POSTULATES AND THEOREMS OF QUANTUM MECHANICS:

Postulate I:

(a). Any state of a dynamical system of N particles is described as fully as possible by a wavefunction, $\Psi(x_1, y_1, z_1, x_2, y_2, z_2, x_3, y_3, z_3, \ldots, t)$ which contains all information one can know about the system.

(b). The quantity $|\Psi|^2\, dx_1\, dy_1\, dz_1\, dx_2\, dy_2\, dz_2\, dx_3\, dy_3\, dz_3 \ldots$ gives the probability that

particle 1 will be found within the small volume between x_1 and $x_1 + dx_1$,
y_1 and $y_1 + dy_1$,
z_1 and $z_1 + dz_1$, while at the same time
particle 2 will be found within the small volume between x_2 and $x_2 + dx_2$,
y_2 and $y_2 + dy_2$,
z_2 and $z_2 + dz_2$, while at the same time
particle 3 will be found within the small volume between x_3 and $x_3 + dx_3$,
y_3 and $y_3 + dy_3$,
z_3 and $z_3 + dz_3$, *etc.*

Postulate II:

The equation of motion for $\Psi(x,t)$, which determines governs the time development of a system, is given by the time-dependent Schrödinger equation:

$$i\hbar \frac{\partial \psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x,t)}{\partial x^2} + V(x)\psi(x,t),$$

where $\hbar = h/2\pi = 1.0545887 \times 10^{-27}$ erg-sec $= 1.0545887 \times 10^{-27}\ \text{g}\cdot\text{cm}^2\cdot\text{s}^{-1}$, $i = \sqrt{-1}$, m is the mass of the particle, and V(x) provides the potential energy as a function of the position coordinate, x. This equation must be suitably modified for a particle moving in more than one dimension, or for more than one particle.

Postulate III:

For every observable property, A, of a system there exists a linear, Hermitian operator, $\hat{A}$. This operator is found by writing the corresponding classical mechanical operator in terms of coordinates and their conjugate momenta, and then converting to operators by replacing the coordinates x, y, and z with operators $\hat{x}$, $\hat{y}$, and $\hat{z}$ which are simply multiplication by the coordinate, and by replacing the conjugate momenta p_x, p_y, and p_z by the differential operators $\hat{p}_x = \hbar/i\ \partial/\partial x$, $\hat{p}_y = \hbar/i\ \partial/\partial y$, and $\hat{p}_z = \hbar/i\ \partial/\partial z$.

An operator, $\hat{A}$, is linear if $\hat{A}(c_1 f + c_2 g) = c_1\hat{A}f + c_2\hat{A}g$ for all functions f and g which are acceptable candidates for wavefunctions. Here c_1 and c_2 are arbitrary constants.

An operator, $\hat{A}$, is Hermitian if
$\int f^*\hat{A}\, g\, d\tau = \int (\hat{A}f)^*\, g\, d\tau$ for all functions f and g which are acceptable candidates for wavefunctions, and the integration is taken over all space.

Postulate IV:

The only possible values that can result from measurements of the property A are the eigenvalues, a_i, of the equation

$$\hat{A}\phi_i = a_i\, \phi_i.$$

Furthermore, if a system is in a state, Ψ, which can be written as a superposition of the eigenfunctions of $\hat{A}$ as

$$\Psi = \sum_n c_n\, \phi_n \quad ,$$

then the probability that a measurement of the property A will result in the observation of the value a_i is given by

$$P(a_i) = |c_n|^2 .$$

Theorem I:

Eigenvalues of a Hermitian operator are real.

This simple statement insures that observable quantities such as position, momentum, kinetic energy, angular momentum, etc. are real, and that you can never ask a question like "Where is the particle?" and get an answer like "3.5i miles soutwest of Seattle."

PROOF: Suppose $\hat{A}$ has an eigenfunction ϕ_i with eigenvalue a_i:

$$\hat{A}\phi_i = a_i\,\phi_i$$

Multiply by ϕ_i^* and integrate, obtaining

$$\begin{aligned}\int \phi_i^*\hat{A}\phi_i\,d\tau &= \int \phi_i^* a_i\,\phi_i\,d\tau \\ &= a_i \int \phi_i^*\,\phi_i\,d\tau \\ &= a_i\end{aligned}$$

Since $\hat{A}$ is Hermitian, however, $\int f^*\hat{A}g\,d\tau = \int (\hat{A}f)^* g\,d\tau = \left[\int g^*(\hat{A}f)\,d\tau\right]^*$.

Letting $f = g = \phi_i$, $\int \phi_i{}^*\hat{A}\phi_i\,d\tau = \left[\int \phi_i^*(\hat{A}\phi_i)\,d\tau\right]^*$.

But from above we know $\int \phi_i{}^*\hat{A}\phi_i\,d\tau = a_i$, so $\left[\int \phi_i^*(\hat{A}\phi_i)\,d\tau\right]^* = a_i$.

Therefore, $\left[\int \phi_i^*(a_i\phi_i)\,d\tau\right]^* = a_i$

and $\left[a_i\right]^* = a_i$,

or $a_i^* = a_i$, which can only be true if a_i is a real number.

Theorem II:

The eigenfunctions of a linear Hermitian operator are, or in the case of a degeneracy can be chosen to be, orthonormal.

PROOF:

(1). Normality

Given a set of eigenfunctions of $\hat{A}$ which are not normalized, $\hat{A}\phi_i = a_i\,\phi_i$ and $\int \phi_i^*\phi_i\,d\tau \neq 1$, we can always define new functions, $\phi_i' \equiv c_i\,\phi_i$ such that the new functions ϕ_i' are normalized. This is done simply by choosing

$$c_i = 1/[\int \phi_i^*\phi_i\,d\tau]^{1/2}$$

Proof that this new ϕ_i' is normalized: Do the integral!

$$\int \phi_i'^*\phi_i'\,d\tau = \int (c_i\phi_i)^*(c_i\phi_i)\,d\tau = (c_i^*\,c_i)\int \phi_i^*\phi_i\,d\tau$$

$$= [1/[\int \phi_i^*\phi_i\,d\tau]^{1/2}\,] \times [1/[\int \phi_i^*\phi_i\,d\tau]^{1/2}\,] \times \int \phi_i^*\phi_i\,d\tau = [\int \phi_i^*\phi_i\,d\tau] / [\int \phi_i^*\phi_i\,d\tau] = 1$$

(2). Orthogonality

A. Let $\hat{A}\phi_i = a_i\,\phi_i$ and $\hat{A}\phi_j = a_j\,\phi_j$, where a_i and a_j are different eigenvalues.

Then, multiplying by ϕ_i^* and integrating we get:

$$\int \phi_i^*\hat{A}\phi_j\,d\tau = \int \phi_i^*\,a_j\,\phi_j\,d\tau = a_j\int \phi_i^*\,\phi_j\,d\tau$$

But $\hat{A}$ is Hermitian, so $\int \phi_i^*\hat{A}\phi_j\,d\tau = \int (\hat{A}\phi_i)^*\phi_j\,d\tau = \int (a_i\phi_i)^*\phi_j\,d\tau = \int a_i^*\phi_i^*\phi_j\,d\tau = a_i^*\int \phi_i^*\phi_j\,d\tau$

Following the equalities around we see that $a_j\int \phi_i^*\,\phi_j\,d\tau = a_i^*\int \phi_i^*\phi_j\,d\tau$, or

$$(a_j - a_i^*)\int \phi_i^*\phi_j\,d\tau = 0.$$

Furthermore, since a_i is real, $a_i = a_i^*$, and $(a_j - a_i)\int \phi_i^*\phi_j\,d\tau = 0$.

This can be true if either $(a_j - a_i) = 0$ or $\int \phi_i^*\phi_j\,d\tau = 0$.

For <u>different</u> eigenvalues ($a_j \neq a_i$), $\int \phi_i^*\phi_j\,d\tau$ must be 0, and the eigenfunctions are orthogonal.

B. If $\hat{A}\phi_i = a\,\phi_i$ and $\hat{A}\phi_j = a\,\phi_j$, we have a degenerate situation and the proof given above does not apply. However, in this case we may choose any linear combination of ϕ_i and ϕ_j, and if will still be an eigenfunction of $\hat{A}$. This is because $\hat{A}$ is a linear operator such that

$$\hat{A}(c_i\phi_i + c_j\phi_j) = c_i\,\hat{A}\phi_i + c_j\,\hat{A}\phi_j$$

$$= c_i\,a\,\phi_i + c_j\,a\,\phi_j$$

$$= a\,(c_i\phi_i + c_j\,\phi_j)$$

Choose our new linear combination of ϕ_i and ϕ_j as:

$\phi_1 = \phi_i$ and $\phi_2 = \phi_j + c\,\phi_i$, and choose c so that ϕ_2 is orthogonal to ϕ_1.

$$0 = \int \phi_1^*\phi_2\, d\tau = \int \phi_i^*(\phi_j + c\,\phi_i)\, d\tau = \int \phi_i^*\phi_j\, d\tau + c\int \phi_i^*\phi_i\, d\tau .$$

Solving for c, one obtains:

$c = -\int \phi_i^*\phi_j\, d\tau \,/\! \int \phi_i^*\phi_i\, d\tau$. This ratio of integrals must take on a defined value unless the denominator is zero, but the denominator cannot be zero if ϕ_i is any function other than 0. Thus, the two functions ϕ_i and ϕ_j can always be combined so that a new pair of functions, ϕ_1 and ϕ_2, are found which are orthogonal.

If a set of three degenerate eigenfunctions {ϕ_i, ϕ_j, and ϕ_k} exist, one simply defines new functions

$\phi_1 \equiv \phi_i$,
$\phi_2 = \phi_j + c\,\phi_i$,
and $\phi_3 = \phi_k + d\,\phi_i + e\,\phi_k$
and imposes the requirements $\int \phi_1^*\phi_2\, d\tau = 0$, $\int \phi_1^*\phi_3\, d\tau = 0$, and $\int \phi_2^*\phi_3\, d\tau = 0$, leading to three equations in the three unknown coefficients c, d, and e.

This process may be generalized to an arbitrarily large set of degenerate eigenfunctions, and is called the Schmidt orthogonalization procedure.

Theorem III:

Any arbitrary, well-behaved function, f, which satisfies the same boundary conditions as the eigenfunctions of a linear Hermitian operator, Â, can be expanded in terms of the complete set of eigenfunctions of Â as:

$$f = \sum_i c_i \phi_i$$

where $c_i = \int \phi_i^* f \, d\tau$.

PROOF: We will not prove that the set of eigenfunctions ϕ_i form a complete set — that is for the mathematicians to do. Instead we will solve for the coefficients, c_i . Begin with the assumption:

$$f = \sum_i c_i \phi_i$$

Multiply by ϕ_i^* and integrate over all space, to obtain

$$\int \phi_j^* f \, d\tau = \sum_i \int \phi_j^* c_i \phi_i \, d\tau = \sum_i c_i \int \phi_j^* \phi_i \, d\tau = \sum_i c_i \delta_{ji} = c_j$$

Thus, $c_j = \int \phi_j^* f \, d\tau$, or $c_i = \int \phi_i^* f \, d\tau$, and

$$f = \sum_i \left[\int \phi_j^* f \, d\tau \right] \phi_i .$$

Theorem IV:

If there exists a common, complete set of eigenfunctions for two linear Hermitian operators, $\hat{A}$ and $\hat{B}$, then these two operators commute such that

$$\hat{A}(\hat{B}f) = \hat{B}(\hat{A}f)$$

for any function, f.

NOTE: In general, operators do not commute. Thus, for example, the operators for position, $\hat{x}$, and momentum, $\hat{p}_x = \hbar/i\, \partial/\partial x$, do not commute:

$$\hat{x}\,\hat{p}_x f = x\,\frac{\hbar}{i}\,\frac{df}{dx} = \frac{\hbar}{i}\,x\,\frac{df}{dx}$$

but

$$\hat{p}_x\,\hat{x} f = \frac{\hbar}{i}\,\frac{d(xf)}{dx} = \frac{\hbar}{i}\left[\frac{dx}{dx}f + x\,\frac{df}{dx}\right] = \frac{\hbar}{i}\left[f + x\,\frac{df}{dx}\right].$$

From this we see that

$$(\hat{p}_x\,\hat{x} - \hat{x}\,\hat{p}_x) f = \frac{\hbar}{i}\left[f + x\,\frac{df}{dx}\right] - \frac{\hbar}{i}\,x\,\frac{df}{dx} = \frac{\hbar}{i} f.$$

This may be nicely summarized by

$$(\hat{p}_x\,\hat{x} - \hat{x}\,\hat{p}_x) f = \frac{\hbar}{i} f.$$

PROOF: By hypothesis, there is a common set of eigenfunctions of $\hat{A}$ and $\hat{B}$, which we may call $\{\phi_i\}$. Letting the eigenvalues be designated by a_i and b_i, we have

$$\hat{A}\phi_i = a_i\,\phi_i \quad \text{and} \quad \hat{B}\phi_i = b_i\,\phi_i$$

We want to prove that $(\hat{A}\,\hat{B} - \hat{B}\,\hat{A})\,\psi = 0$ for <u>any</u> ψ.
First, expand ψ in terms of the eigenfunctions ϕ_i:

$$\psi = \sum_i c_i\,\phi_i\,,$$

where the coefficients c_i are given as $c_i = \int \phi_i^* \psi\, d\tau$. Then

$$\begin{aligned}(\hat{A}\,\hat{B} - \hat{B}\,\hat{A})\,\psi &= (\hat{A}\,\hat{B} - \hat{B}\,\hat{A}) \sum_i c_i\,\phi_i \\ &= \sum_i (\hat{A}\,\hat{B} - \hat{B}\,\hat{A})\, c_i\,\phi_i\end{aligned}$$

$$= \sum_i c_i (\hat{A}\hat{B} - \hat{B}\hat{A})\phi_i$$

$$= \sum_i c_i (\hat{A}\hat{B}\phi_i - \hat{B}\hat{A}\phi_i)$$

$$= \sum_i c_i (\hat{A}\, b_i\phi_i - \hat{B}\, a_i\phi_i)$$

$$= \sum_i c_i (\, b_i\, \hat{A}\phi_i - a_i\, \hat{B}\phi_i)$$

$$= \sum_i c_i (\, b_i\, a_i\, \phi_i - a_i\, b_i\, \phi_i)$$

$$= \sum_i c_i (\, b_i\, a_i - a_i\, b_i)\, \phi_i$$

$$= \sum_i c_i (\, b_i\, a_i - b_i\, a_i)\, \phi_i$$

$$= 0$$

This is true for any ψ which is an acceptable wavefunction, so the operators $\hat{A}$ and $\hat{B}$ commute.

NOTE: For this theorem to hold, the operators $\hat{A}$ and $\hat{B}$ must share a complete set of eigenfunctions, not just a few.

Theorem V:

If $\hat{A}$ and $\hat{B}$ commute, then there exists a common, complete set of eigenfunctions for them.

NOTE: This does not mean that every eigenfunction of $\hat{A}$ must be an eigenfunction of $\hat{B}$. In the case of degeneracy, the right linear combination of eigenfunctions of $\hat{A}$ must be found for them to also be eigenfunctions of $\hat{B}$.

PROOF: Let $\hat{B}\phi_i = b_i\,\phi_i$.
Then, $\hat{A}\hat{B}\phi_i = \hat{A}b_i\,\phi_i = b_i\,\hat{A}\phi_i$.
But if $\hat{A}$ and $\hat{B}$ commute, then $\hat{A}\hat{B}\phi_i = \hat{B}\hat{A}\phi_i$,
so $\hat{B}\hat{A}\phi_i = b_i\,\hat{A}\phi_i$
or $\hat{B}(\hat{A}\phi_i) = b_i\,(\hat{A}\phi_i)$.
So both ϕ_i and $(\hat{A}\phi_i)$ are eigenfunctions of $\hat{B}$ with the same eigenvalue, b_i.

CASE A: The eigenfunctions of $\hat{B}$ are nondegenerate. This means that only one eigenfunction exists which has the eigenvalue b_i. Of course, multiples of this eigenfunction are also eigenfunctions with the same eigenvalue, but are just not properly normalized.

Because no other eigenfunctions exist which give the same eigenvalue, our eigenfunction $(\hat{A}\phi_i)$ must be a multiple of ϕ_i , which could be written as $a_i\,\phi_i$. Thus we have $\hat{A}\phi_i = a_i\,\phi_i$, and our function ϕ_i is simultaneously both an eigenfunction of $\hat{A}$ and $\hat{B}$.

CASE B: The eigenfunctions of $\hat{B}$ are degenerate. This means that there are a set of eigenfunctions of $\hat{B}$ which all have the same eigenvalue, b_i. We can designate these different eigenfunctions with the same eigenvalue as $\phi_i^{(1)}$, $\phi_i^{(2)}$, $\phi_i^{(3)}$, $\phi_i^{(4)}$, *etc*. Our result above, $\hat{B}(\hat{A}\phi_i) = b_i\,(\hat{A}\phi_i)$, is true for any one of these $\phi_i^{(k)}$, giving a corresponding equation for each of the degenerate eigenfunctions of $\hat{B}$, $\phi_i^{(k)}$:

$$\hat{B}(\hat{A}\,\phi_i^{(1)}) = b_i\,(\hat{A}\,\phi_i^{(1)})$$
$$\hat{B}(\hat{A}\,\phi_i^{(2)}) = b_i\,(\hat{A}\,\phi_i^{(2)})$$
$$\hat{B}(\hat{A}\,\phi_i^{(3)}) = b_i\,(\hat{A}\,\phi_i^{(3)})$$
$$\hat{B}(\hat{A}\,\phi_i^{(4)}) = b_i\,(\hat{A}\,\phi_i^{(4)})$$
etc.

Thus, both the set $\{\phi_i^{(1)}, \phi_i^{(2)}, \phi_i^{(3)}, \phi_i^{(4)}, \ldots\}$ and the set $\{\hat{A}\phi_i^{(1)}, \hat{A}\phi_i^{(2)}, \hat{A}\phi_i^{(3)}, \hat{A}\phi_i^{(4)}, \ldots\}$ provide a set of eigenfunctions of $\hat{B}$. Since there are only N linearly independent eigenfunctions of $\hat{B}$, we must assume that our set $\{\hat{A}\phi_i^{(1)}, \hat{A}\phi_i^{(2)}, \hat{A}\phi_i^{(3)}, \hat{A}\phi_i^{(4)}, \ldots\}$ consists of linear combinations of the set of functions $\{\phi_i^{(1)}, \phi_i^{(2)}, \phi_i^{(3)}, \phi_i^{(4)}, \ldots\}$. Thus we may write:

$$\hat{A}\phi_i^{(1)} = A_{11}\,\phi_i^{(1)} + A_{12}\,\phi_i^{(2)} + A_{13}\,\phi_i^{(3)} + \ldots$$
$$\hat{A}\phi_i^{(2)} = A_{21}\,\phi_i^{(1)} + A_{22}\,\phi_i^{(2)} + A_{23}\,\phi_i^{(3)} + \ldots$$
$$\hat{A}\phi_i^{(3)} = A_{31}\,\phi_i^{(1)} + A_{32}\,\phi_i^{(2)} + A_{33}\,\phi_i^{(3)} + \ldots$$
$$\hat{A}\phi_i^{(4)} = A_{41}\,\phi_i^{(1)} + A_{42}\,\phi_i^{(2)} + A_{43}\,\phi_i^{(3)} + \ldots$$
etc.

This may also be written in the matrix form:

$$\hat{A}\begin{pmatrix}\phi_i^{(1)}\\ \phi_i^{(2)}\\ \phi_i^{(3)}\\ \phi_i^{(4)}\end{pmatrix} = \begin{pmatrix}A_{11} & A_{12} & A_{13} & A_{14}\\ A_{21} & A_{22} & A_{23} & A_{24}\\ A_{31} & A_{32} & A_{33} & A_{34}\\ A_{41} & A_{42} & A_{43} & A_{44}\end{pmatrix}\begin{pmatrix}\phi_i^{(1)}\\ \phi_i^{(2)}\\ \phi_i^{(3)}\\ \phi_i^{(4)}\end{pmatrix}$$

If the theorem we wish to prove is true, then we should be able to find some linear combination of the degenerate eigenfunctions of $\hat{B}$, $\{\phi_i^{(1)}, \phi_i^{(2)}, \phi_i^{(3)}, \phi_i^{(4)}, \ldots\}$, which is also an eigenfunction of $\hat{A}$. Calling the new eigenfunctions of $\hat{A}$ χ_j, with eigenvalues a_j, we then obtain:

$\hat{A}\chi_j = a_j\,\chi_j$, or writing χ_j out in full as a linear combination of the $\phi_i^{(k)}$,

$$\hat{A}[C_{j1}\,\phi_i^{(1)} + C_{j2}\,\phi_i^{(2)} + C_{j3}\,\phi_i^{(3)} + \ldots] = a_j[C_{j1}\,\phi_i^{(1)} + C_{j2}\,\phi_i^{(2)} + C_{j3}\,\phi_i^{(3)} + \ldots].$$

However, we know the action of $\hat{A}$ on each $\phi_i^{(k)}$ gives a linear combination of these functions, as

$$\hat{A}\phi_i^{(k)} = A_{k1}\,\phi_i^{(1)} + A_{k2}\,\phi_i^{(2)} + A_{k3}\,\phi_i^{(3)} + \ldots,$$

so we seek the linear combination, $C_{j1}\,\phi_i^{(1)} + C_{j2}\,\phi_i^{(2)} + C_{j3}\,\phi_i^{(3)} + \ldots$, which gives:

$$[C_{j1}\,\hat{A}\phi_i^{(1)} + C_{j2}\,\hat{A}\phi_i^{(2)} + \hat{A}C_{j3}\,\phi_i^{(3)} + \ldots] = a_j[C_{j1}\,\phi_i^{(1)} + C_{j2}\,\phi_i^{(2)} + C_{j3}\,\phi_i^{(3)} + \ldots]$$

or

$$\begin{aligned}&[C_{j1}\,(A_{11}\,\phi_i^{(1)} + A_{12}\,\phi_i^{(2)} + A_{13}\,\phi_i^{(3)} + \ldots)\\ &+ C_{j2}\,(A_{21}\,\phi_i^{(1)} + A_{22}\,\phi_i^{(2)} + A_{23}\,\phi_i^{(3)} + \ldots)\\ &+ C_{j3}\,(A_{31}\,\phi_i^{(1)} + A_{32}\,\phi_i^{(2)} + A_{33}\,\phi_i^{(3)} + \ldots) + \ldots] = a_j[C_{j1}\,\phi_i^{(1)} + C_{j2}\,\phi_i^{(2)} + C_{j3}\,\phi_i^{(3)} + \ldots]\end{aligned}$$

Since the $\phi_i^{(k)}$ functions are orthonormal and everything else in this expression is a constant, we can simplify things by multiplying by $\phi_i^{(k)*}$ and integrating over all space to obtain:

$[C_{j1}\,A_{1k} + C_{j2}\,A_{2k} + C_{j3}\,A_{3k} + \ldots] = a_j[C_{jk}]$, one equation for each value of k.

Writing out several of these equations, for k=1, 2, 3, ... we get:

$$\begin{aligned}C_{j1}\,(A_{11}-a_j) + C_{j2}\,A_{21} + C_{j3}\,A_{31} + \ldots &= 0\\ C_{j1}\,A_{12} + C_{j2}\,(A_{22}-a_j) + C_{j3}\,A_{32} + \ldots &= 0\\ C_{j1}\,A_{13} + C_{j2}\,A_{23} + C_{j3}\,(A_{33}-a_j) + \ldots &= 0\end{aligned}$$

etc.

These equations may be written in matrix form, as

$$\begin{pmatrix} A_{11}-a_j & A_{21} & A_{31} & A_{41} \\ A_{12} & A_{22}-a_j & A_{32} & A_{42} \\ A_{13} & A_{23} & A_{33}-a_j & A_{43} \\ A_{14} & A_{24} & A_{34} & A_{44}-a_j \end{pmatrix} \begin{pmatrix} C_{j1} \\ C_{j2} \\ C_{j3} \\ C_{j4} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} ,$$

or more compactly, as $(\underline{\underline{A}}^t - a_j \underline{\underline{I}})\, \underline{C}_j = \underline{0}$, where $\underline{\underline{A}}^t$ is the transpose of the $\underline{\underline{A}}$ matrix. Thus, by finding the eigenvalues and eigenvectors of the $\underline{\underline{A}}^t$ matrix, the degenerate eigenfunctions of $\hat{B}$ can be formed into linear combinations which are also eigenfunctions of $\hat{A}$. This proves our theorem.

Theorem VI: **The Variational Theorem**

If ψ_0 is the true ground state wavefunction corresponding to a mechanical system specified by the Hamiltonian, $\hat{H}$, with exact ground state energy E_0, and if ϕ is a trial wavefunction which obeys the same boundary conditions as ψ_0, then the variational integral, W, obeys the following inequality:

$$W \equiv \frac{\int \phi^* \hat{H} \phi \, d\tau}{\int \phi^* \phi \, d\tau} \geq E_0 \ .$$

PROOF: By theorem III, ϕ can be expanded (in principle) in terms of the complete set of solutions to the true Hamiltonian (even though we do not know these wavefunctions). Denoting the exact solutions to $\hat{H}\psi = E\psi$ by ψ_n, with energies E_n, this gives

$$\phi = \sum c_n \psi_n \text{, where } \hat{H}\psi_n = E_n \psi_n .$$

Then, evaluating $\int \phi^* \hat{H} \phi \, d\tau$ we obtain

$$\int \phi^* \hat{H} \phi \, d\tau = \int [\sum c_n \psi_n]^* \hat{H} [\sum c_m \psi_m] \, d\tau$$

$$= \sum\sum c_n^* c_m \int \psi_n^* \hat{H} \psi_m \, d\tau$$

$$= \sum\sum c_n^* c_m \int \psi_n^* E_m \psi_m \, d\tau$$

$$= \sum\sum c_n^* c_m E_m \int \psi_n^* \psi_m \, d\tau$$

$$= \sum\sum c_n^* c_m E_m \delta_{n,m}$$

$$= \sum c_n^* c_n E_n$$

$$= \sum |c_n|^2 E_n$$

But we may also evaluate $\int \phi^* \phi \, d\tau$ in this way,

$$\int \phi^* \phi \, d\tau = \int [\sum c_n \psi_n]^* [\sum c_m \psi_m] \, d\tau$$

$$= \sum\sum c_n^* c_m \int \psi_n^* \psi_m \, d\tau$$

$$= \sum\sum c_n^* c_m \delta_{n,m}$$

$$= \sum c_n^* c_n = \sum |c_n|^2$$

Now, let's write out $\int \phi^* \hat{H} \phi \, d\tau$ in detail:

$$\int \Phi^* \hat{H} \phi \, d\tau = \sum |c_n|^2 E_n$$

$$= |c_0|^2 E_0 + |c_1|^2 E_1 + |c_2|^2 E_2 + |c_3|^2 E_3 + \ldots$$

If we choose the labels of the levels such that E_0 is the ground state energy, E_1 is the first excited state energy, *etc*., then we can be sure that

$$E_0 \le E_1 \le E_2 \le E_3 \ldots$$

This implies that $E_1 \ge E_0$, $E_2 \ge E_0$, $E_3 \ge E_0$, *etc.*

Because the energies are ordered in this way, we can be sure that

$$\int \phi^* \hat{H} \phi \, d\tau = \sum |c_n|^2 E_n = |c_0|^2 E_0 + |c_1|^2 E_1 + |c_2|^2 E_2 + |c_3|^2 E_3 + \ldots$$

$$\ge |c_0|^2 E_0 + |c_1|^2 E_0 + |c_2|^2 E_0 + |c_3|^2 E_0 + \ldots$$

$$\ge E_0 (|c_0|^2 + |c_1|^2 + |c_2|^2 + |c_3|^2 + \ldots)$$

$\ge E_0 \sum |c_n|^2$, or

$\int \phi^* \hat{H} \phi \, d\tau \ge E_0 \int \phi^* \phi \, d\tau$, which after division by $\int \phi^* \phi \, d\tau$ gives:

$$W \equiv \frac{\int \phi^* \hat{H} \phi \, d\tau}{\int \phi^* \phi \, d\tau} \ge E_0 \; .$$

COROLLARY: If the trial wavefunction ϕ is chosen to be orthogonal to the exact ground state wavefunction, ψ_0, then

$$W \equiv \frac{\int \phi^* \hat{H} \phi \, d\tau}{\int \phi^* \phi \, d\tau} \geq E_1 \ .$$

PROOF: All follows as before, up to the point

$$\int \Phi^* \hat{H} \phi \ d\tau \quad = |c_0|^2 E_0 + |c_1|^2 E_1 + |c_2|^2 E_2 + |c_3|^2 E_3 + \ldots$$

and $$\int \Phi^* \phi \ d\tau \ = |c_0|^2 + |c_1|^2 + |c_2|^2 + |c_3|^2 \ + \ldots$$

However, since the coefficients, c_n are given from Theorem III as $c_n = \int \psi_n^* \ \phi \ d\tau$, our assumption that ϕ is orthogonal to the true ground state wavefunction, ψ_0, then makes $c_0 = 0$, giving

$$\int \Phi^* \hat{H} \phi \ d\tau \quad = |c_1|^2 E_1 + |c_2|^2 E_2 + |c_3|^2 E_3 + \ldots$$

and $$\int \Phi^* \phi \ d\tau \ = |c_1|^2 + |c_2|^2 + |c_3|^2 \ + \ldots$$

Using the inequalities $E_1 \geq E_0$, $E_2 \geq E_0$, $E_3 \geq E_0$, *etc*., we can now write

$$\begin{aligned}
\int \Phi^* \hat{H} \phi \ d\tau \quad &= |c_1|^2 E_1 + |c_2|^2 E_2 + |c_3|^2 E_3 + \ldots \\
&\geq |c_1|^2 E_1 + |c_2|^2 E_1 + |c_3|^2 E_1 + \ldots \\
&\geq E_1 \left(|c_1|^2 + |c_2|^2 + |c_3|^2 + \ldots\right) \\
&\geq E_1 \left(|c_1|^2 + |c_2|^2 + |c_3|^2 + \ldots\right) \\
&\geq E_1 \sum |c_n|^2 \text{, or}
\end{aligned}$$

$$W \equiv \frac{\int \phi^* \hat{H} \phi \, d\tau}{\int \phi^* \phi \, d\tau} \geq E_1 \ .$$

By extension, if one chooses a trial wavefunction which is orthogonal to ψ_0 and ψ_1, one obtains

$$W \equiv \frac{\int \phi^* \hat{H} \phi \, d\tau}{\int \phi^* \phi \, d\tau} \geq E_2 \ , \ etc.$$

Theorem VII: **The Uncertainty Principle**

For a system described by the wavefunction ψ, the uncertainties in the properties described by the linear, Hermitian operators $\hat{A}$ and $\hat{B}$ obey the relationship

$$\Delta A \Delta B \geq \frac{1}{2}\left|\int \psi^*[\hat{A},\hat{B}]\psi\, d\tau\right|.$$

PROOF: First, define the operators $\delta\hat{A}$ and $\delta\hat{B}$ as

$$\delta\hat{A} \equiv \hat{A} - \langle\hat{A}\rangle$$
$$\delta\hat{B} \equiv \hat{B} - \langle\hat{B}\rangle.$$

Next, consider the integral

$$I = \int_{all\ space} |(\alpha\,\delta\hat{A} - i\delta\hat{B})\psi|^2 d\tau,$$

where α is an arbitrary real number. This integral is clearly greater than or equal to zero, since the integrand is never negative. Expanding this integral, we obtain

$$\int_{all\ space} \{(\alpha\,\delta\hat{A} - i\delta\hat{B})\psi\}^*\{(\alpha\,\delta\hat{A} - i\delta\hat{B})\psi\}d\tau \geq 0\ .$$

Separating the first factor in the integrand out into pieces, we obtain

$$\alpha \int_{all\ space} \{(\delta\hat{A})\psi\}^*\{(\alpha\,\delta\hat{A} - i\delta\hat{B})\psi\}d\tau \ + \ i\int_{all\ space} \{(\delta\hat{B})\psi\}^*\{(\alpha\,\delta\hat{A} - i\delta\hat{B})\psi\}d\tau \geq 0\,.$$

Next, we can make use of the fact that $\hat{A}$ and $\hat{B}$ are Hermitian (and as a consequence, $\delta\hat{A}$ and $\delta\hat{B}$ are also Hermitian) to give

$$\int_{all\ space} \psi^*\{\alpha\,\delta\hat{A}(\alpha\,\delta\hat{A} - i\delta\hat{B})\psi\}d\tau \ + \int_{all\ space} \psi^*\{i(\delta\hat{B})(\alpha\,\delta\hat{A} - i\delta\hat{B})\psi\}d\tau \geq 0\,.$$

The two integrals may now be recombined to give

$$\int_{all\ space} \psi^*(\alpha\,\delta\hat{A} + i\delta\hat{B})(\alpha\,\delta\hat{A} - i\delta\hat{B})\psi\, d\tau \geq 0,$$

which may be recognized as an expectation value and written as

$$\langle(\alpha\,\delta\hat{A} + i\delta\hat{B})(\alpha\,\delta\hat{A} - i\delta\hat{B})\rangle \geq 0\,.$$

Multiplying out the operators, separating the result into terms, and factoring out the constants then gives

$$\alpha^2\langle(\delta\hat{A})^2\rangle - i\alpha\langle\delta\hat{A}\delta\hat{B}\rangle + i\alpha\langle\delta\hat{B}\delta\hat{A}\rangle + \langle(\delta\hat{B})^2\rangle \geq 0 .$$

The two terms involving $\langle\delta\hat{A}\delta\hat{B}\rangle$ and $\langle\delta\hat{B}\delta\hat{A}\rangle$ may be combined to give

$$\alpha^2\langle(\delta\hat{A})^2\rangle - i\alpha\langle\delta\hat{A}\delta\hat{B} - \delta\hat{B}\delta\hat{A}\rangle + \langle(\delta\hat{B})^2\rangle \geq 0 .$$

Defining $\hat{C} \equiv -i\,[\hat{A}, \hat{B}]$, it is easy to show that $\hat{C} = -i\,[\delta\hat{A}, \delta\hat{B}]$. (This is obvious because the constants $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ commute with anything. Then the expression obtained above becomes

$$\alpha^2\langle(\delta\hat{A})^2\rangle + \alpha\langle\hat{C}\rangle + \langle(\delta\hat{B})^2\rangle \geq 0 .$$

Now, we can rewrite this expression in another form by completing the square, to give

$$\langle(\delta\hat{A})^2\rangle\left(\alpha + \frac{\langle\hat{C}\rangle}{2\langle(\delta\hat{A})^2\rangle}\right)^2 + \langle(\delta\hat{B})^2\rangle - \frac{\langle\hat{C}\rangle^2}{4\langle(\delta\hat{A})^2\rangle} \geq 0 .$$

This inequality must hold true for all real values of α. Setting

$$\alpha = -\frac{\langle\hat{C}\rangle}{2\langle(\delta\hat{A})^2\rangle} ,$$

we obtain the inequality

$$\langle(\delta\hat{B})^2\rangle \geq \frac{\langle\hat{C}\rangle^2}{4\langle(\delta\hat{A})^2\rangle} ,$$

which rearranges to

$$\langle(\delta\hat{A})^2\rangle\langle(\delta\hat{B})^2\rangle \geq \frac{\langle\hat{C}\rangle^2}{4} .$$

Recognizing that $\langle(\delta\hat{A})^2\rangle = \langle(\hat{A} - \langle\hat{A}\rangle)^2\rangle = \langle\hat{A}^2 - 2\,\hat{A}\langle\hat{A}\rangle + \langle\hat{A}\rangle^2\rangle$, and noting that $\langle\hat{A}\rangle$ and $\langle\hat{A}\rangle^2$ are constants, which may be pulled out of the integral, we obtain

$$\langle(\delta\hat{A})^2\rangle = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2.$$

Likewise,

$$\langle(\delta\hat{B})^2\rangle = \langle\hat{B}^2\rangle - \langle\hat{B}\rangle^2.$$

Thus,

$$\langle(\delta\hat{A})^2\rangle = (\Delta A)^2$$

and

$$\langle(\delta\hat{B})^2\rangle = (\Delta B)^2.$$

Our previous equation then becomes

$$(\Delta A)^2 (\Delta B)^2 \geq \frac{\langle \hat{C} \rangle^2}{4} \ .$$

Taking the square root and substituting $\hat{C} = -i\,[\delta\hat{A}, \delta\hat{B}]$, we obtain:

$$\Delta A \, \Delta B \geq \frac{1}{2} | \langle [\delta\hat{A}, \delta\hat{B}] \rangle | \ .$$

If we now expand out the commutator, using $\delta\hat{A} = \hat{A} - \langle\hat{A}\rangle$ and $\delta\hat{B} = \hat{B} - \langle\hat{B}\rangle$, and noting that the expectation values $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ are constants and commute with everything, we then obtain

$$\Delta A \, \Delta B \geq \frac{1}{2} | \langle [\hat{A}, \hat{B}] \rangle | \ .$$

This is the uncertainty principle in its most general form, which is valid for any two linear Hermitian operators $\hat{A}$ and $\hat{B}$. For the specific case of position, $\hat{x}$, and momentum, $\hat{p}$, we may use the fact that these operators commute to give the constant $i\hbar$, which may be pulled out of the expectation value to give

$$\boxed{\Delta x \, \Delta p_x \geq \frac{\hbar}{2} \ .}$$

This is the classic form of the Heisenberg uncertainty principle, relating position and momentum.