GENERAL STRUCTURE OF QUANTUM MECHANICS.

In Quantum Mechanics **dynamical variables** such as position, momentum, energy and angular momentum, are represented by **Hermitian operators**. Such operators are **self-adjoint** and have **real eigenvalues**, **real expectation values** and **orthonormal eigenfunctions**; these eigenfunctions constitute **complete sets**: any wave function can be expanded as a linear combination of eigenstates of any of these Hermitian operators. We have already seen and used all these facts for the energy operator, the Hamiltonian $\widehat{H}$, and its eigenfuntions. We will now generalise everything and prove all the above statements. The first 4 sections below are introductory or go beyond what I expect you to know for examinations; but sections (5) to (10) are essential.

(1) The adjoint operator, vectors and matrices.

Consider an operator $\widehat{O}$ and any two wave functions $\varphi(x)$ and $\psi(x)$; the adjoint operator $\widehat{O}^\dagger$ is defined by this equation:

$$\int_{-\infty}^{+\infty} \varphi^*(x)\widehat{O}\psi(x)\,dx \stackrel{\text{df}}{=} \int_{-\infty}^{+\infty} (\widehat{O}^\dagger\varphi(x))^*\psi(x)\,dx \tag{1}$$

In words, the adjoint is what the operator becomes when, instead of operating on the wave function to its right (ψ), it is turned over to operate on the wave function originally to its left (φ *not* φ^*). For the version of QM we are using this reversal is usually accomplished by integration by parts.

To motivate the dagger notation † and to show that it's really the analogue of the matrix dagger operation let's look at a complex $N \times N$ matrix $\mathbf{M}$ and two complex vectors $\mathbf{U}$ and $\mathbf{V}$ in N-dimensions. They can be represented by their components in some coordinate system as column vectors:

$$\mathbf{U} = \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ \vdots \\ U_N \end{pmatrix} \qquad \mathbf{V} = \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ \vdots \\ V_N \end{pmatrix} \qquad \text{and} \quad \mathbf{M} = \begin{pmatrix} M_{11} & M_{12} & \dots & M_{1N} \\ M_{21} & M_{22} & \dots & M_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ M_{N1} & M_{N2} & \dots & M_{NN} \end{pmatrix} \tag{2}$$

Recall that for matrices and vectors the adjoint operation † (often called Hermitian conjugation) takes 'the complex conjugate transpose':

$$\text{For a vector:} \quad \mathbf{U}^\dagger = (U_1^* \; U_2^* \; \dots U_N^*) \tag{3}$$

$$\text{For a matrix:} \quad \mathbf{M}^\dagger = (\mathbf{M}^*)^T \quad \text{ie.} \quad (\mathbf{M}^\dagger)_{ab} = M_{ba}^* \tag{4}$$

To form a number from $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{M}$ we form the product $\mathbf{U}^\dagger\mathbf{MV}$. Now look at the following manipulations (with the analogous operator equation written to the right):

$$\mathbf{U}^\dagger\mathbf{MV} = \sum_{a=1}^{N} U_a^*(\mathbf{MV})_a \qquad \leftrightarrow \qquad \int_{-\infty}^{+\infty} \varphi^*(x)\widehat{O}\psi(x)\,dx \tag{5}$$

$$= \sum_{a=1}^{N}\sum_{b=1}^{N} U_a^* M_{ab} V_b \qquad \text{But } U_a^* \text{ etc. are just numbers \& may be reordered at will, giving}$$

$$= \sum_{a=1}^{N}\sum_{b=1}^{N} M_{ba}^T U_a^* V_b \qquad \text{where we used the definition of the transpose: } M_{ba}^T = M_{ab}$$

$$= \sum_{b=1}^{N}(\sum_{a=1}^{N} M_{ba}^{*T} U_a)^* V_b$$

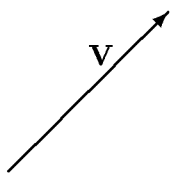
$$= \sum_{b=1}^{N}(\mathbf{M}^\dagger\mathbf{U})_b^* V_b \qquad \leftrightarrow \qquad \int_{-\infty}^{+\infty} (\widehat{O}^\dagger\varphi(x))^*\psi(x)\,dx \tag{6}$$

Notice the correspondence between the component label 'a' of vectors in $V_a, a = 1 \ldots N$ and the argument 'x' in $\psi(x)$, suggesting that $\psi(x)$ are the components of some sort of *'state vector'*: the vector and matrix component labels a take on discrete values and are summed in products from 1 to N; the component label x and the corresponding component $\psi(x)$ of the state vector are continuous and are therefore integrated from $-\infty$ to $+\infty$; the sum $\sum_a$ is the analogue of the $\int dx$. It should be emphasised that $\widehat{O}$ is not a matrix in Schrödinger's QM (there are no rows and columns in $\widehat{p}_x = -i\hbar\,\partial/\partial x$), so the above is an analogy, albeit a very close one. To find the adjoint $\widehat{O}^\dagger$ requires turning the left side of eq.(1) into the right side, usually by integration by parts, and identifying $\widehat{O}^\dagger$ in the result - see examples in part (6) below.

For students not particularly interested in the Dirac notation, skip to section (6) now; only part (5) on matrix mechanics will be essential for understanding the course.

(2) Dirac's state vector, bras and kets.

In the previous section we saw a close analogy between the component label 'a' of vectors in $V_a, a = 1 \ldots N$ and the argument 'x' in $\psi(x), x = -\infty \ldots + \infty$. This suggests that, just as $\mathbf{V}$ is a vector with components V_a, so $\psi(x)$ are the components of some sort of *'state vector'*, $|\psi\rangle$: the vector and matrix component labels a take on discrete values and are summed in products from 1 to N; the component label x and the corresponding component $\psi(x)$ of the state vector are continuous and are therefore integrated from $-\infty$ to$+\infty$ - the state vector space is infinite-dimensional, with the vectors $|\psi\rangle$ having infinitely many components $\psi(x), x = -\infty \ldots + \infty$. The abstract notation for a vector $\mathbf{V}$ in a finite-dimensional vector space has an analogue in in Dirac's abstract notation $|\psi\rangle$ for the wave function 'vector', usually called the **state vector**, or Dirac's **ket**. Vectors live in an N-dimensional vector space, kets live in an infinite dimensional complex vector space, or Hilbert space. The complex conjugate space (the dual space) is populated with **bras**, $\langle\psi|$, whose components are $\psi^*(x)$. The point of inventing a new notation here is the same as in vector calculus: in 3-dimensions we can picture a vector as a directed line without any reference to a coordinate system:

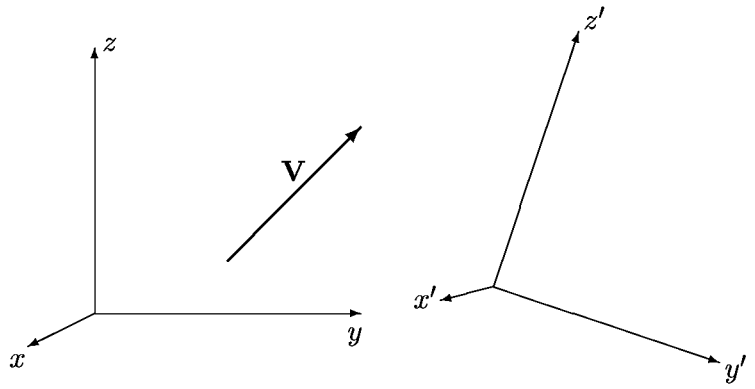


The abstract vector notation, $\mathbf{V}$ (or $\vec{V}$) enables us to carry out vector manipulations and obtain results such as the vector identity,

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A}.\mathbf{C})\,\mathbf{B} - (\mathbf{A}.\mathbf{B})\,\mathbf{C}, \tag{7}$$

which is true in any coordinate system whatever.

Two choices of coordinate axes gives two examples of alternative, but equivalent, ways of *representing* $\mathbf{V}$:



It's clear from the diagram that for the *same* vector $\mathbf{V}$ the components $V_a, a = 1,2,3$ are different in the two coordinate systems. Thus we have the abstract vector,$\mathbf{V}$, an absolute entity, and its concrete representation in terms of its coordinates $V_a, a = 1,2,3$. The latter depends on the coordinate system, having different numerical values in the primed system. (Take a simple example to illustrate this: suppose the y-axis is chosen along the unit vector $\mathbf{V}$ and the primed system is obtained by rotating about the z-axis by 90 degrees so that

$$\begin{array}{lccc} & x' = y, & y' = -x, & z' = z. \\ \text{Thus,} & V_x = 0, & V_y = 1, & V_z = 0, \\ \text{but} & V_{x'} = 1, & V_{y'} = 0, & V_{z'} = 0, \end{array} \tag{8}$$

showing clearly that the components have different values in the two systems, yet represent the same vector $\mathbf{V}$.) Note that the components of a vector are its projections along the coordinate axes.

Thus it is in a similar sense to $\mathbf{V}$ that the ket $|\psi\rangle$ stands for a vector in the Hilbert space - this is the abstract descriptor of the quantum state. And just as the components V_a, etc. form a concrete coordinate-dependent representation of the vector, so the wave function $\psi(x)$ forms a particular representation (the Schrödinger representation) of the state vector. In fact we have already met a different, but physically equivalent, representation of the state vector: this is the momentum space wave function $\phi(p)$. It was Dirac who recognised that there are many other representations possible.

(3) Orthonormality interpreted in Dirac notation.

We all know that the (length)^2 of a vector $\mathbf{V}$ is a particularly interesting quantity because it's independent of the coordinate system:

$$\mathbf{V}^2 = \mathbf{V}.\mathbf{V} = \sum_{a=1}^{N} V_a V_a = (\text{length})^2 \text{ of the vector } \mathbf{V} \tag{9}$$

Clearly its quantum mechanical analogue is the normalisation integral:

$$\langle\psi|\psi\rangle = \int_{-\infty}^{+\infty} |\psi(x)|^2\, dx = \int_{-\infty}^{+\infty} \psi^*(x)\psi(x)\, dx = 1 = (\text{length})^2 \text{ of the state vector } |\psi\rangle. \quad (10)$$

You now see why Dirac used the names bra and ket: physically measurable quantities such as the above and expectation values in general always make up a complete bracket, $\langle\ \rangle$. Note also that the normalisation implies that **the state vector is a unit vector**. Thus we see again the correspondence: the components of $\mathbf{V}$ are $V_a, a = 1...N$ for N-dimensions, which are discrete and therefore summed over; the components of the state vector $|\psi\rangle$ are $\psi(x), x = -\infty \ldots + \infty$, are continuous and are therefore integrated over (an integral is really just the limit of a sum as the variable summed over is evaluated at ever finer intervals). Note a further similarity: just as the dot product is coordinate system independent, so the quantum mechanical one is representation independent: the momentum space wave function is also normalised to one as it should to be an alternative representation of the state vector:

$$\langle\psi|\psi\rangle = \int_{-\infty}^{+\infty} |\phi(p)|^2\, dp = \int_{-\infty}^{+\infty} \phi^*(p)\phi(p)\, dp = 1 = (\text{length})^2 \text{ of the state vector } |\psi\rangle. \quad (11)$$

We are also led to see the *orthogonality* of two wave functions, say $\varphi(x)$ and $\psi(x)$, in a new light by analogy with the dot product of two vectors:

$$\mathbf{U}.\mathbf{V} = \sum_{a=1}^{N} U_a V_a = \text{projection of the vector } \mathbf{V} \text{ onto the vector } \mathbf{U}. \quad (12)$$

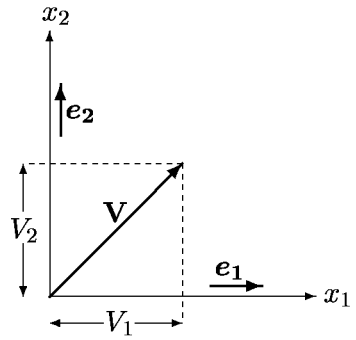
This has as its quantum mechanical analogue the so-called overlap integral:

$$\langle\varphi|\psi\rangle = \int_{-\infty}^{+\infty} \varphi^*(x)\psi(x)\, dx = \text{projection of the state vector } |\psi\rangle \text{ onto the state vector } |\varphi\rangle. \quad (13)$$

Thus, when when we say two wave functions are orthogonal if this integral vanishes, we are really saying that the state vectors are at right angles in Hilbert space.

(4) Basis vectors and the expansion theorem.

We noted above that the components of a vector are its projections onto the coordinate axes; we illustate this in 2-dimensions:



The projections V_a onto the x_a-axes are the dot products of the vector with the unit vectors or **basis vectors** $\boldsymbol{e_a}$ along the axes:

$$V_a = \boldsymbol{V}.\boldsymbol{e_a} \quad = \quad \sum_{b=1}^{N} V_b(\boldsymbol{e_a})_b$$

$$= \sum_{b=1}^{N} V_b e_{ab}, \tag{14}$$

where $(\boldsymbol{e_a})_b = e_{ab} = \delta_{ab}$ is the b-th component of the unit vector along the x_a axis. In the usual 3-dimensional notation the unit vectors along the $x-, y-, z-$axes are:

$$\begin{aligned} \boldsymbol{e_1 = i} &= (1,0,0), \\ \boldsymbol{e_2 = j} &= (0,1,0), \\ \boldsymbol{e_3 = k} &= (0,0,1). \end{aligned} \tag{15}$$

These basis vectors are of unit length and are also perpendicular to each other and therefore satisfy the **orthonormality** (normalised to 1 & orthogonal) relations:

$$\boldsymbol{e_a . e_b} = \delta_{ab} \tag{16}$$

Finally, we have the fact that *any vector* can be expanded in terms of the basis vectors:

$$\boldsymbol{V} = \sum_{a=1}^{N} V_a \, \boldsymbol{e_a} \tag{17}$$

In $N = 3$-dimensions this is the familiar expression

$$\boldsymbol{V} = V_x \, \boldsymbol{i} + V_y \, \boldsymbol{j} + V_z \, \boldsymbol{k} \tag{18}$$

All this should remind us of the **expansion theorem** in QM: the basis vectors are analogous to the eigenstates $\psi_n(x)$ and the components (V_a) correspond to the coefficients c_n:

$$\Psi(x,0) = \sum_n c_n \, \psi_n(x) \tag{19}$$

However, this is not the precise analogy of eqs.17 & 18, rather this is a particular representation (Schrödinger's) of the general state vector expansion in terms of **eigenvectors** $|\psi_n\rangle \equiv |n\rangle$, the basis vectors of QM:

$$|\Psi\rangle = \sum_n c_n \, |\psi_n\rangle \tag{20}$$

Furthermore, the orthonormality of basis vectors $\boldsymbol{e_a}$ is matched by the **orthonormality of the eigenstates**:

$$\langle \psi_m | \psi_n \rangle = \int_{-\infty}^{+\infty} \psi_m^*(x) \, \psi_n(x) \, dx \tag{21}$$

$$= \delta_{mn} \tag{22}$$

$$\text{ie. } \langle m | n \rangle = \delta_{mn} \tag{23}$$

We therefore have gained a more geometrical interpretation of the orthonormality of eigenstates, just as the name suggests, and of the expansion theorem: the eigenstates $\psi_n(x)$ of an operator in QM are the representation of basis vectors $|\psi_n\rangle$. They have unit length and are mutually orthogonal in the infinite-dimensional vector space inhabited by the state vectors of QM. The fact that for ordinary vectors we can choose different coordinate systems with their different basis vectors for representing vectors in general, has its correspondence in QM: we can choose different eigenvectors of different operators to use in the expansion theorem. So far in the course you have seen expansions in terms of energy eigenstates, $|\psi_n\rangle$,

$$\widehat{H} \, |\psi_n\rangle = E_n \, |\psi_n\rangle \tag{24}$$

(Note that this translates into $\widehat{H}\,\psi_n(x) = E_n\,\psi_n(x)$ when we project onto the basis vectors $|x\rangle$ - see above and the note below.) Shortly we shall see examples where the eigenstates used are those of the angular momentum operator.

A Comment: The projection of the general state vector $|\Psi\rangle$ onto some other basis gives the wave functions in various representations. In Schrödinger's version we use the projection, or representation, $\Psi(x,0)$, which can be obtained from $|\Psi\rangle$ by projecting onto basis vectors $|x\rangle$ which have the orthonormality relations appropriate to a continuous variable:

$$\langle x|x'\rangle = \delta(x-x') \qquad \text{(analogue of} \qquad \boldsymbol{e_a}.\boldsymbol{e_b} = \delta_{ab}) \tag{25}$$

where the Dirac delta function $\delta(x-x')$ is zero everywhere except at $x = x'$ where it is infinite, but in such a way that

$$\int_{-\infty}^{+\infty} \delta(x-x') = 1. \tag{26}$$

This completely defines the delta function! The familiar wave functions are therefore given by

$$\begin{aligned} \Psi(x,0) &= \langle x|\Psi\rangle & (27)\\ &= \sum_n c_n \langle x|\psi_n\rangle & (28)\\ &= \sum_n c_n \psi_n(x) & (29) \end{aligned}$$

ie. we have simply projected the general state vector equation into its components $\psi_n(x)$ along a different set of basis vectors $|x\rangle$ from the original basis set $|\psi_n\rangle$; the projection of these latter basis vectors along the $|x\rangle$ basis vectors is the set of eigenstates $\psi_n(x)$.

---

(5) Matrix mechanics.

Another choice within the general scheme is the one originally used in matrix mechanics, the form of quantum mechanics originally developed by Heisenberg - indeed it was the first form used and was shown to be equivalent to the form we use in this course by Schrödinger shortly after he developed wave mechanics. Heisenberg represented QM entirely by its observables, the expectation values, using matrices. The reason we discuss this topic is because the formalism is used extensively in dealing with spin in QM.

Given the eigenfuncions $\psi_a(x)$ of some suitable operator ($\widehat{O}$, say), the generalised expansion theorem allows us to write, for any wave function $\Psi(x)$,

$$\Psi(x) = \sum_a c_a\,\psi_a(x). \tag{30}$$

The expectation value of any operator $\widehat{A}$ is then

$$\begin{aligned} \langle\widehat{A}\rangle &= \int \Psi^*\widehat{A}\Psi\,dx & (31)\\ &= \sum_a\sum_b c_a^* c_b \int \psi_a^*\widehat{A}\psi_b\,dx & (32)\\ &= \mathbf{C}^\dagger\,\mathbf{A}\,\mathbf{C} & (33) \end{aligned}$$

(a) The expansion coefficients have been assembled into a column vector, $\mathbf{C}$ - you can think of

the entries as the components of the original wave function Ψ in the basis of eigenstates ψ_n:

$$\mathbf{C} = \begin{pmatrix} C_1 \\ C_2 \\ \vdots \\ \vdots \\ C_N \end{pmatrix} \qquad \mathbf{C}^\dagger = (\, C_1^* \; C_2^* \; \ldots\ldots \; C_N^*) \tag{34}$$

(In some cases N is finite, eg. the finite square well energy eigenstates, the spin eigenstates of an electron; in others N is infinite, eg. the infinite square well and harmonic oscillator energy eigenstates.) The vector $\mathbf{C}$ tells us everything we can know about the wave function Ψ and plays the role of the wave function in matrix mechanics.
(b) Similarly, the $N \times N$ matrix $\mathbf{A}$ contains all the information about the expectation value of $\widehat{A}$ for all possible situations, ie. for any possible wave function Ψ, now represented by the vector $\mathbf{C}$:

$$\mathbf{A}_{ab} = A_{ab} = \int \psi_a^* \, \widehat{A} \, \psi_b \, dx \tag{35}$$

We can also show that the rules of matrix multiplication apply, ie. the matrix representation of the operator $\widehat{A}\widehat{B}$ is indeed the product $\mathbf{AB}$ of the matrices representing $\widehat{A}$ and $\widehat{B}$ separately. The proof requires the completeness condition, which I don't intend to discuss:

$$\sum_a \psi_a^*(x) \, \psi_a(y) = \delta(x - y)$$

An important special case occurs when the eigenfunctions used in the above are eigenstates of the operator $\widehat{A}$ itself,

$$\widehat{A} \, \psi_a = \lambda_a \, \psi_a \tag{36}$$

where the eigenvalue is λ_a ($= E_a$ for energy eigenstates when $\widehat{A} = \widehat{H}$). In that case the matrix $\mathbf{A}$ is **diagonal**:

$$\begin{aligned} A_{ab} &= \int \psi_a^* \, \widehat{A} \, \psi_b \, dx \\ &= \lambda_b \int \psi_a^* \, \psi_b \, dx, \quad \text{for eigenstates of } \widehat{A}. \\ &= \lambda_b \, \delta_{ab}, \end{aligned} \tag{37}$$

where, in the last step, we used the orthonormality of eigenstates. In its full glory the matrix would look like this:

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \ldots & A_{1N} \\ A_{21} & A_{22} & A_{23} & \ldots & A_{2N} \\ A_{31} & A_{32} & A_{33} & \ldots & A_{3N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & A_{N3} & \ldots & A_{NN} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \lambda_N \end{pmatrix} \tag{38}$$

To summarise: All the consequences of quantum mechanics can be obtained from its matrix formulation: all information about the wave function is contained in the column vector $\mathbf{C}$ and all operators are represented by matrices $\mathbf{A}$; measurements yield only expectation values which are calculated from matrix products such as:

$$\langle \widehat{A} \rangle = \mathbf{C}^\dagger \, \mathbf{A} \, \mathbf{C} \tag{39}$$

The matrix **A** is diagonal in a basis of eigenstates of $\widehat{A}$, the diagonal elements being the eigenvalues of the operator.

The key point for this course is that this formalism is *essential* when one deals with purely quantum phenomena where it is not known what the appropriate variables, or coordinates (such as x) actually are. This is the case for **spin**; it is also the case for the so-called internal symmetries of elementary particle physics: strangeness, isotopic spin, flavour and colour symmetries involving groups such as SU(3).

An explicit example: SPIN-1/2.

A particle such as an electron has an intrinsic spin angular momentum $\hbar/2$, which is not understood classically because experiments show that the electron is very nearly a point particle with no spatial extension. However (see later in the course under Angular Momentum) we can still find the *matrices* representing its spin angular momentum operator $\widehat{\mathbf{S}}$; they are given by the Pauli matrices, σ_i:

$$\mathbf{S}_i = \frac{\hbar}{2}\sigma_i, \quad i = 1,2,3 \ \text{ (or, in an often used equivalent labelling, } \quad i = x,y,z), \tag{40}$$

$$\text{where:} \qquad \sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{41}$$

We note in passing that *matrices* satisfy precisely the same commutation relations as the quantum mechanical angular momentum *operators* themselves:

$$\text{for the QM operators:} \quad [\widehat{S}_x, \widehat{S}_y] = i\hbar\,\widehat{S}_z; \qquad \text{for the matrices:} \quad [\mathbf{S}_x, \mathbf{S}_y] = i\hbar\,\mathbf{S}_z \tag{42}$$

The most general wave function is represented by a two-component vector **C**, which we now call χ with $\alpha = C_1, \beta = C_2$:

$$\chi = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \qquad \chi^\dagger = (\,\alpha^* \ \beta^*), \quad \text{with normalisation: } \sum_a |c_a|^2 = |\alpha|^2 + |\beta|^2 = 1 \tag{43}$$

The eigenstates of the diagonal matrix $\mathbf{S}_z = \hbar/2\,\sigma_z$ will then be interpreted as states with definite values of the $z-$component of spin, the eigenvalues[1] $s_z = +\hbar/2$ and $s_z = -\hbar/2$:

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{44}$$

Notice the normalisation $\chi_+^\dagger\,\chi_+ = 1$ and $\chi_-^\dagger\,\chi_- = 1$ as it should be since we have the normalisation condition $\sum_a |c_a|^2 = 1$; and the orthogonality $\chi_+^\dagger\,\chi_- = 0$, as we expect for the different eigenstates of a Hermitian operator.

---

[1] Later in the course we will write these eigenvalues in terms of the quantum numbers $m_s = \pm 1/2$, the subscript s denoting 'spin':

$$s_z = \hbar\, m_s = \pm\frac{\hbar}{2}$$

(6) Self-adjoint & Hermitian operators defined & exemplified.

We now return to our starting point, the definition of the adjoint $\widehat{O}^\dagger$ of an operator $\widehat{O}$:

$$\int_{-\infty}^{+\infty} \varphi^* \widehat{O}\psi \, dx \stackrel{\text{df}}{=} \int_{-\infty}^{+\infty} (\widehat{O}^\dagger \varphi)^* \psi \, dx \tag{45}$$

for any two wave functions $\varphi(x)$ and $\psi(x)$; the case that most concerns us is of course, $\varphi(x) = \psi(x)$.
First notice two properties of the † operation which precisely match those of the corresponding matrix operation:
(a) For any complex number α the † operation acts as complex conjugation,

$$(\alpha \, \widehat{A})^\dagger = \alpha^* \, \widehat{A}^\dagger \tag{46}$$

(b) For any two operators,

$$(\widehat{A}\,\widehat{B})^\dagger = \widehat{B}^\dagger \, \widehat{A}^\dagger \tag{47}$$

Proof: The result (a) is obvious; (b) is fairly obvious. This proof repeatedly uses the definition of the adjoint and puts in *every* step:

$$\begin{aligned}
\int_{-\infty}^{+\infty} \varphi^* \widehat{A}\widehat{B}\psi \, dx &= \int_{-\infty}^{+\infty} \varphi^* \widehat{A}(\widehat{B}\psi) \, dx \\
&= \int_{-\infty}^{+\infty} \varphi^* \widehat{A}\overline{\psi} \, dx \qquad \text{where} \quad \overline{\psi} \equiv \widehat{B}\psi \\
&= \int_{-\infty}^{+\infty} (\widehat{A}^\dagger \varphi)^* \overline{\psi} \, dx \qquad \text{using def. of} \quad \widehat{A}^\dagger, \\
&= \int_{-\infty}^{+\infty} (\widehat{A}^\dagger \varphi)^* \widehat{B}\psi \, dx \\
&= \int_{-\infty}^{+\infty} \overline{\varphi}^* \widehat{B}\psi \, dx \qquad \text{where} \quad \overline{\varphi} \equiv \widehat{A}^\dagger \varphi \\
&= \int_{-\infty}^{+\infty} (\widehat{B}^\dagger \overline{\varphi})^* \, \psi \, dx \\
&= \int_{-\infty}^{+\infty} (\widehat{B}^\dagger \widehat{A}^\dagger \varphi)^* \, \psi \, dx \qquad \textbf{Q.E.D.}
\end{aligned} \tag{48}$$

We end our preparation with a definition of great importance in QM:
A **self-adjoint, or Hermitian operator** is one for which:

$$\widehat{O}^\dagger = \widehat{O} \qquad \text{Def. of } \textbf{self-adjoint} \text{ or } \textbf{Hermitian operator} \tag{49}$$

To understand the foregoing general discussion better let us find the adjoint of some simple operators.
(i) Since x is a real number, the result (a) above implies,

$$\widehat{x}^\dagger = x^* = x \tag{50}$$

(ii) The differential operator, $\partial/\partial x$; we move it to the left by integrating by parts and discarding terms vanishing at $\pm\infty$ because all wave functions vanish there:

$$\begin{aligned}
\int_{-\infty}^{+\infty} \varphi^* \, \frac{\partial}{\partial x} \psi \, dx &= [\varphi^* \, \psi]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} (\frac{\partial}{\partial x}\varphi^*) \, \psi \, dx \qquad \text{after integrating by parts;} \\
&= -\int_{-\infty}^{+\infty} (\frac{\partial}{\partial x}\varphi)^* \, \psi \, dx \quad \text{discarding terms at } \pm\infty, \text{ \& using } x \text{ real.}
\end{aligned} \tag{51}$$

Hence we discover that the differential operator is *not self-adjoint*, but anti-self-adjoint:

$$(\frac{\partial}{\partial x})^\dagger = -\frac{\partial}{\partial x} \tag{52}$$

Clearly, however, we can reverse this minus sign with a facor of i, so that the momentum operator *is self-adjoint*:

$$\widehat{p}_x^\dagger = (-i\hbar\frac{\partial}{\partial x})^\dagger = (-i\hbar)^*(\frac{\partial}{\partial x})^\dagger = -i\hbar\frac{\partial}{\partial x} \qquad \text{ie.} \quad \widehat{p}_x^\dagger = \widehat{p}_x \tag{53}$$

(iii) For real potentials, $V(x,t)$, the Hamiltonian is also Hermitian:

$$\widehat{H}^\dagger = \widehat{H} \tag{54}$$

This can either be proved by doing the integral twice by parts for

$$\widehat{p}_x^2 = -\hbar^2\frac{\partial^2}{\partial x^2} \quad \text{in the Hamiltonian} \quad \widehat{H} = \frac{\widehat{p}^2}{2m} + V(x,t),$$

thereby generating *two* minus signs and discarding terms at $\pm\infty$; or one can use the result (b) proved above, $(\widehat{A}\widehat{B})^\dagger = \widehat{B}^\dagger\widehat{A}^\dagger$, with $\widehat{A} = \widehat{B} = \widehat{p}_x$ and $\widehat{p}_x^\dagger = \widehat{p}_x$.

(7) Self-adjoint & Hermitian operators in QM.

What part do self-adjoint (or Hermitian) operators play in QM?
Hermitian operators have real expectation values and so are the only candidates for physical observables in QM.
Proof: For an operator $\widehat{A}$ the expectation value in a state $\psi(x)$ is

$$\begin{aligned} \langle\widehat{A}\rangle &= \int_{-\infty}^{+\infty}\psi^*\widehat{A}\psi\,dx \\ &= \int_{-\infty}^{+\infty}(\widehat{A}^\dagger\psi)^*\psi\,dx. \quad \text{using def. of } \widehat{A}^\dagger \end{aligned} \tag{55}$$

Taking the complex conjugate of this last line gives an expression for the complex conjugate of the expectation value:

$$\begin{aligned} \langle\widehat{A}\rangle^* &= \int_{-\infty}^{+\infty}(\widehat{A}^\dagger\psi)\psi^*\,dx \\ &= \int_{-\infty}^{+\infty}\psi^*\widehat{A}^\dagger\psi\,dx \\ &= \langle\widehat{A}\rangle \qquad \textit{only if } \widehat{A}^\dagger = \widehat{A}. \end{aligned} \tag{56}$$

(Note that the second step follows because the wave functions ψ^* and $(\widehat{A}^\dagger\psi)$ are just functions and can be interchanged if convenient.) This is the physical motivation for the important
POSTULATE: In QM all physical observables are represented by Hermitian operators.

We now consider the eigenfunctions $\psi_i(x)$ and the corresponding eigenvalues λ_i of Hermitian operators:

$$\widehat{A}\,\psi_i(x) = \lambda_i\,\psi_i(x) \tag{57}$$

(i) The eigenvalues of a Hermitian operator are real, and
(ii) the eigenfunctions of a Hermitian operator are orthonormal.

Proof: We can prove both by evaluating the following integral in two ways: one directly; the other by using Hermiticity of the operator.

$$\begin{aligned}\langle\widehat{A}\rangle = \int_{-\infty}^{+\infty} \psi_i^*\, \widehat{A}\psi_j\, dx &= \lambda_j \int_{-\infty}^{+\infty} \psi_i^*\, \psi_j\, dx \qquad (58)\\ &= \int_{-\infty}^{+\infty} (\widehat{A}^\dagger\psi_i)^*\, \psi_j\, dx \quad \text{from def. of } \widehat{A}^\dagger\\ &= \int_{-\infty}^{+\infty} (\widehat{A}\,\psi_i)^*\, \psi_j\, dx \quad \text{from Hermiticity } \widehat{A}^\dagger = \widehat{A}\\ &= \lambda_i^* \int_{-\infty}^{+\infty} \psi_i^*\, \psi_j\, dx \qquad (59)\end{aligned}$$

In deriving both the first and last lines we have used the assumption that the ψ_i are eigenstates of $\widehat{A}$. Putting these lines together gives the condition:

$$(\lambda_j - \lambda_i^*) \int_{-\infty}^{+\infty} \psi_i^*\, \psi_j\, dx = 0 \qquad (60)$$

To prove (i) we put $j = i$:

$$(\lambda_i - \lambda_i^*) \int_{-\infty}^{+\infty} |\psi_i|^2\, dx = 0 \qquad (61)$$

Since the integrand $|\psi_i|^2$ is positive the integral cannot vanish (indeed we always normalise the eigenfunctions by making *this* integral= 1); therefore the other factor must vanish,

$$\lambda_i^* = \lambda_i, \quad \textbf{ie. the eigenvalues are real.} \qquad \textbf{Q.E.D.(i)} \qquad (62)$$

To prove (ii) we now use the reality just proved, $\lambda_i^* = \lambda_i$, with $j \neq i$:

$$(\lambda_i - \lambda_j) \int_{-\infty}^{+\infty} \psi_i^*\, \psi_j\, dx = 0 \qquad (63)$$

Clearly, for $i = j$ the equation is automatically satisfied; while for $i \neq j$, assuming that there is no degeneracy [2], then $\lambda_i \neq \lambda_j$, which means the integral has to vanish. This is **orthogonality**:

$$\int_{-\infty}^{+\infty} \psi_i^*(x)\, \psi_j(x)\, dx = 0 \text{ for } i \neq j. \qquad \text{ORTHOGONALITY} \qquad \textbf{Q.E.D.(ii)} \qquad (64)$$

Combining this with normalisation to one gives **orthonormality**:

$$\int_{-\infty}^{+\infty} \psi_i^*(x)\, \psi_j(x)\, dx = \delta_{ij} \quad \text{ORTHONORMALITY} \qquad (65)$$

Again note that we have seen these results and the proof before for the eigenvalues ($\lambda_i = E_i$) and eigenstates of the Hamiltonian - the energy eigenstates. Now we see that the crucial ingredient was Hermiticity, a property essential for a dynamical observable in quantum mechanics.

(8) The expansion theorem and the measurement postulate.

[2] Degeneracy is the occurrence of *more than one eigenfunction with the same eigenvalue.* It can be shown that it's always possible to construct a set of mutually orthogonal eigenfunctions from linear combinations of degenerate ones - see Bransden, Joachain, p.199.

The eigenstates of Hermitian operators also form complete sets of functions, which means that any wave function Ψ may be expressed as a linear superposition of these eigenfunctions:

$$\Psi = \sum_i c_i \psi_i \tag{66}$$

where the c_i are constants. The measurement postulate then states that for a system prepared in the above state Ψ, the only possible outcome of a single measurement of the observable corresponding to the operator $\widehat{A}$ is one of the eigenvalues λ_j, with probability $|c_j|^2$
Immediately after an ideal measurement yielding the value λ_j, the state of the system 'collapses' to the corresponding eigenstate:

$$\Psi_{after} = \psi_j \tag{67}$$

If the system is left undisturbed until a subsequent measurement is made, the result of that measurement is the same eigenvalue, λ_j, with probability 1.

We note that the requirement that Ψ be a normalised state and the orthonormality of the eigenstates ψ_i shows us that

$$\sum_i |c_i|^2 = 1 \tag{68}$$

$$\text{with average, or expectation value} \quad \langle \widehat{A} \rangle = \sum_i |c_i|^2 \lambda_i \tag{69}$$

The **Proofs** are the same as for the energy operator:

$$\text{Normalisation:} \quad 1 = \int |\Psi|^2 \, dx \tag{70}$$

$$= \sum_i \sum_j c_i^* c_j \int \psi_i^* \psi_j \, dx \quad \text{from expansion theorem,} \tag{71}$$

$$= \sum_i |c_i|^2 \quad \text{using orthonormality.} \quad \textbf{Q.E.D.} \tag{72}$$

Likewise the expectation value can be written out using the expansion theorem:

$$\langle \widehat{A} \rangle \equiv \int \Psi^* \widehat{A} \Psi \, dx \tag{73}$$

$$= \sum_i \sum_j c_i^* c_j \int \psi_i^* \widehat{A} \psi_j \, dx \quad \text{from expansion theorem,} \tag{74}$$

$$= \sum_i \sum_j c_i^* c_j \lambda_j \int \psi_i^* \psi_j \, dx \quad \text{from } \widehat{A}\psi_j = \lambda_j \psi_j \tag{75}$$

$$= \sum_i |c_i|^2 \lambda_i \quad \text{using orthonormality.} \quad \textbf{Q.E.D.} \tag{76}$$

(9) Compatible observables.

So far we have studied one special case of the above general development, with the operator being the Hamiltonian $\widehat{H}$ which, as we have seen, is indeed Hermitian. The actual form of the eigenfunctions varies according to the potential, but we have seen that all the properties enunciated above are satisfied. For the energy eigenfunctions of symmetric wells, $V(-x) = V(x)$, we have also seen another phenomenon: these energy eigenstates are simultaneously eigenstates of parity; and we showed that the symmetry of the potential under mirror reflection, $x \to -x$, implies that the parity operator and the Hamiltonian commute, $[\widehat{H}, \widehat{\mathcal{P}}] = 0$.

We now show that this can be generalised to any pair of commuting observables. That this is so is not too surprising when we remember the generalised Heisenberg uncertainty relation for two observables A and B (proof in Bransden, Joachain, p.208):

$$\Delta A \, \Delta B \geq \frac{1}{2} \, |\langle [\widehat{A}, \widehat{B}] \rangle| \tag{77}$$

Not only does this relation show that two non-commuting observables (such as $\widehat{p}_x$ and x) cannot *both* be measured to 100% precision, but it also shows that if the observables *do* commute, such measurement is possible. Thus for two commuting observables a quantum state can be an eigenstate of *both observables*.

Proof: (i) Assume $[\widehat{A}, \widehat{B}] = 0$; we will show that eigenstates of $\widehat{A}$ are also eigenstates of $\widehat{B}$.
Let ψ be an eigenstate of $\widehat{A}$:

$$\widehat{A} \, \psi = \lambda \, \psi \tag{78}$$

Then,

$$\begin{aligned} \widehat{B}(\widehat{A} \, \psi) &= \lambda \, (\widehat{B} \, \psi) \;\; \text{since } \psi \text{ is an eigenfunction of } \widehat{A} \\ &= \widehat{A}(\widehat{B} \, \psi) \;\; \text{since } \; \widehat{A} \, \widehat{B} = \widehat{B} \, \widehat{A} \end{aligned} \tag{79}$$

Comparing the two quantities on the right shows that $\widehat{B}\psi$ is also an eigenfunction of $\widehat{A}$ with the *same* eigenvalue λ as ψ itself. Assuming the eigenfunctions are unique (ie. not degenerate - see previous footnote), we conclude that $\widehat{B}\psi$ can only be a constant (call it μ) times ψ:

$$\widehat{B} \, \psi = \mu \, \psi, \tag{80}$$

which just states that ψ is also an eigenstate of $\widehat{B}$, with eigenvalue μ —- Q.E.D.
Proof: (ii) The converse is also true: if the observables have a common eigenstate, then they commute. Let ψ be the common eigenfunction:

$$\begin{aligned} \widehat{A} \, \widehat{B} \psi &= \lambda \, \widehat{A} \, \psi \quad \text{since } \; \widehat{A} \, \psi = \lambda \, \psi \\ &= \lambda \, \mu \, \psi \quad \text{since } \; \widehat{B} \, \psi = \mu \, \psi \\ &= \mu \, \lambda \, \psi \quad \text{since the constants can be interchanged} \\ &= \widehat{B} \, \widehat{A} \psi \quad \text{by reversing the above steps} \end{aligned} \tag{81}$$

But since this is true for *all* the eigenstates and these eigenstates make up a complete set for expanding an *arbitrary* wave function then it must be true for the operators alone:

$$\text{ie. } \widehat{A} \, \widehat{B} = \widehat{B} \, \widehat{A} \qquad \text{ie. } [\widehat{A}, \widehat{B}] = 0. \tag{82}$$

(The degenrate case is treated by Bransden, Joachain, p.207, showing that the results still hold in that case.)
The arguments just given can be extended to as many commuting observables as can be found, so that a complete set of simultaneous eigenfunctions exists. For a given quantum system, the largest set of mutually commuting observables is known as a **compatible** or **a complete set of commuting observables**. Their physical significance is that the set of all their eigenvalues specifies the eigenstates completely. The one example we will study, albeit superficially, are the states of the electron in the Hydrogen atom: the set includes the energy, the parity, the total angular momentum, its $z-$component, the orbital angular momentum and its z-component, the electron spin and its z-component, etc. The list extends further the more physical effects are included.

(10) Recovering classical mechanics from QM - Ehrenfest's Theorem.

So far we have single-mindedly pursued QM without reference to the impressive successes of classical mechanics - in the guise of Newton's second law (Newton II) - in describing the motion of everyday objects around us as well as the dynamics of the solar system and stellar systems in general. Against this must be put the overwhelming successes of quantum mechanics in describing microscopic phenomena: atoms, their stability and energy levels, the periodic table, the interaction of atoms with radiation; molecules and the forces binding them, particularly the covalent exchange force which we discovered in the context of the H_2^+ ion, and crucial in all chemistry, especially organic chemistry; the behaviour of solids and electrons in solids, including superconductivity; the bizarre phenomenon of superfluidity of liquid helium at low temperatures; our understanding of white dwarfs, neutron stars and supernovae; and finally, the entire field of nuclear and elementary particle physics.

With the accumulated experience gained from countless experiments since the time of Galileo, we have every reason to believe that Newtonian laws are an excellent approximation for macroscopic systems; but we also recognise that quantum mechanics must be a more accurate form whose need is only recognised once we begin to probe atomic and sub-atomic phenomena. We therefore expect to find Newton II hidden somewhere inside QM. Let us first try to guess what form Newton II might take when derived from QM. In QM the variables x and p_x are actually represented by operators $\widehat{x}$ and $\widehat{p}_x$ and quantum states do not have unique values for position and momentum. For any *single* measurement on a quantum system we can only predict the *probabilities* of certain outcomes; in an ensemble measurement we determine all the possible outcomes and their probabilities, thereby measuring the average or expectation values, $\langle\widehat{\mathbf{r}}\rangle$ and $\langle\widehat{\mathbf{p}}\rangle$. Since these seem to be the only *measurable* quantities related to position and momentum, it seems reasonable to expect these to correspond to the *classical* quantities. With this guess we might expect the x-component of Newton II,

$$\frac{dp_x}{dt} = F_x(\mathbf{r},t) = -\frac{\partial V(\mathbf{r},t)}{\partial x} \tag{83}$$

to emerge from QM in the form with the replacements $\mathbf{r} \to \langle\widehat{\mathbf{r}}\rangle$ and $p_x \to \langle\widehat{p}_x\rangle$,

$$\frac{d\,\langle\widehat{p}_x\rangle}{dt} = F_x(\langle\widehat{\mathbf{r}}\rangle,t) = -\frac{\partial V(\langle\widehat{\mathbf{r}}\rangle,t)}{\partial\langle x\rangle}. \tag{84}$$

As we shall see, this form emerges only in the approximation of well-localised particles in a force field with a gentle enough spatial variation. Throughout we shall work in 1-dimension to simplify the discussion, although the results we obtain are quite general. In 1-D the expected equation above therefore becomes:

$$\frac{d\,\langle\widehat{p}_x\rangle}{dt} = F_x(\langle\widehat{x}\rangle,t) = -\frac{\partial V(\langle\widehat{x}\rangle,t)}{\partial\,\langle x\rangle}. \tag{85}$$

(a) The general quantum equation of motion: Heisenberg's equation.

In the following we shall use the TDSE and its complex conjugate for a general Hamiltonian,

$$\widehat{H}\,\Psi(x,t) = i\hbar\frac{\partial\,\Psi(x,t)}{\partial\,t} \qquad \text{and the complex conjugate,} \qquad (\widehat{H}\,\Psi(x,t))^* = -i\hbar\frac{\partial\,\Psi^*(x,t)}{\partial\,t} \tag{86}$$

(Of course x and t are real.) We shall also make use of the fact that the Hamiltonian is Hermitian (for a real potential) to rewrite the complex conjugate equation:

$$\text{For} \quad \widehat{H}^\dagger = \widehat{H}, \quad (\widehat{H}\,\Psi(x,t))^* = -i\hbar\frac{\partial\,\Psi^*(x,t)}{\partial\,t} \quad \text{becomes} \quad (\widehat{H}^\dagger\,\Psi(x,t))^* = -i\hbar\frac{\partial\,\Psi^*(x,t)}{\partial\,t} \tag{87}$$

We can then use the *general definition*, eq.(1), of the adjoint operation †, with $\varphi(x) \to \Psi(x,t)$, $\widehat{O} \to \widehat{H}$ and $\psi \to \overline{\Psi}(x,t)$ to make the following transformation:

$$\int_{-\infty}^{+\infty} (\widehat{H}^\dagger \Psi(x,t))^* \, \overline{\Psi}(x,t)\, dx = \int_{-\infty}^{+\infty} \Psi^*(x,t) \widehat{H} \overline{\Psi}(x,t)\, dx. \tag{88}$$

In applying this below we will identify $\overline{\Psi}$ as $\widehat{O}\Psi$.

We first consider the time-dependence of an operator $\widehat{O}$.

(i) A particularly simple example is the momentum operator $\widehat{p}_x$, which, from the representation $\widehat{p}_x = -i\hbar\partial/\partial x$, clearly has no explicit time-dependence; thus, $\partial\widehat{p}_x/\partial t = 0$. Similarly for the Hamiltonian of a free particle, $\widehat{H}_{free} = \widehat{p}^2/2m$ and for a particle in a time-independent field, $\widehat{H} = \widehat{p}^2/2m + V(x)$.

(ii) But for a time-dependent potential, $\widehat{H} = \widehat{p}^2/2m + V(x,t)$ there *is* explicit time-dependence: $\partial\widehat{H}/\partial t = \partial V(x,t)/\partial t \neq 0$. Suppose, for example, the particle is electrically charged and is placed in a spatially uniform electric field E pointing in the $x-$direction. The particle feels a force in the $x-$direction, $F_x = qE$, where q is its charge and a corresponding potential $V(x) = -x\,qE$ (remember $F_x = -\partial V/\partial x$). This gives a Hamiltonian, $\widehat{H} = \widehat{p}^2/2m - x\,qE$, which also has no explicit time dependence if the electric field is not varied, so $\partial\widehat{H}/\partial t = 0$. But if the electric field is made to oscillate periodically, $E = E(t) = E_0 \cos\omega t$; the Hamiltonian acquires *explicit time-dependence*, $\widehat{H} = \widehat{p}^2/2m - x\,qE_0 \cos\omega t$, and hence $\partial\widehat{H}/\partial t = x\omega qE_0 \sin\omega t \neq 0$.

We now consider the time-dependence of the same operator's *expectation value*. This acquires time dependence from *two* distinct contributions: the explicit time-dependence of the operator $\widehat{O}$ exemplified above, and the time-dependence of the wave function $\Psi(x,t)$. Thus,

$$\begin{aligned}
\frac{d\langle\widehat{O}\rangle}{dt} &= \frac{d}{dt}\int \Psi^*(x,t)\,\widehat{O}\,\Psi(x,t)\,dx \qquad (89)\\
&= \int \frac{\partial\Psi^*(x,t)}{\partial t}\,\widehat{O}\,\Psi(x,t)\,dx + \int \Psi^*(x,t)\,\widehat{O}\,\frac{\partial\Psi(x,t)}{\partial t}\,dx + \int \Psi^*(x,t)\,\frac{\partial\widehat{O}}{\partial t}\,\Psi(x,t)\,dx,\\
&= \frac{1}{i\hbar}\left\{-\int (\widehat{H}\Psi)^*\,\widehat{O}\,\Psi\,dx + \int \Psi^*\,\widehat{O}\,(\widehat{H}\Psi)\,dx\right\} + \langle\,\frac{\partial\widehat{O}}{\partial t}\,\rangle \quad \text{using the TDSE and (TDSE)*},\\
&= \frac{1}{i\hbar}\left\{-\int (\widehat{H}^\dagger\Psi)^*\,\widehat{O}\,\Psi\,dx + \int \Psi^*\,\widehat{O}\,(\widehat{H}\Psi)\,dx\right\} + \langle\,\frac{\partial\widehat{O}}{\partial t}\,\rangle \quad \text{since} \quad \widehat{H} \ \text{is Hermitian},\\
&= \frac{1}{i\hbar}\left\{-\int \Psi^*\,\widehat{H}\widehat{O}\,\Psi\,dx + \int \Psi^*\,\widehat{O}\widehat{H}\Psi\,dx\right\} + \langle\,\frac{\partial\widehat{O}}{\partial t}\,\rangle \quad \text{using definition of} \quad \widehat{H}^\dagger,\\
&= -\frac{1}{i\hbar}\int \Psi^*\left\{\widehat{H}\widehat{O} - \widehat{O}\widehat{H}\right\}\Psi\,dx + \langle\,\frac{\partial\widehat{O}}{\partial t}\,\rangle \quad \text{taking out common factors \& a minus sign},\\
&= \frac{i}{\hbar}\int \Psi^*\,[\widehat{H},\widehat{O}]\,\Psi\,dx + \langle\,\frac{\partial\widehat{O}}{\partial t}\,\rangle. \qquad (90)
\end{aligned}$$

The final result is **Heisenberg's equation of motion:**

$$\frac{d\langle\widehat{O}\rangle}{dt} = \frac{i}{\hbar}\,\langle\,[\widehat{H},\widehat{O}]\,\rangle + \langle\,\frac{\partial\widehat{O}}{\partial t}\,\rangle. \tag{91}$$

In this form, QM already begins to resemble classical mechanics. Problems may be solved using Heisenberg's equation instead of Schrödinger's, and using physical observables $\langle\widehat{O}\rangle$ instead of the wave function. An example will be given in a problem sheet.

Heisenberg's equation shows, for an operator $\widehat{O}$, the consequences of commutation with the Hamiltonian (which itself is intimately related to a *symmetry* of the Hamiltonian): provided the operator has no explicit time dependence, commutation with the Hamiltonian, $[\widehat{H},\widehat{O}] = 0$ implies the conservation of the expectation value $\langle\,\widehat{O}\,\rangle$, ie. $d\,\langle\,\widehat{O}\,\rangle/dt = 0$. Thus, as we have seen

already, for a free particle, $[\widehat{H}, \widehat{\mathbf{p}}] = 0$, $\widehat{\mathbf{p}}$ has no explicit time dependence, and so momentum is conserved *on average*, $d\langle\,\widehat{\mathbf{p}}\,\rangle/dt = 0$. This chain of connections is at the very heart of fundamental physics: a symmetry of the dynamics (ie. of the Hamiltonian) under some transformation represented by an operator $\widehat{O}$, leads to commutation with the Hamiltonian, $[\widehat{H}, \widehat{O}] = 0$, which leads, via Heisenberg's equation, to a conservation law, $d\langle\,\widehat{O}\,\rangle/dt = 0$. In modern elementary particle physics theories *begin* with symmetries, which are believed to be the truly fundamental principles in nature.

(b) Ehrenfest's Theorem & how to recover Newton II.

Ehrenfest's theorem corresponds to the special choice $\widehat{O} = \widehat{p}_x$. Heisenberg's equation now simplifies because $\widehat{p}_x$ has no explicit time dependence and the commutator gives:

$$\begin{aligned}
[\widehat{H}, \widehat{p}_x]\psi(x) &= \left[\frac{\widehat{p}_x^2}{2m} + V(x,t), \widehat{p}_x\right]\psi(x) \qquad (92)\\
&= \frac{1}{2m}[\widehat{p}_x^2, \widehat{p}_x]\psi(x) + [V(x,t), \widehat{p}_x]\psi(x).\\
&\quad \text{With the first commutator vanishing and using} \quad \widehat{p}_x = -i\hbar\frac{\partial}{\partial x},\\
&= -i\hbar\left\{V(x,t)\frac{\partial\psi(x)}{\partial x} - \frac{\partial(V(x,t)\psi(x))}{\partial x}\right\}\\
&= -i\hbar\left\{V(x,t)\,\frac{\partial\psi(x)}{\partial x} - \frac{\partial V(x,t)}{\partial x}\,\psi(x) - V(x,t)\,\frac{\partial\psi(x)}{\partial x}\right\}\\
&= \left(i\hbar\,\frac{\partial V(x,t)}{\partial x}\right)\psi(x). \qquad (93)
\end{aligned}$$

Since this holds for arbitrary $\psi(x)$, the operator identity follows:

$$[\widehat{H}, \widehat{p}_x] = i\hbar\,\frac{\partial V(x,t)}{\partial x}, \qquad (94)$$

leading to **Ehrenfest's Theorem:** [3]

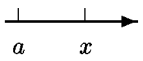
$$\frac{d\,\langle\widehat{p}_x\rangle}{dt} = -\langle\,\frac{\partial V(x,t)}{\partial x}\,\rangle = \langle\,F_x(x,t)\,\rangle. \qquad (95)$$

This is not quite the expected form and so does not quite correspond to the identification of expectation values of operators with their classical counterparts. To see the conditions under which this identification *is* possible we carry out a Taylor expansion [4] about the average position $\langle x\rangle$,

$$F(x) = F(\langle x\rangle) + (x - \langle x\rangle)\,F'(\langle x\rangle) + \frac{1}{2}(x - \langle x\rangle)^2\,F''(\langle x\rangle) + \ldots, \qquad (96)$$

[3] Ehrenfest's Theorem shows very clearly the conservation of momentum in the absence of external forces, $V = 0$, in much the same way as in Newtonian mechanics, but now for the classical momentum, $\mathbf{p}$.

[4]



We have chosen $a = \langle x\rangle$ in the Taylor expansion

$$f(x) = f(a) + \frac{(x-a)}{1!}\,f'(a) + \frac{(x-a)^2}{2!}\,f''(a) + \ldots, \quad \text{where} \quad f'(a) = \left(\frac{d\,f(x)}{dx}\right)_{x=a}$$

where we omitted the component subscript x to avoid confusion. Now we evaluate the average, or expectation value, using the fact that the average of a constant such as $\langle x \rangle$ or $F(\langle x \rangle)$ is the constant itself:

$$\begin{aligned} \langle F(x) \rangle &= F(\langle x \rangle) + (\langle x \rangle - \langle x \rangle) F'(\langle x \rangle) + \frac{1}{2} \langle (x - \langle x \rangle)^2 \rangle F''(\langle x \rangle) + \ldots, \\ &= F(\langle x \rangle) + \frac{1}{2} (\Delta x)^2 F''(\langle x \rangle) + \ldots, \end{aligned} \tag{97}$$

where the uncertainty, or variance is obtained from:

$$\begin{aligned} \langle (x - \langle x \rangle)^2 \rangle &= \langle\, x - 2x\langle x \rangle + \langle x \rangle^2 \,\rangle \\ &= \langle x \rangle - 2\langle x \rangle \langle x \rangle + \langle x \rangle^2 \\ &= \langle x \rangle - \langle x \rangle^2 \\ &= (\Delta x)^2 \end{aligned} \tag{98}$$

Thus, we obtain the classical equation only when both Δx is small, ie. *the particle is well localised*, and when the $F''(\langle x \rangle)$ is small, ie. *there is only a gradual spatial variation of the force field.* It is under these conditions that we recover the expected classical equation of motion with the classical dynamical variables corresponding to the expectation values of the corresponding quantum operators.

APPENDIX

HEISENBERG'S GENERALISED UNCERTAINTY RELATION.

Take two Hermitian operators representing physical observables: $\widehat{A}$ and $\widehat{B}$. We start by defining the two related operators which are also Hermitian,

$$\widehat{U} = \widehat{A} - \langle \widehat{A} \rangle \tag{99}$$

$$\widehat{V} = \widehat{B} - \langle \widehat{B} \rangle. \tag{100}$$

These are much simpler than the original operators since,

$$\langle \widehat{U} \rangle = 0 \quad \text{and} \quad \langle \widehat{U}^2 \rangle = (\Delta A)^2 \tag{101}$$

$$\langle \widehat{V} \rangle = 0 \quad \text{and} \quad \langle \widehat{V}^2 \rangle = (\Delta B)^2 \tag{102}$$

$$\text{while,} \quad [\widehat{U}, \widehat{V}] = [\widehat{A}, \widehat{B}]. \tag{103}$$

These all follow because the expectation values $\langle \widehat{U} \rangle$, etc are just numerical constants whose expectation values are themselves and which commute with everything.

The next step is to consider an operator $\widehat{U} + i\lambda\widehat{V}$, where λ is a *real number*, and the function Φ obtained by operating on Ψ:

$$\Phi \equiv (\widehat{U} + i\lambda\widehat{V})\Psi \tag{104}$$

Since $|\Phi|^2$ is real and positive, this function has a positive norm,

$$0 \leq \int \Phi^* \Phi \, dx \equiv I(\lambda) \tag{105}$$

$$= \int \{(\widehat{U} + i\lambda\widehat{V})\Psi\}^* (\widehat{U} + i\lambda\widehat{V})\Psi \, dx \tag{106}$$

$$= \int \{(\widehat{U} - i\lambda\widehat{V})^\dagger \Psi\}^* (\widehat{U} + i\lambda\widehat{V})\Psi \, dx \quad \text{using definition of the adjoint operator,} \tag{107}$$

$$= \int \Psi^* (\widehat{U} - i\lambda\widehat{V})(\widehat{U} + i\lambda\widehat{V})\Psi \, dx \quad \text{moving the operator over to the right,} \tag{108}$$

$$= \int \Psi^* (\widehat{U}^2 + \lambda^2\widehat{V}^2 + i\lambda\widehat{U}\widehat{V} - i\lambda\widehat{V}\widehat{U})\Psi \, dx \quad \text{expanding the operator product.} \tag{109}$$

In the third line we used the fact that $\widehat{U}$ and $\widehat{V}$ are Hermitian and that λ is real to write

$$(\widehat{U} - i\lambda\widehat{V})^\dagger = (\widehat{U} + i\lambda\widehat{V}),$$

and in the fourth line we moved this operator over to the right, an operation which requires us to take the adjoint again; this just removes the dagger (look at the first equation on p.1 to see this).We now recognise the first two terms as $(\Delta A)^2$ and $(\Delta B)^2$. The last two terms combine to make the commutator, $[\widehat{U}, \widehat{V}] = [\widehat{A}, \widehat{B}]$, so that we end up with the inequality

$$I(\lambda) = (\Delta A)^2 + \lambda^2(\Delta B)^2 + i\lambda\langle[\widehat{A}, \widehat{B}]\rangle \geq 0 \tag{110}$$

Now this is true for all real λ, but the most restrictive inequality is obtained by looking for the smallest value of $I(\lambda)$, the minimum of the parabola

$$\frac{dI(\lambda)}{d\lambda} = 2\lambda(\Delta B)^2 + i\langle[\widehat{A}, \widehat{B}]\rangle = 0 \tag{111}$$

This gives a value for λ,

$$\lambda = -i\frac{\langle[\widehat{A}, \widehat{B}]\rangle}{2(\Delta B)^2}, \tag{112}$$

which we substitute into the original inequality to give the most stringent inequality we can find,

$$(\Delta A)^2(\Delta B)^2 \geq \frac{1}{4}\left(i\langle[\widehat{A}, \widehat{B}]\rangle\right)^2 \tag{113}$$

Finally we note that in all cases we have met so far the commutator is *always* imaginary [5], so that the product $i\,[\widehat{A}, \widehat{B}]$ is always real and its square positive. We can therefore take the square root, yielding the generalised Heisenberg uncertainty relation:

$$\Delta A \Delta B \geq \frac{1}{2}|\langle[\widehat{A}, \widehat{B}]\rangle| \tag{114}$$

---

[5] In fact the result follows from the assumptions that all the operators involved are observables, ie. are Hermitian. Thus, writing

$$[\widehat{A}, \widehat{B}] = \alpha\,\widehat{C},$$

where $\widehat{A}$, $\widehat{B}$ and $\widehat{C}$ are Hermitian and α is a complex constant, and using the fact that

$$(\widehat{A}\widehat{B})^\dagger = \widehat{B}^\dagger\widehat{A}^\dagger = \widehat{B}\widehat{A},$$

we find that

$$\begin{aligned} [\widehat{A}, \widehat{B}]^\dagger &= (\alpha\widehat{C})^\dagger \\ &= \alpha^*\widehat{C}^\dagger \\ &= \alpha^*\widehat{C} \end{aligned}$$

But this is also equal to

$$\begin{aligned} [\widehat{A}, \widehat{B}]^\dagger &= [\widehat{B}, \widehat{A}] \\ &= -[\widehat{A}, \widehat{B}] \\ &= -\alpha\widehat{C} \end{aligned}$$

Hence, comparing these last two equations we discover that $\alpha = -\alpha^*$ is pure imaginary. Q.E.D.