Theory Option

Quantum Mechanics

Prof. J.J. Binney

Oxford University

TT 2002

1 Dirac notation

1.1 Quantum amplitudes

According to quantum mechanics the future is to an extent unpredictable and the best we can do is to calculate probabilities for the outcomes of measurements. This unpredictability isn't peculiar to QM – the existence of classical chaos ensures that something similar is true in classical physics and real life. What *is* amazing is *how* the probability P of any event E is calculated in QM: $p = |A|^2$, where the complex number A is the **amplitude** for the event E. While probabilities are widely used throughout the natural and social sciences, as far as I know, probability amplitudes are unique to quantum mechanics and nobody has the faintest idea why they arise.

Imagine that you're investigating some physical system: some particles in an ion trap, a drop of liquid He, the e.m. field in a resonant cavity. What do you know about the state of this system? You have two types of knowledge: (1) a specification of the physical nature of the system (e.g., size & shape of the resonant cavity), and (2) information about the current dynamical state of the system. Information of type (1) goes to define the system's Hamiltonian H. Information of type (2) is more subtle. It must consist of predictions for the outcomes of measurements you could make on the system. Since these outcomes are inherently uncertain, your information must relate to the probabilities of different outcomes, and will ideally consist of values of the relevant probability amplitudes. For example, your knowledge might consist of amplitudes for the various possible outcomes of a measurement of energy, or of a measurement of momentum.

Given the amplitudes for a certain set of events, it is often possible to calculate amplitudes for other events. For example, if I know the amplitudes for finding the spin of an electron to be either parallel or antiparallel to the z axis, I can calculate from these two complex numbers the amplitude for the spin to be parallel to any other axis. On the other hand, if I know only the amplitude for the spin to be parallel to the z axis, I can't calculate the amplitude to be parallel to *any* other axis. Generalizing from this example, we have the concept of a **complete set of amplitudes**: the set contains enough information to enable one to calculate amplitudes for the outcome of *any* measurement whatsoever. Hence, such a set gives a complete specification of the physical state of the system. A complete set of amplitudes is generally understood to be a minimal set in the sense that none of the amplitudes can be calculated from the others.

Dirac introduced the symbol $|\psi\rangle$, pronounced 'ket psi', to denote a complete set of amplitudes for the system. If the system consists of a spin-zero particle trapped in a potential well, $|\psi\rangle$ could consist of the amplitudes a_n that the energy is E_n ($n = 1, 2, \ldots$) the nth eigenvalue of the Hamiltonian H, or it might consist of the amplitudes $\psi(x)$ that the particle is found at x ($-\infty < x < \infty$), or it might consist of the amplitudes $\widetilde{\psi}(p)$ that the momentum is measured to be p ($-\infty < p < \infty$). Using the abstract symbol $|\psi\rangle$ enables us to think about the system without committing ourselves to what complete set of amplitudes we are going to use. Moreover, the symbol $|\psi\rangle$ encapsulates the crucial concept of a **quantum state**, which is independent of the particular set of amplitudes that we choose to quantify it with in the same way that a geometrical point is independent of the coordinates (x, y, z), (r, θ, ϕ) or whatever by which we refer to it. The idea of a quantum state is fundamental to several branches of physics.

The interference phenomena that are the hallmark of quantum mechanics occur because amplitudes must sometimes be added: if an outcome can be achieved in two different ways and we do not monitor the way in which it is achieved, we add the amplitudes associated with each way to get the overall

amplitude for the outcome. The classical example is two-slit diffraction: the amplitude for a particle to get from $\mathbf{r}_1$ on one side of the screen to $\mathbf{r}_2$ on the other is $A_1 + A_2$, where A_i is the amplitude to go via the ith slit. In view of this additivity, we write

$$|\psi_3\rangle = |\psi_1\rangle + |\psi_2\rangle \tag{1.1}$$

to mean that every amplitude in the complete set $|\psi_3\rangle$ is the sum of the corresponding amplitudes in the complete sets $|\psi_1\rangle$ and $|\psi_2\rangle$. Since amplitudes are complex numbers, for any complex number α we can define

$$|\psi'\rangle = \alpha|\psi\rangle \tag{1.2}$$

to mean that every amplitude in the set $|\psi'\rangle$ is α times the corresponding amplitude in $|\psi\rangle$.

1.2 Vector spaces and their adjoints

Mathematicians say that objects that you can add and multiply by complex numbers inhabit a **vector space**. We now take some time to review the general theory of vector spaces. The ground we cover is as relevant to special relativity as it is to quantum mechanics.

For any vector space V it is natural to choose a set of **basis vectors**, that is, a set of vectors $|i\rangle$ that is large enough for it to be possible to express any given vector $|\psi\rangle$ as a linear combination of the set's members:

$$|\psi\rangle = \sum_i a_i|i\rangle. \tag{1.3}$$

The set should be minimal in the sense that none of its members can be expressed as a linear combination of the remaining ones.

An important role is played by complex-valued linear functions on V. Let $\langle f|$ (pronounced 'bra f') be such a function, i.e., an object such that for and any $|\psi\rangle$, $\langle f|\psi\rangle$ is a complex number and for any complex numbers α, β

$$\langle f|(\alpha|\psi\rangle + \beta|\phi\rangle) = \alpha\langle f|\psi\rangle + \beta\langle f|\phi\rangle. \tag{1.4}$$

Then these functions can be added and multiplied by complex numbers because for given $\langle f|$ and $\langle g|$ we can define $\langle h| = \langle f| + \langle g|$ and $\langle p| = \alpha\langle f|$ by

$$\langle h|\psi\rangle = \langle f|\psi\rangle + \langle g|\psi\rangle \quad ; \quad \langle p|\psi\rangle = \alpha\langle f|\psi\rangle. \tag{1.5}$$

Since the functions can be added and multiplied by complex numbers, they themselves form a vector space V', the **adjoint space** of V.

The **dimension** of a vector space is the number of vectors required to make up a basis for the space. We now show that V and V' have the same dimension. Let $|i\rangle$, $(i = 1, N)$ be a basis for V. Then a linear function $\langle f|$ on V is fully defined once we have given the N numbers $\langle f|i\rangle$ – to see that this is true, we use (1.3) and the linearity of $\langle f|$ to calculate $\langle f|\psi\rangle$ for an arbitrary vector $|\psi\rangle$:

$$\langle f|\psi\rangle = \sum_{i=1}^{N} a_i\langle f|i\rangle. \tag{1.6}$$

This result implies that we can define N functions $\langle j|$ $(j = 1, N)$ through the equations

$$\langle j|i\rangle = \delta_{ij}, \tag{1.7}$$

where δ_{ij} is as usual 1 if $i = j$ and zero otherwise. Now consider the following linear combination of these bras:

$$\langle F| \equiv \sum_{j=1}^{N} \langle f|j\rangle\langle j|. \tag{1.8}$$

It is trivial to check that for any i we have $\langle F|i\rangle = \langle f|i\rangle$, and from this it follows that $\langle F| = \langle f|$ because we have already agreed that a bra is fully specified by the values it takes on the basis vectors. Since we have now shown that *any* bra can be expressed as a linear combination of the N bras of (1.7), and the latter are manifestly linear independent, it follows that the dimensionality of V' is N, the dimensionality of V.

In summary, we have established that with every N-dimensional vector space V there is associated the N-dimensional adjoint space of linear functions on V. Moreover, we have shown that once we have chosen a basis for V, there is an associated basis for V'.

1.2.1 Changes of basis From our experience of 3-dimensional space we know that it is sometimes expedient to swap the basis vectors that we employ for others that are better suited to the problem in hand (i.e., rotate our coordinates). So let's investigate what happens if we make a unitary transformation to a new basis for V. That is, we define new basis vectors by

$$|n'\rangle = \sum_i U_{ni}|i\rangle \tag{1.9a}$$

where the complex numbers U_{ni} satisfy the unitarity condition

$$\sum_i U_{ni}U^*_{mi} = \delta_{nm} \quad \leftrightarrow \quad \mathbf{U}^{T*} = \mathbf{U}^{-1}. \tag{1.9b}$$

The bras associated with the kets $|n'\rangle$ will be linear combinations of the old basis bras, so we write out their defining equation in the form

$$\delta_{mn} = \langle m'|n'\rangle = \left(\sum_i B_{mi}\langle i|\right)\left(\sum_j U_{nj}|j\rangle\right) = \sum_i B_{mi}U_{ni}. \tag{1.10}$$

Comparing this equation with (1.9b), we deduce that $B_{mi} = U^*_{mi}$, i.e., that under (1.9) the basis bras transform by the rule

$$\langle m'| = \sum_i U^*_{mi}\langle i| \tag{1.11}$$

Now let's ask how the expansion coefficients a_i of equation (1.3) transform under (1.9):

$$\begin{aligned} |\psi\rangle &= \sum_i a_i|i\rangle = \sum_n a'_n|n'\rangle = \sum_{nj} a'_n U_{nj}|j\rangle \\ \Rightarrow \quad a_i &= \sum_n a'_n U_{ni} \quad \Rightarrow \quad a'_n = \sum_i U^*_{ni}a_i, \end{aligned} \tag{1.12}$$

where we have again used (1.9b). There is a corresponding rule for the transformation of bras:

$$\begin{aligned} \langle\phi| &= \sum_i b_i\langle i| = \sum_n b'_n\langle n'| = \sum_{nj} b'_n U^*_{nj}\langle j| \\ \Rightarrow \quad b_i &= \sum_n b'_n U^*_{ni} \quad \Rightarrow \quad b'_n = \sum_i U_{ni}b_i. \end{aligned} \tag{1.13}$$

These rules are such that the complex number $\langle\phi|\psi\rangle$ comes out the same no matter which basis we use for its evaluation:

$$\begin{aligned} \langle\phi|\psi\rangle &= \left(\sum_i b_i\langle i|\right)\left(\sum_j a_j|j\rangle\right) = \sum_i b_i a_i \\ \langle\phi|\psi\rangle &= \left(\sum_m b'_m\langle m'|\right)\left(\sum_n a'_n|n'\rangle\right) = \sum_m b'_m a'_m = \sum_{mij} U_{mi}b_i U^*_{mj}a_j = \sum_i b_i a_i, \end{aligned} \tag{1.14}$$

where the unitarity of **U** has again been used.[1]

Equations (1.9a) and (1.11) suggest that with any ket $|\psi\rangle$ we associate a bra $\langle\psi|$ by the rule

$$|\psi\rangle = \sum_i a_i|i\rangle \quad \Leftrightarrow \quad \langle\psi| = \sum_i a_i^*\langle i| \tag{1.15}$$

In addition to being consistent with the rule for transforming basis vectors, this rule implies that $\langle\psi|\psi\rangle = \sum_i |a_i|^2 \geq 0$ so that we can speak of $\langle\psi|\psi\rangle$ as the 'length squared' of $|\psi\rangle$. Notice that if $\sigma = \sum_i c_i|i\rangle$, then with (1.14) the rule (1.15) implies that

$$\langle\psi|\sigma\rangle = \sum_i a_i^* c_i, \tag{1.16}$$

which we identify as the dot product of $|\psi\rangle$ with $|\sigma\rangle$.

1.3 Operators and observables

After this orgy of algebra, let's get back to physics. QM asserts that with every possible measurement there is associated a linear operator on V, i.e., a beast Q that turns kets into other kets, $|\phi\rangle = Q|\psi\rangle$, in a linear way: $Q\sum_i a_i|i\rangle = \sum_i a_i Q|i\rangle$. The operator is claimed to be Hermitian in the sense that for any two kets $|\psi\rangle$, $|\phi\rangle$, the complex numbers $\langle\psi|Q|\phi\rangle$ and $\langle\phi|Q|\psi\rangle$ are complex conjugates of one another:

$$\langle\psi|Q|\phi\rangle = (\langle\phi|Q|\psi\rangle)^* \qquad \text{(Hermitian } Q\text{)}. \tag{1.17}$$

By expanding the input and output kets in our chosen basis we discover that Q is completely determined by the array of numbers (**matrix elements**)

$$Q_{ij} \equiv \langle i|Q|j\rangle. \tag{1.18}$$

In fact

$$|\phi\rangle = \sum_i b_i|i\rangle = \sum_j a_j Q|j\rangle \quad \Rightarrow \quad b_i = \sum_j a_j\langle i|Q|j\rangle = \sum_j Q_{ij}a_j. \tag{1.19}$$

The matrix Q_{ij} is Hermitian: $Q_{ij} = Q_{ji}^*$.

It would be natural to imagine that the act of measurement, which disturbs the system, changed its amplitude set from $|\psi\rangle$ to $|\phi\rangle = Q|\psi\rangle$. This is not what QM asserts, however. Rather it asserts that when we make the measurement, $|\psi\rangle$ 'collapses' into one of the eigenkets of Q, i.e., a ket $|q_i\rangle$ that satisfies the eigenvalue equation

$$Q|q_i\rangle = q_i|q_i\rangle, \tag{1.20}$$

where q_i is a number (which we can prove is real). In principle the normalization of an eigenvector is arbitrary, but we always insist that they have unit length-square:

$$\langle q_i|q_i\rangle = 1. \tag{1.21}$$

We may assume that the eigenvectors are mutually orthogonal, so that with (1.21) we have

$$\langle q_i|q_j\rangle = \delta_{ij}. \tag{1.22}$$

QM asserts that if $|\psi\rangle \to |q_i\rangle$, the measuring instrument shows the number q_i, and just after the measurement the state of our system is $|q_i\rangle$. The amplitude that we measure q_i is claimed to be

$$a_i = \langle q_i|\psi\rangle, \tag{1.23a}$$

[1] Multiplying (1.9) by U_{nj}^* and summing over n we find that $\sum_{ni} U_{nj}^* U_{ni} U_{mi}^* = U_{mj}^*$ which implies that $\sum_n U_{nj}^* U_{ni} = \delta_{ij}$, which is the relation required in (1.14).

which is precisely the coefficient of $|q_i\rangle$ if we use the eigenvectors as a basis in which to expand $|\psi\rangle$:

$$|\psi\rangle = \sum_i a_i |q_i\rangle \tag{1.23b}$$

The condition that we measure *some* value for Q is that the resulting probabilities add up to one:

$$1 = \sum_i P_i = \sum_i |a_i|^2 = \langle\psi|\psi\rangle, \tag{1.24}$$

so the ket that describes our system has to have unit length-squared.

The ket $|q_i\rangle$ describes the state of the system in which it is certain that a measurement of Q will yield the value q_i. Hence, it is directly connected to the complete set of amplitudes $\langle q_j|q_i\rangle$, $(j = 1, N)$.

Suppose we know the amplitudes $a_i = \langle q_i|\psi\rangle$ and we want to calculate the probability that a measurement of some quantity R will yield the value r_n. Then we have to calculate the amplitude $\langle r_n|\psi\rangle$, where $R|r_n\rangle = r_n|r_n\rangle$. Multiplying (1.23b) by $\langle r_n|$ we find

$$\langle r_n|\psi\rangle = \sum_i \langle r_n|q_i\rangle\langle q_i|\psi\rangle \tag{1.25}$$

Substituting $a'_n \equiv \langle r_n|\psi\rangle$, $a_i \equiv \langle q_i|\psi\rangle$ and $U^*_{ni} = \langle r_n|q_i\rangle$ this becomes identical with (1.12).

Exercise (1):
Show that the matrix $U^*_{ni} \equiv \langle r_n|q_i\rangle$ is unitary.

Hence a unitary transformation in the space of kets corresponds to the process of calculating the amplitudes for a new set of events given the amplitudes for some other complete set of events. This conclusion is very remarkable because it implies that the amplitude for *any* event can be expressed as a *linear* combination of some basic set of amplitudes; a priori it's far from obvious that all functional relationships between amplitudes are linear.

1.3.1 The identity operator and functions of operators Equations (1.23) imply that for any $|\psi\rangle$ we have $|\psi\rangle = \sum_i |q_i\rangle\langle q_i|\psi\rangle$, so

$$\sum_i |q_i\rangle\langle q_i| = I, \quad \text{the identity operator.} \tag{1.26}$$

For any operator Q we may write

$$Q = \sum_i |q_i\rangle q_i \langle q_i| \tag{1.27}$$

as we may verify by operating with each side on the expansion (1.23b) of an arbitrary ket $|\psi\rangle$. If $f(x)$ is any function of the real variable x, we define the operator $f(Q)$ by

$$f(Q) \equiv \sum_i |q_i\rangle f(q_i)\langle q_i|. \tag{1.28}$$

If f is a polynomial or has a convergent Taylor series $f(x) = \sum_n f_n x^n$, we can alternatively define $f(Q) = \sum_n f_n Q^n$. It is straightforward to show that this less general definition of $f(Q)$ is equivalent to (1.28).

1.3.2 Hermitian adjoints The complex conjugate of $\langle\phi|\psi\rangle$ is $\langle\psi|\phi\rangle$. But if Q is an operator, what is the complex conjugate of the number $\langle\phi|Q|\psi\rangle$? If Q is a Hermitian operator, the answer is $\langle\psi|Q|\phi\rangle$. But if it is not Hermitian, there is an operator $Q^\dagger$, called the **Hermitian adjoint** of Q such that for any $|\phi\rangle$, $|\psi\rangle$

$$\langle\psi|Q^\dagger|\phi\rangle = (\langle\phi|Q|\psi\rangle)^*. \tag{1.29}$$

In the special case in which $|\phi\rangle$, $|\psi\rangle$ are two members of a complete set of states, this equation becomes a relation between the matrix elements of Q and $Q^\dagger$:

$$Q^\dagger_{ij} = Q^*_{ji}, \tag{1.30}$$

i.e., the matrix elements of $Q^\dagger$ are the transposed complex conjugates of those of Q. Let R be another operator. What then is the Hermitian adjoint of the product QR? To find out we substitute QR for Q in (1.30):

$$(QR)^\dagger_{ij} = (QR)^*_{ji} = \sum_k Q^*_{jk} R^*_{ki} = \sum_k R^\dagger_{ik} Q^\dagger_{kj} = (R^\dagger Q^\dagger)_{ij}. \tag{1.31}$$

Thus, to dagger a product we reverse the terms and dagger the individual operators. By induction it is now easy to show that $(ABC\ldots Z)^\dagger = Z^\dagger \ldots C^\dagger B^\dagger A^\dagger$ If we agree that the Hermitian adjoint of a complex number is its complex conjugate and that $|\psi\rangle^\dagger = \langle\psi|$ and $\langle\psi|^\dagger = |\psi\rangle$, then we can consider the basic rule (1.29) to be an application of the rule we have derived about reversing the order and daggering the components of a product.

1.3.3 Commutators The commutator of two operators A, B is defined to be

$$[A, B] \equiv AB - BA. \tag{1.32}$$

Three invaluable rules are

$$\begin{aligned} [A + B, C] &= [A, C] + [B, C] \\ AB &= BA + [A, B] \\ [AB, C] &= [A, C]B + A[B, C]. \end{aligned} \tag{1.33}$$

All three rules are trivial to prove by explicitly writing out the contents of $[,]$. With these rules it is unnecessary subsequently ever to write out the contents of a commutator – thus eliminating the commonest source of error and tedium for beginners. Notice the similarity of the third rule to the standard rule for differentiating a product:[2] $\mathrm{d}(ab)/\mathrm{d}c = (\mathrm{d}a/\mathrm{d}c)b + a(\mathrm{d}b/\mathrm{d}c)$. The rule is easily generalized by induction to the rule

$$[ABC\ldots, Z] = [A, Z]BC\ldots + A[B, Z]C\ldots + AB[C, Z]\ldots \tag{1.34}$$

1.4 The position representation

The x coordinate of a particle is associated with an operator $\hat{x}$. The set of numbers that are eigenvalues of an operator is called the **spectrum** of the operator. The spectrum of $\hat{x}$ consists of the real numbers $-\infty < x < \infty$. $|x\rangle$ is the state in which the particle is certainly at x and satisfies the eigenvalue equation $\hat{x}|x\rangle = x|x\rangle$. If the particle is in some state $|\psi\rangle$, the amplitude to find the particle at x is the complex number $\langle x|\psi\rangle$. The functional dependence of this number on x is what's called the wavefunction $\psi(x)$:

$$\psi(x) = \langle x|\psi\rangle.$$

Since the particle has to be somewhere, the sum of the probabilities $|\psi(x)|^2$ has to be unity. Since the values of x are continuous rather than discrete, this implies

$$1 = \int \mathrm{d}x\, |\psi(x)|^2 = \int \mathrm{d}x\, \langle\psi|x\rangle\langle x|\psi\rangle, \tag{1.35}$$

which should be compared with the corresponding expression (1.24) for the case of discrete eigenvalues.

The generalization of (1.26) to the case of an operator such as $\hat{x}$ that has continuous eigenvalues is

$$\int \mathrm{d}x\, |x\rangle\langle x| = I. \tag{1.36}$$

Multiplying this by $\langle x'|$ and $|\psi\rangle$ we have

$$\langle x'|\psi\rangle = \int \mathrm{d}x\, \langle x'|x\rangle\langle x|\psi\rangle. \tag{1.37}$$

[2] This similarity arises because a commutator (like a Poisson bracket in classical mechanics) is intimately associated with a 'Lie derivative' in differential geometry.

Comparing this equation with the definition of the Dirac delta function

$$\psi(x') = \int \mathrm{d}x\, \delta(x - x')\psi(x) \tag{1.38}$$

we see that

$$\langle x'|x\rangle = \delta(x - x'). \tag{1.39}$$

This result makes good sense physically because $\langle x'|x\rangle$ is the amplitude to find the particle at x' given that it is known to be at x. Clearly this vanishes for $x' \neq x$ and it has to be extremely large for $x' = x$ if it is to satisfy the normalization condition (1.35).

Pre- and post-multiplying (1.36) by $\langle\phi|$ and $|\psi\rangle$ we obtain the important formula

$$\langle\phi|\psi\rangle = \int \mathrm{d}x\, \langle\phi|x\rangle\langle x|\psi\rangle = \int \mathrm{d}x\, \phi^*(x)\psi(x). \tag{1.40}$$

Many quantum-mechanical calculations are carried out using the numbers $\psi(x)$ as the complete set of amplitudes. Such calculations are said to be conducted in the **position representation**. In this representation, the functional form $\psi()$ is to be identified with $|\psi\rangle$ and the complex number $\psi(x)$ with the amplitude $\langle x|\psi\rangle$. The operator $\hat{x}$ turns the function that evaluates to $\psi(x)$ into the one that evaluates to $x\psi(x)$; that is, it operates on functions by multiplying them by their argument.

In the position representation x-momentum, p, operates on wavefunctions by differentiation:

$$\langle x|\hat{p}|\psi\rangle = (\hat{p}\psi)(x) = -\mathrm{i}\hbar\frac{\partial\psi}{\partial x} \tag{1.41}$$

We find the wavefunction $u_p(x) \equiv \langle x|p\rangle$ of a state of well defined momentum from the defining equation of $|p\rangle$

$$\hat{p}|p\rangle = p|p\rangle \quad \rightarrow \quad \langle x|\hat{p}|p\rangle = -i\hbar\frac{\partial u_p}{\partial x} = p\langle x|p\rangle = pu_p(x) \tag{1.42}$$

The solution of this differential equation is

$$u_p(x) = A\mathrm{e}^{\mathrm{i}px/\hbar}. \tag{1.43}$$

The constant A may be determined as follows. By analogy with (1.39) we require $\langle p'|p\rangle = \delta(p - p')$. When we use (1.36) to insert an identity operator into this expression, it becomes

$$\delta(p - p') = \int \mathrm{d}x\, \langle p'|x\rangle\langle x|p\rangle = |A|^2 \int \mathrm{d}x\, \mathrm{e}^{\mathrm{i}(p-p')x/\hbar} = 2\pi\hbar|A|^2\delta(p - p'), \tag{1.44}$$

where we have taken the value of the integral from the theory of Fourier transforms. Thus $|A|^2 = h^{-1}$ and the correctly normalized wavefunction of a particle of momentum p is

$$u_p(x) = \frac{1}{\sqrt{h}}\mathrm{e}^{\mathrm{i}px/\hbar}. \tag{1.45}$$

In the position representation the Hamiltonian for a particle acts by a combination of differentiation and multiplication:

$$\langle x|H|\psi\rangle = (H\psi)(x) = \left(\frac{p^2}{2m} + V\right)\psi = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi(x) \tag{1.46}$$

1.5 The energy representation

$|E_n\rangle$ is the state in which a measurement of H certainly returns the value E_n. Its defining equation is

$$H|E_n\rangle = E_n|E_n\rangle, \tag{1.47}$$

which is called the **time-independent Schrödinger equation (TISE)**. H has a special status in non-relativistic QM because it isn't just something you can measure; it governs the time evolution of the system through the **time-dependent Schrödinger equation (TDSE)**:

$$\mathrm{i}\hbar\frac{\partial|\psi\rangle}{\partial t} = H|\psi\rangle. \tag{1.48}$$

Since physics is all about predicting what will happen, this is the fundamental equation of non-relativistic QM. The ket of *any* system must satisfy this equation. By contrast the TISE (1.47) is satisfied only for states of well-defined energy. These states evolve in time in an exceptionally simple way

$$\mathrm{i}\hbar\frac{\partial|E_n\rangle}{\partial t} = H|E_n\rangle = E_n|E_n\rangle \quad \Rightarrow \quad |E_n,t\rangle = |E_n,0\rangle\mathrm{e}^{-\mathrm{i}E_nt/\hbar} \tag{1.49}$$

That is, the passage of time simply changes the phase of the ket at a rate $E_n/\hbar$. Let's use this result to calculate the time evolution of an arbitrary state $|\psi\rangle$. In the energy representation the state is

$$|\psi,t\rangle = \sum_n a_n(t)|E_n,t\rangle \tag{1.50}$$

Substituting this expansion into the TDSE (1.48) we find

$$\mathrm{i}\hbar\frac{\partial|\psi\rangle}{\partial t} = \sum_n \mathrm{i}\hbar\left(\dot{a}_n|E_n\rangle + a_n\frac{\partial|E_n\rangle}{\partial t}\right) = \sum_n a_nH|E_n\rangle. \tag{1.51}$$

The right side cancels with the second term in the middle, so we have $\dot{a}_n = 0$. Since the a_n are constant, on eliminating $|E_n,t\rangle$ between equations (1.49) and (1.50), we find that the evolution of $|\psi\rangle$ is simply given by

$$|\psi,t\rangle = \sum_n a_n\mathrm{e}^{-\mathrm{i}E_nt/\hbar}|E_n,0\rangle. \tag{1.52}$$

States of well-defined energy are unphysical and never occur in Nature because they are incapable of changing in any way, and hence it is impossible to get a system into such a state. But they play an extremely important role in QM because they provide the almost trivial solution (1.52) to the governing equation (1.48).

1.6 Conserved quantities

Let Q be an operator that commutes with the Hamiltonian: $[Q,H] = 0$. Then we may show that for any state $|\psi\rangle$ the expectation value of Q is constant in time, or a 'conserved quantity':

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\langle\psi|Q|\psi\rangle = -\langle\psi|HQ|\psi\rangle + \langle\psi|QH|\psi\rangle = \langle\psi|[Q,H]|\psi\rangle = 0, \tag{1.53}$$

where we have used both the TDSE (1.48) and its Hermitian adjoint. Moreover Q^2 also commutes with H so $\langle\psi|(\Delta Q)^2|\psi\rangle = \langle\psi|Q^2|\psi\rangle - (\langle\psi|Q|\psi\rangle)^2$ is also constant. If initially ψ is a state of well-defined Q, i.e., $|\psi\rangle = |q_i\rangle$ for some i, then $\langle\psi|(\Delta Q)^2|\psi\rangle = 0$ at all times. Hence, if $[Q,H] = 0$, a state of well defined Q evolves into another such state, so the value of Q can be known precisely at all times. The value q_i is then said to be a **good quantum number**.

2 Harmonic oscillators

Consideration of the mechanics of a harmonic oscillator provides a particularly convenient example of how the energy representation works in practice. Harmonic oscillators are, moreover, of enormous importance for physics because most of condensed-matter physics and quantum electrodynamics centre on weakly perturbed harmonic oscillators and use the results we will derive. The Hamiltonian is

$$H = \frac{1}{2m}\{p^2 + (m\omega x)^2\}. \tag{2.1}$$

We introduce the dimensionless operator

$$A \equiv \frac{m\omega x + \mathrm{i}p}{\sqrt{2m\hbar\omega}} \tag{2.2a}$$

This operator isn't Hermitian. Bearing in mind that x and p *are* Hermitian, we see that its adjoint is

$$A^\dagger = \frac{m\omega x - \mathrm{i}p}{\sqrt{2m\hbar\omega}} \tag{2.2b}$$

The commutator of A and $A^\dagger$ is

$$\begin{aligned}[A^\dagger, A] &= \frac{1}{2m\hbar\omega}[m\omega x - \mathrm{i}p, m\omega x + \mathrm{i}p] \\ &= \frac{1}{2m\hbar\omega}\{m^2\omega^2[x,x] + \mathrm{i}m\omega\big([x,p] - [p,x]\big) + [p,p]\} \\ &= -1\end{aligned} \tag{2.3}$$

The product $A^\dagger A$ is

$$A^\dagger A = \frac{1}{2m\hbar\omega}\{(m\omega x)^2 + 2m\omega[x,p] + p^2\} = \frac{H}{\hbar\omega} - \tfrac{1}{2}. \tag{2.4}$$

We now multiply both sides of the defining relation of $|E_n\rangle$ by $A^\dagger$:

$$\begin{aligned}E_n A^\dagger|E_n\rangle &= A^\dagger H|E_n\rangle = \big(HA^\dagger + [A^\dagger, H]\big)|E_n\rangle \\ &= \big(HA^\dagger + \hbar\omega[A^\dagger, A^\dagger A + \tfrac{1}{2}]\big)|E_n\rangle = \big(HA^\dagger + \hbar\omega A^\dagger[A^\dagger, A]\big)|E_n\rangle \\ &= \big(HA^\dagger - \hbar\omega A^\dagger)\big)|E_n\rangle\end{aligned} \tag{2.5}$$

Hence

$$(E_n + \hbar\omega)\big(A^\dagger|E_n\rangle\big) = H\big(A^\dagger|E_n\rangle\big). \tag{2.6}$$

Provided $|b\rangle \equiv A^\dagger|E_n\rangle$ has non-zero length-squared this shows that $|b\rangle$ is an eigenket of H with eigenvalue $E_n + \hbar\omega$. The length-square of $|b\rangle$ is

$$\big|A^\dagger|E_n\rangle\big|^2 = \langle E_n|AA^\dagger|E_n\rangle = \langle E_n|(H + \tfrac{1}{2}\hbar\omega)|E_n\rangle = E_n + \tfrac{1}{2}\hbar\omega \tag{2.7}$$

Now squeezing H between $\langle E_n|$ and $|E_n\rangle$ we find with (2.1) that

$$E_n = \langle E_n|H|E_n\rangle = \frac{1}{2m\omega}\big(\big|p|E_n\rangle\big|^2 + m^2\omega^2\big|x|E_n\rangle\big|^2\big) \geq 0 \tag{2.8}$$

Thus the energy eigenvalues are non-negative, so $\big|A^\dagger|E_n\rangle\big|^2 > 0$ and by repeated application of $A^\dagger$ we can construct an infinite series of eigenstates with energy $E_n + k\hbar\omega$ for $k = 0, 1, \ldots$.

Similarly, we can show that provided $A|E_n\rangle$ has non-zero length-squared, it is an eigenket of H for energy $E_n - \hbar\omega$. Since we know that all eigenvalues are non-negative, for some energy E_0, $A|E_0\rangle$ must vanish. Equating to zero the length-squared of this vector we obtain an equation for E_0:

$$0 = \big|A|E_0\rangle\big|^2 = \langle E_0|(H - \tfrac{1}{2}\hbar\omega)|E_0\rangle = E_0 - \tfrac{1}{2}\hbar\omega. \tag{2.9}$$

This completes the proof that the eigenvalues of H are $\hbar\omega \times (\frac{1}{2}, \frac{3}{2}, \ldots, \frac{2r+1}{2})$.

Let $|r\rangle$ denote the state of energy $\frac{2r+1}{2}\hbar\omega$. In this notation the ground state is $|0\rangle$ and we obtain its wavefunction by writing its defining equation, $A|0\rangle = 0$, in the position representation. From (1.41) and (2.2) this equation reads

$$\frac{1}{\sqrt{2m\hbar\omega}}\left(m\omega x + \hbar\frac{\partial}{\partial x}\right)\langle x|0\rangle = 0. \tag{2.10}$$

This is a linear, first-order differential equation. Its integrating factor is $\exp(m\omega x^2/2\hbar)$, so the correctly normalized wavefunction is

$$\langle x|0\rangle = \frac{1}{(a^2\pi)^{1/4}}\,\mathrm{e}^{-x^2/2a^2}, \quad \text{where} \quad a = \sqrt{\frac{\hbar}{m\omega}}. \tag{2.11}$$

Notice that this solution is unique, so the ground state is non-degenerate. We obtain the wavefunctions of excited states by applying powers of the differential operator $A^\dagger$ to $\langle x|0\rangle$. Equation (2.7) enables us to find the normalization constant α in the equation $|n+1\rangle = \alpha A^\dagger|n\rangle$; it implies that $\alpha^2 = n+1$. the generalization of equation (2.9) enables us to determine β in $|n-1\rangle = \beta A|n\rangle$, and we have finally

$$|n+1\rangle = \frac{1}{\sqrt{n+1}}A^\dagger|n\rangle \quad ; \quad |n-1\rangle = \frac{1}{\sqrt{n}}A|n\rangle. \tag{2.12}$$

The normalization constant is always the square root of the largest value of n appearing in the equation. As a specific example

$$\langle x|1\rangle = \frac{1}{\sqrt{2m\hbar\omega}}\left(m\omega x - \hbar\frac{\partial}{\partial x}\right)\langle x|0\rangle = \frac{\sqrt{2}}{(a^2\pi)^{1/4}}\frac{x}{a}\,\mathrm{e}^{-x^2/2a^2}. \tag{2.13}$$

Notice that the product $A^\dagger A$ is Hermitian and (2.12) implies that $A^\dagger A|n\rangle = n|n\rangle$, so its eigenvalue tells you the number of excitations the oscillator has. Hence $N \equiv A^\dagger A$ is called the **number operator**.

Exercise (2):
Calculate $\langle x|2\rangle$.

Let's use these results to find the mean-square displacement $\langle n|x|n\rangle$ when the oscillator's in its nth excited state and check that for large n our result agrees with classical physics as the **correspondence principle** requires. Adding equations (2.2) we express x as a linear combination of A and $A^\dagger$

$$x = \sqrt{\frac{\hbar}{2m\omega}}(A + A^\dagger), \tag{2.14}$$

so

$$\langle n|x^2|n\rangle = \frac{\hbar}{2m\omega}\langle n|(A + A^\dagger)^2|n\rangle. \tag{2.15}$$

When we multiply out the bracket on the right we need only retain the terms that involve equal numbers of As and $A^\dagger$s. Thus

$$\langle n|x^2|n\rangle = \frac{\hbar}{2m\omega}\langle n|(AA^\dagger + A^\dagger A)|n\rangle = \frac{\hbar}{2m\omega}(2n+1) = \frac{E_n}{m\omega^2}, \tag{2.16}$$

where we have used (2.3). Classically, the average value of x^2 is proportional to the average potential energy, which is just half the total energy. Hence, averaging the Hamiltonian (2.1) we conclude that classically $\overline{x^2} = E/(m\omega^2)$, which agrees with the QM result for *any* value of n, not just large n as the correspondence principle requires.

Exercise (3):
Use (2.14) to write out the first 16 entries of the matrix x_{mn}. Similarly calculate part of the matrix p_{mn}

Exercise (4):
Show that for any two $N \times N$ matrices A, B, $\mathrm{trace}[A, B] = 0$. Comment on this result in the light of the results of the last exercise and the **canonical commutation relation** $[x, p] = \mathrm{i}\hbar$.

2.1 Motion in a magnetic field

The formalism we've just developed for a harmonic oscillator enables us to solve an important, and you might have thought unconnected, problem: the motion of a particle of mass m and charge Q in a uniform magnetic field of flux density B. The Hamiltonian is

$$H = \frac{1}{2m}(\mathbf{p} - Q\mathbf{A})^2, \tag{2.17}$$

where $\mathbf{A}$ is the vector potential. Let the z-axis be parallel to $\mathbf{B}$ and choose the gauge in which $\mathbf{A} = \frac{1}{2}B(-y, x, 0)$. Then H is

$$\begin{aligned} H &= \frac{1}{2m}\Big\{\big(p_x + \tfrac{1}{2}QBy\big)^2 + \big(p_y - \tfrac{1}{2}QBx\big)^2 + p_z^2\Big\} \\ &= \tfrac{1}{2}\hbar\omega(\pi_x^2 + \pi_y^2) + \frac{p_z^2}{2m}, \end{aligned} \tag{2.18a}$$

where $\omega = QB/m$ is the Larmor frequency and we have defined the dimensionless operators

$$\pi_x \equiv \frac{p_x + \frac{1}{2}QBy}{\sqrt{QB\hbar}} \quad ; \quad \pi_y \equiv \frac{p_y - \frac{1}{2}QBx}{\sqrt{QB\hbar}}. \tag{2.18b}$$

H has broken into two parts. The $p_z^2/2m$ is just the Hamiltonian of a free particle in one dimension and isn't interesting. The part H_{xy} involving the π_i is essentially the Hamiltonian of a harmonic oscillator because

$$[\pi_x, \pi_y] = \frac{1}{2\hbar}\big([y, p_y] - [p_x, x]\big) = \mathrm{i}. \tag{2.19}$$

The ladder operators are

$$\left.\begin{aligned} A &= \frac{1}{\sqrt{2}}(\pi_x + \mathrm{i}\pi_y) \\ A^\dagger &= \frac{1}{\sqrt{2}}(\pi_x - \mathrm{i}\pi_y) \end{aligned}\right\} \quad \Rightarrow \quad [A, A^\dagger] = \mathrm{i}[\pi_y, \pi_x] = 1, \tag{2.20}$$

and in terms of them H_{xy} is

$$H_{xy} = \hbar\omega(A^\dagger A + \tfrac{1}{2}) \tag{2.21}$$

It follows that the energy levels are $E = \hbar\omega(\frac{1}{2}, \frac{3}{2}, \ldots)$. These discrete energy levels for a charged particle in a uniform magnetic field are known as **Landau levels**.

To find the associated wavefunctions, we write the ground state's defining equation in the position representation

$$A|0\rangle = 0 \quad \leftrightarrow \quad \left\{\hbar\Big(\frac{\partial}{\partial x} + \mathrm{i}\frac{\partial}{\partial y}\Big) + \tfrac{1}{2}QB(x + \mathrm{i}y)\right\}\langle x|0\rangle = 0. \tag{2.22}$$

We transform to new coordinates $u \equiv x + \mathrm{i}y$, $v \equiv x - \mathrm{i}y$ and find that (2.22) becomes

$$2\hbar\frac{\partial\langle x|0\rangle}{\partial v} + \tfrac{1}{2}QBu\langle x|0\rangle = 0. \tag{2.23}$$

Solving this first-order linear o.d.e. we find

$$\langle x|0\rangle = g(u)\,\mathrm{e}^{-(QB/4\hbar)uv} \tag{2.24}$$

where $g(u)$ is an arbitrary function. Choosing $g = 1$ the wavefunction of the ground state becomes[3]

$$\langle x|0\rangle = \mathrm{e}^{-(QB/4\hbar)(x^2+y^2)}. \tag{2.25}$$

A particle in this state is localized near the origin. Since the underlying physics [unlike the Hamiltonian (2.18)] is invariant under displacements within the xy plane, there must be other ground-state kets that are localized elsewhere. You can investigate this in the first problem set.

[3] Aficiondos of functions of a complex variable may be nervous about the derivation of (2.25). They should reassure themselves by checking that it *really* does solve (2.22).

3 Angular momentum

If you rotate the axes you are using, the components that describe a fixed physical vector have to change. In fact, if the axis around which we rotate our axes by an infinitesimal angle $\delta\phi$ is parallel to the unit vector $\mathbf{n}$, then the components of any fixed vector $\mathbf{B}$ have to change such that

$$\mathbf{B} \to \mathbf{B}' = \mathbf{B} - \delta\phi(\mathbf{n} \times \mathbf{B}). \tag{3.1}$$

$\mathbf{B}'$ is proportional to $\mathbf{B}$ so we must be able to express $\mathbf{B}'$ as a matrix $\mathbf{R}(\mathbf{n}, \phi)$ times $\mathbf{B}$. To find $\mathbf{R}$ we write out (3.1) in components

$$\mathbf{B}' = \mathbf{R}\mathbf{B} = \mathbf{B} - \delta\phi \begin{pmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{pmatrix} \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix}. \tag{3.2}$$

Next we make the linear dependence of $\mathbf{B}' - \mathbf{B}$ upon $\mathbf{n}$ manifest by rewriting (3.2) as

$$\mathbf{R}\mathbf{B} = \{\mathbf{I} + \frac{\mathrm{i}\delta\phi}{\hbar}(n_x\mathbf{J}_x + n_y\mathbf{J}_y + n_z\mathbf{J}_z)\}\mathbf{B} \tag{3.3a}$$

where we have defined three Hermitian matrices $\mathbf{J}_i$ by

$$J_x \equiv \hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix} \quad ; \quad J_y \equiv \hbar \begin{pmatrix} 0 & 0 & \mathrm{i} \\ 0 & 0 & 0 \\ -\mathrm{i} & 0 & 0 \end{pmatrix} \quad ; \quad J_z \equiv \hbar \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.3b}$$

It is straightforward to show that these matrices satisfy the commutation relations

$$[\mathbf{J}_i, \mathbf{J}_j] = \mathrm{i}\hbar \sum_k \epsilon_{ijk}\mathbf{J}_k, \tag{3.4}$$

where ϵ_{ijk} is the object that changes sign if any two subscripts are interchanged and has $\epsilon_{xyz} = 1$. The physical content of these commutation relations is that when we rotate our axes by $\delta\phi_1$ about the x axis and then rotate by $\delta\phi_2$ about the y axis the resulting vector $\mathbf{B}'$ differs from the result $\mathbf{B}''$ obtained by performing the same rotations in the opposite order by an amount

$$\begin{aligned} \mathbf{B}' - \mathbf{B}'' &= \left\{ \left(\mathbf{I} + \frac{\mathrm{i}\delta\phi_2}{\hbar}\mathbf{J}_y\right)\left(\mathbf{I} + \frac{\mathrm{i}\delta\phi_1}{\hbar}\mathbf{J}_x\right) - \left(\mathbf{I} + \frac{\mathrm{i}\delta\phi_1}{\hbar}\mathbf{J}_x\right)\left(\mathbf{I} + \frac{\mathrm{i}\delta\phi_2}{\hbar}\mathbf{J}_y\right) \right\} \mathbf{B} \\ &= -\frac{\delta\phi_1\delta\phi_2}{\hbar^2}[\mathbf{J}_y, \mathbf{J}_x]\mathbf{B} = \frac{\mathrm{i}\delta\phi_1\delta\phi_2}{\hbar}\mathbf{J}_z\mathbf{B} \end{aligned} \tag{3.5}$$

which is the change in $\mathbf{B}$ induced by a rotation through an angle $\delta\phi_1\delta\phi_2$ about the z axis.

Suppose now that we rotate our *system* by angle $\delta\phi$ about $\mathbf{n}$ while continuing to use the same axes. This will cause the components of vectors that rotate with the system to change by the same amount as if we had kept our system fixed and rotated our axes by $-\delta\phi$ about $\mathbf{n}$. Thus the matrices $\mathbf{J}_i$ describe how vectors fixed to a system change when the system is rotated.

The discussion above is purely classical in nature, although cancelling factors of $\hbar$ have crept in with the definition of the $\mathbf{J}_i$. Turning to QM, we must expect the amplitudes of a complete set to change when we rotate the system because the state of the system is changed by the rotation. Rotation through a small angle $\delta\phi$ will leave states almost unchanged, so for infinitesimal angles the new ket will be

$$|\psi'\rangle = R(\mathbf{n}, \delta\phi)|\psi\rangle = \left(I + \frac{\mathrm{i}\delta\phi}{\hbar}\mathbf{n}.\mathbf{S}\right)|\psi\rangle, \tag{3.6}$$

where the S_i are three operators to be determined. Since the difference between rotating the system by $\delta\phi_1$ about x and then $\delta\phi_2$ about y must be the same as the change induced by a rotation by $\delta\phi_1\delta\phi_2$ about z, these operators, like the $\mathbf{J}_i$, must satisfy

$$[S_i, S_j] = \mathrm{i}\hbar \sum_k \epsilon_{ijk}S_k. \tag{3.7}$$

Since the normalization of $|\psi\rangle$ should be unaffected by rotation, the operators must be Hermitian:

$$\begin{aligned} 1 = \langle\psi'|\psi'\rangle &= \langle\psi|\left(I - \frac{\mathrm{i}\delta\phi}{\hbar}\mathbf{n}.\mathbf{S}^\dagger\right)\left(I + \frac{\mathrm{i}\delta\phi}{\hbar}\mathbf{n}.\mathbf{S}\right)|\psi\rangle \\ &\simeq 1 + \frac{\mathrm{i}\delta\phi}{\hbar}(\mathbf{n}.\mathbf{S} - \mathbf{n}.\mathbf{S}^\dagger)|\psi\rangle. \end{aligned} \tag{3.8}$$

Since they are Hermitian, it is likely that they are associated with observable quantities, and we must ask 'what are the possible values that these observables can take' – or in mathematical language, what are the eigenvalues of the S_i. Remarkably, we can establish what these eigenvalues are using only the commutation relations (3.8).

We first show that the operator $S^2 \equiv S_x^2 + S_y^2 + S_z^2$ commutes with all three S_k:

$$\begin{aligned} [S_z, S^2] &= [S_z, S_x^2] + [S_z, S_y^2] = [S_z, S_x]S_x + S_x[S_z, S_x] + [S_z, S_y]S_y + S_y[S_z, S_y] \\ &= \mathrm{i}\hbar(S_yS_x + S_xS_y) - \mathrm{i}\hbar(S_xS_y + S_yS_x) = 0. \end{aligned} \tag{3.9}$$

Hence there is a complete set of mutual eigenkets of S^2 and, say, S_z. Let $|\beta, m\rangle$ be a member of this set, the labels being such that

$$S^2|\beta, m\rangle = \beta\hbar^2|\beta, m\rangle \quad ; \quad S_z|\beta, m\rangle = m\hbar|\beta, m\rangle. \tag{3.10}$$

Notice that no claim is being made here: the eigenvalues contain explicit factors of $\hbar$ to ensure that the labels β and m are dimensionless. We define

$$S_\pm = S_x \pm \mathrm{i}S_y. \tag{3.11}$$

These objects clearly commute with S^2, while their commutation relation with S_z are

$$[S_\pm, S_z] = [S_x, S_z] \pm \mathrm{i}[S_y, S_z] = -\mathrm{i}\hbar S_y \mp \hbar S_x = \mp\hbar S_\pm. \tag{3.12}$$

Since $S_\pm$ commutes with S^2, the ket $S_\pm|\beta, m\rangle$ is an eigenket of S^2 with eigenvalue $\beta\hbar^2$. Operating with S_z on this ket we find

$$S_zS_\pm|\beta, m\rangle = (S_\pm S_z + [S_z, S_\pm])\,|\beta, m\rangle = (m \pm 1)\hbar S_\pm|\beta, m\rangle. \tag{3.13}$$

Thus, $S_z|\beta, m\rangle$ is also a member of the complete set of states $\{|\beta, m\rangle\}$ but its eigenvalue with respect to S_z differs from that of $|\beta, m\rangle$ by $\pm\hbar$. Therefore, we may write

$$S_\pm|\beta, m\rangle = \alpha_\pm\hbar|\beta, m \pm 1\rangle. \tag{3.14}$$

We determine the dimensionless constant $\alpha_\pm$ by taking the length-squared of both sides of (3.14). Bearing in mind that $S_\pm^\dagger = S_\mp$ we find

$$\begin{aligned} |\alpha_\pm^2|\hbar^2 &= \langle\beta, m|S_\mp S_\pm|\beta, m\rangle = \langle\beta, m|(S_x \mp \mathrm{i}S_y)(S_x \pm \mathrm{i}S_y)|\beta, m\rangle \\ &= \langle\beta, m|(S^2 - S_z^2 \mp \hbar S_z)|\beta, m\rangle = \hbar^2\{\beta - m(m \pm 1)\}, \end{aligned} \tag{3.15}$$

so

$$\alpha_\pm = \sqrt{\beta - m(m \pm 1)}. \tag{3.16}$$

From the definition of S^2 and the fact that the S_i are Hermitian operators so $\langle\psi|S_i^2|\psi\rangle = \left|S_i|\psi\rangle\right|^2 \geq 0$, we infer that $m^2 \leq \beta$. So notwithstanding (3.14) it cannot be possible to create states with ever larger eigenvalues of S_z by repeated application of S_+. All that can stop us doing this is the vanishing of α_+ when we reach some maximum eigenvalue $m_{\max}$ that satisfies

$$\beta - m_{\max}(m_{\max} + 1) = 0. \tag{3.17}$$

Similarly, α_- must vanish for a smallest value of m that satisfies

$$\beta - m_{\min}(m_{\min} - 1) = 0. \tag{3.18}$$

Eliminating β between (3.17) and (3.18) we obtain a relation between $m_{\max}$ and $m_{\min}$ that we can treat as a quadratic equation for $m_{\min}$. Solving this equation we find that

$$m_{\min} = \tfrac{1}{2}\{1 \pm (2m_{\max} + 1)\}. \tag{3.19}$$

The plus sign yields a value of $m_{\min}$ that is incompatible with our requirement that $m_{\min} \leq m_{\max}$, so we must have $m_{\min} = -m_{\max}$. To simplify the notation, we define $s \equiv m_{\max}$, so that (3.17) becomes $\beta = s(s+1)$ and $-s \leq m \leq s$. Finally, we note that since an integer number of applications of S_- will take us from $|\beta, s\rangle$ to $|\beta, -s\rangle$, $2s$ must be an integer. In summary, the eigenvalues of S^2 are $s(s+1)\hbar^2$ with $2s = 0, 1, 2, \ldots$ and for each value of s the eigenvalues $m\hbar$ of S_z are $\hbar(s, s-1, \ldots, -s)$. Henceforth we simplify the labelling of kets by defining $|s, m\rangle$ to be what has hitherto been denoted $|\beta, m\rangle$ with $\beta = s(s+1)$.

3.1 Spin-half

Consider a system for which $s = \frac{1}{2}$. The possible eigenvalues of S_z are then $\pm\frac{1}{2}\hbar$. Let $|+\rangle \equiv |\frac{1}{2}, +\frac{1}{2}\rangle$ and $|-\rangle \equiv |\frac{1}{2}, -\frac{1}{2}\rangle$. Then any state for which $s = \frac{1}{2}$ can be written as a linear combination $|\psi\rangle = a|+\rangle + b|-\rangle$ and the operators S_i are associated with the 2×2 matrices [cf (1.18)]

$$\begin{matrix} S_x = \begin{pmatrix} \langle +|S_x|+\rangle & \langle +|S_x|-\rangle \\ \langle -|S_x|+\rangle & \langle -|S_x|-\rangle \end{pmatrix} \\ S_y = \begin{pmatrix} \langle +|S_y|+\rangle & \langle +|S_y|-\rangle \\ \langle -|S_y|+\rangle & \langle -|S_y|-\rangle \end{pmatrix} \end{matrix} \quad ; \quad S_z = \begin{pmatrix} \langle +|S_z|+\rangle & \langle +|S_z|-\rangle \\ \langle -|S_z|+\rangle & \langle -|S_z|-\rangle \end{pmatrix}. \tag{3.20}$$

The elements of the matrix S_z are trivially evaluated because $|\pm\rangle$ are the eigenkets of S_z with eigenvalues $\pm\frac{1}{2}\hbar$. To evaluate the other two matrices we replace S_x by $\frac{1}{2}(S_+ + S_-)$, S_y by $\frac{1}{2\mathrm{i}}(S_+ - S_-)$ and use the relations $S_+|-\rangle = \hbar|+\rangle$ and $S_-|+\rangle = \hbar|-\rangle$ which follow from (3.14) and (3.16). The result of these operations is

$$\begin{matrix} S_x = \frac{1}{2}\hbar\sigma_x = \frac{1}{2}\hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ S_y = \frac{1}{2}\hbar\sigma_y = \frac{1}{2}\hbar \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \end{matrix} \quad ; \quad S_z = \frac{1}{2}\hbar\sigma_z = \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.21}$$

where the **Pauli matrices** σ_i have been defined.

From equation (3.6) we can derive a differential equation for the evolution of $|\psi\rangle$ when the system that it describes is rotated around the direction $\mathbf{n}$:

$$-\mathrm{i}\hbar\frac{\partial}{\partial\phi}|\psi\rangle = \mathbf{n}.\mathbf{S}|\psi\rangle. \tag{3.22}$$

It is interesting to solve this equation for the case of a spin-half system. We orient our z axis parallel to $\mathbf{n}$ and replace $\mathbf{n}.\mathbf{S}$ by $\frac{1}{2}\hbar\sigma_z$ to obtain

$$\frac{\partial}{\partial\phi}|\psi\rangle = \tfrac{\mathrm{i}}{2}\sigma_z|\psi\rangle. \tag{3.23}$$

It is easy to verify that this equation is solved by

$$|\psi\rangle = \exp\left(\frac{\mathrm{i}\phi}{2}\sigma_z\right)|\psi 0\rangle, \tag{3.24a}$$

where

$$\exp\left(\frac{\mathrm{i}\phi}{2}\sigma_z\right) \equiv I + \left(\frac{\mathrm{i}\phi}{2}\sigma_z\right) + \tfrac{1}{2!}\left(\frac{\mathrm{i}\phi}{2}\sigma_z\right)^2 + \cdots \tag{3.24b}$$

and $|\psi 0\rangle$ is the value of $|\psi\rangle$ for $\phi = 0$. Now $\sigma_i^2 = I$, the identity matrix, so

$$\begin{aligned}\exp\left(\frac{\mathrm{i}\phi}{2}\sigma_z\right) &= I\left\{1 - \tfrac{1}{2!}\left(\frac{\phi}{2}\right)^2 + \cdots\right\} + \mathrm{i}\sigma_z\left\{\left(\frac{\phi}{2}\right) - \tfrac{1}{3!}\left(\frac{\phi}{2}\right)^3 + \cdots\right\} \\ &= \cos(\phi/2)I + \mathrm{i}\sin(\phi/2)\sigma_z \\ &= \begin{pmatrix} \mathrm{e}^{\mathrm{i}\phi/2} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\phi/2}\end{pmatrix}.\end{aligned} \tag{3.25}$$

For $\phi = 2\pi$ this becomes $-I$ rather than I, so when the system is rotated through 2π, $|\psi\rangle$ does not return to its original value but minus it. In other words, a spin-half system is not restored to its original state by rotation through 2π. This is a surprising result and contrasts with the case of integer-spin systems, which (as we will see) *are* returned to their original state by rotation through 2π. Does the unexpected behaviour of spin-half kets under rotation imply that spin-half systems don't exist? In fact, electrons and quarks are all spin-half objects, so spin-half objects are extremely common! We find their behaviour under rotations disconcerting because the fields associated with spin-half objects never attain macroscopic values: the quantum uncertainty in the value of a spin-half field is always on the same order as the value of the field itself. This contrasts with the case of integer-spin fields, such as those associated with electromagnetism and gravity, which can attain macroscopic values: values that are vastly greater than their quantum uncertainties. Our intuition about what happens when a system is rotated has grown out of our experience of fields with macroscopic values, so we consider that things return to their original state after rotation by 2π. If we had hands-on experience of spin-half objects, we would recognize that this is not generally true.

3.2 Spin one

In the case that $s = 1$, three values of m are possible, $-1, 0, 1$, so the S_i are represented by 3×3 matrices. The calculation of these matrices proceeds exactly as for spin-half, the main difference being that (3.14) and (3.16) now yield

$$S_+|-1\rangle = \sqrt{2}|0\rangle\ ;\ S_+|0\rangle = \sqrt{2}|1\rangle\ ;\ S_-|+\rangle = \sqrt{2}|0\rangle\ ;\ S_-|0\rangle = \sqrt{2}|-1\rangle \tag{3.26}$$

The result is

$$\begin{aligned} S_x &= \frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&1&0\\1&0&1\\0&1&0\end{pmatrix} \\ S_y &= \frac{\hbar}{\sqrt{2}}\begin{pmatrix}0&-\mathrm{i}&0\\\mathrm{i}&0&-\mathrm{i}\\0&\mathrm{i}&0\end{pmatrix}\end{aligned} \quad ; \quad S_z = \hbar\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}. \tag{3.27}$$

Exercise (5):
Show that $S_x = \mathbf{B}\mathbf{J}_x\mathbf{B}^\dagger$, where $\mathbf{J}_x$ is the matrix defined by (3.4) and

$$\mathbf{B} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{\mathrm{i}}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ -\frac{1}{\sqrt{2}} & \frac{\mathrm{i}}{\sqrt{2}} & 0 \end{pmatrix}$$

Hence show that S_x would be the matrix associated with rotations of vectors about the x axis if we used new coordinates $X_1 = (x + \mathrm{i}y)/\sqrt{2}$, $X_2 = z$, $X_3 = -(x - \mathrm{i}y)/\sqrt{2}$.

One may readily verify that the matrices (3.27) satisfy $S_i^3 = \hbar^2 S_i$ and thus from (3.22) show that the effect on $|\psi\rangle$ of rotating the system about the ith axis is to multiply it by

$$\begin{aligned} D(\phi) \equiv \exp\left(\frac{\mathrm{i}\phi S_i}{\hbar}\right) &= I + \mathrm{i}\phi\frac{S_i}{\hbar} + \tfrac{1}{2}(\mathrm{i}\phi)^2\left(\frac{S_i}{\hbar}\right)^2 + \tfrac{1}{3!}(\mathrm{i}\phi)^3\frac{S_i}{\hbar} + \cdots \\ &= I + \mathrm{i}\sin\phi\frac{S_i}{\hbar} + (\cos\phi - 1)\left(\frac{S_i}{\hbar}\right)^2. \end{aligned} \tag{3.28}$$

Since there is nothing inherently special about the coordinate directions, this equation must remain valid if S_i is replaced by the spin matrix associated with an arbitrary direction, $\mathbf{n}.\mathbf{S}$. In contrast to the rule (3.25) for rotating a spin-half system, (3.28) implies that $|\psi\rangle$ returns to its original value on rotation through 2π.

3.3 Orbital angular momentum

Suppose we decide to use the position representation to describe the quantum state of a single particle. Then the state is represented by the function of position $\psi(\mathbf{x}) \equiv \langle \mathbf{x}|\psi\rangle$. We use spherical polar coordinates and ask how $\psi(r,\theta,\phi)$ changes when we rotate the system. Rotation changes the angular dependence of the wavefunction so $\psi \to \psi'$ where $\psi'(r,\theta',\phi') = \psi(r,\theta,\phi)$, where (θ',ϕ') are the coordinates of the position into which the rotation moves the point with coordinates (θ,ϕ). Since the spherical harmonics Y_l^m form a complete set of functions on a sphere, any function ψ can be expanded in the form

$$\psi(r,\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{lm}(r)\mathrm{Y}_l^m(\theta,\phi). \tag{3.29}$$

If ψ is such that $c_{lm} \neq 0$ for only a single value of l, so that it is of the form $\psi(r,\theta,\phi) = \sum_{m=-l}^{l} c_m(r)\mathrm{Y}_l^m(\theta,\phi)$, then one may show that

$$\psi'(r,\theta',\phi') = \sum_{m=-l}^{l} c'_m(r)\mathrm{Y}_l^m(\theta',\phi'). \tag{3.30}$$

That is, the property of being expandable using only a single value of l is unchanged by rotation. Quantum states in which the total orbital angular momentum $L^2 = |\mathbf{x}\times\mathbf{p}|^2$ has a well-defined value have this property, and the corresponding value of l is such that $L^2\psi = l(l+1)\hbar^2\psi$. Moreover, the new expansion coefficients c'_m are related to the old ones c_m by an $(2l+1)\times(2l+1)$ matrix D:

$$c'_m = \sum_{n=-l}^{l} D_{mn}c_n. \tag{3.31}$$

D depends on the angles $(\vartheta_x,\vartheta_y,\vartheta_z)$ through which the system is rotated. In the case $l=1$ and $\vartheta_j = \phi\delta_{ji}$, the 3×3 matrix D is that given by (3.28). The 5×5 matrix D for the case of $l=2$ is given by precisely the same formula with S_i replaced by the value that S_i takes for $s=2$.

Exercise (6):
Write down the 5×5 matrix S_z for $s=2$.

To summarize, if we work in the position representation and use spherical harmonics to expand the angular dependence of a wavefunction, we get numbers c_{lm} that transform into one another under rotations. For any given l they transform according to a rule that can be derived from the fundamental commutation relations (3.7). Precisely half of all possible transformation rules can be encountered in this way, because only integer values of $s=l$ occur. In real systems all the rules occur because wavefunctions, like vector fields, are generically multi-component objects, and a half-integer spin rule may be required to transform the components of the wavefunction when the system is rotated.

The angular momentum of a system is determined by how its ket transforms when the system is rotated. The contribution to this transformation rule from the spatial dependence $\langle\mathbf{x}|\psi\rangle$ is called orbital angular momentum, while that associated with the multi-component character of $|\psi\rangle$ is called spin. One often reserves the letters L and S for orbital angular momentum and spin, respectively, and uses the letter J to denote angular momentum of any type.

3.4 Addition of angular momenta

Box 1: Kets for composite systems

When we put together two pieces to make a single system, we have to *multiply* their kets to form the ket of the whole system. The reason that physical addition leads to logical multiplication is that the probability of finding the whole system in some state is the *product* of the probability that the first subsystem is in the given state and the probability that the second subsystem is in its given state, and we ensure that this multiplicative relationship holds by expressing the amplitude for the whole system to be in a given state as the product of the amplitudes of the parts to be in their allotted states. What we do for a single amplitude we must do for a complete set of amplitudes, so we multiply the kets of the subsystems to form the ket of the whole.

To give a concrete example, a hydrogen atom consists of a proton and an electron. To specify the state of the atom, we have to specify the state $|p\rangle$ of the proton *and* the state $|e\rangle$ of the electron. So the state of the atom is $|a\rangle = |p\rangle|e\rangle$. If $|a'\rangle \equiv |p'\rangle|e'\rangle$ is the state that the atom will be left in after we've measured Q to have the value q, then the amplitude to measure q is

$$\langle a'|a\rangle = (\langle e'|\langle p'|)(|p\rangle|e\rangle) = \langle p'|p\rangle\langle e'|e\rangle.$$

Imagine that we have two gyros in a box and that we know that the first gyro has total angular momentum j_1, while the second gyro has total angular momentum j_2. Without loss of generality we may assume $j_1 \geq j_2$. A ket describing the state of the first gyro is of the form

$$|\psi_1\rangle = \sum_{m=-j_1}^{j_1} c_m |j_1, m\rangle, \tag{3.32a}$$

while the state of the second is

$$|\psi_2\rangle = \sum_{m=-j_2}^{j_2} d_m |j_2, m\rangle, \tag{3.32b}$$

and from the discussion in Box 1 it follows that the state of the box is

$$|\psi\rangle = |\psi_1\rangle|\psi_2\rangle. \tag{3.33}$$

The coefficients c_m and d_m are the amplitudes to find the individual gyros in particular orientations with respect to the z axis. For example, if both gyros are almost certainly aligned parallel to the z axis, we will have $|c_{j_1}| = |d_{j_2}| = 1$ and $c_{m_1} = d_{m_2} = 0$ for $m_1 \neq j_1$ and $m_2 \neq j_2$. The operators of interest are the operators J_i^2, J_{iz} and $J_{i\pm}$ of the ith gyro and the corresponding operators of the box. The operators J_z and $J_\pm$ for the box are simply sums of the corresponding operators for the gyros

$$J_z = J_{1z} + J_{2z} \quad ; \quad J_\pm = J_{1\pm} + J_{2\pm}. \tag{3.34}$$

Operators belonging to different systems always commute, so $[J_{1i}, J_{2j}] = 0$ for any values of i, j. The operator for the square of the box's angular momentum is

$$J^2 = (\mathbf{J}_1 + \mathbf{J}_2)^2 = J_1^2 + J_2^2 + 2\mathbf{J}_1.\mathbf{J}_2. \tag{3.35}$$

Now

$$J_{1+}J_{2-} = (J_{1x} + \mathrm{i}J_{1y})(J_{2x} - \mathrm{i}J_{2y}) = (J_{1x}J_{2x} + J_{1y}J_{2y}) + \mathrm{i}(J_{1y}J_{2x} - J_{1x}J_{2y}). \tag{3.36}$$

The expression for $J_{1-}J_{2+}$ can be obtained by swapping the labels 1 and 2 on the r.h.s., so[4]

$$J_{1+}J_{2-} + J_{1-}J_{2+} + 2J_{1z}J_{2z} = 2\mathbf{J}_1.\mathbf{J}_2. \tag{3.37}$$

Using this expression to eliminate $\mathbf{J}_1.\mathbf{J}_2$ from (3.35) we obtain

$$J^2 = J_1^2 + J_2^2 + J_{1+}J_{2-} + J_{1-}J_{2+} + 2J_{1z}J_{2z}. \tag{3.38}$$

[4] Recall that J_{1i} commutes with J_{2j} for all ij.

While the total angular momenta of the individual gyros are fixed, that of the box is variable because it depends on the mutual orientation of the two gyros: if the latter are parallel, the squared angular momentum in the box might be expected to have quantum number $j_1 + j_2$, while if they are antiparallel, the box's angular momentum might be expected to have quantum number $j_1 - j_2$. We shall show that this conjecture is true by explicitly calculating the values of the coefficients c_m and d_m for which the box is in an eigenstate of both J^2 and J_z. We start by examining the state $|j_1, j_1\rangle|j_2, j_2\rangle$ in which both gyros are parallel to the z axis. It is easy to see that this object is an eigenket of J_z with eigenvalue $j_1 + j_2$. We use (3.38) to show that it is also an eigenket of J^2:

$$\begin{aligned} J^2|j_1, j_1\rangle|j_2, j_2\rangle &= (J_1^2 + J_2^2 + J_{1+}J_{2-} + J_{1-}J_{2+} + 2J_{1z}J_{2z})|j_1, j_1\rangle|j_2, j_2\rangle \\ &= \hbar^2\{j_1(j_1+1) + j_2(j_2+1) + 2j_1j_2\}|j_1, j_1\rangle|j_2, j_2\rangle, \end{aligned} \tag{3.39}$$

where we have used the equation $J_{i+}|j_i, j_i\rangle = 0$, which follows from (3.16). It is straightforward to show that the expression in $\{.\}$ in (3.39) equals $j(j+1)$ with $j = j_1 + j_2$. Hence we may write

$$|j_1 + j_2, j_1 + j_2\rangle = |j_1, j_1\rangle|j_2, j_2\rangle. \tag{3.40}$$

Now that we have found one mutual eigenket for the box of J^2 and J_z we can easily find others by applying J_- to reorient the angular momentum of the box away from the z axis. Again setting $j = j_1 + j_2$ we find with equation (3.16)

$$\begin{aligned} & J_-|j, j\rangle = (J_{1-} + J_{2-})|j_1, j_1\rangle|j_2, j_2\rangle \\ \Rightarrow\ & \sqrt{j(j+1) - j(j-1)}|j, j-1\rangle = \sqrt{j_1(j_1+1) - j_1(j_1-1)}|j_1, j_1 - 1\rangle|j_2, j_2\rangle \\ & \qquad + \sqrt{j_2(j_2+1) - j_2(j_2-1)}|j_1, j_1\rangle|j_2, j_2 - 1\rangle \\ \Rightarrow\ & \sqrt{2j}\,|j, j-1\rangle = \sqrt{2j_1}\,|j_1, j_1 - 1\rangle|j_2, j_2\rangle + \sqrt{2j_2}\,|j_1, j_1\rangle|j_2, j_2 - 1\rangle. \end{aligned} \tag{3.41}$$

A further application of J_- to the left side of this equation and of $J_{1-} + J_{2-}$ to the right side would produce an expression for $|j, j-2\rangle$ and so on.

We seek an expression for $|j-1, j-1\rangle$. It is trivial to verify that any ket of the form $\alpha|j_1, m_1\rangle + \beta|j_2, m_2\rangle$ is an eigenket of J_z with eigenvalue $(m_1 + m_2)\hbar$. We require $m_1 + m_2 = j_1 + j_2 - 1$, so either $m_1 = j_1 - 1$ and $m_2 = j_2$, or $m_1 = j_1$ and $m_2 = j_2 - 1$. Equation (3.41) shows that $|j, j-1\rangle$ involves precisely these two cases, and must be orthogonal to $|j-1, j-1\rangle$ because it has a different eigenvalue with respect to J^2. So the ket we seek is the unique linear combination of the kets appearing in (3.41) that is orthogonal to the linear combination that appears there. That is,

$$|j-1, j-1\rangle = \sqrt{j_2/j}|j_1, j_1 - 1\rangle|j_2, j_2\rangle - \sqrt{j_1/j}|j_1, j_1\rangle|j_2, j_2 - 1\rangle. \tag{3.42}$$

All the kets in the sequence $|j-1, m\rangle$ $m = j-2, \ldots$ can be constructed by applying J_- to this equation rather than to (3.41) as we did above. Moreover, $|j-2, j-2\rangle$, like $|j, j-2\rangle$ and $|j-1, j-2\rangle$, will be a linear combination of $|j_1, j_1 - 2\rangle|j_2, j_2\rangle$, $|j_1, j_1 - 1\rangle|j_2, j_2 - 1\rangle$ and $|j_1, j_1\rangle|j_2, j_2 - 2\rangle$ and must be orthogonal to $|j, j-2\rangle$ and $|j-1, j-2\rangle$ for which we already have expressions. Hence we can determine which linear combination is required, and then generate the remaining kets of the series $|j-2, m\rangle$ by applying J_- to it.

Figure 1 helps to organize the results of this calculation. States of the box with well determined angular momentum are marked by dots. The radius of each semi-circle is proportional to j', where $j'(j'+1)\hbar^2$ is the eigenvalue of the kets with respect to J^2. The height of each ket above the centre of the circles is proportional to m. The left panel shows the case $j_1 = 2$, $j_2 = 1$, while the right panel is for $j_1 = 1$, $j_2 = \frac{1}{2}$.

The numbers $C(j, m; j_1, j_2, m_1, m_2) \equiv \langle j, m|j_1, m_1\rangle|j_2, m_2\rangle$ that we have been evaluating are called **Clebsch–Gordon** coefficients. They have a simple physical interpretation: $C(j, m; j_1, j_2, m_1, m_2)$ is the amplitude that on opening the box when it's in a state of well determined angular momentum we will find the first and second gyros to be oriented with $m_1\hbar$ and $m_2\hbar$ of their spins parallel to the z axis. For example, equation (3.41) implies that $C(3, 2; 2, 1, 1, 1) = \sqrt{2/3}$, so if a box that contains a

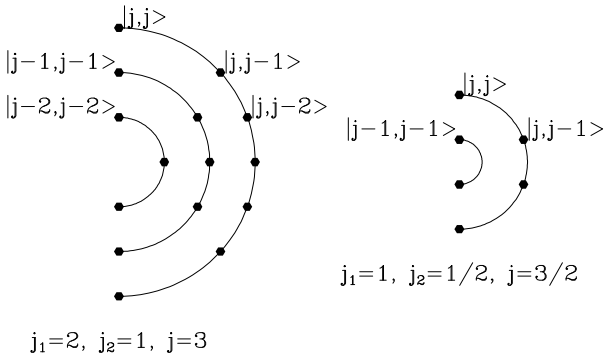

Figure 1 The left panel shows states obtained by adding a system of angular momentum $j_2 = 1$ to one with $j_1 = 2$, while the right panel is for $j_1 = 1$ and $j_2 = \frac{1}{2}$.

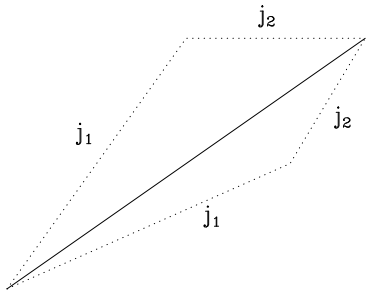

Figure 2 Interpretation of Clebsch–Gordon coefficients in terms of vectors. The full line has length $\sqrt{3(3+1)}$ and its vertical component has length 2. The dotted lines labelled j_1 have length $\sqrt{2(2+1)}$ and vertical components of length 2 and 1.

spin-two gyro and a spin-one gyro has spin-three, there is a probability 2/3 that on opening the box the second gyro will be nearly parallel to the z axis and the second significantly inclined, and only a probability 1/3 of finding the reverse arrangement that is depicted by the upper dotted lines in Fig. 2. The classical interpretation is that the two gyros precess around the fixed angular-momentum vector of the box, and that the two configurations for which the Clebsch–Gordon coefficients give amplitudes are two of the states through which the precession carries the system. This picture is intuitive and of some value, but should not be taken too seriously. For one thing, the rules for adding angular momentum are independent of any statement about the Hamiltonian, and therefore carry no implication about the time evolution of the system. The gyros may or may not precess, depending on whether they are dynamically coupled.

4 Perturbation theory

It is rarely possible to solve exactly for the dynamics of a system of experimental interest. In these circumstances we use some kind of approximation to tweak the solution to some model system that is as close as possible to the system of interest and yet is simple enough to have analytically solvable dynamics. That is we treat the difference between the experimental system and the model system as a 'perturbation' of the model. Perturbation theory in this sense was an important part of mathematical physics before quantum mechanics appeared on the scene – in fact the development of Hamiltonian mechanics was driven by people who were using perturbation theory to understand the dynamics of the solar system. Interestingly, while perturbation theory in classical mechanics remains an eclectic branch of knowledge that is understood only by a select few, perturbation theory in quantum mechanics is a part of main-stream undergraduate syllabuses. There are two reasons for this. First, analytically

soluble models are even rarer in quantum than in classical physics, so more systems have to be modelled approximately. Second, in quantum mechanics perturbation theory is a good deal simpler and works rather better than in classical mechanics.

4.1 Time-independent perturbations

Let H be the Hamiltonian of the experimental system and H_0 the Hamiltonian of the model system. We hope that $\Delta \equiv H - H_0$ is small and define

$$H_\beta = H_0 + \beta\Delta. \tag{4.1}$$

We can think of H_β as the Hamiltonian of an apparatus that has a knob on it labelled 'β'; when the knob is turned to $\beta = 0$, the apparatus is the model system, and as the knob is turned round to $\beta = 1$, the apparatus is gradually deformed into the system of experimental interest.

We seek the eigenkets $|E\rangle$ and eigenvalues E of H_β as functions of β. Since the Hamiltonian of the apparatus is a continuous function of β, we conjecture that the $|E\rangle$ and E are continuous functions of β too. In fact, we conjecture that they are analytic functions of β so they can be expanded as power series

$$|E\rangle = |a\rangle + \beta|b\rangle + \beta^2|c\rangle + \cdots \quad ; \quad E = E_a + \beta E_b + \beta^2 E_c + \cdots \tag{4.2}$$

Beware of these conjectures! Much interesting physics is associated with phenomena in which a small change in one variable can produce a large change in another (phase changes, narrow resonances, caustics, ...). In classical physics perturbation theory is bedevilled by such phenomena. In quantum mechanics these conjectures do rather better, but they are still not trustworthy.

We plug our conjectured forms (4.2) into the TISE

$$(H_0 + \beta\Delta)\big(|a\rangle + \beta|b\rangle + \beta^2|c\rangle + \big) = \big(E_a + \beta E_b + \beta^2 E_c + \big)\ \big(|a\rangle + \beta|b\rangle + \beta^2|c\rangle + \big). \tag{4.3}$$

Since we require the equality to hold for any value of β, we can equate to zero the coefficient of every power of β

$$\begin{array}{ll} \beta^0: & H_0|a\rangle = E_a|a\rangle \\ \beta^1: & H_0|b\rangle + \Delta|a\rangle = E_a|b\rangle + E_b|a\rangle \\ \beta^2: & H_0|c\rangle + \Delta|b\rangle = E_a|c\rangle + E_b|b\rangle + E_c|a\rangle \end{array} \tag{4.4}$$

The first equation simply states that E_a and $|a\rangle$ are an eigenvalue and eigenket of H_0. Physically, $|a\rangle$ is the state that we will find the system in if we slowly turn the knob back to zero after making a measurement of the energy. To determine E_b we multiply the second equation through by $\langle a|$:

$$\langle a|H_0|b\rangle + \langle a|\Delta|a\rangle = E_a\langle a|b\rangle + E_b. \tag{4.5}$$

Now $\langle a|H_0|b\rangle = (\langle b|H_0|a\rangle)^* = E_a\langle a|b\rangle$. Cancelling this with the identical term on the right, we are left with

$$E_b = \langle a|\Delta|a\rangle. \tag{4.6}$$

This important result makes intuitive sense: the first-order change in the energy is just the expectation value of the change in the Hamiltonian when the system is in its unperturbed state.

To extract the second-order change in E we multiply the third of equations (4.4) by $\langle a|$. Cancelling $\langle a|H_0|c\rangle$ on $E_a\langle a|c\rangle$ by strict analogy with what we just did, we obtain

$$E_c = \langle a|\Delta|b\rangle - E_b\langle a|b\rangle. \tag{4.7}$$

To proceed further we have to determine $|b\rangle$, the first-order change in the state vector. We express $|b\rangle$ as a linear combination of the eigenkets $|E_n\rangle$ of H_0:

$$|b\rangle = \sum_k b_k|E_k\rangle \tag{4.8}$$

In the second of equations (4.4) we replace $|b\rangle$ by this expansion and multiply through by $\langle E_m| \neq \langle a|$ to find

$$b_m = \frac{\langle E_m|\Delta|a\rangle}{E_a - E_m}. \tag{4.9}$$

This expression determines the coefficient of every ket in (4.8) except the coefficient of $|a\rangle$, which we know is identical with one of the $|E_k\rangle$. Fortunately we can argue that the coefficient of $|a\rangle$ can be taken to be zero from the requirement that $|E\rangle = |a\rangle + \beta|b\rangle + \mathrm{O}(\beta^2)$ remain correctly normalized. We can draw a useful analogy with changing a two-dimensional vector so that the condition $|\mathbf{r}| = 1$ is preserved; clearly we have to move $\mathbf{r}$ on the unit circle and the first-order change in $\mathbf{r}$ is necessarily perpendicular to the original value of $\mathbf{r}$. The QM normalization condition implies that as β increases $|E\rangle$ moves on a hypersphere in state space and $\langle b|a\rangle = 0$. So we exclude $|a\rangle$ from the sum in (4.8) and use this expression with the b_k given by (4.9) in (4.7) to find that the second-order change in E is

$$E_c = \sum_{k\neq a} \frac{\langle a|\Delta|E_k\rangle\langle E_k|\Delta|a\rangle}{E_a - E_k}. \tag{4.10}$$

4.1.1 Normal Stark effect Let's apply the theory we've developed so far to a hydrogen atom that has been placed in an electric field. In this case H_0 is the Hamiltonian of an isolated atom, and $\Delta = e\mathcal{E}z$, where e is the charge on an electron, $\mathcal{E}$ is the strength of the field, and z is the component of the vector from proton to electron in the direction of the field. Let the atom be in its ground state. Then the first-order energy change in E is $E_b = e\mathcal{E}\langle E_1|z|E_1\rangle$. This vanishes because the ground-state ket $|E_1\rangle$ has well defined (even) parity, so the integrand of the integral to which $\langle E_1|z|E_1\rangle$ gives rise is an odd function of $\mathbf{r}$ and the integral vanishes. The change in E is therefore dominated by the second-order term E_c. Since the ground state has zero angular momentum l, the non-vanishing matrix elements[5] in our expression (4.10) for this all have $l = 1$, $m = 0$. Hence

$$E_c = e^2\mathcal{E}^2 \sum_{n>1} \frac{\langle 100|z|n10\rangle\langle n10|z|100\rangle}{E_{10} - E_{n1}}, \tag{4.11}$$

where the numbers in the kets are the values of n, l, m. It is easy to understand physically why the change in E is proportional to $\mathcal{E}^2$: the field polarizes the atom, generating a polarization P that is $\propto \mathcal{E}$. This polarization is encoded in the coefficients b_k, which are indeed proportional to $\mathcal{E}$. The energy of the dipole P in a field is $P\mathcal{E}$, so the energy change caused by the field is proportional to $\mathcal{E}^2$.

4.1.2 Anomalous Stark effect Consider now the shift in the energy of the $n = 2$, $l = 0$ state of Hydrogen when an electric field is applied. The sum over k in (4.10) now includes the term

$$\frac{\langle 200|z|210\rangle\langle 210|z|200\rangle}{E_{20} - E_{21}},$$

which is infinite if you neglect the very small Lamb shift, because the top is non-zero and the difference of energies on the bottom vanishes. It hardly seems likely that a negligible field will produce an arbitrarily large change in the energy of the first excited state of H. So what did we do wrong?

Our error was to assume at the outset that a small stimulus produces a small response. Our infinite contribution to E_c can be traced to (4.9), which diverges as $E_m \to E_a$. That is, the change in the wavefunction that a given field produces is inversely proportional to the energy difference between the original state $|a\rangle$ and the state $|E_m\rangle$ we are pushing the system towards. This is an entirely reasonable result, analogous to what happens as we push a marble that lies at the bottom of a bowl: the distance the marble moves before coming into equilibrium depends on the curvature of the bowl. In the limit that the curvature goes to zero, and the bottom of the bowl becomes flat, an infinitesimal force will move the marble arbitrarily far, because all locations have the same energy. So we conclude that when

[5] Here we exploit a result from group theory, the **Wigner–Eckart theorem**, which tells us that if the operator V is a component of a vector, then $\langle j|V|j'\rangle = 0$ unless states with angular momentum $j'' = 1$ can be made by adding systems with angular momentum j and j' – we discuss the addition of angular momenta below.

the system's initial energy is a degenerate eigenvalue E_d of H, a tiny stimulus *is* liable to produce a big change in the *state* (but not the energy) of the system. Disaster will attend an attempt to calculate this abrupt change of state by the approach we have been developing.

So must we just give up in despair? No, because we can see that the only states that are going to acquire a non-negligible amplitude during the abrupt change are ones that have the same energy as E_d. That is, the state to which the system abruptly moves can be expressed as a linear combination of the kets belonging to E_d. Unless we are very unlucky, there will be only a small number of these (4 in the problem of H on which we are working). What we have to do is to diagonalize the matrix Δ_{ij} formed by Δ squeezed between all pairs of these kets. The eigenkets of Δ in this small subspace will be states of well-defined energy in the slightly perturbed system. As β is ramped up from zero to unity their energies will diverge from E_d. We conjecture that in the instant that β departs from zero, the system's state jumps to the eigenket with the lowest energy, and subsequently stays in this state as β increases. If this conjecture is correct, we should be able to use the perturbation theory we have developed provided we use as basis kets ones that diagonalize Δ as well as H_0.

So let's diagonalize $e\mathcal{E}z$ in the 4-dimensional subspace of Hydrogen kets with $n = 2$. We order the kets 1 $|200\rangle$, 2 $|210\rangle$, 3 $|211\rangle$, 4 $|21{-}1\rangle$. Then the matrix of Δ looks like this

$$\Delta_{ij} = e\mathcal{E}\begin{pmatrix} 0 & a & 0 & 0 \\ a^* & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{where} \quad a = \langle 200|z|210\rangle \tag{4.12}$$

It is easy to show that the eigenvalues of this matrix are $\pm|a|$ and 0, while appropriate eigenkets are $2^{-1/2}(1, \pm 1, 0, 0)$, $(0, 0, 1, 0)$ and $(0, 0, 0, 1)$. We conclude that as soon as the slightest perturbation is switched on, the system is in the state $|\psi\rangle = 2^{-1/2}\big(|200\rangle - |210\rangle\big)$ and we use this state to determine E_b. We find

$$\begin{aligned} E_b &= \tfrac{1}{2}e\mathcal{E}\big(\langle 200| - \langle 210|\big)z\big(|200\rangle - |210\rangle\big) = -\tfrac{1}{2}e\mathcal{E}\big(\langle 200|z|210\rangle + \langle 210|z|200\rangle\big) \\ &= xxa_0e\mathcal{E} \end{aligned} \tag{4.13}$$

From our discussion of the normal Stark effect we know that a change in E that is proportional to $\mathcal{E}$ requires the dipole moment P of an atom to be independent of $\mathcal{E}$. Thus an H atom in the $n = 2$ state has a permanent electric dipole. In classical physics this result is to be expected because the orbit of the electron would in general be elliptical, and the time-averaged charge density along the ellipse would be higher at apocentre than at pericentre, because the electron lingers at apocentre and rushes through pericentre. Hence the centre of charge would lie on the opposite side of the geometrical centre of the ellipse from the focus, where the proton's cancelling charge lies. Thus, if the electron's orbit were a perfect Kepler ellipse, the atom would have a permanent electric dipole moment parallel to the orbit's major axis. Any deviation of the radial force field from $F \propto r^{-2}$ will cause the major axis of the ellipse to precess, and therefore the time-averaged polarization of the atom to be zero. In H the force-field deviates verify little from an inverse-square law, and precession is unimportant in even a weak electric field.

4.2 Time-dependent perturbation theory

We now describe two approaches to obtaining approximate solutions to the TISE.

4.2.1 Fermi golden rule Consider the evolution of a system that is initially in a state that is nearly, but not quite, in a stationary state. Specifically, at $t = 0$ it is in the Nth eigenstate of a Hamiltonian H_0 that differs by only a small, time-independent operator V from the true Hamiltonian H:

$$H = H_0 + V. \tag{4.14}$$

Inspired by (1.52) we expand the solution to the TDSE for this H in the form

$$|\psi\rangle = \sum_n a_n(t)\mathrm{e}^{-\mathrm{i}E_nt/\hbar}|E_n\rangle, \tag{4.15}$$

where $|E_n\rangle$ is a (time-independent) eigenket of H_0 with eigenvalue E_n. Substituting this form into the TDSE we have

$$\begin{aligned} \mathrm{i}\hbar\frac{\partial|\psi\rangle}{\partial t} &= (H_0+V)|\psi\rangle = \sum_n \big(E_n|E_n\rangle + V|E_n\rangle\big)a_n \mathrm{e}^{-\mathrm{i}E_n t/\hbar} \\ &= \sum_n \big(\mathrm{i}\hbar\dot{a}_n + E_n a_n\big)\mathrm{e}^{-\mathrm{i}E_n t/\hbar}|E_n\rangle. \end{aligned} \tag{4.16}$$

We simplify this by multiplying through by $\langle E_k|$:

$$\mathrm{i}\hbar\dot{a}_k \mathrm{e}^{-\mathrm{i}E_k t/\hbar} = \sum_n a_n \mathrm{e}^{-\mathrm{i}E_n t/\hbar}\langle E_k|V|E_n\rangle. \tag{4.17}$$

This constitutes a set of linear ordinary differential equations for the $a_n(t)$ which must be solved subject to the boundary conditions $a_N(0)=1$ and $a_n(0)=0$ for $n\neq N$. Hence, at the earliest times the term on the right of (4.17) with $n=N$ will dominate the equation of motion of a_k with $k\neq N$, and we have the approximation

$$\dot{a}_k \simeq -\frac{\mathrm{i}}{\hbar}\mathrm{e}^{-\mathrm{i}(E_N-E_k)t/\hbar}\langle E_k|V|E_N\rangle. \tag{4.18}$$

Integrating from $t=0$ we find

$$a_k(t) = \frac{\langle E_k|V|E_N\rangle}{E_N-E_k}\left[\mathrm{e}^{-\mathrm{i}(E_N-E_k)t'/\hbar}\right]_0^t, \tag{4.19}$$

so the probability that after t the system has made the transition to the kth eigenstate of H_0 is

$$\begin{aligned} P_k(t) = |a_k|^2 &= \frac{|\langle E_k|V|E_N\rangle|^2}{(E_N-E_k)^2}\left\{2-2\cos\left(\frac{(E_N-E_k)t}{\hbar}\right)\right\} \\ &= 4|\langle E_k|V|E_N\rangle|^2\frac{\sin^2\left((E_N-E_k)t/2\hbar\right)}{(E_N-E_k)^2} \end{aligned} \tag{4.20}$$

For a time of order $h/(E_k-E_N)$ this expression grows like t^2. Subsequently it oscillates.

In applications of interest H_0 has a large number of eigenvalues E_k within an interval of width h/t of E_N, and we are typically interested in the probability that the system has made the transition to *any* one of these states. Hence we sum the P_k over k. Let there be $\rho(E)\,\mathrm{d}E$ eigenvalues in the interval $(E+\mathrm{d}E, E)$. Then the total transition probability is

$$\begin{aligned} \sum_k P_k(t) &= 4\int \mathrm{d}E\,\rho(E)\,|\langle E|V|E_N\rangle|^2\,\frac{\sin^2\left((E_N-E)t/2\hbar\right)}{(E_N-E)^2} \\ &= \frac{2}{\hbar}\int \mathrm{d}x\,\rho(E_N-2\hbar x)\,|\langle E_N-2\hbar x|V|E_N\rangle|^2\,\frac{\sin^2(xt)}{x^2}, \end{aligned} \tag{4.21}$$

where we've introduced a new variable, $x=(E_N-E)/2\hbar$. The function $f_t(x)\equiv\sin^2(xt)/x^2$ is dominated by a bump around the origin that is of height t^2 and width $2\pi/t$. Hence, the area under the bump is proportional to t and in the limit of large t, $f_t(x)\propto t\delta(x)$. We find the constant of proportionality by differentiating $\int \mathrm{d}x\,f_t$ with respect to t:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty}\mathrm{d}x\,f_t(x) = \int_{-\infty}^{\infty}\mathrm{d}x\,\frac{\sin(2xt)}{x} = \pi. \tag{4.22}$$

Replacing $\sin^2(xt)/x^2$ by $\pi t\delta(x)$ in (4.21) and integrating over x, we have finally

$$\sum_k P_k = \frac{2\pi t}{\hbar}\rho(E_N)\,|\langle \mathrm{out}|V|\mathrm{in}\rangle|^2. \tag{4.23}$$

This simple result is **Fermi's golden rule** of perturbation theory (which was actually first given by Dirac).

Scattering Let's use the golden rule to study scattering by a localized potential $V(\mathbf{x})$. In this case H_0 is the Hamiltonian, $p^2/2m$, of a free particle. Initially our particle is in a state of well defined momentum $\hbar\mathbf{k}$, and after some time V may have scattered the particle into a state of similar energy that has momentum $\hbar\mathbf{k}'$. From §1.4 we know that the wavefunction of a particle with momentum $\hbar\mathbf{k}$ is proportional to $\mathrm{e}^{\mathrm{i}\mathbf{k}.\mathbf{x}}$ but we have to think carefully about the normalization. If our particle is free to roam over all space, it has negligible probability of straying into the region in which $V(\mathbf{x}) \neq 0$ and being scattered. So we confine the particle in a box of side L with periodic boundary conditions. Then the correctly normalized wavefunctions of the initial and final states are

$$\langle \mathbf{x}|\mathrm{in}\rangle = \frac{1}{L^{3/2}}\mathrm{e}^{\mathrm{i}\mathbf{k}.\mathbf{x}} \quad \text{and} \quad \langle \mathbf{x}|\mathrm{out}\rangle = \frac{1}{L^{3/2}}\mathrm{e}^{\mathrm{i}\mathbf{k}'.\mathbf{x}}. \tag{4.24}$$

Each state occupies $(2\pi/L)^3$ of k-space, so the density of states with energy $(E+\mathrm{d}E, E)$ and $\mathbf{k}'$ in solid angle $\mathrm{d}^2\Omega$ is

$$\begin{aligned}\rho(E) = \frac{\mathrm{d}N}{\mathrm{d}E} &= \frac{(L/2\pi)^3 k'^2 \mathrm{d}k'\,\mathrm{d}^2\Omega}{\hbar^2 k' \mathrm{d}k'/m} = \Big(\frac{L}{2\pi}\Big)^3 \frac{m}{\hbar^2} k'\,\mathrm{d}^2\Omega \\ &= \Big(\frac{L}{2\pi}\Big)^3 \frac{m^{3/2}}{\hbar^3}\sqrt{2E}\,\mathrm{d}^2\Omega.\end{aligned} \tag{4.25}$$

Substituting (4.24) and (4.25) into (4.23) we find that the scattering rate is

$$\dot{P} = \frac{m^{3/2}\sqrt{2E}}{(2\pi)^2\hbar^4 L^3}\,\mathrm{d}^2\Omega \left| \int \mathrm{d}^3\mathbf{x}\,\mathrm{e}^{\mathrm{i}(\mathbf{k}-\mathbf{k}').\mathbf{x}} V(\mathbf{x}) \right|^2. \tag{4.26}$$

As expected this vanishes in the limit $L \to \infty$. We factor out this L dependence by defining the **scattering cross section** σ to be

$$\begin{aligned}\sigma \equiv \frac{\dot{P}}{\text{incoming flux}} &= \frac{\dot{P}}{L^{-3}v} = \frac{L^3\dot{P}}{\sqrt{2E/m}} \\ &= \frac{m^2}{(2\pi)^2\hbar^4}\,\mathrm{d}^2\Omega \left| \int \mathrm{d}^3\mathbf{x}\,\mathrm{e}^{\mathrm{i}(\mathbf{k}-\mathbf{k}').\mathbf{x}} V(\mathbf{x}) \right|^2.\end{aligned} \tag{4.27}$$

This is the **Born scattering cross section**. It is valid provided V is weak in the sense that the exact wavefunction would not differ greatly from the plane-wave approximations we have used even in the region 'shadowed' by the scattering potential V.

4.2.2 Adiabatic principle Consider the case in which the Hamiltonian depends on time – an example would be an atom exposed to a slowly increasing electric field. The system's ket still satisfies (1.48). Equation (4.15) provides an appropriate trial solution for the case of constant Hamiltonian eigenvalues. We modify this to allow for time-variation of the E_n:

$$|\psi, t\rangle = \sum_n a_n(t) \exp\Big(-\frac{\mathrm{i}}{\hbar}\int_0^t \mathrm{d}t'\,E_n(t')\Big)|E_n(t)\rangle, \tag{4.28}$$

where $E_n(t)$ is the nth eigenvalue of $H(t)$ and $|E_n(t)\rangle$ is the corresponding eigenket. When we substitute this expansion into the TDSE, we find

$$\begin{aligned}\mathrm{i}\hbar\frac{\partial|\psi\rangle}{\partial t} = H|\psi\rangle &= \sum_n a_n \exp\Big(-\frac{\mathrm{i}}{\hbar}\int_0^t \mathrm{d}t'\,E_n(t')\Big) H(t)|E_n(t)\rangle \\ &= \sum_a \Big(\{\mathrm{i}\hbar\dot{a}_n + a_n E_n(t)\}|E_n(t)\rangle + \mathrm{i}\hbar a_n \frac{\partial|E_n\rangle}{\partial t}\Big) \exp\Big(-\frac{\mathrm{i}}{\hbar}\int_0^t \mathrm{d}t'\,E_n(t')\Big).\end{aligned} \tag{4.29}$$

Exploiting the fact that $|E_n(t)\rangle$ is an eigenket of $H(t)$ we can cancel a term from each side and are left with

$$0 = \sum_n \Big(\dot{a}_n|E_n(t)\rangle + a_n\frac{\partial|E_n\rangle}{\partial t}\Big) \exp\Big(-\frac{\mathrm{i}}{\hbar}\int_0^t \mathrm{d}t'\,E_n(t')\Big). \tag{4.30}$$

Now we use the perturbation theory developed in §4.1 to expand $|E_n(t+\delta t)\rangle$ as a linear combination of the complete set $\{|E_n(t)\rangle\}$. That is, we write

$$|E_n(t+\delta t)\rangle - |E_n(t)\rangle = \sum_{m\neq n} b_{nm}|E_m(t)\rangle, \tag{4.31}$$

where from (4.9) we have

$$b_{nm} = \frac{\langle E_m(t)|\delta H|E_n(t)\rangle}{E_n(t) - E_m(t)} \tag{4.32}$$

with δH the change in H between t and $t+\delta t$. Dividing (4.31) by δt and substituting the result into (4.30) we find

$$0 = \sum_n \left(\dot{a}_n|E_n(t)\rangle + a_n \sum_{m\neq n} \frac{\langle E_m(t)|\dot{H}|E_n(t)\rangle}{E_n(t) - E_m(t)}|E_m\rangle \right) \exp\Big(-\frac{\mathrm{i}}{\hbar}\int_0^t \mathrm{d}t'\, E_n(t')\Big). \tag{4.33}$$

When we multiply through by $\langle E_k(t)|$ this yields

$$\dot{a}_k = -\sum_{n\neq k} a_n(t)\frac{\langle E_k(t)|\dot{H}|E_n(t)\rangle}{E_n(t) - E_k(t)} \exp\Big(-\frac{\mathrm{i}}{\hbar}\int_0^t \mathrm{d}t'\,\{E_n(t') - E_k(t')\}\Big). \tag{4.34}$$

Although we have used first-order perturbation theory, our working so far has been exact because we can make δH as small as we please by taking δt to be small. Now we introduce an approximation by supposing that H is a slowly varying function of time in the sense that it changes by very little in the time $\hbar/\min(|E_n - E_k|)$. In this approximation, the right side of (4.34) is a product of a slowly varying function of time times an approximately sinusoidal term that oscillates much more rapidly. When we integrate this expression to get the change in a_k, the integral vanishes rather precisely. Hence, if initially $a_k = 1$ for some k, it will remain unity throughout the evolution. That is, if a system is initially in the kth state of well-defined energy, it will stay in this state when the Hamiltonian is slowly changed. This **adiabatic principle** is extraordinarily useful, because it allows us to solve the TDSE for slowly-varying as well as for constant Hamiltonians.

The adiabatic principle enables us to calculate things such as forces very easily. Consider for example the force exerted on its confining walls by a particle of mass m that is in the nth energy level of an infinite square-well potential of width a. The particle's energy is $E_n = \hbar^2\pi^2n^2/(2ma^2)$ and the work done by the particle in pushing the walls apart by δa is $-(\partial E_n/\partial a)\delta a = 2(E_n/a)\delta a$. Equating this to $F\delta a$, where F is the force that the particle exerts on the walls, we obtain $F = 2E_n/a$.