have

$$\mu_x = \mu_0 \left( |\mu_x^+\rangle \langle \mu_x^+ | - |\mu_x^-\rangle \langle \mu_x^- | \right), \quad (2.17)$$

which can be transformed to the $|\pm\rangle$ basis by substituting Eq. (2.15). This gives

$$\mu_x = \mu_0 \left( e^{-i\beta_1} |+\rangle \langle - | + e^{i\beta_1} |-\rangle \langle + | \right). \quad (2.18)$$

Similarly, we find

$$\mu_y = \mu_0 \left( e^{-i\beta_2} |+\rangle \langle - | + e^{i\beta_2} |-\rangle \langle + | \right), \quad (2.19)$$
and of course we have

$$\mu_z = \mu_0 \left( |+\rangle \langle +| - |-\rangle \langle -| \right). \tag{2.20}$$

Next, a relation can be found between the phases $e^{i\beta_1}$ and $e^{i\beta_2}$ by imagining another Stern-Gerlach experiment, like Fig. 2.1 except that we measure $\mu_y$ in the second magnet instead of $\mu_z$. Again, we can take it as an experimental result that the two values $\mu_y = \pm \mu_0$ are measured with equal probability. Therefore we have

$$\frac{1}{2} = |(\mu_x + |y\rangle \langle y|)|^2 = \frac{1}{2} [1 + \cos(\beta_2 - \beta_1)], \tag{2.21}$$

where we use Eqs. (2.15) and (2.16). But this implies $\beta_2 = \beta_1 \pm \pi/2$, or

$$e^{i\beta_2} = \pm ie^{i\beta_1}. \tag{2.22}$$

Thus, the two unknown phases in Eqs. (2.18) and (2.19) reduce to one unknown phase and an unknown sign.

The remaining unknown phase, say, $e^{i\beta_1}$, can be pinned down by choosing a phase convention for the ket $|\rangle$; for example, we can make the matrix elements of either $\mu_x$ or $\mu_y$ in the $|\pm\rangle$ basis purely real, but not both, because by Eq. (2.22), if the matrix elements of one of these operators is real, those of the other operator must be purely imaginary. Here we choose to absorb the phase $e^{i\beta_1}$ into the definition of the ket $|\rangle$, so that Eqs. (2.18), (2.19) and (2.20) become

$$\mu_x = \mu_0 \left( |+\rangle \langle +| - |-\rangle \langle -| \right),$$

$$\mu_y = \pm \mu_0 \left( -i|+\rangle \langle -| + i|-\rangle \langle +| \right),$$

$$\mu_z = \mu_0 \left( |+\rangle \langle +| - |-\rangle \langle -| \right). \tag{2.23}$$

We notice that we cannot change the phase of $|+\rangle$ without messing up equations such as (2.15) or (2.16). Thus, all the freedom in phase conventions has been used up. To state results equivalent to Eqs. (2.23), we can write out the matrices representing the three operators in the $|\pm\rangle$ basis,

$$[\mu_x] = \mu_0 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad [\mu_y] = \pm \mu_0 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad [\mu_z] = \mu_0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.24}$$

where the square brackets $[ ]$ indicate the matrix representing the operator enclosed. Of course, when we use this notation, it must be understood which basis is being employed. We recognize the Pauli matrices in these results.
We see that to within a final unknown sign, the three Pauli matrices are determined solely from the postulates of quantum mechanics, the experimental results, and some phase conventions. But there is the question of the final, unknown sign. Sakurai has some discussion of this point, but he never definitively settles the issue. We will look at this question more carefully in a homework problem.

Now let us return to the question of degeneracies. So far we have simply assumed that the eigenspaces of $\mu_x$ (and $\mu_y, \mu_z$) are nondegenerate. How would we know if this were not true? If we were given a Hilbert space and a Hermitian operator as a purely mathematical problem, then the answer could be obtained purely by mathematics. But here we are building up the Hilbert space out of the results of physical measurements, and there must be physical meaning to any degeneracies which might exist. As we will now show, the answer to this question involves the notion of compatible or commuting observables.

Consider an idealized measurement such as illustrated in Fig. 2.3. A system which is known to be in a pure state $|\psi_0\rangle$ is first subjected to a measurement of observable $A$. Out of the several possible outcomes, all are thrown away, except for $a_n$. A device which selects an eigenvalue (or range of eigenvalues) in this manner is called a filter. The system after the measurement of $A = a_n$ is described by $|\psi_1\rangle$; this system is passed to a device which measures observable $B$, and all outcomes except $B = b_m$ are thrown away. The state of the system after the second measurement is described by $|\psi_2\rangle$.

According to the postulates, the probability of measuring $A = a_n$ in the first apparatus is

$$\text{Prob}(a_n) = \frac{\langle \psi_0 | P_{An} | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle},$$  \hspace{1cm} (2.25)

where $P_{An}$ is the projection operator onto the eigenspace of $A$ corresponding to eigenvalue $a_n$. Also, the ket describing the state which emerges from the first filter is

$$|\psi_1\rangle = P_{An} |\psi_0\rangle.$$  \hspace{1cm} (2.26)
Next let us compute the probability of measuring first \( A = a_n \) and then \( B = b_m \). This is the (conditional) probability of measuring \( B = b_m \) given that we had \( A = a_n \), multiplied times the probability of measuring \( A = a_n \) in the first place. We denote the probability of this compound measurement by \( \text{Prob}(a_n; b_m) \); it is given by

\[
\text{Prob}(a_n; b_m) = \frac{\langle \psi_1 | P_{Bm} | \psi_1 \rangle \langle \psi_0 | P_{An} | \psi_0 \rangle}{\langle \psi_1 | \psi_1 \rangle \langle \psi_0 | \psi_0 \rangle},
\]

(2.27)

where \( P_{Bm} \) is the projection operator onto the eigenspace of \( B \) with eigenvalue \( b_m \). But we also have

\[
\langle \psi_1 | \psi_1 \rangle = \langle \psi_0 | P_{An}^\dagger P_{An} | \psi_0 \rangle = \langle \psi_0 | P_{An} | \psi_0 \rangle,
\]

(2.28)

where we have used Eq. (2.26) and the fact that \( P_{An} \) is Hermitian and idempotent. Thus, the denominator of the first factor in Eq. (2.27) cancels the numerator of the second factor, and we have

\[
\text{Prob}(a_n; b_m) = \frac{\langle \psi_0 | P_{An} P_{Bm} P_{An} | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle}.
\]

(2.29)

Similarly, if we reverse the order of the measurements, measuring first \( B = b_m \) and next \( A = a_n \), we obtain the probability

\[
\text{Prob}(b_m; a_n) = \frac{\langle \psi_0 | P_{Bm} P_{An} P_{Bm} | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle}.
\]

(2.30)

These two probabilities are not equal, in general.

Now it can be shown that if \([A, B] = 0\), then \([P_{An}, P_{Bm}] = 0\), so in this case we have

\[
P_{An} P_{Bm} P_{An} = P_{An}^2 P_{Bm} = P_{An} P_{Bm} = P_{Bm} P_{An} P_{Bm},
\]

(2.31)

so that both probabilities are equal,

\[
\text{Prob}(a_n; b_m) = \text{Prob}(b_m; a_n) = \frac{\langle \psi_0 | P_{An} P_{Bm} | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle}.
\]

(2.32)

Conversely, it can be shown that if the probabilities are equal for all initial states \( |\psi_0\rangle \), then \([P_{An}, P_{Bm}] = 0\); and if this commutator vanishes for all \( n \) and \( m \), then \([A, B] = 0\).

Altogether, we can say that probability of a compound measurement is independent of the order of the measurements, for all initial states and all outcomes of the two measurements, if and only if the observables commute. This is the physical meaning of commuting observables. Commuting observables are also said to be compatible.

Let us now return to the question of degeneracies. Referring to Fig. 2.3, how do we know if the eigenvalues of the operator \( A \) are degenerate? If an eigenvalue \( a_n \) is degenerate, then the eigenspace \( \mathcal{E}_n \) corresponding to this eigenvalue is multidimensional, so we are
asking about the dimensionality of the subspaces $E_n$. The answer is obtained by searching for other observables $B$ which commute with $A$, to see if one of them will ‘resolve’ the degeneracy of $a_n$, that is, produce more than one outcome when a measurement of $B$ is made subsequent to the measurement $A = a_n$. If such an observable can be found, then the states emerging from the $B$-apparatus lie in the simultaneous eigenspaces of the operators $A$ and $B$, which are subspaces of the eigenspace $E_n$ of $A$. A vector lying in one of these subspaces is obtained by applying the projectors $P_{An}$ and $P_{Bm}$ (in either order, since they commute) to an arbitrary ket, such as $|\psi_0\rangle$ in the figure. In this case, the order of the degeneracy of $a_n$ is at least equal to the number of outcomes of the subsequent $B$-measurement, since each of the simultaneous eigenspaces is at least one-dimensional. However, these eigenspaces may also be degenerate. To find out if they are, we can search for another operator $C$, which commutes with both $A$ and $B$, which will resolve the simultaneous eigenspaces of $A$ and $B$ into smaller subspaces. The process continues until no more resolutions are possible; then the set of observables $(A, B, C, \ldots)$ constitutes a complete set of commuting observables, or CSCO for short. At this point we can declare that the simultaneous eigenspaces of the CSCO are nondegenerate.

Often we choose not to go all the way to a complete set of commuting observables, because we are only interested in some subset of the degrees of freedom of a system. For example, in our discussions of the Stern-Gerlach apparatus, we have been ignoring the spatial degrees of freedom of the silver atom, as well as the internal degrees of freedom representing the motion of the electrons around the nucleus, the motion of the nucleons in the nucleus, the quarks in the nucleons, etc. If we were to include the spatial degrees of freedom of the silver atoms as well as the spin, our Hilbert space would be infinite-dimensional (actually, $2 \times \infty$).

We turn now to some simple observations of a statistical nature regarding the postulates of quantum mechanics. Let us denote the probability of measuring $A = a_n$ on a normalized, pure state $|\psi\rangle$ by $f_n$, so that

$$f_n = \langle \psi | P_n | \psi \rangle,$$

where $P_n$ is the projector onto the $n$-th eigenspace of $A$. Then the average value of $a$, in the ordinary sense of statistics, is just the sum of the $a_n$’s weighted by the probabilities $f_n$. This quantity can, however, be expressed in terms of Hilbert space operations,

$$\langle a \rangle = \sum_n f_n a_n = \langle \psi | \sum_n a_n P_n | \psi \rangle = \langle \psi | A | \psi \rangle,$$

where we use Eq. (1.101). We will also write $\langle A \rangle = \langle a \rangle$ for this quantity. Similarly, the variance $\Delta a^2$, defined in the usual way in statistics, can be expressed in terms of Hilbert
space operations:
\[
\Delta a^2 = \sum_n f_n a_n^2 - \left( \sum_n f_n a_n \right)^2 = \langle \psi | A_1^2 | \psi \rangle,
\]  \hspace{1cm} (2.35)

where
\[
A_1 = A - \langle A \rangle.
\]  \hspace{1cm} (2.36)

We will also write \( \Delta A^2 = \Delta a^2 \). The proof of Eq. (2.35) is straightforward. But this equation has an immediate consequence. Let us ask for the conditions under which \( \Delta a^2 = 0 \), that is, the conditions under which measurements of a quantum mechanical system will yield a single value with 100% probability, with no dispersion. We write
\[
\Delta a^2 = 0 = \langle \psi | A_1^2 | A_1 | \psi \rangle = \langle \phi | \phi \rangle,
\]  \hspace{1cm} (2.37)

where \( |\phi\rangle = A_1 |\psi\rangle \). But by Eq. (1.27), this holds if and only if \( |\phi\rangle = 0 \), or,
\[
A |\psi\rangle = \langle A \rangle |\psi\rangle.
\]  \hspace{1cm} (2.38)

In other words, a quantum measurement of an observable \( A \) produces no dispersion if and only if the state \( |\psi\rangle \) is an eigenstate of \( A \).

A related subject is that of the generalized uncertainty principle, which places lower bounds on the products of the dispersions of two observables, \( A \) and \( B \). We first quote the result,
\[
\Delta A^2 \Delta B^2 \geq \frac{1}{4} |\langle [A, B] \rangle|^2,
\]  \hspace{1cm} (2.39)

where expectation values are taken with respect to some state \( |\psi\rangle \). To prove this, we define
\[
|\alpha\rangle = A_1 |\psi\rangle,
|\beta\rangle = B_1 |\psi\rangle,
\]  \hspace{1cm} (2.40)

where
\[
A_1 = A - \langle A \rangle,
B_1 = B - \langle B \rangle,
\]  \hspace{1cm} (2.41)

and use the Schwarz inequality (1.28) in the form,
\[
\langle \alpha | \alpha \rangle \langle \beta | \beta \rangle \geq |\langle \alpha | \beta \rangle|^2.
\]  \hspace{1cm} (2.42)

By Eq. (2.35), the left hand side of this is \( \Delta A^2 \Delta B^2 \). As for the right hand side, we have
\[
\langle \alpha | \beta \rangle = \langle \psi | A_1 B_1 | \psi \rangle = \frac{1}{2} \langle \psi | [A_1, B_1] | \psi \rangle + \frac{1}{2} \langle \psi | \{A_1, B_1\} | \psi \rangle,
\]  \hspace{1cm} (2.43)
where we have written the product $A_1 B_1$ as one half the sum of the commutator and the anticommutator. Since $A_1$ and $B_1$ are Hermitian, the commutator is anti-Hermitian and the anticommutator is Hermitian. Therefore the first term on the right hand side of Eq. (2.43) is purely imaginary and the second is purely real; the two terms on the right hand side are the real and imaginary parts of the expression on the left. Therefore when we take the absolute value squared of the left hand side, we obtain the inequality,

$$\left|\langle \alpha | \beta \rangle \right|^2 \geq \frac{1}{4} \left| \langle \psi | [A_1, B_1] | \psi \rangle \right|^2.$$  \hfill (2.44)

But since $A$ and $A_1$ differ only by the $c$-number $\langle A \rangle$, which commutes with everything, and likewise for $B$ and $B_1$, the commutator in Eq. (2.44) can be replaced simply by $[A, B]$. Combining all the pieces, we obtain Eq. (2.39). Of course, the best known example of this generalized uncertainty principle is the case $A = x$, $B = p$, $[A, B] = i \hbar$, and therefore (after taking the square root)

$$\Delta x \Delta p \geq \frac{\hbar}{2}.$$  \hfill (2.45)