Operators in Quantum Mechanics

Richard Chapling*

v1.1 2 December 2016

One of the most striking features of quantum mechanics, that caused much confusion in its early days, is that doing operations in different orders gives different results. As any mathematician (or, indeed, shoe-and-sock-wearing person) will tell you, this is actually quite a standard state of affairs, and not actually that alarming.

As we know, the wavefunction describes the characteristics of a physical system: it has a deterministic, linear time evolution given by the Schrödinger equation, but measuring the system is described by a probabilistic process.[1] In theory, the wavefunction contains all information about the physical system, but since in general it is a continuous function, and hence contains countably many bits of information, we would not realistically be able to extract every piece of information from it. The situation is in fact worse than this: because not all operators commute, we obtain further restrictions on the amount of certain information we are able to extract: this is the *uncertainty principle* described below.

1 The Postulates of Quantum Mechanics

We summarise the basis of the theory of Quantum Mechanics here:

States The state of the system is described by a vector ψ in a Hilbert space $\mathcal{H}$.

Observables To each observable quantity A (position, momentum, angular momentum, energy, charge, &c.) there corresponds a Hermitian operator $\hat{A}$ on $\mathcal{H}$, although we will subsequently write A for both.

Probability The probability of measuring ψ as being in state ϕ is $|(\phi,\psi)|^2/(\phi,\phi)(\psi,\psi)$. After the system is measured to be in state ϕ, it remains in ϕ. In particular, the only possible values of an observable A are given by the eigenvalues of the operator $\hat{A}$, and when we measure that A has value a, the state of the system becomes an eigenstate of $\hat{A}$ with eigenvalue a.

Average The average value of an operator $\hat{A}$ in a state ψ is given by $(\psi,\hat{A}\psi)/(\psi,\psi)$.

Time evolution The vector evolves (in a way that preserves total probability) via the Schrödinger equation. $i\hbar\partial_t\psi = H\psi$.

Something very useful to note is that this description has said nothing about what ψ actually is, apart from that it lives in something called a Hilbert space. This means that we are able to talk about quantum mechanics in more generality with these ideas than just working with the wavefunction all the time.

2 Hilbert Space

The following is the natural type of space in which to do quantum mechanics:

Definition 1 (Hilbert space). Let $\mathcal{H}$ be a vector space over $\mathbb{C}$, $\langle\cdot,\cdot\rangle : \mathcal{H}\times\mathcal{H}\to\mathbb{C}$ an inner product on $\mathcal{H}$, and $\mathcal{H}$ be complete in the topology induced by the norm $\|\psi\| = \langle\psi,\psi\rangle^{1/2}$. Then $\mathcal{H}$ is called a *Hilbert space*.

* Trinity College, Cambridge

[1] Reconciling these two apparently disparate behaviours is called the Measurement Problem, and is probably the most controversial aspect of quantum mechanics that people actually worry about. However, most physicists are quite happy to use a weird theory if it gives the right results, even if they don't really understand why it does what it does.

Examples You already know plenty of examples of Hilbert spaces; normally the main subtlety is in checking the completeness.

- $\mathbb{C}^n$ with inner product $\langle x, y\rangle = \sum_{i=1}^n \bar{x}_i y_i$ is a very simple example of a Hilbert space, which is used in quantum information theory.
- ℓ^2 is the set of complex sequences (z_k) so that $\sum_k |z_k|^2$ is finite. It is a Hilbert space when given the inner product $\langle z, w\rangle = \sum_{k=1}^\infty \bar{z}_n w_n$.

A notable absentee here that you may have expected is something like $C[0,1]$ with the inner product given by $\langle f, g\rangle = \int_0^1 \bar{f} g$, which you may have discussed when talking about Sturm–Liouville theory. This space is unfortunately not complete: it is easy to make up a sequence of functions with norms converging to zero, that do not tend to the zero function.[2]

3 Operators

In this course, an *operator* is a linear function $\mathcal{H} \to \mathcal{H}$.[3] A lot of the basic properties of operators are familiar from matrix theory: for example, they form a complex vector space $L(H)$,[4] and have a composition operation that looks like multiplication, but is not in general commutative:

$$AB \neq BA.$$

We also know plenty of examples of operators: all matrices are operators, (and it is sometimes helpful to pretend that operators are infinite-dimensional matrices), derivatives are operators, multiplying by a number is an operator, ...

Definition 2 (Adjoint). Let $A \in L(H)$. If there is $B \in L(H)$ so that for all $\phi, \psi \in \mathcal{H}$,

$$\langle \phi, A\psi\rangle = \langle B\phi, \psi\rangle, \tag{1}$$

then B is called the *adjoint* of A. (It is easy to check that it is unique.) We normally write the adjoint as $A^\dagger$.

If $A^\dagger = A$, then A is called *self-adjoint*, or in this course, *Hermitian*.

If $AA^\dagger = A^\dagger A$, A is called *normal*.

In quantum mechanics, every observable quantity (such as energy, position, momentum and angular momentum) has a corresponding self-adjoint operator; hence the self-adjoint operator itself is also referred to as an *observable*.

Normal operators look rather peculiar, and perhaps even extraneous: who cares if a matrix commutes with its adjoint. The value lies in the fact that self-adjoint, skew-adjoint and unitary matrices are all normal, so we can prove properties about them simultaneously.

3.1 Commutators

The most important new object that concerns operators is the *commutator*, which is another sort of "product" on L, in that it takes two operators A, B and produces another, their commutator $[A, B]$. This is defined by

$$[A, B] = AB - BA.$$

Three properties of the commutator suffice for this course:

1. Obviously it is antisymmetric: $[A, B] = -[B, A]$. Therefore we also have $[A, A] = 0$.

[2] The search for a space where this inner product does give a complete norm was only ended by Lebesgue's new theory of integration: the Riemann integral and its successors simply aren't able to do this. This was one of the main reasons for the widespread adoption of Lebesgue's integral as the standard one used in mathematics.

[3] It is possible to take linear functions between different Hilbert spaces, or indeed, different sorts of spaces, but the most interesting ones are the endomorphisms, that map a space to itself: think of all the extra stuff we can do with square matrices!

[4] Warning: this notation is not standard.

2. It is linear in each argument: $[A, B + C] = [A, B] + [A, C]$ (the other case following from the first property).
3. We have the following identity, which is not a million miles from Leibniz's rule for the derivative:[5]

$$[A, BC] = [A, B]C + B[A, C]; \tag{2}$$

this may be proven by expanding the right-hand side.

Equipped with these, we can calculate essentially any commutator we like from more basic ones.
The most important commutator in quantum mechanics is that between position and momentum:

$$[x, p] = i\hbar. \tag{3}$$

Hence, for example, we have

$$[x, H] = \left[x, \frac{p^2}{2m} + V(x)\right] = \frac{1}{2m}[x, p^2] = \frac{1}{2m}([x, p]p + p[x, p]) = \frac{i\hbar}{m}p. \tag{4}$$

In general, $[x, p] = i\hbar$, and then the inductive step

$$[x^k, p] = x^{k-1}[x, p] + [x^{k-1}, p]x = i\hbar x^{k-1} + [x^{k-1}, p]x,$$

implies the general formula

$$[x^n, p] = i\hbar n x^{n-1}.$$

4 Eigenvalues and Eigenstates

As with a matrix, λ is an *eigenvalue* of the operator A if there is at least one $\psi \neq 0$ so that

$$A\psi = \lambda\psi;$$

ψ is then called an *eigenstate* of A, since it is a state vector.

The three things you know about the eigensystem of symmetric (and Hermitian) matrices are actually properties of normal operators:

Theorem 3. *Let A be normal. Then*

1. $Ax = \lambda x \iff A^\dagger x = \bar{\lambda} x$.
2. *Eigenvectors corresponding to different eigenvalues are orthogonal.*
3. *Special cases:*
   a) *If A is self-adjoint, the eigenvalues of A are real.*
   b) *If A is skew-adjoint ($A^\dagger = -A$), the eigenvalues of A are pure imaginary.*
   c) *If A is unitary, the eigenvalues of A have modulus 1.*

These are all straightforward consequences of the simple identity $\|Ax\| = \|A^\dagger x\|$ for normal operators, along with the other, all-too-familiar proofs.

In this course, we also assume that the eigenstates of a Hermitian operator can be chosen to form a basis of $\mathcal{H}$, so that if the eigenstates are labelled $e_{\lambda,n}$, we have[6]

$$\psi = \sum_{\lambda,n} \langle e_{\lambda,n}, \psi\rangle e_{\lambda,n}.$$

[5] For non-commutative multiplication, Leibniz's rule takes the form $d(BC) = (dB)C + B(dC)$.
[6] As with any Fourier series, equality here is pointwise if ψ is nice enough.

5 Uncertainty Principle

The uncertainty principle referred to above is a direct consequence of non-commutativity. We shall prove it here in a general form. In particular, recall the *variance* or uncertainty of A measured for a state ψ is given by

$$(\Delta A)_\psi^2 = \langle (A-a)^2 \rangle_\psi = \langle A^2 \rangle_\psi - \langle A \rangle_\psi^2, \tag{5}$$

where $\langle C \rangle_\psi = \int_{\mathbb{R}} \bar{\psi} C \psi$ for any operator C. If A and B are Hermitian, then $i[A,B]$ is also Hermitian: taking $C = A + itB$ with t real, we have

$$C^\dagger C = (A^\dagger - itB^\dagger)(A+itB) = A^\dagger A + it(A^\dagger B - B^\dagger A) + t^2 B^\dagger B.$$

Applying the Hermiticity,

$$C^\dagger C - (A^2 + t^2 B^2) = ti(AB - BA) = ti[A,B].$$

The left-hand side is Hermitian, so the right-hand side is too.

We can use the same idea to produce an uncertainty relation between A and B: let $C' = (A-a) + it(B-b)$, where the small letters are the expectations of the variables denoted by the capitals, $a = \langle A \rangle_\psi$ &c. Then we can define the polynomial

$$p(t) = \langle C\psi, C\psi \rangle = \langle C^\dagger C \rangle_\psi = \Big\langle \Big((A-a) - it(B-b) \Big) \Big((A-a) + it(B-b) \Big) \Big\rangle_\psi . \tag{6}$$

$p(t)$ is the length of a vector, so it must be nonnegative.

Expanding out,

$$p(t) = \Big\langle (A-a)^2 + it\Big((A-a)(B-b) - (B-b)(A-a) \Big) + t^2 (B-b)^2 \Big\rangle_\psi .$$

Now applying the linearity of expectation,

$$p(t) = \langle (A-a)^2 \rangle_\psi + t \langle i[A-a, B-b] \rangle_\psi + t^2 \langle (B-b)^2 \rangle_\psi$$

The middle term must be real, since every other term in the equation is. We also have $[A-a, B-b] = [A,B]$, because the commutator is linear and the constants commute with everything. Hence

$$p(t) = (\Delta A)_\psi^2 + t \langle i[A,B] \rangle_\psi + t^2 (\Delta B)_\psi^2 .$$

Since p is nonnegative, the discriminant of p must be nonpositive, that is,

$$(\langle i[A,B] \rangle_\psi)^2 - 4(\Delta A)_\psi^2 (\Delta B)_\psi^2 \leqslant 0,$$

and square rooting gives

$$(\Delta A)_\psi^2 (\Delta B)_\psi^2 \geqslant \tfrac{1}{2} \left| \langle [A,B] \rangle_\psi \right|, \tag{7}$$

the *General Uncertainty Principle*.

Taking $A = x$ and $B = p$, we obtain the famous *Heisenberg Uncertainty Principle*,

$$(\Delta x)_\psi (\Delta p)_\psi \geqslant \tfrac{1}{2} \left| \langle [x,p] \rangle_\psi \right| = \tfrac{1}{2} \hbar, \tag{8}$$

beloved of physics coffee mug enthusiasts.

6 A Small Return to the Classical: Ehrenfest's Theorem

It may be of concern that we have dealt exclusively with quantum effects, with essentially no appeal to classical physics. The natural question is of what the relationship between the two very different-seeming theories is. While we can't answer this question in general here,[7] we can give one quite comforting example: the average values of operators follow classical trajectories. More precisely,

Theorem 4 (Ehrenfest's Theorem). *Let $\psi(t)$ be the wavefunction of a quantum system with potential $V(x)$. Then*

$$\frac{d}{dt} \langle x \rangle_{\psi(t)} = \frac{1}{m} \langle p \rangle_{\psi(t)} \tag{9}$$

$$\frac{d}{dt} \langle p \rangle_{\psi(t)} = -\langle V' \rangle_{\psi(t)} \tag{10}$$

[7] And indeed, there are several forests'-worth of discussion on the various issues involved in this.

Proof. The proof is essentially an application of the product rule, the Schrödinger equation, and its conjugate,

$$-i\hbar\partial_t\psi^* = \left(-\frac{\hbar^2}{2m}\partial_x^2 + V\right)\psi^*.$$

Then

$$\begin{aligned}
\frac{d}{dt}\langle x\rangle_{\psi(t)} &= \frac{d}{dt}\int \psi^*(t)x\psi(t) \\
&= \int \partial_t\psi^*(t)x\psi(t) + \int \psi^*(t)x\partial_t\psi(t) \\
&= \langle \partial_t\psi(t), x\psi(t)\rangle + \langle \psi(t), x\partial_t\psi(t)\rangle \\
&= -\frac{1}{i\hbar}\langle H\psi(t), x\psi(t)\rangle + \frac{1}{i\hbar}\langle \psi(t), xH\psi(t)\rangle \\
&= \frac{1}{i\hbar}\langle \psi(t), [x,H]\psi(t)\rangle \\
&= \frac{1}{2mi\hbar}\langle \psi(t), [x,p^2]\psi(t)\rangle \\
&= \frac{1}{2mi\hbar}\langle \psi(t), 2i\hbar p\psi(t)\rangle \\
&= \frac{1}{m}\langle p\rangle_{\psi(t)}.
\end{aligned}$$

The p result is similar, using $[p,H] = [p,V(x)] = -i\hbar V'(x)$. □

The proof suggests to us the following more general result:

Theorem 5 (Heisenberg). *Let $\psi(t)$ be the wavefunction of a quantum system with potential $V(x)$, $A(t)$ an operator. Then*

$$\frac{d}{dt}\langle A(t)\rangle_{\psi(t)} = \langle [A(t),H]\rangle_{\psi(t)} + \langle \partial_t A(t)\rangle_{\psi(t)}. \tag{11}$$

The proof is as before, with an extra term for the time-derivative of A.

Ehrenfest's equantions should remind you of Hamilton's equations,

$$\frac{dq}{dt} = \frac{\partial H}{\partial p} = \frac{p}{m}, \qquad \frac{dp}{dt} = -\frac{\partial H}{\partial q} = -V'(q). \tag{12}$$

In other words, the centre of mass follows a classical trajectory, and the quantum contribution is of perturbation about this trajectory. There is also a classical version of Heisenberg's theorem,[8] which uses an object called the *Poisson bracket*;[9] this is defined by

$$\{A,B\} = \sum_k \frac{\partial A}{\partial q_k}\frac{\partial B}{\partial p_k} - \frac{\partial B}{\partial q_k}\frac{\partial A}{\partial p_k}.$$

for two functions A, B of p and q.

This gives the oddly familiar-looking

$$\{q,p\} = 1,$$

and time-evolution is described by *Liouville's equation*,

$$\frac{dA}{dt} = \{A,H\} + \frac{\partial A}{\partial t}.$$

[8] Or rather, Heisenberg's result is the quantum analogue of the following result, which preceded it.

[9] Discussed in Classical Dynamics in Part II.

7 The Harmonic Oscillator

The potential most beloved of physicists is that of the *harmonic oscillator*,

$$V(x) = \frac{1}{2}m\omega^2 x^2. \tag{13}$$

This has many nice properties, some of which are explored on the example sheets. We want to find the energy eigenstates.

The time-independent Schrödinger equation is

$$-\frac{\hbar^2}{2m}\psi'' + \frac{1}{2}m\omega^2 x^2\psi = E\psi. \tag{14}$$

We change variables, setting $y = \sqrt{m\omega/\hbar}x$ and $\mathcal{E} = 2E/\hbar\omega$. The differential equation simplifies to

$$-\frac{d^2\psi}{dx^2} + y^2\psi = \mathcal{E}\psi. \tag{15}$$

This doesn't look enormously promising: we still can't solve this!

We have two options at this point: we can guess a solution, or try and do the same trick as on the first example sheet, where we factorise the Hamiltonian into two first-order operators, which are much easier to use.

7.1 Guessing a Solution

Put $\mathcal{E} = 1$. Then it so happens that

$$-\frac{d^2\psi}{dy^2} + (y^2 - 1)\psi = 0. \tag{16}$$

is solved by

$$\psi_0(y) = Ce^{-y^2/2}. \tag{17}$$

(Check!). It is easy to show using the Uncertainty Principle that this solution has the lowest possible energy, so it is a ground state. Energy levels are non-degenerate in one dimension, so it is *the* ground state.

To get to the rest of the eigenstates and energy levels, we make a sensible assumption, that the character of the decay of the solutions doesn't depend on the energy (since $y^2\psi$ is the dominant term in the equation for large y, this seems reasonable). Thus we look for solutions in the form

$$\psi(y) = f(y)e^{-y^2/2}.$$

Inserting this into the Schrödinger equation, we find that f satisfies (check!)

$$\frac{d^2f}{dy^2} - 2y\frac{df}{dy} + (\mathcal{E} - 1)f = 0.$$

It is then clear that if $\mathcal{E} = 1$, $f(y) = 1$ is a solution. Otherwise, we now have to use the method of Frobenius.[10] It is clear that $y = 0$ is a regular point, since the leading coefficient is 1, and the rest are analytic nearby. Multiplying to make the equation homogeneous,

$$\left(y^2\frac{d^2f}{dy^2}\right) - 2y^2\left(y\frac{df}{dy}\right) + (\mathcal{E} - 1)y^2 f = 0.$$

We now substitute

$$f(y) = \sum_{n=0}^{\infty} y^n,$$

[10] Recall IA Differential Equations

which gives

$$\begin{aligned} 0 &= \sum_n \left[n(n-1) - 2ny^2 + (\mathcal{E}-1)y^2\right] a_n y^n \\ &= \sum_n \left[[n(n-1)]a_n - [2(n-2) - (\mathcal{E}-1)]a_{n-2}\right] y^n, \end{aligned}$$

rearranging to get the same coefficients. Hence

$$a_n = \frac{2n - \mathcal{E} - 3}{n(n-1)} a_{n-2},$$

or

$$a_{n+2} = \frac{2n + 1 - \mathcal{E}}{(n+2)(n+1)} a_n. \tag{18}$$

We firstly see that there are two independent solutions: an odd one and an even one. Our next concern is the large-y behaviour. Suppose the power series we have found does not terminate. Then for large n, the recurrence relation to first order

$$a_{n+2} \sim \frac{2}{n} a_n.$$

But this recurrence is that satisfied by e^{2x^2}: the large-y behaviour of the two functions is the same, and this implies that the overall ψ will not be normalisable. Hence we must have the series terminate, so there is a nonnegative integer N so that $\mathcal{E} = 2N + 1$. Unwinding the substitution we made, we conclude that the energy levels are given by

$$E_N = \left(N + \frac{1}{2}\right)\hbar\omega, \qquad N \in \{0, 1, 2, \dots\}, \tag{19}$$

and the eigenfunctions are polynomials of degree N, odd or even if N is odd or even respectively.

What are these polynomials?

7.2 The Hermite Polynomials

The solutions to (16) with $\mathcal{E} = 2N + 1$ are therefore of the form

$$\psi_N(y) = c_N h_N(y) e^{-y^2/2}, \tag{20}$$

where h_N has degree N and parity $h_N(-y) = (-1)^N h_N(y)$, and c_N is the normalisation constant Because (16) is a Sturm–Liouville equation on $(-\infty, \infty)$ and these solutions have different eigenvalues, they are all mutually orthogonal, i.e. if $m \neq n$,

$$0 = \int_{-\infty}^{\infty} \overline{\psi_m}\psi_n = \int_{-\infty}^{\infty} h_m(y) h_n(y) e^{-y^2}\, dy, \tag{21}$$

so the h_N are orthogonal polynomials with respect to the weight e^{-y^2}.

This fact is all we need to calculate them, using, for example, the Gram–Schmidt proceedure: we find that with the most common convention,

$$\begin{aligned} h_0(y) &= 1 \\ h_1(y) &= 2y \\ h_2(y) &= 4y^2 - 2 \\ h_3(y) &= 8y^3 - 12y \\ &\vdots \end{aligned}$$

These are not normalised with respect to the inner product: instead, they are designed so that they have the exponential generating function

$$G(x,t) = \sum_{n=0}^{\infty} h_n(x) \frac{t^n}{n!} = e^{2xt - t^2}; \tag{22}$$

one can check that this satisfies the differential equation $G_{xx} - 2xG_x + 2tG_t = 0$, which contains all of the other differential equations as coefficients of t.

7.2.1 Hermite Polynomials and the Fourier Transform

One nice property of the harmonic oscillator wavefunctions ψ_n that is often not mentioned is that they are eigenfunctions of the Fourier transform: to see this, we can use the generating function. We use the unitary transform with simple phase, which is best for quantum mechanics:

$$\mathcal{F}(f)(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iky} f(y)\, dy.$$

The transform of the polynomials can then be computed by

$$\begin{aligned}
\sum_{n=0}^{\infty} \frac{t^n}{n!} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iky} h_n(y) e^{-y^2/2}\, dy &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ity} G(y,t) e^{-y^2/2}\, dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ity} e^{2yt - t^2 - y^2/2}\, dy \\
&= e^{-k^2/2 - 2kit + t^2} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-(y - t + ik/2)^2/2}\, dy \\
&= G(k, -it) e^{-k^2/2},
\end{aligned}$$

so equating coefficients,

$$\mathcal{F}(h_n(y) e^{-y^2/2})(k) = (-i)^n h_n(k) e^{-k^2/2}. \tag{23}$$

Since the Fourier transform is a unitary operator with $\mathcal{F}^4 = 1$, these are the only possible eigenvalues.

Another useful fact, which we do not prove here, is that these wavefunctions span $L^2(\mathbb{R})$, the space of integrable functions on $\mathbb{R}$ with $\int_{\mathbb{R}} |f|^2 < \infty$. Hence we may obtain an orthonormal basis of $L^2(\mathbb{R})$ which *diagonalises* the Fourier transform.[11]

7.2.2 Rodriguez-like formulae

From the generating function (22), we can find a Rodriguez-like formula for h_n:

$$h_n(y) = (-1)^n e^{y^2} \partial_y^n e^{-y^2},$$

and we have the operator identity

$$-e^{y^2} \partial_y e^{-y^2} = e^{y^2/2} (y - \partial_y) e^{-y^2/2},$$

so we have

$$h_n(y) = e^{y^2/2} (y - \partial_y)^n e^{-y^2/2}. \tag{24}$$

That's interesting: we can go from one polynomial to another using a differential operator. Can we do the same for the solutions to the original problem?

7.3 Factorising the Hamiltonian

The obvious place to start hunting for such a differential operator is in the Hamiltonian, since it has a differential operator in it, albeit second-order. We can use the trick we learnt on the first example sheet: we suspect that there is a differential operator a so that $H = ca^\dagger a + d$ for some numeric constant c. An obvious candidate for a is some linear combination of p and x, since $H = \frac{1}{2m} p^2 + \frac{m\omega^2}{2} x^2$, so let $a = \lambda p + \mu x$. Then, we have

$$\begin{aligned}
a^\dagger a &= (\lambda^* p^\dagger + \mu^* x^\dagger)(\lambda p + \mu x) \\
&= (\lambda^* p + \mu^* x)(\lambda p + \mu x) \\
&= |\lambda|^2 p^2 + \lambda^* \mu p x + \lambda \mu^* x p + |\mu|^2 x^2 \\
&= |\lambda|^2 p^2 + |\mu|^2 x^2 + (\lambda^* \mu + \lambda \mu^*) p x + \lambda \mu^* [x, p] \\
&= |\lambda|^2 p^2 + |\mu|^2 x^2 + 2\Re(\lambda \mu^*) p x + \lambda \mu^* i\hbar
\end{aligned}$$

[11] For more details of this, see the excellent book *Fourier Series and Integrals*, by Dym and McKean, § 2.5.

We conclude that

$$c\,|\lambda^2| = \frac{1}{2m}, \quad c\,|\mu|^2 = \frac{m\omega^2}{2}, \quad \Re(\lambda\mu^*) = 0.$$

One obvious way to satisfy these is to take

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{ip}{m\omega}\right) \qquad a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{ip}{m\omega}\right),$$

and then

$$H = \hbar\omega(a^\dagger a + \tfrac{1}{2}).$$

The commutation relations for x and p give corresponding ones for a and $a^\dagger$:

$$\begin{aligned}[a, a^\dagger] &= \frac{m\omega}{2\hbar}\left[x + \frac{ip}{m\omega}, x - \frac{ip}{m\omega}\right] \\ &= \frac{m\omega}{2\hbar}\left(2\frac{-i}{m\omega}[x,p]\right) = 1;\end{aligned}$$

if we also introduce the self-adjoint operator

$$N = a^\dagger a$$

(the *number* operator), then we also have

$$[N, a] = -a, \qquad [N, a^\dagger] = a^\dagger.$$

These identities are the key to what follows. a is called the *annihilation operator*, and $a^\dagger$ the *creation operator*, for reasons that will become clear: essentially, $a^\dagger$ can be used to change an eigenstate into one with higher energy, and conversely, a into one with lower energy.

It is clear that the eigenstates of H are the same as those of N, and the eigenvalues are linearly related. Suppose we have a normalised eigenstate ψ with eigenvalue λ. Then

$$\lambda = \langle\psi, N\psi\rangle = \langle\psi, a^\dagger a\psi\rangle = \langle a\psi, a\psi\rangle = \|a\psi\|^2 \geqslant 0,$$

so the eigenvalues are nonnegative, and zero if and only if $a\psi = 0$.

Ah, but this is only a first-order equation:

$$a\psi = 0 \iff 0 = (ip + m\omega x)\psi = \left(\hbar\frac{d}{dx} + m\omega x\right)\psi,$$

which we know has solution

$$\psi(x) = Ce^{-(m\omega/\hbar)x^2/2}.$$

Now, if ψ has N-eigenvalue λ, then $a\psi$ is also an eigenstate:

$$Na\psi = ([N, a] + aN)\psi = (-a + a\lambda)\psi = (\lambda - 1)a\psi$$

Therefore if the eigenvalue of N is not an integer, we can make a state with number eigenvalue as negative as we like by applying a enough times. This is obviously no good, since we know that the eigenvalues are nonnegative for normalisable states! Therefore the eigenvalues of N are positive integers.

On the other hand, each positive integer is an eigenvalue: since $[N, a^\dagger] = a^\dagger$, if ψ has eigenvalue λ, $a^\dagger\psi$ has eigenvalue $\lambda + 1$. Also, the norm of the state $a^\dagger\psi$ is

$$\left\|a^\dagger\psi\right\|^2 = \langle a^\dagger\psi, a^\dagger\psi\rangle = \langle\psi, aa^\dagger\psi\rangle = \langle\psi, N + [a, a^\dagger]\psi\rangle = (\lambda + 1).$$

Putting all this together, we find:

1. The eigenvalues of H are $\hbar\omega(n + 1/2)$, for $n \in \{0, 1, 2, \dots\}$

2. The ground state is

$$\psi_0(x) = \left(\frac{m\omega}{2\pi\hbar}\right)^{1/4} \exp\left(-\frac{m\omega}{\hbar}\frac{x^2}{2}\right) \tag{25}$$

3. The higher normalised eigenstates are

$$\psi_n(x) = \frac{1}{\sqrt{n!}}(a^\dagger)^n \psi_0. \tag{26}$$

All of which took considerably less work (and especially algebra) than working with the differential equations!

8 Angular Momentum

In three dimensions, position and momentum are given by vectors, rather than just real numbers. Hence in Quantum Mechanics we need to use vector operators to represent them. The easiest way to do this is to look at each component separately, so we have three position operators x_1, x_2, x_3, and three momentum operators p_1, p_2, p_3. Analogously to one dimension, the momentum operators are represented by derivatives, $p_k = -i\hbar\partial_{x_k}$, or in vectorial terms, $\mathbf{p} = -i\hbar\nabla$. One then has the commutators

$$[x_i, x_j] = 0 \tag{27}$$
$$[p_i, p_j] = 0 \tag{28}$$
$$[x_i, p_j] = -i\hbar x_i \partial_{x_j} + i\hbar\partial_{x_j} x_i = i\hbar\delta_{ij}. \tag{29}$$

The classical definition of angular momentum is $\mathbf{L} = \mathbf{x} \times \mathbf{p}$, and we carry this over to the quantum case by considering each individual component: we have the three operators, for angular momentum in the three directions,

$$L_i = \varepsilon_{ijk} x_j p_k. \tag{30}$$

We find the commutator of two of these in the usual way, although we need some fairly heroic Levi-Civita manipulation:

$$\begin{aligned}
[L_i, L_j] &= [\varepsilon_{ikl} x_k p_l, \varepsilon_{jmn} x_m p_n] \\
&= \varepsilon_{ikl}\varepsilon_{jmn}[x_k p_l, x_m p_n] \\
&= \varepsilon_{ikl}\varepsilon_{jmn}([x_k p_l, x_m]p_n + x_m[x_k p_l, p_n]) \\
&= \varepsilon_{ikl}\varepsilon_{jmn}(-i\hbar\delta_{lm} x_k p_n + i\hbar\delta_{kn} x_m p_l) \\
&= i\hbar(\varepsilon_{ikl}\varepsilon_{jmk} x_m p_l - \varepsilon_{ikl}\varepsilon_{jln} x_k p_n) \\
&= i\hbar\left((\delta_{im}\delta_{jl} - \delta_{ij}\delta_{km}) x_m p_l - (\delta_{in}\delta_{jl} - \delta_{ij}\delta_{kn}) x_k p_n\right) \\
&= i\hbar(x_i p_j - x_j p_i) \\
&= i\hbar\varepsilon_{ijk} L_k
\end{aligned}$$

This, of course, means that we can't find more than one component of the angular momentum with certainty.

On the other hand, we can look at the total angular momentum, i.e. the scalar quantity $L^2 := L_1^2 + L_2^2 + L_3^2 = L_i L_i$. It turns out that this commutes with the individual components of $\mathbf{L}$:

$$[L^2, L_3] = [L_i L_i, L_3] = L_i[L_i, L_3] + [L_i, L_3]L_i = i\hbar\varepsilon_{i3k}(L_j L_k + L_k L_j) = 0$$

since it is the contraction of a symmetric tensor with an antisymmetric one. Therefore, we *can* choose states that have both definite L_3 (called *magnetic quantum number*) and definite L^2 (called *azimuthal quantum number*).[12]

[12] The magnetic quantum number is so-called because particles with different values of m can be distinguished by their movement in a magnetic field.

L^2 contains two derivatives, so we expect it to relate to the Laplacian. We can do this by expanding:

$$\begin{aligned}
L_iL_i &= \varepsilon_{ijk}\varepsilon_{ilm}x_jp_kx_lp_m \\
&= (\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl})x_j(x_lp_k + [p_k, x_l])p_m \\
&= (\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl})x_j(x_lp_k - i\hbar\delta_{kl})p_m \\
&= x_j(x_jp_k - i\hbar\delta_{jk})p_k - x_j(x_kp_k - i\hbar\delta_{kk})p_j \\
&= r^2p^2 + 2i\hbar(x\cdot p) - x_kx_jp_kp_j \\
&= r^2p^2 + 2i\hbar(x\cdot p) - x_k(p_kx_j + i\hbar\delta_{jk})p_j \\
&= r^2p^2 + i\hbar(x\cdot p) - (x\cdot p)^2
\end{aligned}$$

It is not quite clear yet, but with the representation $p = -i\hbar\nabla$, we find that $x\cdot p = -i\hbar r\partial_r$, and so after rearrangement we have

$$-\hbar^2\nabla^2 = -\hbar^2\left(\frac{1}{r}\partial_r r\partial_r + \frac{1}{r}\partial_r\right) + \frac{1}{r^2}L^2 = -\hbar^2\left(\partial_r^2 + \frac{2}{r}\partial_r\right) + \frac{1}{r^2}L^2.$$

The bracket, of course, is the radial part of the Laplacian, which can also be written as $r^{-2}\partial_r r^2\partial_r$ or $r^{-1}\partial_r^2 r$.

In particular, this implies that L^2 contains only angular derivatives. Similarly, we find that $L_3 = -i\hbar\partial_\phi$. It follows that if $V = V(r)$ is radial, then $[H, L^2] = [H, L_3] = 0$, so we can find a simultaneous eigenstate of all three operators. In the language of partial differential equations, we want to separate variables, initially to $\psi(r,\theta,\phi) = R(r)Y(\theta,\phi)$. Then

$$2mr^2\left(-\frac{\hbar^2}{2m}\frac{1}{rR}\partial_r^2 rR + V(r) - E\right) = -\frac{L^2Y}{Y}.$$

The left-hand side depends only on r, the right on θ and ϕ, so they are both equal to a constant, which we write as $\hbar^2\ell(\ell+1)$ for reasons that will become apparent. The Schrödinger equation therefore becomes the radial equation,

$$-\frac{\hbar^2}{2m}\frac{1}{r}\partial_r^2 rR + \frac{\hbar^2\ell(\ell+1)}{2mr^2}R + VR = ER \tag{31}$$

and the angular equation

$$L^2Y = \hbar^2\ell(\ell+1)Y.$$

The functions Y are universal in that they are the same for any radial potential; they are called *spherical harmonics*.

The radial equation (31) can be simplified further by taking $\chi = rR$, reducing it to a one-dimensional Schrödinger equation with a modified potential,

$$-\frac{\hbar^2}{2m}\partial_r^2\chi + \left(\frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2} + V\right)\chi = E\chi;$$

this is analogous to the introduction of a centripetal term in classical mechanics.[13]

[13] That this looks odd is probably due to the Hamiltonian being used, rather than the more usual Lagrangian. Remember the Lagrangian has very little application in Quantum Mechanics.