6 The basic formalism of quantum mechanics

6.1 Spaces of functions

We will not be too precise about the space of functions to which physically relevant wavefunctions belong in any given context. One possible example is the set of infinitely differentiable complex-valued functions $\psi(x)$ of one real variable x. This forms a *complex vector space* which we denote by $C^\infty(R)$. Another is the space of normalisable wavefunctions, which again forms a complex vector space, $L^2(R)$.

Exercise Check these are indeed vector spaces.

6.2 The inner product

We can define a natural inner product on the space of normalisable wavefunctions by

$$(\psi_1, \psi_2) = \int_{-\infty}^{\infty} \psi_1^*(x)\psi_2(x)dx\,. \tag{6.1}$$

Exercise: Show that this is well-defined whenever ψ_1 and ψ_2 are normalisable. [Hint: look ahead to Thm. 6.6 below.]

We extend the definition of the inner product $(\,,\,)$ to any ψ_1 and ψ_2 for which the integral is well-defined, whether or not ψ_1 or ψ_2 is normalisable.

6.2.1 Properties of the inner product

1. $(\psi_1, \psi_2) = (\psi_2, \psi_1)^*$

2. The inner product is *anti-linear* in the first entry and *linear* in the second:

$$(a_1\psi_1 + a_2\psi_2, \phi) = a_1^*(\psi_1, \phi) + a_2^*(\psi_2, \phi)\,,$$
$$(\psi, a_1\phi_1 + a_2\phi_2) = a_1(\psi, \phi_1) + a_2(\psi, \phi_2)\,.$$

3. The inner product is *positive definite* on continuous wavefunctions, i.e.

$$(\psi, \psi) \geq 0 \qquad \text{and } (\psi, \psi) = 0 \text{ if and only if } \psi = 0\,.$$

To see this, note that

$$(\psi, \psi) = \int_{-\infty}^{\infty} \psi^*(x)\psi(x)dx = \int_{-\infty}^{\infty} |\psi(x)|^2 dx\,.$$

6.3 Operators

Recall that we defined an *operator* $\hat{O}$ to be any *linear* map from a space of functions[18] (for example $C^\infty(R)$ or $L^2(R)$) to itself, i.e. any map such that

$$\hat{O}(a_1\psi_1 + a_2\psi_2) = a_1\hat{O}\psi_1 + a_2\hat{O}\psi_2 \tag{6.2}$$

for all complex numbers a_1, a_2 and all $\psi_1, \psi_2 \in C^\infty(R)$.

6.3.1 Some examples of operators

- Finite differential operators $\sum_{i=0}^{n} p_i(x)\frac{\partial^i}{\partial x^i}$, where the $p_i(x)$ are polynomials in x. (This class includes $\hat{p}$ and $\hat{x}$, defined below.)
- The translation operator

$$S_a : \psi(x) \to \psi_a(x) = \psi(x-a)\,.$$

- The parity operator

$$P : \psi(x) \to \psi_P(x) = \psi(-x)\,.$$

6.4 Hermitian operators

We define the *hermitian conjugate* $A^\dagger$ of an operator A to be the operator such that

$$(A^\dagger\psi_1, \psi_2) = (\psi_1, A\psi_2) \tag{6.3}$$

for all normalisable wavefunctions ψ_1, ψ_2. An operator is *hermitian* if $A = A^\dagger$.

Exercise Verify the identities

$$(a_1A_1 + a_2A_2)^\dagger = a_1^*A_1^\dagger + a_2^*A_2^\dagger\,,$$
$$(AB)^\dagger = B^\dagger A^\dagger\,.$$

6.4.1 Classical states and dynamical variables

By the *state* of a physical system we mean its mathematical representation in a given theory. In classical mechanics, the states are points in a $2n$-dimensional space, *phase space*. For example, a system of n point particles in 1D is described by $2n$ coordinates: $(x_1, \ldots, x_n, p_1, \ldots, p_n)$, where x_i is the position of the ith particle

[18] *Non-examinable comment: In our discussions below, we generally take the space of functions to be the space $L^2(R)$ of normalisable physical wavefunctions satisfying Eqn. (3.11). Strictly speaking, a completely rigorous discussion ought to take account of subtleties which arise from this choice. For example, the momentum operator is not actually defined on all functions in $L^2(R)$, but only those that are differentiable. Similarly, the position operator is not defined on all functions in $L^2(R)$, but only those normalisable functions $\psi(x)$ for which $x\psi(x)$ is also normalisable.

However, a fully rigorous treatment would go well beyond the scope of this course. We shall follow standard practice in introductory treatments of quantum mechanics and simply assume, whenever we have an equation in which the position, momentum or other operators act on physical wavefunctions, that the wavefunctions are chosen such that the action of the operators is well defined. [You should make the same assumption in tackling problem sheets or exam questions.] *

positions and p_i its momentum. Classical dynamical variables – for instance the energy

$$E = \sum_{i=1}^{n} \frac{p_i^2}{2m_i} + V(x_1, \ldots, x_n) \tag{6.4}$$

– are defined by functions on phase space. Note that the classical state of a finite number of particles can be described by a finite number of parameters.

In principle, the classical state and (hence) the value of all classical dynamical variables can be measured with arbitrary precision.

6.4.2 Quantum states and observables

The possible quantum mechanical states of a physical system at any given time t are normalised wavefunctions. In particular, the state of a single particle in 1D is given by a wavefunction $\psi(x,t)$ such that $\int_{-\infty}^{\infty} |\psi(x,t)|^2 dx = 1$. We need infinitely many parameters to specify a normalised wavefunction, and thus to specify the state of a quantum system – even a single particle in 1D.

All quantum dynamical variables or *observables* – quantities we can measure – are represented by hermitian operators, and vice versa. Examples of hermitian operators defining quantum observables are the position, momentum and energy operators

$$\hat{p} = -i\hbar \frac{\partial}{\partial x} \tag{6.5}$$

$$\hat{x} = x \qquad \text{(i.e. multiplication by } x\text{)} \tag{6.6}$$

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \tag{6.7}$$

Another example is the parity operator

$$P : \psi(x) \to \psi_P(x) = \psi(-x)\,. \tag{6.8}$$

Exercise (important!): Verify that each of the operators (6.5-6.8) is hermitian.

6.5 Some theorems about hermitian operators

Theorem 6.1. *The eigenvalues of a hermitian operator are real.*

Proof. Let A be hermitian and ψ be a normalised eigenfunction with eigenvalue a: $A\psi = a\psi$. We have

$$\begin{aligned}(\psi, A\psi) &= (\psi, a\psi) = a(\psi, \psi) = a\,,\\ (\psi, A\psi) &= (A^\dagger\psi, \psi) = (A\psi, \psi)\\ &= (a\psi, \psi) = a^*(\psi, \psi) = a^*\,.\end{aligned}$$

Hence $a = a^*$. ∎

Theorem 6.2. *Let A be a hermitian operator and ψ_1, ψ_2 be normalised eigenfunctions with different eigenvalues a_1, a_2. Then ψ_1 and ψ_2 are orthogonal.*

Proof. We have $A\psi_1 = a_1\psi_1$ and $A\psi_2 = a_2\psi_2$, and by theorem 6.1 we have that a_1, a_2 are real.

$$\begin{aligned}a_1(\psi_1, \psi_2) &= a_1^*(\psi_1, \psi_2) = (a_1\psi_1, \psi_2)\\ &= (A\psi_1, \psi_2) = (A^\dagger\psi_1, \psi_2)\\ &= (\psi_1, A\psi_2) = (\psi_1, a_2\psi_2)\\ &= a_2(\psi_1, \psi_2)\,.\end{aligned}$$

Hence $(a_1 - a_2)(\psi_1, \psi_2) = 0$, and since $a_1 \neq a_2$ we have $(\psi_1, \psi_2) = 0$. ∎

Our discussion is complicated by the fact that the eigenfunctions of hermitian operators (i.e. of quantum observables) are *not* necessarily all normalisable. For example:

- The momentum operator $\hat{p} = -i\hbar\frac{\partial}{\partial x}$ has eigenfunctions $\exp(\frac{i}{\hbar}px)$ with eigenvalue p.
- The position operator $\hat{x}$ has eigenfunctions $\delta(x - q)$ with eigenvalue q: we have $\hat{x}\delta(x - q) = x\delta(x - q) = q\delta(x - q)$.
- The energy operator $\hat{H} = \frac{1}{2m}\hat{p}^2 + V(x)$, for a finite square well potential $V(x)$, has normalisable eigenfunctions (corresponding to bound states) and unnormalisable eigenfunctions (corresponding to scattering solutions).

Theorems 6.1 and 6.2 extend to these non-normalisable eigenfunctions of $\hat{p}, \hat{x}$ and $\hat{H}$: these are the only operators with continuous sets of eigenfunctions that we consider in this course.

In general, a hermitian operator may have both a set of normalisable eigenfunctions with discrete eigenvalues and a set of non-normalisable eigenfunctions with continuous eigenvalues, and the theorems apply to both sets.[19]

[19]*Non-examinable technical note: it is well beyond our scope here but, in fact, it turns out one can find a more general notion of normalisability which covers both sets of eigenfunctions, and more general versions of the theorems can be precisely framed in terms of this condition. This definition includes bound states and scattering solutions to the time-independent SE, but not solutions which blow up exponentially. A discussion can be found in, for example, Messiah, "Quantum Mechanics", vol. 1, chap V.9.*

Theorem 6.3. *The discrete and continuous sets of eigenfunctions of any hermitian operator together form a complete orthogonal basis of the physical wavefunctions, i.e. of the normalisable complex-valued functions $\psi(x)$ of one real variable x.*

Proof. This is quite hard to prove in complete generality. We will assume it without proof in this course. ■

Note: We say the hermitian operator A has a *degenerate* eigenvalue λ if it has more than one linearly independent eigenfunction with eigenvalue λ. If A has degenerate eigenvalues, we define an orthogonal basis of its eigenfunctions by choosing orthogonal bases for the eigenfunctions associated with each degenerate eigenvalue λ.

Corollary 6.3.1. *Let A be a hermitian operator with a discrete set of normalised eigenfunctions $\{\psi_i\}_{i=1}^N$ (we include $N = \infty$ as a possibility) and a continuous set of eigenfunctions $\{\psi_\alpha\}_{\alpha\in\Delta}$, where the indexing set Δ is some sub-interval of the real numbers. Then any wavefunction ψ can be written as*

$$\psi = \sum_{i=1}^{N} a_i\psi_i + \int_\Delta a_\alpha\psi_\alpha d\alpha\,, \tag{6.9}$$

where $a_i = (\psi_i, \psi)$ and a_α are complex numbers.

Comment In fact, by appropriately normalising, we can also ensure that $a_\alpha = (\psi_\alpha, \psi)$.

Corollary 6.3.2. *Let A be a hermitian operator with (only) a discrete set of orthonormalised eigenfunctions $\{\psi_i\}_{i=1}^\infty$. Then any wavefunction ψ can be written as*

$$\psi = \sum_{i=1}^{\infty}(\psi_i, \psi)\psi_i\,. \tag{6.10}$$

Proof. Theorem 6.3 implies that

$$\psi = \sum_{i=1}^{\infty} a_i\psi_i \tag{6.11}$$

for some complex numbers a_i. Taking the inner product with ψ_j, and applying theorem 6.2, we see $(\psi_j, \psi) = a_j$. ■

6.6 Quantum measurement postulates

- **Postulate 1** Every quantum observable O is represented by a hermitian operator $\hat{O}$.
- **Postulate 2** The possible outcomes of a measurement of O are the eigenvalues of $\hat{O}$.
- **Postulate 3** If $\hat{O}$ has (only) a discrete set of normalised eigenfunctions $\{\psi_i\}_{i=1}^{\infty}$ with corresponding distinct eigenvalues $\{\lambda_i\}$, and a measurement of O is carried out on a particle with normalised wavefunction

$$\psi = \sum_{i=1}^{\infty} a_i \psi_i \,,$$

  then the probability of outcome λ_i is $|a_i|^2$.
- **Postulate 4** (the *projection postulate*) If a measurement of the observable O is carried out on a particle with normalised wavefunction $\psi(x,t)$ at time t and the outcome λ_i is obtained, the wavefunction instantaneously after the measurement becomes $\psi_i(x)$. [This is sometimes referred to as the "collapse of the wavefunction".] The wavefunction then evolves according to the Schrödinger equation, with initial state $\psi_i(x)$ at time t, until the next measurement.

Notes

- It follows from postulates 2 and 3 that the total probability of all possible outcomes is

$$\sum_i |a_i|^2 = \sum_i (a_i\psi_i, a_i\psi_i) = \sum_{ij} (a_i\psi_i, a_j\psi_j) = (\psi,\psi) = 1 \,. \tag{6.12}$$

  So the postulates are consistent: the sum of the probabilities of all possible outcomes is 1, and so if you carry out a measurement you will certainly get *some* outcome and you will only get one outcome. (We already verified this in the case of the Born rule for position measurements.)
- It follows from postulates 3 and 4 that if an observable O is measured twice, with infinitesimal time separating the two measurements, then if the first outcome is λ_i the second will, with probability one, also be λ_i.

  Exercise: Check this.

  This has the important consequence that quantum measurements resemble classical measurements in at least one sense: they establish a property of the system that can be repeatedly verified. If we measure something, and then quickly measure it again, we get the same answer. If this was not true, it would be hard to find any good reason for the use of the term "measurement" in quantum mechanics.

- If the wavefunction ψ is an eigenfunction ψ_i of $\hat{O}$, the measurement outcome will be λ_i with probability one. For example, a stationary state obeying $\hat{H}\psi = E\psi$ will always give outcome E if the energy is measured.
- But unless the wavefunction ψ is an eigenfunction of the measured observable, the measurement outcome is not definitely predictable. In contrast to classical mechanics, a quantum observable does not generally have a definite value on a quantum state.
- We can extend postulates 3 and 4 to the case when $\hat{O}$ has degenerate eigenvalues. If $\psi = \sum_i a_i\psi_i$ is measured, where $\{\psi_i\}_{i=1}^{\infty}$ are orthonormalised eigenfunctions of $\hat{O}$ and $\{\psi_i\}_{i\in I}$ are a complete set of orthonormalised eigenfunctions with the same eigenvalue λ, the probability of outcome λ is $\sum_{i\in I}|a_i|^2$, and the state resulting after a measurement with outcome λ is (up to normalisation) $\sum_{i\in I} a_i\psi_i$.
- The projection postulate is so called because it implies that the the post-measurement wavefunction $\psi_i(x,t)$ is obtained from the pre-measurement wavefunction $\psi(x,t)$ by the action of the *projection operator* P_i defined by $P_i : \psi \to (\psi_i, \psi)\psi_i$, up to normalisation. We call P_i a projection since it maps any state onto its component in a particular linear subspace, namely the subspace spanned by ψ_i – an action analogous to, for instance, the projection of a 3D vector (x, y, z) onto its x-component $(x, 0, 0)$.

6.7 Expectation values

Consider a measurement of the observable A on the state ψ. If the corresponding hermitian operator $\hat{A}$ has only a discrete set of normalisable eigenfunctions $\{\psi_i\}$, the possible outcomes are the corresponding eigenvalues λ_i, and the outcome probabilities are $p_i = |(\psi, \psi_i)|^2$. The *expectation value* of the measured outcome, in the standard statistical sense of the term, is thus

$$\begin{aligned}\sum_i p_i\lambda_i &= \sum_i |(\psi,\psi_i)|^2\lambda_i \\ &= (\sum_i (\psi,\psi_i)\psi_i, \lambda_j \sum_j (\psi,\psi_j)\psi_j) \\ &= (\psi, \hat{A}\psi)\,. \end{aligned} \tag{6.13}$$

We write $\langle\hat{A}\rangle_\psi$ for the expectation value of a measurement of A on the state ψ. Thus we have

$$\langle\hat{A}\rangle_\psi = (\psi, \hat{A}\psi)\,. \tag{6.14}$$

We can similarly justify this definition of expectation value for the position operator $\hat{x}$ from the Born rule. Recall that the probability of obtaining a position measurement outcome in the interval $[x, x + dx]$ is given by $|\psi(x)|^2 dx$. The expectation value of a position measurement is thus

$$\int_{-\infty}^{\infty} x|\psi(x)|^2 dx = \int_{\infty}^{\infty} \psi^*(x) x \psi(x) dx = (\psi, \hat{x}\psi)\,.$$

This encourages us to take Eqn. (6.14) as a general definition of expectation value for any observable, whether its eigenvalues are discrete, continuous or a combination. This general definition too can be similarly justified, given a careful treatment of the eigenfunctions of general observables and their normalisation: this is beyond our scope here, but it is important to note that the definition turns out to agree with the statistical definition of expectation in all cases. In particular, the definitions

$$\langle \hat{p} \rangle_\psi = (\psi, \hat{p}\psi)\,, \qquad \langle \hat{H} \rangle_\psi = (\psi, \hat{H}\psi)\,,$$

often give simple ways of calculating the statistical expectation of measurements of momentum and energy for general wavefunctions ψ, since the right hand sides are often given by simple integrals.

Note that the expectation value is *linear* with respect to real scalars: i.e.

$$\langle a\hat{A} + b\hat{B} \rangle_\psi = a\langle \hat{A} \rangle_\psi + b\langle \hat{B} \rangle_\psi\,, \tag{6.15}$$

for any hermitian operators A, B and any real numbers a, b. We restrict to a, b real here because the interpretation of $\langle \hat{A} \rangle_\psi$ as an expectation value of an observable requires that A is hermitian, since observables are always represented by hermitian operators. A complex multiple of a hermitian operator is not generally hermitian: if A is hermitian then $(aA)^\dagger = a^* A^\dagger$.

6.8 Commutation relations

We define the *commutator* of two operators by

$$[A, B] = AB - BA\,. \tag{6.16}$$

It is easy to verify the identities

$$\begin{aligned} [A, A] &= 0\,, \\ [A, B] &= -[B, A]\,, \\ [A, BC] &= [A, B]C + B[A, C] \\ [AB, C] &= A[B, C] + [A, C]B\,. \end{aligned} \tag{6.17}$$

Note that the commutator $[A, B]$ is a sum of products of operators, and thus itself an operator. Note also that it depends linearly on both entries.

The commutator plays a crucial role in describing symmetries of quantum mechanical systems, as we will see when we consider angular momentum. It also gives a way of calibrating how close two operators are to having simultaneously determinable eigenvalues: see the following note and the later discussion of the uncertainty principle.

Definition We say two hermitian operators A and B are *simultaneously diagonalisable* if the space of normalisable wavefunctions has a complete basis of joint eigenfunctions $\{\psi_i\}$, i.e. of eigenfunctions such that $A\psi_i = a_i\psi_i$ and $B\psi_i = b_i\psi_i$ for some real numbers a_i, b_i.

Theorem 6.4. *Two hermitian operators A and B are simultaneously diagonalisable if and only if $[A, B] = 0$.*

Proof. If A and B are simultaneously diagonalisable, the space of normalisable wavefunctions has a complete basis of joint eigenfunctions $\{\psi_i\}$. Now for any such eigenfunction

$$[A, B]\psi_i = AB\psi_i - BA\psi_i = (a_ib_i - b_ia_i)\psi_i = 0\,. \tag{6.18}$$

If the basis $\{\psi_i\}$ is complete, any wavefunction ψ can be written as $\psi = \sum_i c_i\psi_i$, and we have that

$$[A, B]\psi = [A, B]\sum_i c_i\psi_i = \sum_i c_i[A, B]\psi_i = 0\,. \tag{6.19}$$

So we have the operator equation $[A, B] = 0$, where 0 is the zero operator, which maps any wavefunction to the zero function.

Conversely, if $[A, B] = 0$ and $A\psi_i = a_i\psi_i$, then $0 = [A, B]\psi_i = A(B\psi_i) - a_i(B\psi_i)$, so that $B\psi_i$ is also an eigenfunction of A with eigenvalue a_i. Thus B maps the eigenspace of A with eigenvalue a_i to itself. If we write E for this eigenspace, and $B|_E$ for the operator B restricted to E, then clearly $B|_E$ is a hermitian operator on E. Thus, by theorem 6.3, we can find a basis of E in which B acts diagonally. Since this is true for all eigenspaces of A, we can find a complete basis of simultaneous eigenfunctions of A and B. ∎

6.8.1 The canonical commutation relations

Recall that the position and momentum operators are $\hat{p} = -i\hbar\frac{\partial}{\partial x}$ and $\hat{x} = x$ (multiplication by x). We can work out their commutator by considering their action on a general wavefunction ψ:

$$\begin{aligned}
\hat{x}\hat{p}\psi &= -ix\hbar\frac{\partial\psi}{\partial x} \\
\hat{p}\hat{x}\psi &= -i\hbar\frac{\partial}{\partial x}(x\psi) \\
&= -i\hbar\psi - ix\hbar\frac{\partial\psi}{\partial x} \\
\text{so}\quad [\hat{x}, \hat{p}]\psi &= \hat{x}\hat{p}\psi - \hat{p}\hat{x}\psi = i\hbar\psi
\end{aligned}$$

and as this is true for all ψ we have

$$[\hat{x}, \hat{p}] = i\hbar\,. \tag{6.20}$$

It can be shown (though not in this course: see Part II Principles of Quantum Mechanics) that these commutation relations essentially characterise the position and momentum operators. That is, any pair of operators satisfying these relations is equivalent (in a sense that can be made precise) to position and momentum.

6.9 Heisenberg's Uncertainty Principle

We define the *uncertainty* $\Delta_\psi A$ in a measurement of A on the state ψ by

$$\begin{aligned}(\Delta_\psi A)^2 &= \langle (A - \langle A\rangle_\psi)^2 \rangle_\psi \\ &= \langle A^2\rangle_\psi - (\langle A\rangle_\psi)^2\,.\end{aligned} \tag{6.21}$$

Note that Theorem 6.1 implies that the expectation value and the uncertainty are always real, as we would expect if they are physically meaningful.

Exercise Verify that $(\Delta_\psi A)^2$ is the statistical variance of the probability distribution for the possible outcomes of the measurement of A on ψ, and $\Delta_\psi A$ is the distribution's standard deviation.

Lemma 6.5. *The uncertainty $\Delta_\psi A \geq 0$ and $\Delta_\psi A = 0$ if and only if ψ is an eigenfunction of A.*

Proof. We can write

$$\begin{aligned}(\Delta_\psi A)^2 &= \langle (A - \langle A\rangle_\psi)^2 \rangle_\psi \\ &= ((A - \langle A\rangle_\psi)\psi, (A - \langle A\rangle_\psi)\psi)\,.\end{aligned}$$

Write $\phi = (A - \langle A\rangle_\psi)\psi$. We have $(\phi, \phi) \geq 0$, with equality only if $\phi = 0$, i.e.

$$A\psi = \langle A\rangle_\psi \psi\,, \tag{6.22}$$

which implies that ψ is an eigenfunction of A. Conversely, if ψ is an eigenfunction of A then $(\psi\,,\, A\psi) = \langle A\rangle_\psi$, and so $A\psi = \langle A\rangle_\psi \psi$. In other words A has eigenvalue $\langle A\rangle_\psi$, and so $\Delta_\psi A = 0$. ∎

Theorem 6.6 (Schwarz's inequality)**.** *If ϕ, ψ are any two normalisable wavefunctions, then $|(\phi, \psi)|^2 \leq (\phi, \phi)(\psi, \psi)$. We have equality if and only if $\phi = a\psi$ for some complex number a.*

Proof. For any a, $0 \leq (\phi - a\psi, \phi - a\psi)$. If we take $a = \frac{(\psi,\phi)}{(\psi,\psi)}$, we see

$$\begin{aligned}0 \leq\ & (\phi, \phi) - \frac{|(\psi, \phi)|^2}{(\psi, \psi)} - \frac{|(\psi, \phi)|^2}{(\psi, \psi)} + \frac{|(\psi, \phi)|^2}{(\psi, \psi)} \\ &= (\phi, \phi) - \frac{|(\psi, \phi)|^2}{(\psi, \psi)}\,.\end{aligned}$$

Multiplying through by (ψ, ψ) we obtain the inequality. We have equality only if $\phi - a\psi = 0$, for the value of a specified above. ∎

Theorem 6.7 (the *generalised uncertainty relations*). *If A and B are any two observables, and ψ is any state, then*

$$\Delta_\psi A \Delta_\psi B \geq \frac{1}{2}|(\psi, [A,B]\psi)|\,. \tag{6.23}$$

Proof. We have

$$\begin{aligned}(\Delta_\psi A)^2 &= \langle (A - \langle A\rangle_\psi)^2\rangle_\psi \\ &= ((A-\langle A\rangle_\psi)\psi, (A-\langle A\rangle_\psi)\psi) \\ (\Delta_\psi B)^2 &= ((B-\langle B\rangle_\psi)\psi, (B-\langle B\rangle_\psi)\psi)\,.\end{aligned}$$

So, writing $A' = A - \langle A\rangle_\psi$, $B' = B - \langle B\rangle_\psi$, we have

$$\begin{aligned}(\Delta_\psi A)^2(\Delta_\psi B)^2 &= (A'\psi, A'\psi)(B'\psi, B'\psi) \\ &\geq |(A'\psi, B'\psi)|^2 \qquad (\text{ from Thm. 6.6}) \\ &= |(\psi, A'B'\psi)|^2 \qquad (\text{as } A' \text{ is hermitian})\,.\end{aligned} \tag{6.24}$$

Now: (1) $A'B' = \frac{1}{2}([A',B'] + \{A',B'\})$, where the *anti-commutator* $\{A',B'\} = A'B' + B'A'$;

(2) $[A',B'] = [A,B]$;

(3) $(\psi, \{A',B'\}\psi)$ is real, since

$$(\psi, \{A',B'\}\psi) = ((\{A',B'\})^\dagger\psi, \psi) = ((\{A',B'\})\psi, \psi) = (\psi, \{A',B'\}\psi)^*\,. \tag{6.25}$$

A similar argument shows that $(\psi, [A',B']\psi)$ is imaginary.

So

$$|(\psi, A'B'\psi)|^2 = \frac{1}{4}(|(\psi, \{A',B'\}\psi)|^2 + |(\psi, [A',B']\psi)|^2)\,. \tag{6.26}$$

Combining (6.24) and (6.26), we have that

$$(\Delta_\psi A)^2(\Delta_\psi B)^2 \geq \frac{1}{4}|(\psi, [A,B]\psi)|^2\,. \tag{6.27}$$

Taking the square root gives (6.23). ∎

Corollary 6.7.1. *(the Heisenberg uncertainty principle for position and momentum).*

$$(\Delta_\psi x)(\Delta_\psi p) \geq \frac{1}{2}\hbar \tag{6.28}$$

Proof. Taking $A = \hat{x}$ and $B = \hat{p}$, we have $[A,B] = i\hbar$, and the result follows from Thm. 6.23. ∎

Thus, the smaller the uncertainty in position, $\Delta_\psi x$, the greater the minimum possible uncertainty in momentum, $\Delta_\psi p$, and vice versa.

Lemma 6.8. *If*

$$\hat{x}\psi = ia\hat{p}\psi \tag{6.29}$$

for some real parameter a, then $(\Delta_\psi x)(\Delta_\psi p) = \frac{1}{2}\hbar$.

Proof. If $\hat{x}\psi = ia\hat{p}\psi$ then we have

$$\begin{aligned}(\psi, \{\hat{x}, \hat{p}\}\psi) &= (\psi, \hat{x}\hat{p}\psi) + (\psi, \hat{p}\hat{x}\psi) \\ &= (\hat{x}\psi, \hat{p}\psi) + (\hat{p}\psi, \hat{x}\psi) \\ &= (ia\hat{p}\psi, \hat{p}\psi) + (\hat{p}\psi, ia\hat{p}\psi) \\ &= (-ia + ia)(\hat{p}\psi, \hat{p}\psi) = 0\,,\end{aligned}$$

which is the condition for the first term on the RHS of (6.26) to vanish. We also have that $\langle\hat{x}\rangle_\psi = ia\langle\hat{p}\rangle_\psi$ and, since both expectations are real, this implies that $\langle\hat{x}\rangle_\psi = \langle\hat{p}\rangle_\psi = 0$. Hence

$$(\hat{x} - \langle\hat{x}\rangle_\psi)\psi = ia(\hat{p} - \langle\hat{p}\rangle_\psi)\psi\,,$$

which means we have equality in the Schwarz's inequality (6.24) used to derive (6.28). ■

Lemma 6.9. *The condition (6.29) holds if and only if $\psi(x) = C\exp(-bx^2)$ for some constants b, C.*

Proof. If $\hat{x}\psi = ia\hat{p}\psi$ for some real a, we have that $x\psi = a\hbar\frac{\partial}{\partial x}\psi$ and so $\psi(x) = C\exp(-bx^2)$ for some real $b = -\frac{a\hbar}{2}$, and because we have equality in (6.28) we know the uncertainty is minimised. Conversely, any wavefunction of the form $\psi(x) = C\exp(-bx^2)$ satisfies $\hat{x}\psi = ia\hat{p}\psi$ for some real a. ■

Note: For the wavefunction to be normalisable, we require $b > 0$ and $C \neq 0$. We can take $C = |C| > 0$ by multiplying $\psi(x)$ by a phase factor (which does not alter any physical quantity: the probabilities of outcomes for any measurement are unaffected).

Exercise: Show that the condition (6.29) for minimum uncertainty is necessary as well as sufficient. Hence show that the normalisable minimum uncertainty states are precisely the wavefunctions defined by Gaussian functions.

6.9.1 What does the uncertainty principle tell us?

The uncertainty principle is a mathematical statement relating the uncertainties of x and p (or more generally A and B), which are quantities defined for a given state ψ. We can say, for example, that for any state ψ with $\Delta_\psi x = \delta$, we have $\Delta_\psi p \geq \frac{\hbar}{2\delta}$.

Heisenberg originally suggested that the uncertainty principle can be understood simply by observing that a measurement of A creates uncertainty by disturbing the value of any observable B that does not commute with A. This is not a valid argument! There are two problems with it.

First, according to quantum mechanics, it is not generally the case that there is a definite fixed pre-measurement value of either A or B, which is somehow disturbed in the process of our attempting to measure it. Unless the wavefunction is an eigenfunction of A or B, the result of any measurement is indeterminate.

Second, the mathematical derivation of the uncertainty principle does not require us to consider measurements of A or B actually taking place. The quantity $\Delta_\psi A$ is mathematically defined whether or not we choose to carry out a measurement of A. Even if we choose to interpret it as referring to a possible measurement, we cannot interpret the mathematics as referring to *successive* measurements of A

and B. $(\Delta_\psi A)$ and $(\Delta_\psi B)$ are the standard deviations for the outcomes of measurements of A and B, but these hypothetical measurements are *alternative* possible measurements on the same state ψ, *not* actual measurements carried out one after the other. If we measured, say, first A and then B, the first measurement would collapse the wavefunction onto an eigenfunction of A, and the second measurement would hence not generally be a measurement on the original state ψ.

6.10 Ehrenfest's theorem

Theorem 6.10 (Ehrenfest's theorem)**.** *The expectation value $\langle A\rangle_\psi$ of an operator A in the state ψ evolves by*

$$\frac{d}{dt}\langle A\rangle_\psi = \frac{i}{\hbar}\langle[\hat{H}, A]\rangle_\psi + \langle\frac{\partial A}{\partial t}\rangle_\psi\,. \tag{6.30}$$

Proof. We have

$$\begin{aligned}
\frac{d\langle A\rangle_\psi}{dt} &= \frac{d}{dt}\int_{-\infty}^{\infty}\psi^* A\psi dx \\
&= \int_{-\infty}^{\infty}(\frac{\partial\psi^*}{\partial t}A\psi + \psi^*\frac{\partial A}{\partial t}\psi + \psi^* A\frac{\partial\psi}{\partial t})dx \\
&= \langle\frac{\partial A}{\partial t}\rangle_\psi + \frac{i}{\hbar}\int_{-\infty}^{\infty}((\hat{H}\psi)^* A\psi - \psi^* A(\hat{H}\psi))dx \\
&= \langle\frac{\partial A}{\partial t}\rangle_\psi + \frac{i}{\hbar}\int_{-\infty}^{\infty}(\psi^*\hat{H}A\psi - \psi^* A(\hat{H}\psi))dx \\
&= \frac{i}{\hbar}\langle[\hat{H}, A]\rangle_\psi + \langle\frac{\partial A}{\partial t}\rangle_\psi\,.
\end{aligned} \tag{6.31}$$

■

6.10.1 Applications of Ehrenfest's theorem

For $\hat{H} = \frac{\hat{p}^2}{2m} + V(x)$, we have

$$\begin{aligned}
[\hat{H}, \hat{p}] &= [V(x), \hat{p}] \\
&= [V(x), -i\hbar\frac{\partial}{\partial x}] \\
&= i\hbar\frac{dV}{dx}
\end{aligned} \tag{6.32}$$

$$\begin{aligned}
[\hat{H}, \hat{x}] &= [\frac{\hat{p}^2}{2m}, \hat{x}] \\
&= \frac{1}{2m}2[\hat{p}, \hat{x}]\hat{p} = \frac{-i\hbar}{m}\hat{p}
\end{aligned} \tag{6.33}$$

$$[\hat{H}, \hat{H}] = 0\,. \tag{6.34}$$

Since none of these operators is explicitly time-dependent, we have that $\frac{\partial \hat{H}}{\partial t} = \frac{\partial \hat{x}}{\partial t} = \frac{\partial \hat{p}}{\partial t} = 0$ and so the $\langle \frac{\partial A}{\partial t} \rangle_\psi$ term on the RHS of (6.30) vanishes in each case, giving

$$\begin{aligned} \frac{d}{dt}\langle \hat{p} \rangle_\psi &= -\langle \frac{dV}{dx} \rangle_\psi \,, \\ \frac{d}{dt}\langle \hat{x} \rangle_\psi &= \frac{1}{m}\langle \hat{p} \rangle_\psi \,, \\ \frac{d}{dt}\langle \hat{H} \rangle_\psi &= 0 \,. \end{aligned} \tag{6.35}$$

These are quantum versions of the classical laws $\frac{d}{dt}x = \frac{1}{m}p$ (which follows from $p = mv$), $\frac{d}{dt}p = -\frac{dV}{dx}$ (which follows from $F = ma$), and $\frac{d}{dt}E = 0$ (conservation of total energy).

* Ehrenfest's theorem and the harmonic oscillator. *

6.11 *The harmonic oscillator revisited

By considering commutation relations, we can give a much nicer and more illuminating derivation of the energy spectrum of the harmonic oscillator. This derivation forms part of the material for the Part II Principles of Quantum Mechanics course. It is non-examinable material, in the sense that if you are asked to derive the energy spectrum (without any method being stipulated) then the derivation given earlier is a perfectly adequate answer. However, the derivation below is simpler and slicker, and of course it also may be used in this context.

Recall that the harmonic oscillator hamiltonian is

$$\begin{aligned}\hat{H} &= \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 \\ &= \frac{1}{2m}(\hat{p} + im\omega\hat{x})(\hat{p} - im\omega\hat{x}) + \frac{i\omega}{2}[\hat{p},\hat{x}] \\ &= \frac{1}{2m}(\hat{p} + im\omega\hat{x})(\hat{p} - im\omega\hat{x}) + \frac{\hbar\omega}{2} \end{aligned} \tag{6.36}$$

Define the operator $a = \frac{1}{\sqrt{2m}}(\hat{p} - im\omega\hat{x})$. Since $\hat{p}$ and $\hat{x}$ are hermitian, we have $a^\dagger = \frac{1}{\sqrt{2m}}(\hat{p} + im\omega\hat{x})$, and

$$\hat{H} = a^\dagger a + \frac{1}{2}\hbar\omega\,. \tag{6.37}$$

We have the following commutation relations:

$$[a, a^\dagger] = \frac{1}{2m}(-im\omega)2[\hat{x},\hat{p}] = \hbar\omega\,, \tag{6.38}$$

$$[\hat{H}, a] = [a^\dagger a, a] = [a^\dagger, a]a = -\hbar\omega a\,, \tag{6.39}$$

$$[\hat{H}, a^\dagger] = [a^\dagger a, a^\dagger] = a^\dagger[a, a^\dagger] = \hbar\omega a^\dagger\,. \tag{6.40}$$

Suppose now that ψ is a harmonic oscillator eigenfunction of energy E:

$$\hat{H}\psi = E\psi\,.$$

We then have

$$\hat{H}a\psi = [\hat{H}, a]\psi + a\hat{H}\psi = (E - \hbar\omega)a\psi \tag{6.41}$$

$$\hat{H}a^\dagger\psi = [\hat{H}, a^\dagger]\psi + a^\dagger\hat{H}\psi = (E + \hbar\omega)a^\dagger\psi\,, \tag{6.42}$$

so that $a\psi$ and $a^\dagger\psi$ are eigenfunctions of energy $(E - \hbar\omega)$ and $(E + \hbar\omega)$.

We can use this to prove by induction that $a^n\psi$ and $(a^\dagger)^n\psi$ are eigenfunctions of energy $(E - n\hbar\omega)$ and $(E + n\hbar\omega)$. For example,

$$\hat{H}a^n\psi = \hat{H}a(a^{n-1}\psi) = (E_{n-1} - \hbar\omega)a^n\psi\,, \tag{6.43}$$

where E_r is the energy eigenvalue of $a^r\psi$. Since $E_0 = E$, it follows by induction that $E_n = (E - n\hbar\omega)$.

In particular, if it were true that $a^n\psi \neq 0$ for all n, there would be eigenfunctions of arbitrarily low energy, and so there would be no ground state.

However, given any physical wavefunction ψ, we have that

$$\begin{aligned}\langle \hat{H} \rangle_\psi &= \int_{-\infty}^{\infty} \psi^* \left(\frac{-\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2 x^2 \psi\right) dx \\ &\geq 0\,,\end{aligned}$$

since both terms are non-negative. (Important note: this argument can obviously be generalised to show that, if we have any potential V such that $V(x) \geq 0$ for all x, then $\langle \hat{H} \rangle_\psi \geq 0$ for all states ψ.)

So there cannot be negative energy eigenfunctions. Thus there must be a lowest energy (i.e. ground state) eigenfunction ψ_0 such that

$$0 = a\psi_0 = \frac{1}{\sqrt{2m}}(\hat{p} - im\omega\hat{x})\psi_0\,, \tag{6.44}$$

which implies

$$-i\hbar\frac{d\psi_0}{dx} = im\omega x\psi_0 \tag{6.45}$$

and hence

$$\psi_0(x) = C\exp\left(-\frac{m\omega x^2}{2\hbar}\right)\,, \tag{6.46}$$

which is indeed the ground state wavefunction we previously obtained. (This time, though, we derived it much more simply, by introducing the operators a, $a^\dagger$ and considering their commutation relations.)

Since $\hat{H} = a^\dagger a + \frac{\hbar\omega}{2}$ and $a\psi_0 = 0$, we have $\hat{H}\psi_0 = \frac{\hbar\omega}{2}\psi_0$, giving us the previously obtained value of $\frac{\hbar\omega}{2}$ for the ground state energy. We have also obtained a closed form expression (6.46) for the ground state and hence for the excited states,

$$(a^\dagger)^n\psi_0 = C\left(\frac{1}{\sqrt{2m}}(\hat{p} + im\omega\hat{x})\right)^n \exp\left(-\frac{m\omega x^2}{2\hbar}\right)\,, \tag{6.47}$$

and we see immediately that their energies are $(n + \frac{1}{2})\hbar\omega$, as previously obtained by a less direct argument.

We can similarly show that there cannot be eigenfunctions ψ' with energies taking values other than $(n + \frac{1}{2})\hbar\omega$). If there were, then $a^m\psi'$ cannot vanish for any m, since ψ_0 is the unique wavefunction annihilated by a. So there would be negative energy eigenfunctions, which contradicts the result shown above.

With a little more thought, we can similarly also show that the eigenspaces must all be non-degenerate: i.e. there is (up to scalar multiplication) only one eigenfunction of each energy.

The derivation of the harmonic oscillator spectrum in this subsection illustrates an important general feature: symmetries or regularities in a quantum mechanical spectrum (such as the regular spacing of the harmonic oscillator energy levels) suggest the existence of a set of operators whose commutation relations define the symmetry or explain the regularity (in this case, the operators $\hat{H}$, a and $a^\dagger$). *