

# **Quantum Mechanics:**

**A graduate level course**

Richard Fitzpatrick  
Associate Professor of Physics  
The University of Texas at Austin

# Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introduction</b>                                       | <b>5</b>  |
| 1.1      | Major sources . . . . .                                   | 5         |
| <b>2</b> | <b>Fundamental concepts</b>                               | <b>6</b>  |
| 2.1      | The breakdown of classical physics . . . . .              | 6         |
| 2.2      | The polarization of photons . . . . .                     | 7         |
| 2.3      | The fundamental principles of quantum mechanics . . . . . | 9         |
| 2.4      | Ket space . . . . .                                       | 10        |
| 2.5      | Bra space . . . . .                                       | 14        |
| 2.6      | Operators . . . . .                                       | 17        |
| 2.7      | The outer product . . . . .                               | 19        |
| 2.8      | Eigenvalues and eigenvectors . . . . .                    | 20        |
| 2.9      | Observables . . . . .                                     | 21        |
| 2.10     | Measurements . . . . .                                    | 24        |
| 2.11     | Expectation values . . . . .                              | 25        |
| 2.12     | Degeneracy . . . . .                                      | 26        |
| 2.13     | Compatible observables . . . . .                          | 27        |
| 2.14     | The uncertainty relation . . . . .                        | 28        |
| 2.15     | Continuous spectra . . . . .                              | 31        |
| <b>3</b> | <b>Position and momentum</b>                              | <b>33</b> |
| 3.1      | Introduction . . . . .                                    | 33        |
| 3.2      | Poisson brackets . . . . .                                | 33        |
| 3.3      | Wave-functions . . . . .                                  | 37        |
| 3.4      | Schrödinger's representation - I . . . . .                | 39        |
| 3.5      | Schrödinger's representation - II . . . . .               | 43        |
| 3.6      | The momentum representation . . . . .                     | 46        |
| 3.7      | The uncertainty relation . . . . .                        | 48        |
| 3.8      | Displacement operators . . . . .                          | 50        |
| <b>4</b> | <b>Quantum dynamics</b>                                   | <b>55</b> |
| 4.1      | Schrödinger's equations of motion . . . . .               | 55        |
| 4.2      | Heisenberg's equations of motion . . . . .                | 59        |

|          |   |            |
|----------|---|------------|
| 4.3      | Ehrenfest's theorem . . . . .                             | 61         |
| 4.4      | Schrödinger's wave-equation . . . . .                     | 65         |
| <b>5</b> | <b>Angular momentum</b>                                   | <b>71</b>  |
| 5.1      | Orbital angular momentum . . . . .                        | 71         |
| 5.2      | Eigenvalues of angular momentum . . . . .                 | 74         |
| 5.3      | Rotation operators . . . . .                              | 78         |
| 5.4      | Eigenfunctions of orbital angular momentum . . . . .      | 81         |
| 5.5      | Motion in a central field . . . . .                       | 84         |
| 5.6      | Energy levels of the hydrogen atom . . . . .              | 86         |
| 5.7      | Spin angular momentum . . . . .                           | 89         |
| 5.8      | Wave-function of a spin one-half particle . . . . .       | 91         |
| 5.9      | Rotation operators in spin space . . . . .                | 93         |
| 5.10     | Magnetic moments . . . . .                                | 96         |
| 5.11     | Spin precession . . . . .                                 | 97         |
| 5.12     | Pauli two-component formalism . . . . .                   | 99         |
| 5.13     | Spin greater than one-half systems . . . . .              | 105        |
| 5.14     | Addition of angular momentum . . . . .                    | 110        |
| <b>6</b> | <b>Approximation methods</b>                              | <b>120</b> |
| 6.1      | Introduction . . . . .                                    | 120        |
| 6.2      | The two-state system . . . . .                            | 120        |
| 6.3      | Non-degenerate perturbation theory . . . . .              | 122        |
| 6.4      | The quadratic Stark effect . . . . .                      | 124        |
| 6.5      | Degenerate perturbation theory . . . . .                  | 129        |
| 6.6      | The linear Stark effect . . . . .                         | 132        |
| 6.7      | Fine structure . . . . .                                  | 134        |
| 6.8      | The Zeeman effect . . . . .                               | 140        |
| 6.9      | Time-dependent perturbation theory . . . . .              | 144        |
| 6.10     | The two-state system . . . . .                            | 146        |
| 6.11     | Spin magnetic resonance . . . . .                         | 149        |
| 6.12     | The Dyson series . . . . .                                | 150        |
| 6.13     | Constant perturbations . . . . .                          | 154        |
| 6.14     | Harmonic perturbations . . . . .                          | 158        |
| 6.15     | Absorption and stimulated emission of radiation . . . . . | 159        |

|          |   |            |
|----------|---|------------|
| 6.16     | The electric dipole approximation . . . . . | 162        |
| 6.17     | Energy-shifts and decay-widths . . . . .    | 165        |
| <b>7</b> | <b>Scattering theory</b>                    | <b>170</b> |
| 7.1      | Introduction . . . . .                      | 170        |
| 7.2      | The Lipmann-Schwinger equation . . . . .    | 170        |
| 7.3      | The Born approximation . . . . .            | 175        |
| 7.4      | Partial waves . . . . .                     | 178        |
| 7.5      | The optical theorem . . . . .               | 181        |
| 7.6      | Determination of phase-shifts . . . . .     | 182        |
| 7.7      | Hard sphere scattering . . . . .            | 184        |
| 7.8      | Low energy scattering . . . . .             | 186        |
| 7.9      | Resonances . . . . .                        | 188        |

# 1 Introduction

## 1.1 Major sources

The textbooks which I have consulted most frequently while developing course material are:

*The principles of quantum mechanics*, P.A.M. Dirac, 4th Edition (revised), (Oxford University Press, Oxford, UK, 1958).

*The Feynman lectures on physics*, R.P. Feynman, R.B. Leighton, and M. Sands, Volume III (Addison-Wesley, Reading MA, 1965).

*Quantum mechanics*, E. Merzbacher, 2nd Edition (John Wiley & Sons, New York NY, 1970).

*Modern quantum mechanics*, J.J. Sakurai, (Benjamin/Cummings, Menlo Park CA, 1985).

## 2 Fundamental concepts

### 2.1 The breakdown of classical physics

The necessity for a departure from classical mechanics is clearly demonstrated by:

1. The anomalous stability of atoms and molecules: According to classical physics, an electron orbiting a nucleus should lose energy by emission of synchrotron radiation, and gradually spiral in towards the nucleus. Experimentally, this is not observed to happen.
2. The anomalously low specific heats of atoms and molecules: According to the equipartition theorem of classical physics, each degree of freedom of an atomic or molecular system should contribute  $R/2$  to its molar specific heat, where  $R$  is the ideal gas constant. In fact, only the translational and some rotational degrees of freedom seem to contribute. The vibrational degrees of freedom appear to make no contribution at all (except at high temperatures). Incidentally, this fundamental problem with classical physics was known and appreciated in the middle of the nineteenth century. Stories that physicists at the start of the twentieth century thought that classical physics explained everything, and that there was nothing left to discover, are largely apocryphal (see Feynman, Vol. I, Cha. 40).
3. The ultraviolet catastrophe: According to classical physics, the energy density of an electromagnetic field in vacuum is infinite due to a divergence of energy carried by short wave-length modes. Experimentally, there is no such divergence, and the total energy density is finite.
4. Wave-particle duality: Classical physics can deal with waves *or* particles. However, various experiments (*e.g.*, light interference, the photo-electric effect, electron diffraction) show quite clearly that waves sometimes act as if they were streams of particles, and streams of particles sometimes act as if they were waves. This is completely inexplicable within the framework of classical physics.

## 2.2 The polarization of photons

It is known experimentally that when plane polarized light is used to eject photoelectrons there is a preferred direction of emission of the electrons. Clearly, the polarization properties of light, which are more usually associated with its wave-like behaviour, also extend to its particle-like behaviour. In particular, a polarization can be ascribed to each individual photon in a beam of light.

Consider the following well-known experiment. A beam of plane polarized light is passed through a polaroid film, which has the property that it is only transparent to light whose plane of polarization lies perpendicular to its optic axis. Classical electromagnetic wave theory tells us that if the beam is polarized perpendicular to the optic axis then all of the light is transmitted, if the beam is polarized parallel to the optic axis then none of the light is transmitted, and if the light is polarized at an angle  $\alpha$  to the axis then a fraction  $\sin^2 \alpha$  of the beam is transmitted. Let us try to account for these observations at the individual photon level.

A beam of light which is plane polarized in a certain direction is made up of a stream of photons which are each plane polarized in that direction. This picture leads to no difficulty if the plane of polarization lies parallel or perpendicular to the optic axis of the polaroid. In the former case, none of the photons are transmitted, and, in the latter case, all of the photons are transmitted. But, what happens in the case of an obliquely polarized incident beam?

The above question is not very precise. Let us reformulate it as a question relating to the result of some experiment which we could perform. Suppose that we were to fire a single photon at a polaroid film, and then look to see whether or not it emerges from the other side. The possible results of the experiment are that either a whole photon, whose energy is equal to the energy of the incident photon, is observed, or no photon is observed. Any photon which is transmitted though the film must be polarized perpendicular to the optic axis. Furthermore, it is impossible to imagine (in physics) finding part of a photon on the other side of the film. If we repeat the experiment a great number of times then, on average, a fraction  $\sin^2 \alpha$  of the photons are transmitted through the film, and a fraction

$\cos^2 \alpha$  are absorbed. Thus, we conclude that a photon has a *probability*  $\sin^2 \alpha$  of being transmitted as a photon polarized in the plane perpendicular to the optic axis, and a probability  $\cos^2 \alpha$  of being absorbed. These values for the probabilities lead to the correct classical limit for a beam containing a large number of photons.

Note that we have only been able to preserve the individuality of photons, in all cases, by abandoning the determinacy of classical theory, and adopting a fundamentally *probabilistic* approach. We have no way of knowing whether an individual obliquely polarized photon is going to be absorbed by or transmitted through a polaroid film. We only know the probability of each event occurring. This is a fairly sweeping statement, but recall that the state of a photon is fully specified once its energy, direction of propagation, and polarization are known. If we imagine performing experiments using monochromatic light, normally incident on a polaroid film, with a particular oblique polarization, then the state of each individual photon in the beam is completely specified, and there is nothing left over to uniquely determine whether the photon is transmitted or absorbed by the film.

The above discussion about the results of an experiment with a single obliquely polarized photon incident on a polaroid film answers all that can be legitimately asked about what happens to the photon when it reaches the film. Questions as to what decides whether the photon is transmitted or not, or how it changes its direction of polarization, are illegitimate, since they do not relate to the outcome of a possible experiment. Nevertheless, some further description is needed in order to allow the results of this experiment to be correlated with the results of other experiments which can be performed using photons.

The further description provided by quantum mechanics is as follows. It is supposed that a photon polarized obliquely to the optic axis can be regarded as being partly in a state of polarization parallel to the axis, and partly in a state of polarization perpendicular to the axis. In other words, the oblique polarization state is some sort of *superposition* of two states of parallel and perpendicular polarization. Since there is nothing special about the orientation of the optic axis in our experiment, we must conclude that any state of polarization can be regarded as a superposition of two mutually perpendicular states of polarization.

When we make the photon encounter a polaroid film, we are subjecting it to an observation. In fact, we are observing whether it is polarized parallel or perpendicular to the optic axis. The effect of making this observation is to force the photon entirely into a state of parallel or perpendicular polarization. In other words, the photon has to jump suddenly from being partly in each of these two states to being entirely in one or the other of them. Which of the two states it will jump into cannot be predicted, but is governed by probability laws. If the photon jumps into a state of parallel polarization then it is absorbed. Otherwise, it is transmitted. Note that, in this example, the introduction of indeterminacy into the problem is clearly connected with the act of observation. In other words, the indeterminacy is related to the inevitable disturbance of the system associated with the act of observation.

### **2.3 The fundamental principles of quantum mechanics**

There is nothing special about the transmission and absorption of photons through a polaroid film. Exactly the same conclusions as those outlined above are obtained by studying other simple experiments, such as the interference of photons (see Dirac, Sect. I.3), and the Stern-Gerlach experiment (see Sakurai, Cha. 1; Feynman, Cha. 5). The study of these simple experiments leads us to formulate the following fundamental principles of quantum mechanics:

1. Dirac's razor: Quantum mechanics can only answer questions regarding the outcome of possible experiments. Any other questions lie beyond the realms of physics.
2. The principle of superposition of states: Any microscopic system (*i.e.*, an atom, molecule, or particle) in a given state can be regarded as being partly in each of two or more other states. In other words, any state can be regarded as a superposition of two or more other states. Such superpositions can be performed in an infinite number of different ways.
3. The principle of indeterminacy: An observation made on a microscopic system causes it to jump into one or more particular states (which are related to

the type of observation). It is impossible to predict into which final state a particular system will jump, however the probability of a given system jumping into a given final state can be predicted.

The first of these principles was formulated by quantum physicists (such as Dirac) in the 1920s to fend off awkward questions such as “How can a system suddenly jump from one state into another?”, or “How does a system decide which state to jump into?”. As we shall see, the second principle is the basis for the mathematical formulation of quantum mechanics. The final principle is still rather vague. We need to extend it so that we can predict which possible states a system can jump into after a particular type of observation, as well as the probability of the system making a particular jump.

## 2.4 Ket space

Consider a microscopic system composed of particles or bodies with specific properties (mass, moment of inertia, *etc.*) interacting according to specific laws of force. There will be various possible motions of the particles or bodies consistent with the laws of force. Let us term each such motion a *state* of the system. According to the principle of superposition of states, any given state can be regarded as a superposition of two or more other states. Thus, states must be related to mathematical quantities of a kind which can be added together to give other quantities of the same kind. The most obvious examples of such quantities are *vectors*.

Let us consider a particular microscopic system in a particular state, which we label  $A$ : *e.g.*, a photon with a particular energy, momentum, and polarization. We can represent this state as a particular vector, which we also label  $A$ , residing in some vector space, where the other elements of the space represent all of the other possible states of the system. Such a space is called a *ket space* (after Dirac). The state vector  $A$  is conventionally written

$$|A\rangle. \tag{2.1}$$

Suppose that state  $A$  is, in fact, the superposition of two different states,  $B$  and

C. This interrelation is represented in ket space by writing

$$|A\rangle = |B\rangle + |C\rangle, \quad (2.2)$$

where  $|B\rangle$  is the vector relating to the state B, *etc.* For instance, state  $|B\rangle$  might represent a photon propagating in the  $z$ -direction, and plane polarized in the  $x$ -direction, and state  $|C\rangle$  might represent a similar photon plane polarized in the  $y$ -direction. In this case, the sum of these two states represents a photon whose plane of polarization makes an angle of  $45^\circ$  with both the  $x$ - and  $y$ -directions (by analogy with classical physics). This latter state is represented by  $|B\rangle + |C\rangle$  in ket space.

Suppose that we want to construct a state whose plane of polarization makes an arbitrary angle  $\alpha$  with the  $x$ -direction. We can do this via a suitably weighted superposition of states B and C. By analogy with classical physics, we require  $\cos \alpha$  of state B, and  $\sin \alpha$  of state C. This new state is represented by

$$\cos \alpha |B\rangle + \sin \alpha |C\rangle \quad (2.3)$$

in ket space. Note that we cannot form a new state by superposing a state with itself. For instance, a photon polarized in the  $y$ -direction superposed with another photon polarized in the  $y$ -direction (with the same energy and momentum) gives the same photon. This implies that the ket vector

$$c_1|A\rangle + c_2|A\rangle = (c_1 + c_2)|A\rangle \quad (2.4)$$

corresponds to the same state that  $|A\rangle$  does. Thus, ket vectors differ from conventional vectors in that their magnitudes, or lengths, are physically irrelevant. All the states of the system are in one to one correspondence with all the possible directions of vectors in the ket space, no distinction being made between the directions of the ket vectors  $|A\rangle$  and  $-|A\rangle$ . There is, however, one caveat to the above statements. If  $c_1 + c_2 = 0$  then the superposition process yields nothing at all: *i.e.*, no state. The absence of a state is represented by the null vector  $|0\rangle$  in ket space. The null vector has the fairly obvious property that

$$|A\rangle + |0\rangle = |A\rangle, \quad (2.5)$$

for any vector  $|A\rangle$ . The fact that ket vectors pointing in the same direction represent the same state relates ultimately to the quantization of matter: *i.e.*, the fact

that it comes in irreducible packets called photons, electrons, atoms, *etc.* If we observe a microscopic system then we either see a state (*i.e.*, a photon, or an atom, or a molecule, *etc.*) or we see nothing—we can never see a fraction or a multiple of a state. In classical physics, if we observe a wave then the amplitude of the wave can take any value between zero and infinity. Thus, if we were to represent a classical wave by a vector, then the magnitude, or length, of the vector would correspond to the amplitude of the wave, and the direction would correspond to the frequency and wave-length, so that two vectors of different lengths pointing in the same direction would represent different wave states.

We have seen, in Eq. (2.3), that any plane polarized state of a photon can be represented as a linear superposition of two orthogonal polarization states in which the weights are real numbers. Suppose that we want to construct a circularly polarized photon state. Well, we know from classical physics that a circularly polarized wave is a superposition of two waves of equal amplitude, plane polarized in orthogonal directions, which are in *phase quadrature*. This suggests that a circularly polarized photon is the superposition of a photon polarized in the  $x$ -direction (state B) and a photon polarized in the  $y$ -direction (state C), with equal weights given to the two states, but with the proviso that state C is  $90^\circ$  out of phase with state B. By analogy with classical physics, we can use *complex numbers* to simultaneously represent the weighting and relative phase in a linear superposition. Thus, a circularly polarized photon is represented by

$$|B\rangle + i|C\rangle \quad (2.6)$$

in ket space. A general elliptically polarized photon is represented by

$$c_1|B\rangle + c_2|C\rangle, \quad (2.7)$$

where  $c_1$  and  $c_2$  are complex numbers. We conclude that a ket space must be a complex vector space if it is to properly represent the mutual interrelations between the possible states of a microscopic system.

Suppose that the ket  $|R\rangle$  is expressible linearly in terms of the kets  $|A\rangle$  and  $|B\rangle$ , so that

$$|R\rangle = c_1|A\rangle + c_2|B\rangle. \quad (2.8)$$

We say that  $|R\rangle$  is *dependent* on  $|A\rangle$  and  $|B\rangle$ . It follows that the state R can be regarded as a linear superposition of the states A and B. So, we can also say that state R is dependent on states A and B. In fact, any ket vector (or state) which is expressible linearly in terms of certain others is said to be dependent on them. Likewise, a set of ket vectors (or states) are termed *independent* if none of them are expressible linearly in terms of the others.

The dimensionality of a conventional vector space is defined as the number of independent vectors contained in the space. Likewise, the dimensionality of a ket space is equivalent to the number of independent ket vectors it contains. Thus, the ket space which represents the possible polarization states of a photon propagating in the z-direction is two-dimensional (the two independent vectors correspond to photons plane polarized in the x- and y-directions, respectively). Some microscopic systems have a finite number of independent states (e.g., the spin states of an electron in a magnetic field). If there are N independent states, then the possible states of the system are represented as an N-dimensional ket space. Some microscopic systems have a denumerably infinite number of independent states (e.g., a particle in an infinitely deep, one-dimensional potential well). The possible states of such a system are represented as a ket space whose dimensions are denumerably infinite. Such a space can be treated in more or less the same manner as a finite-dimensional space. Unfortunately, some microscopic systems have a nondenumerably infinite number of independent states (e.g., a free particle). The possible states of such a system are represented as a ket space whose dimensions are nondenumerably infinite. This type of space requires a slightly different treatment to spaces of finite, or denumerably infinite, dimensions.

In conclusion, the states of a general microscopic system can be represented as a complex vector space of (possibly) infinite dimensions. Such a space is termed a *Hilbert space* by mathematicians.

## 2.5 Bra space

A snack machine inputs coins plus some code entered on a key pad, and (hopefully) outputs a snack. It also does so in a deterministic manner: *i.e.*, the same money plus the same code produces the same snack (or the same error message) time after time. Note that the input and output of the machine have completely different natures. We can imagine building a rather abstract snack machine which inputs ket vectors and outputs complex numbers in a deterministic fashion. Mathematicians call such a machine a *functional*. Imagine a general functional, labeled  $F$ , acting on a general ket vector, labeled  $A$ , and spitting out a general complex number  $\phi_A$ . This process is represented mathematically by writing

$$\langle F(|A\rangle) \rangle = \phi_A. \quad (2.9)$$

Let us narrow our focus to those functionals which preserve the linear dependencies of the ket vectors upon which they operate. Not surprisingly, such functionals are termed *linear functionals*. A general linear functional, labeled  $F$ , satisfies

$$\langle F(|A\rangle + |B\rangle) \rangle = \langle F(|A\rangle) \rangle + \langle F(|B\rangle) \rangle, \quad (2.10)$$

where  $|A\rangle$  and  $|B\rangle$  are any two kets in a given ket space.

Consider an  $N$ -dimensional ket space [*i.e.*, a finite-dimensional, or denumerably infinite dimensional (*i.e.*,  $N \rightarrow \infty$ ), space]. Let the  $|i\rangle$  (where  $i$  runs from 1 to  $N$ ) represent  $N$  independent ket vectors in this space. A general ket vector can be written<sup>1</sup>

$$|A\rangle = \sum_{i=1}^N \alpha_i |i\rangle, \quad (2.11)$$

where the  $\alpha_i$  are an arbitrary set of complex numbers. The only way the functional  $F$  can satisfy Eq. (2.10) for all vectors in the ket space is if

$$\langle F(|A\rangle) \rangle = \sum_{i=1}^N f_i \alpha_i, \quad (2.12)$$

---

<sup>1</sup>Actually, this is only strictly true for finite-dimensional spaces. Only a special subset of denumerably infinite dimensional spaces have this property (*i.e.*, they are complete), but since a ket space must be complete if it is to represent the states of a microscopic system, we need only consider this special subset.

where the  $f_i$  are a set of complex numbers relating to the functional.

Let us define  $N$  basis functionals  $\langle i|$  which satisfy

$$\langle i|(j\rangle) = \delta_{ij}. \quad (2.13)$$

It follows from the previous three equations that

$$\langle F| = \sum_{i=1}^N f_i \langle i|. \quad (2.14)$$

But, this implies that the set of all possible linear functionals acting on an  $N$ -dimensional ket space is itself an  $N$ -dimensional vector space. This type of vector space is called a *bra space* (after Dirac), and its constituent vectors (which are actually functionals of the ket space) are called bra vectors. Note that bra vectors are quite different in nature to ket vectors (hence, these vectors are written in mirror image notation,  $\langle \dots |$  and  $|\dots\rangle$ , so that they can never be confused). Bra space is an example of what mathematicians call a *dual vector space* (i.e., it is dual to the original ket space). There is a one to one correspondence between the elements of the ket space and those of the related bra space. So, for every element  $A$  of the ket space, there is a corresponding element, which it is also convenient to label  $A$ , in the bra space. That is,

$$|A\rangle \xleftrightarrow{\text{DC}} \langle A|, \quad (2.15)$$

where DC stands for *dual correspondence*.

There are an infinite number of ways of setting up the correspondence between vectors in a ket space and those in the related bra space. However, only one of these has any physical significance. For a general ket vector  $A$ , specified by Eq. (2.11), the corresponding bra vector is written

$$\langle A| = \sum_{i=1}^N \alpha_i^* \langle i|, \quad (2.16)$$

where the  $\alpha_i^*$  are the complex conjugates of the  $\alpha_i$ .  $\langle A|$  is termed the dual vector to  $|A\rangle$ . It follows, from the above, that the dual to  $c|A\rangle$  is  $c^*\langle A|$ , where  $c$  is a complex number. More generally,

$$c_1|A\rangle + c_2|B\rangle \xleftrightarrow{\text{DC}} c_1^*\langle A| + c_2^*\langle B|. \quad (2.17)$$

Recall that a bra vector is a functional which acts on a general ket vector, and spits out a complex number. Consider the functional which is dual to the ket vector

$$|B\rangle = \sum_{i=1}^N \beta_i |i\rangle \quad (2.18)$$

acting on the ket vector  $|A\rangle$ . This operation is denoted  $\langle B|(|A\rangle)$ . Note, however, that we can omit the round brackets without causing any ambiguity, so the operation can also be written  $\langle B|A\rangle$ . This expression can be further simplified to give  $\langle B|A\rangle$ . According to Eqs. (2.11), (2.12), (2.16), and (2.18),

$$\langle B|A\rangle = \sum_{i=1}^N \beta_i^* \alpha_i. \quad (2.19)$$

Mathematicians term  $\langle B|A\rangle$  the *inner product* of a bra and a ket.<sup>2</sup> An inner product is (almost) analogous to a scalar product between a covariant and contravariant vector in some curvilinear space. It is easily demonstrated that

$$\langle B|A\rangle = \langle A|B\rangle^*. \quad (2.20)$$

Consider the special case where  $|B\rangle \rightarrow |A\rangle$ . It follows from Eqs. (2.12) and (2.20) that  $\langle A|A\rangle$  is a real number, and that

$$\langle A|A\rangle \geq 0. \quad (2.21)$$

The equality sign only holds if  $|A\rangle$  is the null ket [*i.e.*, if all of the  $\alpha_i$  are zero in Eq. (2.11)]. This property of bra and ket vectors is essential for the probabilistic interpretation of quantum mechanics, as will become apparent later.

Two kets  $|A\rangle$  and  $|B\rangle$  are said to be *orthogonal* if

$$\langle A|B\rangle = 0, \quad (2.22)$$

which also implies that  $\langle B|A\rangle = 0$ .

Given a ket  $|A\rangle$  which is not the null ket, we can define a *normalized ket*  $|\tilde{A}\rangle$ , where

$$|\tilde{A}\rangle = \left( \frac{1}{\sqrt{\langle A|A\rangle}} \right) |A\rangle, \quad (2.23)$$

---

<sup>2</sup>We can now appreciate the elegance of Dirac's notation. The combination of a bra and a ket yields a "bra(c)ket" (which is just a number).

with the property

$$\langle \tilde{A} | \tilde{A} \rangle = 1. \quad (2.24)$$

Here,  $\sqrt{\langle A | A \rangle}$  is known as the *norm* or “length” of  $|A\rangle$ , and is analogous to the length, or magnitude, of a conventional vector. Since  $|A\rangle$  and  $c|A\rangle$  represent the same physical state, it makes sense to require that all kets corresponding to physical states have unit norms.

It is possible to define a dual bra space for a ket space of nondenumerably infinite dimensions in much the same manner as that described above. The main differences are that summations over discrete labels become integrations over continuous labels, Kronecker delta-functions become Dirac delta-functions, completeness must be assumed (it cannot be proved), and the normalization convention is somewhat different. More of this later.

## 2.6 Operators

We have seen that a functional is a machine which inputs a ket vector and spits out a complex number. Consider a somewhat different machine which inputs a ket vector and spits out another ket vector in a deterministic fashion. Mathematicians call such a machine an *operator*. We are only interested in operators which preserve the linear dependencies of the ket vectors upon which they act. Such operators are termed *linear operators*. Consider an operator labeled  $X$ . Suppose that when this operator acts on a general ket vector  $|A\rangle$  it spits out a new ket vector which is denoted  $X|A\rangle$ . Operator  $X$  is linear provided that

$$X(|A\rangle + |B\rangle) = X|A\rangle + X|B\rangle, \quad (2.25)$$

for all ket vectors  $|A\rangle$  and  $|B\rangle$ , and

$$X(c|A\rangle) = cX|A\rangle, \quad (2.26)$$

for all complex numbers  $c$ . Operators  $X$  and  $Y$  are said to be equal if

$$X|A\rangle = Y|A\rangle \quad (2.27)$$

for all kets in the ket space in question. Operator  $X$  is termed the *null operator* if

$$X|A\rangle = |0\rangle \quad (2.28)$$

for all ket vectors in the space. Operators can be added together. Such addition is defined to obey a commutative and associate algebra:

$$X + Y = Y + X, \quad (2.29)$$

$$X + (Y + Z) = (X + Y) + Z. \quad (2.30)$$

Operators can also be multiplied. The multiplication is associative:

$$X(Y|A\rangle) = (XY)|A\rangle = XY|A\rangle, \quad (2.31)$$

$$X(YZ) = (XY)Z = XYZ. \quad (2.32)$$

However, in general, it is *noncommutative*:

$$XY \neq YX. \quad (2.33)$$

So far, we have only considered linear operators acting on ket vectors. We can also give a meaning to their operating on bra vectors. Consider the inner product of a general bra  $\langle B|$  with the ket  $X|A\rangle$ . This product is a number which depends linearly on  $|A\rangle$ . Thus, it may be considered to be the inner product of  $|A\rangle$  with some bra. This bra depends linearly on  $\langle B|$ , so we may look on it as the result of some linear operator applied to  $\langle B|$ . This operator is uniquely determined by the original operator  $X$ , so we might as well call it the same operator acting on  $|B\rangle$ . A suitable notation to use for the resulting bra when  $X$  operates on  $\langle B|$  is  $\langle B|X$ . The equation which defines this vector is

$$(\langle B|X)|A\rangle = \langle B|(X|A\rangle) \quad (2.34)$$

for any  $|A\rangle$  and  $\langle B|$ . The triple product of  $\langle B|$ ,  $X$ , and  $|A\rangle$  can be written  $\langle B|X|A\rangle$  without ambiguity, provided we adopt the convention that the bra vector always goes on the left, the operator in the middle, and the ket vector on the right.

Consider the dual bra to  $X|A\rangle$ . This bra depends antilinearly on  $|A\rangle$  and must therefore depend linearly on  $\langle A|$ . Thus, it may be regarded as the result of some

linear operator applied to  $\langle A|$ . This operator is termed the *adjoint* of  $X$ , and is denoted  $X^\dagger$ . Thus,

$$X|A\rangle \xleftrightarrow{\text{DC}} \langle A|X^\dagger. \quad (2.35)$$

It is readily demonstrated that

$$\langle B|X^\dagger|A\rangle = \langle A|X|B\rangle^*, \quad (2.36)$$

plus

$$(XY)^\dagger = Y^\dagger X^\dagger. \quad (2.37)$$

It is also easily seen that the adjoint of the adjoint of a linear operator is equivalent to the original operator. A *Hermitian* operator  $\xi$  has the special property that it is its own adjoint: *i.e.*,

$$\xi = \xi^\dagger. \quad (2.38)$$

## 2.7 The outer product

So far we have formed the following products:  $\langle B|A\rangle$ ,  $X|A\rangle$ ,  $\langle A|X$ ,  $XY$ ,  $\langle B|X|A\rangle$ . Are there any other products we are allowed to form? How about

$$|B\rangle\langle A|? \quad (2.39)$$

This clearly depends linearly on the ket  $|A\rangle$  and the bra  $|B\rangle$ . Suppose that we right-multiply the above product by the general ket  $|C\rangle$ . We obtain

$$|B\rangle\langle A|C\rangle = \langle A|C\rangle|B\rangle, \quad (2.40)$$

since  $\langle A|C\rangle$  is just a number. Thus,  $|B\rangle\langle A|$  acting on a general ket  $|C\rangle$  yields another ket. Clearly, the product  $|B\rangle\langle A|$  is a linear operator. This operator also acts on bras, as is easily demonstrated by left-multiplying the expression (2.39) by a general bra  $\langle C|$ . It is also easily demonstrated that

$$(|B\rangle\langle A|)^\dagger = |A\rangle\langle B|. \quad (2.41)$$

Mathematicians term the operator  $|B\rangle\langle A|$  the *outer product* of  $|B\rangle$  and  $\langle A|$ . The outer product should not be confused with the inner product,  $\langle A|B\rangle$ , which is just a number.

## 2.8 Eigenvalues and eigenvectors

In general, the ket  $X|A\rangle$  is not a constant multiple of  $|A\rangle$ . However, there are some special kets known as the *eigenkets* of operator  $X$ . These are denoted

$$|x'\rangle, |x''\rangle, |x'''\rangle, \dots, \quad (2.42)$$

and have the property

$$X|x'\rangle = x'|x'\rangle, \quad X|x''\rangle = x''|x''\rangle, \dots, \quad (2.43)$$

where  $x', x'', \dots$  are numbers called *eigenvalues*. Clearly, applying  $X$  to one of its eigenkets yields the same eigenket multiplied by the associated eigenvalue.

Consider the eigenkets and eigenvalues of a Hermitian operator  $\xi$ . These are denoted

$$\xi|\xi'\rangle = \xi'|\xi'\rangle, \quad (2.44)$$

where  $|\xi'\rangle$  is the eigenket associated with the eigenvalue  $\xi'$ . Three important results are readily deduced:

(i) *The eigenvalues are all real numbers, and the eigenkets corresponding to different eigenvalues are orthogonal.* Since  $\xi$  is Hermitian, the dual equation to Eq. (2.44) (for the eigenvalue  $\xi''$ ) reads

$$\langle \xi'' | \xi = \xi''^* \langle \xi'' |. \quad (2.45)$$

If we left-multiply Eq. (2.44) by  $\langle \xi'' |$ , right-multiply the above equation by  $|\xi'\rangle$ , and take the difference, we obtain

$$(\xi' - \xi''^*) \langle \xi'' | \xi' \rangle = 0. \quad (2.46)$$

Suppose that the eigenvalues  $\xi'$  and  $\xi''$  are the same. It follows from the above that

$$\xi' = \xi'^*, \quad (2.47)$$

where we have used the fact that  $|\xi'\rangle$  is not the null ket. This proves that the eigenvalues are real numbers. Suppose that the eigenvalues  $\xi'$  and  $\xi''$  are different. It follows that

$$\langle \xi'' | \xi' \rangle = 0, \quad (2.48)$$

which demonstrates that eigenkets corresponding to different eigenvalues are orthogonal.

(ii) *The eigenvalues associated with eigenkets are the same as the eigenvalues associated with eigenbras.* An eigenbra of  $\xi$ , corresponding to an eigenvalue  $\xi'$  is defined

$$\langle \xi' | \xi = \langle \xi' | \xi'. \quad (2.49)$$

(iii) *The dual of any eigenket is an eigenbra belonging to the same eigenvalue, and conversely.*

## 2.9 Observables

We have developed a mathematical formalism which comprises three types of objects—bras, kets, and linear operators. We have already seen that kets can be used to represent the possible states of a microscopic system. However, there is a one to one correspondence between the elements of a ket space and its dual bra space, so we must conclude that bras could just as well be used to represent the states of a microscopic system. What about the dynamical variables of the system (*e.g.*, its position, momentum, energy, spin, *etc.*)? How can these be represented in our formalism? Well, the only objects we have left over are operators. We, therefore, assume that *the dynamical variables of a microscopic system are represented as linear operators acting on the bras and kets which correspond to the various possible states of the system.* Note that the operators have to be linear, otherwise they would, in general, spit out bras/kets pointing in different directions when fed bras/kets pointing in the same direction but differing in length. Since the lengths of bras and kets have no physical significance, it is reasonable to suppose that non-linear operators are also without physical significance.

We have seen that if we observe the polarization state of a photon, by placing a polaroid film in its path, the result is to cause the photon to jump into a state of polarization parallel or perpendicular to the optic axis of the film. The former state is absorbed, and the latter state is transmitted (which is how we tell them apart). In general, we cannot predict into which state a given photon will jump

(except in a statistical sense). However, we do know that if the photon is initially polarized parallel to the optic axis then it will definitely be absorbed, and if it is initially polarized perpendicular to the axis then it will definitely be transmitted. We also know that after passing through the film a photon must be in a state of polarization perpendicular to the optic axis (otherwise it would not have been transmitted). We can make a second observation of the polarization state of such a photon by placing an identical polaroid film (with the same orientation of the optic axis) immediately behind the first film. It is clear that the photon will definitely be transmitted through the second film.

There is nothing special about the polarization states of a photon. So, more generally, we can say that when a dynamical variable of a microscopic system is measured the system is caused to jump into one of a number of *independent* states (note that the perpendicular and parallel polarization states of our photon are linearly independent). In general, each of these final states is associated with a different result of the measurement: *i.e.*, a different value of the dynamical variable. Note that the result of the measurement must be a *real* number (there are no measurement machines which output complex numbers). Finally, if an observation is made, and the system is found to be a one particular final state, with one particular value for the dynamical variable, then a second observation, made immediately after the first one, will *definitely* find the system in the same state, and yield the same value for the dynamical variable.

How can we represent all of these facts in our mathematical formalism? Well, by a fairly non-obvious leap of intuition, we are going to assert that *a measurement of a dynamical variable corresponding to an operator  $X$  in ket space causes the system to jump into a state corresponding to one of the eigenkets of  $X$* . Not surprisingly, such a state is termed an *eigenstate*. Furthermore, *the result of the measurement is the eigenvalue associated with the eigenket into which the system jumps*. The fact that the result of the measurement must be a real number implies that *dynamical variables can only be represented by Hermitian operators* (since only Hermitian operators are guaranteed to have real eigenvalues). The fact that the eigenkets of a Hermitian operator corresponding to different eigenvalues (*i.e.*, different results of the measurement) are orthogonal is in accordance with our earlier requirement that the states into which the system jumps should be mutually

independent. We can conclude that the result of a measurement of a dynamical variable represented by a Hermitian operator  $\xi$  must be one of the eigenvalues of  $\xi$ . Conversely, every eigenvalue of  $\xi$  is a possible result of a measurement made on the corresponding dynamical variable. This gives us the physical significance of the eigenvalues. (From now on, the distinction between a state and its representative ket vector, and a dynamical variable and its representative operator, will be dropped, for the sake of simplicity.)

It is reasonable to suppose that *if a certain dynamical variable  $\xi$  is measured with the system in a particular state, then the states into which the system may jump on account of the measurement are such that the original state is dependent on them.* This fairly innocuous statement has two very important corollaries. First, immediately after an observation whose result is a particular eigenvalue  $\xi'$ , the system is left in the associated eigenstate. However, this eigenstate is orthogonal to (*i.e.*, independent of) any other eigenstate corresponding to a different eigenvalue. It follows that a second measurement made immediately after the first one must leave the system in an eigenstate corresponding to the eigenvalue  $\xi'$ . In other words, the second measurement is bound to give the same result as the first. Furthermore, *if the system is in an eigenstate of  $\xi$ , corresponding to an eigenvalue  $\xi'$ , then a measurement of  $\xi$  is bound to give the result  $\xi'$ .* This follows because the system cannot jump into an eigenstate corresponding to a different eigenvalue of  $\xi$ , since such a state is not dependent on the original state. Second, it stands to reason that a measurement of  $\xi$  must always yield some result. It follows that no matter what the initial state of the system, it must always be able to jump into one of the eigenstates of  $\xi$ . In other words, a general ket must always be dependent on the eigenkets of  $\xi$ . This can only be the case if the eigenkets form a *complete set* (*i.e.*, they span ket space). Thus, *in order for a Hermitian operator  $\xi$  to be observable its eigenkets must form a complete set.* A Hermitian operator which satisfies this condition is termed an *observable*. Conversely, any observable quantity must be a Hermitian operator with a complete set of eigenstates.

## 2.10 Measurements

We have seen that a measurement of some observable  $\xi$  of a microscopic system causes the system to jump into one of the eigenstates of  $\xi$ . The result of the measurement is the associated eigenvalue (or some function of this quantity). It is impossible to determine into which eigenstate a given system will jump, but it is possible to predict the probability of such a transition. So, what is the probability that a system in some initial state  $|A\rangle$  makes a transition to an eigenstate  $|\xi'\rangle$  of an observable  $\xi$ , as a result of a measurement made on the system? Let us start with the simplest case. If the system is initially in an eigenstate  $|\xi'\rangle$  then the transition probability to a eigenstate  $|\xi''\rangle$  corresponding to a different eigenvalue is zero, and the transition probability to the same eigenstate  $|\xi'\rangle$  is unity. It is convenient to normalize our eigenkets such that they all have unit norms. It follows from the orthogonality property of the eigenkets that

$$\langle \xi' | \xi'' \rangle = \delta_{\xi' \xi''}, \quad (2.50)$$

where  $\delta_{\xi' \xi''}$  is unity if  $\xi' = \xi''$ , and zero otherwise. For the moment, we are assuming that the eigenvalues of  $\xi$  are all different.

Note that the probability of a transition from an initial eigenstate  $|\xi'\rangle$  to a final eigenstate  $|\xi''\rangle$  is the same as the value of the inner product  $\langle \xi' | \xi'' \rangle$ . Can we use this correspondence to obtain a general rule for calculating transition probabilities? Well, suppose that the system is initially in a state  $|A\rangle$  which is not an eigenstate of  $\xi$ . Can we identify the transition probability to a final eigenstate  $|\xi'\rangle$  with the inner product  $\langle A | \xi' \rangle$ ? The straight answer is “no”, since  $\langle A | \xi' \rangle$  is, in general, a complex number, and complex probabilities do not make much sense. Let us try again. How about if we identify the transition probability with the *modulus squared* of the inner product,  $|\langle A | \xi' \rangle|^2$ ? This quantity is definitely a positive number (so it could be a probability). This guess also gives the right answer for the transition probabilities between eigenstates. In fact, it is the correct guess.

Since the eigenstates of an observable  $\xi$  form a complete set, we can express any given state  $|A\rangle$  as a linear combination of them. It is easily demonstrated that

$$|A\rangle = \sum_{\xi'} |\xi'\rangle \langle \xi' | A \rangle, \quad (2.51)$$

$$\langle \mathcal{A} | = \sum_{\xi'} \langle \mathcal{A} | \xi' \rangle \langle \xi' |, \quad (2.52)$$

$$\langle \mathcal{A} | \mathcal{A} \rangle = \sum_{\xi'} \langle \mathcal{A} | \xi' \rangle \langle \xi' | \mathcal{A} \rangle = \sum_{\xi'} |\langle \mathcal{A} | \xi' \rangle|^2, \quad (2.53)$$

where the summation is over all the different eigenvalues of  $\xi$ , and use has been made of Eq. (2.20), and the fact that the eigenstates are mutually orthogonal. Note that all of the above results follow from the extremely useful (and easily proved) result

$$\sum_{\xi'} |\xi' \rangle \langle \xi' | = 1, \quad (2.54)$$

where 1 denotes the identity operator. The relative probability of a transition to an eigenstate  $|\xi'\rangle$ , which is equivalent to the relative probability of a measurement of  $\xi$  yielding the result  $\xi'$ , is

$$P(\xi') \propto |\langle \mathcal{A} | \xi' \rangle|^2. \quad (2.55)$$

The absolute probability is clearly

$$P(\xi') = \frac{|\langle \mathcal{A} | \xi' \rangle|^2}{\sum_{\xi'} |\langle \mathcal{A} | \xi' \rangle|^2} = \frac{|\langle \mathcal{A} | \xi' \rangle|^2}{\langle \mathcal{A} | \mathcal{A} \rangle}. \quad (2.56)$$

If the ket  $|\mathcal{A}\rangle$  is normalized such that its norm is unity, then this probability simply reduces to

$$P(\xi') = |\langle \mathcal{A} | \xi' \rangle|^2. \quad (2.57)$$

## 2.11 Expectation values

Consider an *ensemble* of microscopic systems prepared in the same initial state  $|\mathcal{A}\rangle$ . Suppose a measurement of the observable  $\xi$  is made on each system. We know that each measurement yields the value  $\xi'$  with probability  $P(\xi')$ . What is the mean value of the measurement? This quantity, which is generally referred to as the *expectation value* of  $\xi$ , is given by

$$\langle \xi \rangle = \sum_{\xi'} \xi' P(\xi') = \sum_{\xi'} \xi' |\langle \mathcal{A} | \xi' \rangle|^2$$

$$= \sum_{\xi'} \xi' \langle \mathcal{A} | \xi' \rangle \langle \xi' | \mathcal{A} \rangle = \sum_{\xi'} \langle \mathcal{A} | \xi | \xi' \rangle \langle \xi' | \mathcal{A} \rangle, \quad (2.58)$$

which reduces to

$$\langle \xi \rangle = \langle \mathcal{A} | \xi | \mathcal{A} \rangle \quad (2.59)$$

with the aid of Eq. (2.54).

Consider the identity operator, 1. All states are eigenstates of this operator with the eigenvalue unity. Thus, the expectation value of this operator is always unity: *i.e.*,

$$\langle \mathcal{A} | 1 | \mathcal{A} \rangle = \langle \mathcal{A} | \mathcal{A} \rangle = 1, \quad (2.60)$$

for all  $|\mathcal{A}\rangle$ . Note that it is only possible to normalize a given ket  $|\mathcal{A}\rangle$  such that Eq. (2.60) is satisfied because of the more general property (2.21) of the norm. This property depends on the particular correspondence (2.16), that we adopted earlier, between the elements of a ket space and those of its dual bra space.

## 2.12 Degeneracy

Suppose that two different eigenstates  $|\xi'_a\rangle$  and  $|\xi'_b\rangle$  of  $\xi$  correspond to the *same* eigenvalue  $\xi'$ . These are termed *degenerate* eigenstates. Degenerate eigenstates are necessarily orthogonal to any eigenstates corresponding to different eigenvalues, but, in general, they are not orthogonal to each other (*i.e.*, the proof of orthogonality given in Sect. 2.8 does not work in this case). This is unfortunate, since much of the previous formalism depends crucially on the mutual orthogonality of the different eigenstates of an observable. Note, however, that any linear combination of  $|\xi'_a\rangle$  and  $|\xi'_b\rangle$  is also an eigenstate corresponding to the eigenvalue  $\xi'$ . It follows that we can always *construct* two mutually orthogonal degenerate eigenstates. For instance,

$$|\xi'_1\rangle = |\xi'_a\rangle, \quad (2.61)$$

$$|\xi'_2\rangle = \frac{|\xi'_b\rangle - \langle \xi'_a | \xi'_b \rangle |\xi'_a\rangle}{1 - |\langle \xi'_a | \xi'_b \rangle|^2}. \quad (2.62)$$

This result is easily generalized to the case of more than two degenerate eigenstates. We conclude that it is always possible to construct a complete set of mutually orthogonal eigenstates for any given observable.

### 2.13 Compatible observables

Suppose that we wish to simultaneously measure two observables,  $\xi$  and  $\eta$ , of a microscopic system? Let us assume that we possess an apparatus which is capable of measuring  $\xi$ , and another which can measure  $\eta$ . For instance, the two observables in question might be the projection in the  $x$ - and  $z$ -directions of the spin angular momentum of a spin one-half particle. These could be measured using appropriate Stern-Gerlach apparatuses (see Sakurai, Sect. 1.1). Suppose that we make a measurement of  $\xi$ , and the system is consequently thrown into one of the eigenstates of  $\xi$ ,  $|\xi'\rangle$ , with eigenvalue  $\xi'$ . What happens if we now make a measurement of  $\eta$ ? Well, suppose that the eigenstate  $|\xi'\rangle$  is also an eigenstate of  $\eta$ , with eigenvalue  $\eta'$ . In this case, a measurement of  $\eta$  will definitely give the result  $\eta'$ . A second measurement of  $\xi$  will definitely give the result  $\xi'$ , and so on. In this sense, we can say that the observables  $\xi$  and  $\eta$  *simultaneously* have the values  $\xi'$  and  $\eta'$ , respectively. Clearly, if all eigenstates of  $\xi$  are also eigenstates of  $\eta$  then it is always possible to make a simultaneous measurement of  $\xi$  and  $\eta$ . Such observables are termed *compatible*.

Suppose, however, that the eigenstates of  $\xi$  are not eigenstates of  $\eta$ . Is it still possible to measure both observables simultaneously? Let us again make an observation of  $\xi$  which throws the system into an eigenstate  $|\xi'\rangle$ , with eigenvalue  $\xi'$ . We can now make a second observation to determine  $\eta$ . This will throw the system into one of the (many) eigenstates of  $\eta$  which depend on  $|\xi'\rangle$ . In principle, each of these eigenstates is associated with a different result of the measurement. Suppose that the system is thrown into an eigenstate  $|\eta'\rangle$ , with the eigenvalue  $\eta'$ . Another measurement of  $\xi$  will throw the system into one of the (many) eigenstates of  $\xi$  which depend on  $|\eta'\rangle$ . Each eigenstate is again associated with a different possible result of the measurement. It is clear that if the observables  $\xi$  and  $\eta$  do not possess simultaneous eigenstates then if the value

of  $\xi$  is known (*i.e.*, the system is in an eigenstate of  $\xi$ ) then the value of  $\eta$  is uncertain (*i.e.*, the system is not in an eigenstate of  $\eta$ ), and *vice versa*. We say that the two observables are *incompatible*.

We have seen that *the condition for two observables  $\xi$  and  $\eta$  to be simultaneously measurable is that they should possess simultaneous eigenstates* (*i.e.*, every eigenstate of  $\xi$  should also be an eigenstate of  $\eta$ ). Suppose that this is the case. Let a general eigenstate of  $\xi$ , with eigenvalue  $\xi'$ , also be an eigenstate of  $\eta$ , with eigenvalue  $\eta'$ . It is convenient to denote this simultaneous eigenstate  $|\xi'\eta'\rangle$ . We have

$$\xi|\xi'\eta'\rangle = \xi'|\xi'\eta'\rangle, \quad (2.63)$$

$$\eta|\xi'\eta'\rangle = \eta'|\xi'\eta'\rangle. \quad (2.64)$$

We can left-multiply the first equation by  $\eta$ , and the second equation by  $\xi$ , and then take the difference. The result is

$$(\xi\eta - \eta\xi)|\xi'\eta'\rangle = |0\rangle \quad (2.65)$$

for each simultaneous eigenstate. Recall that the eigenstates of an observable must form a complete set. It follows that the simultaneous eigenstates of two observables must also form a complete set. Thus, the above equation implies that

$$(\xi\eta - \eta\xi)|A\rangle = |0\rangle, \quad (2.66)$$

where  $|A\rangle$  is a general ket. The only way that this can be true is if

$$\xi\eta = \eta\xi. \quad (2.67)$$

Thus, *the condition for two observables  $\xi$  and  $\eta$  to be simultaneously measurable is that they should commute*.

## 2.14 The uncertainty relation

We have seen that if  $\xi$  and  $\eta$  are two noncommuting observables, then a determination of the value of  $\xi$  leaves the value of  $\eta$  uncertain, and *vice versa*. It is

possible to quantify this uncertainty. For a general observable  $\xi$ , we can define a Hermitian operator

$$\Delta\xi = \xi - \langle \xi \rangle, \quad (2.68)$$

where the expectation value is taken over the particular physical state under consideration. It is obvious that the expectation value of  $\Delta\xi$  is zero. The expectation value of  $(\Delta\xi)^2 \equiv \Delta\xi \Delta\xi$  is termed the *variance* of  $\xi$ , and is, in general, non-zero. In fact, it is easily demonstrated that

$$\langle (\Delta\xi)^2 \rangle = \langle \xi^2 \rangle - \langle \xi \rangle^2. \quad (2.69)$$

The variance of  $\xi$  is a measure of the uncertainty in the value of  $\xi$  for the particular state in question (*i.e.*, it is a measure of the width of the distribution of likely values of  $\xi$  about the expectation value). If the variance is zero then there is no uncertainty, and a measurement of  $\xi$  is bound to give the expectation value,  $\langle \xi \rangle$ .

Consider the Schwarz inequality

$$\langle A|A \rangle \langle B|B \rangle \geq |\langle A|B \rangle|^2, \quad (2.70)$$

which is analogous to

$$|\mathbf{a}|^2 |\mathbf{b}|^2 \geq |\mathbf{a} \cdot \mathbf{b}|^2 \quad (2.71)$$

in Euclidian space. This inequality can be proved by noting that

$$(\langle A| + c^* \langle B|)(|A \rangle + c|B \rangle) \geq 0, \quad (2.72)$$

where  $c$  is any complex number. If  $c$  takes the special value  $-\langle B|A \rangle / \langle B|B \rangle$  then the above inequality reduces to

$$\langle A|A \rangle \langle B|B \rangle - |\langle A|B \rangle|^2 \geq 0, \quad (2.73)$$

which is the same as the Schwarz inequality.

Let us substitute

$$|A \rangle = \Delta\xi | \rangle, \quad (2.74)$$

$$|B \rangle = \Delta\eta | \rangle, \quad (2.75)$$

into the Schwarz inequality, where the blank ket  $|\ \rangle$  stands for any general ket. We find

$$\langle(\Delta\xi)^2\rangle\langle(\Delta\eta)^2\rangle \geq |\langle\Delta\xi\Delta\eta\rangle|^2, \quad (2.76)$$

where use has been made of the fact that  $\Delta\xi$  and  $\Delta\eta$  are Hermitian operators. Note that

$$\Delta\xi\Delta\eta = \frac{1}{2}[\Delta\xi, \Delta\eta] + \frac{1}{2}\{\Delta\xi, \Delta\eta\}, \quad (2.77)$$

where the *commutator*,  $[\Delta\xi, \Delta\eta]$ , and the *anti-commutator*,  $\{\Delta\xi, \Delta\eta\}$ , are defined

$$[\Delta\xi, \Delta\eta] \equiv \Delta\xi\Delta\eta - \Delta\eta\Delta\xi, \quad (2.78)$$

$$\{\Delta\xi, \Delta\eta\} \equiv \Delta\xi\Delta\eta + \Delta\eta\Delta\xi. \quad (2.79)$$

The commutator is clearly *anti-Hermitian*,

$$([\Delta\xi, \Delta\eta])^\dagger = (\Delta\xi\Delta\eta - \Delta\eta\Delta\xi)^\dagger = \Delta\eta\Delta\xi - \Delta\xi\Delta\eta = -[\Delta\xi, \Delta\eta], \quad (2.80)$$

whereas the anti-commutator is obviously Hermitian. Now, it is easily demonstrated that the expectation value of a Hermitian operator is a real number, whereas the expectation value of an anti-Hermitian operator is a pure imaginary number. It is clear that the right hand side of

$$\langle\Delta\xi\Delta\eta\rangle = \frac{1}{2}\langle[\Delta\xi, \Delta\eta]\rangle + \frac{1}{2}\langle\{\Delta\xi, \Delta\eta\}\rangle, \quad (2.81)$$

consists of the sum of a purely real and a purely imaginary number. Taking the modulus squared of both sides gives

$$|\langle\Delta\xi\Delta\eta\rangle|^2 = \frac{1}{4}|\langle[\xi, \eta]\rangle|^2 + \frac{1}{4}|\langle\{\Delta\xi, \Delta\eta\}\rangle|^2, \quad (2.82)$$

where use has been made of  $\langle\Delta\xi\rangle = 0$ , *etc.* The final term in the above expression is positive definite, so we can write

$$\langle(\Delta\xi)^2\rangle\langle(\Delta\eta)^2\rangle \geq \frac{1}{4}|\langle[\xi, \eta]\rangle|^2, \quad (2.83)$$

where use has been made of Eq. (2.76). The above expression is termed the *uncertainty relation*. According to this relation, an exact knowledge of the value of  $\xi$  implies no knowledge whatsoever of the value of  $\eta$ , and *vice versa*. The one exception to this rule is when  $\xi$  and  $\eta$  commute, in which case exact knowledge of  $\xi$  does not necessarily imply no knowledge of  $\eta$ .

## 2.15 Continuous spectra

Up to now, we have studiously avoided dealing with observables possessing eigenvalues which lie in a continuous range, rather than having discrete values. The reason for this is because continuous eigenvalues imply a ket space of nondenumerably infinite dimension. Unfortunately, continuous eigenvalues are unavoidable in quantum mechanics. In fact, the most important observables of all, namely position and momentum, generally have continuous eigenvalues. Fortunately, many of the results we obtained previously for a finite-dimensional ket space with discrete eigenvalues can be generalized to ket spaces of nondenumerably infinite dimensions.

Suppose that  $\xi$  is an observable with continuous eigenvalues. We can still write the eigenvalue equation as

$$\xi|\xi'\rangle = \xi'|\xi'\rangle. \quad (2.84)$$

But,  $\xi'$  can now take a continuous range of values. Let us assume, for the sake of simplicity, that  $\xi'$  can take any value. The orthogonality condition (2.50) generalizes to

$$\langle \xi'|\xi''\rangle = \delta(\xi' - \xi''), \quad (2.85)$$

where  $\delta(x)$  denotes the famous Dirac delta-function. Note that there are clearly a nondenumerably infinite number of mutually orthogonal eigenstates of  $\xi$ . Hence, the dimensionality of ket space is nondenumerably infinite. Note, also, that eigenstates corresponding to a continuous range of eigenvalues *cannot* be normalized so that they have unit norms. In fact, these eigenstates have *infinite* norms: *i.e.*, they are infinitely long. This is the major difference between eigenstates in a finite-dimensional and an infinite-dimensional ket space. The extremely useful relation (2.54) generalizes to

$$\int d\xi' |\xi'\rangle \langle \xi'| = 1. \quad (2.86)$$

Note that a summation over discrete eigenvalues goes over into an integral over a continuous range of eigenvalues. The eigenstates  $|\xi'\rangle$  must form a complete set if  $\xi$  is to be an observable. It follows that any general ket can be expanded in

terms of the  $|\xi'\rangle$ . In fact, the expansions (2.51)–(2.53) generalize to

$$|A\rangle = \int d\xi' |\xi'\rangle \langle \xi'|A\rangle, \quad (2.87)$$

$$\langle A| = \int d\xi' \langle A|\xi'\rangle \langle \xi'|, \quad (2.88)$$

$$\langle A|A\rangle = \int d\xi' \langle A|\xi'\rangle \langle \xi'|A\rangle = \int d\xi' |\langle A|\xi'\rangle|^2. \quad (2.89)$$

These results also follow simply from Eq. (2.86). We have seen that it is not possible to normalize the eigenstates  $|\xi'\rangle$  such that they have unit norms. Fortunately, this convenient normalization is still possible for a general state vector. In fact, according to Eq. (2.89), the normalization condition can be written

$$\langle A|A\rangle = \int d\xi' |\langle A|\xi'\rangle|^2 = 1. \quad (2.90)$$

We have now studied observables whose eigenvalues can take a discrete number of values as well as those whose eigenvalues can take any value. There are number of other cases we could look at. For instance, observables whose eigenvalues can only take a finite range of values, or observables whose eigenvalues take on a finite range of values plus a set of discrete values. Both of these cases can be dealt with using a fairly straight-forward generalization of the previous analysis (see Dirac, Cha. II and III).

## 3 Position and momentum

### 3.1 Introduction

So far, we have considered *general* dynamical variables represented by *general* linear operators acting in ket space. However, in classical mechanics the most important dynamical variables are those involving position and momentum. Let us investigate the role of such variables in quantum mechanics.

In classical mechanics, the position  $q$  and momentum  $p$  of some component of a dynamical system are represented as *real numbers* which, by definition, commute. In quantum mechanics, these quantities are represented as *noncommuting* linear Hermitian operators acting in a ket space which represents all of the possible states of the system. Our first task is to discover a quantum mechanical replacement for the classical result  $q p - p q = 0$ . Do the position and momentum operators commute? If not, what is the value of  $q p - p q$ ?

### 3.2 Poisson brackets

Consider a dynamic system whose state at a particular time  $t$  is fully specified by  $N$  independent classical coordinates  $q_i$  (where  $i$  runs from 1 to  $N$ ). Associated with each generalized coordinate  $q_i$  is a classical canonical momentum  $p_i$ . For instance, a Cartesian coordinate has an associated linear momentum, an angular coordinate has an associated angular momentum, *etc.* As is well-known, the behaviour of a classical system can be specified in terms of Lagrangian or Hamiltonian dynamics. For instance, in Hamiltonian dynamics,

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad (3.1)$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad (3.2)$$

where the function  $H(q_i, p_i, t)$  is the energy of the system at time  $t$  expressed in terms of the classical coordinates and canonical momenta. This function is

usually referred to as the *Hamiltonian* of the system.

We are interested in finding some construct of classical dynamics which consists of *products* of dynamical variables. If such a construct exists we hope to generalize it somehow to obtain a rule describing how dynamical variables commute with one another in quantum mechanics. There is, indeed, one well-known construct in classical dynamics which involves products of dynamical variables. The *Poisson bracket* of two dynamical variables  $u$  and  $v$  is defined

$$[u, v] = \sum_{i=1}^N \left( \frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial v}{\partial q_i} \right), \quad (3.3)$$

where  $u$  and  $v$  are regarded as functions of the coordinates and momenta  $q_i$  and  $p_i$ . It is easily demonstrated that

$$[q_i, q_j] = 0, \quad (3.4)$$

$$[p_i, p_j] = 0, \quad (3.5)$$

$$[q_i, p_j] = \delta_{ij}. \quad (3.6)$$

The time evolution of a dynamical variable can also be written in terms of a Poisson bracket by noting that

$$\begin{aligned} \frac{du}{dt} &= \sum_{i=1}^N \left( \frac{\partial u}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial u}{\partial p_i} \frac{dp_i}{dt} \right) \\ &= \sum_{i=1}^N \left( \frac{\partial u}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial H}{\partial q_i} \right) \\ &= [u, H], \end{aligned} \quad (3.7)$$

where use has been made of Hamilton's equations.

Can we construct a quantum mechanical Poisson bracket in which  $u$  and  $v$  are noncommuting operators, instead of functions? Well, the main properties of the classical Poisson bracket are as follows:

$$[u, v] = -[v, u], \quad (3.8)$$

$$[\mathbf{u}, c] = 0, \quad (3.9)$$

$$[\mathbf{u}_1 + \mathbf{u}_2, \mathbf{v}] = [\mathbf{u}_1, \mathbf{v}] + [\mathbf{u}_2, \mathbf{v}], \quad (3.10)$$

$$[\mathbf{u}, \mathbf{v}_1 + \mathbf{v}_2] = [\mathbf{u}, \mathbf{v}_1] + [\mathbf{u}, \mathbf{v}_2] \quad (3.11)$$

$$[\mathbf{u}_1 \mathbf{u}_2, \mathbf{v}] = [\mathbf{u}_1, \mathbf{v}] \mathbf{u}_2 + \mathbf{u}_1 [\mathbf{u}_2, \mathbf{v}], \quad (3.12)$$

$$[\mathbf{u}, \mathbf{v}_1 \mathbf{v}_2] = [\mathbf{u}, \mathbf{v}_1] \mathbf{v}_2 + \mathbf{v}_1 [\mathbf{u}, \mathbf{v}_2], \quad (3.13)$$

and

$$[\mathbf{u}, [\mathbf{v}, \mathbf{w}]] + [\mathbf{v}, [\mathbf{w}, \mathbf{u}]] + [\mathbf{w}, [\mathbf{u}, \mathbf{v}]] = 0. \quad (3.14)$$

The last relation is known as the *Jacobi identity*. In the above,  $\mathbf{u}$ ,  $\mathbf{v}$ ,  $\mathbf{w}$ , *etc.*, represent dynamical variables, and  $c$  represents a number. Can we find some combination of noncommuting operators  $\mathbf{u}$  and  $\mathbf{v}$ , *etc.*, which satisfies all of the above relations?

Well, we can evaluate the Poisson bracket  $[\mathbf{u}_1 \mathbf{u}_2, \mathbf{v}_1 \mathbf{v}_2]$  in two different ways, since we can use either of the formulae (3.12) or (3.13) first. Thus,

$$\begin{aligned} [\mathbf{u}_1 \mathbf{u}_2, \mathbf{v}_1 \mathbf{v}_2] &= [\mathbf{u}_1, \mathbf{v}_1 \mathbf{v}_2] \mathbf{u}_2 + \mathbf{u}_1 [\mathbf{u}_2, \mathbf{v}_1 \mathbf{v}_2] \quad (3.15) \\ &= \{[\mathbf{u}_1, \mathbf{v}_1] \mathbf{v}_2 + \mathbf{v}_1 [\mathbf{u}_1, \mathbf{v}_2]\} \mathbf{u}_2 + \mathbf{u}_1 \{[\mathbf{u}_2, \mathbf{v}_1] \mathbf{v}_2 + \mathbf{v}_1 [\mathbf{u}_2, \mathbf{v}_2]\} \\ &= [\mathbf{u}_1, \mathbf{v}_1] \mathbf{v}_2 \mathbf{u}_2 + \mathbf{v}_1 [\mathbf{u}_1, \mathbf{v}_2] \mathbf{u}_2 + \mathbf{u}_1 [\mathbf{u}_2, \mathbf{v}_1] \mathbf{v}_2 + \mathbf{u}_1 \mathbf{v}_1 [\mathbf{u}_2, \mathbf{v}_2], \end{aligned}$$

and

$$\begin{aligned} [\mathbf{u}_1 \mathbf{u}_2, \mathbf{v}_1 \mathbf{v}_2] &= [\mathbf{u}_1 \mathbf{u}_2, \mathbf{v}_1] \mathbf{v}_2 + \mathbf{v}_1 [\mathbf{u}_1 \mathbf{u}_2, \mathbf{v}_2] \quad (3.16) \\ &= [\mathbf{u}_1, \mathbf{v}_1] \mathbf{u}_2 \mathbf{v}_2 + \mathbf{u}_1 [\mathbf{u}_2, \mathbf{v}_1] \mathbf{v}_2 + \mathbf{v}_1 [\mathbf{u}_1, \mathbf{v}_2] \mathbf{u}_2 + \mathbf{v}_1 \mathbf{u}_1 [\mathbf{u}_2, \mathbf{v}_2]. \end{aligned}$$

Note that the order of the various factors has been preserved, since they now represent *noncommuting* operators. Equating the above two results yields

$$[\mathbf{u}_1, \mathbf{v}_1] (\mathbf{u}_2 \mathbf{v}_2 - \mathbf{v}_2 \mathbf{u}_2) = (\mathbf{u}_1 \mathbf{v}_1 - \mathbf{v}_1 \mathbf{u}_1) [\mathbf{u}_2, \mathbf{v}_2]. \quad (3.17)$$

Since this relation must hold for  $\mathbf{u}_1$  and  $\mathbf{v}_1$  quite independent of  $\mathbf{u}_2$  and  $\mathbf{v}_2$ , it follows that

$$\mathbf{u}_1 \mathbf{v}_1 - \mathbf{v}_1 \mathbf{u}_1 = i \hbar [\mathbf{u}_1, \mathbf{v}_1], \quad (3.18)$$

$$\mathbf{u}_2 \mathbf{v}_2 - \mathbf{v}_2 \mathbf{u}_2 = i \hbar [\mathbf{u}_2, \mathbf{v}_2], \quad (3.19)$$

where  $\hbar$  does not depend on  $u_1, v_1, u_2, v_2$ , and also commutes with  $(u_1 v_1 - v_1 u_1)$ . Since  $u_1$ , etc., are quite general operators, it follows that  $\hbar$  is just a number. We want the quantum mechanical Poisson bracket of two Hermitian operators to be an Hermitian operator itself, since the classical Poisson bracket of two real dynamical variables is real. This requirement is satisfied if  $\hbar$  is a real number. Thus, the quantum mechanical Poisson bracket of two dynamical variables  $u$  and  $v$  is given by

$$[u, v] = \frac{u v - v u}{i \hbar}, \quad (3.20)$$

where  $\hbar$  is a new universal constant of nature. Quantum mechanics agrees with experiments provided that  $\hbar$  takes the value  $h/2\pi$ , where

$$h = 6.6261 \times 10^{-34} \text{ Js} \quad (3.21)$$

is *Planck's constant*. Somewhat confusingly, the notation  $[u, v]$  is conventionally reserved for the commutator  $u v - v u$  in quantum mechanics. We will use  $[u, v]_{\text{quantum}}$  to denote the quantum Poisson bracket. Thus,

$$[u, v]_{\text{quantum}} = \frac{[u, v]}{i \hbar}. \quad (3.22)$$

It is easily demonstrated that the quantum mechanical Poisson bracket, as defined above, satisfies all of the relations (3.8)–(3.14).

The strong analogy we have found between the classical Poisson bracket, defined in Eq. (3.3), and the quantum mechanical Poisson bracket, defined in Eq. (3.22), leads us to make the assumption that the quantum mechanical bracket has the same value as the corresponding classical bracket, at least for the simplest cases. In other words, we are assuming that Eqs. (3.4)–(3.6) hold for quantum mechanical as well as classical Poisson brackets. This argument yields the fundamental commutation relations

$$[q_i, q_j] = 0, \quad (3.23)$$

$$[p_i, p_j] = 0, \quad (3.24)$$

$$[q_i, p_j] = i \hbar \delta_{ij}. \quad (3.25)$$

These results provide us with the basis for calculating commutation relations between general dynamical variables. For instance, if two dynamical variables,  $\xi$

and  $\eta$ , can both be written as a power series in the  $q_i$  and  $p_i$ , then repeated application of Eqs. (3.8)–(3.13) allows  $[\xi, \eta]$  to be expressed in terms of the fundamental commutation relations (3.23)–(3.25).

Equations (3.23)–(3.25) provide the foundation for the analogy between quantum mechanics and classical mechanics. Note that the classical result (that everything commutes) is obtained in the limit  $\hbar \rightarrow 0$ . Thus, *classical mechanics can be regarded as the limiting case of quantum mechanics when  $\hbar$  goes to zero*. In classical mechanics, each pair of generalized coordinate and its conjugate momentum,  $q_i$  and  $p_i$ , correspond to a different classical degree of freedom of the system. It is clear from Eqs. (3.23)–(3.25) that in quantum mechanics the *dynamical variables corresponding to different degrees of freedom all commute*. It is only those variables corresponding to the same degree of freedom which may fail to commute.

### 3.3 Wave-functions

Consider a simple system with one classical degree of freedom, which corresponds to the Cartesian coordinate  $x$ . Suppose that  $x$  is free to take any value (e.g.,  $x$  could be the position of a free particle). The classical dynamical variable  $x$  is represented in quantum mechanics as a linear Hermitian operator which is also called  $x$ . Moreover, the operator  $x$  possesses eigenvalues  $x'$  lying in the *continuous* range  $-\infty < x' < +\infty$  (since the eigenvalues correspond to all the possible results of a measurement of  $x$ ). We can span ket space using the suitably normalized eigenkets of  $x$ . An eigenket corresponding to the eigenvalue  $x'$  is denoted  $|x'\rangle$ . Moreover, [see Eq. (2.85)]

$$\langle x'|x''\rangle = \delta(x' - x''). \quad (3.26)$$

The eigenkets satisfy the extremely useful relation [see Eq. (2.86)]

$$\int_{-\infty}^{+\infty} dx' |x'\rangle \langle x'| = 1. \quad (3.27)$$

This formula expresses the fact that the eigenkets are complete, mutually orthogonal, and suitably normalized.

A state ket  $|A\rangle$  (which represents a general state  $A$  of the system) can be expressed as a linear superposition of the eigenkets of the position operator using Eq. (3.27). Thus,

$$|A\rangle = \int_{-\infty}^{+\infty} dx' \langle x'|A\rangle |x'\rangle \quad (3.28)$$

The quantity  $\langle x'|A\rangle$  is a complex function of the position eigenvalue  $x'$ . We can write

$$\langle x'|A\rangle = \psi_A(x'). \quad (3.29)$$

Here,  $\psi_A(x')$  is the famous *wave-function* of quantum mechanics. Note that state  $A$  is completely specified by its wave-function  $\psi_A(x')$  [since the wave-function can be used to reconstruct the state ket  $|A\rangle$  using Eq. (3.28)]. It is clear that the wave-function of state  $A$  is simply the collection of the weights of the corresponding state ket  $|A\rangle$ , when it is expanded in terms of the eigenkets of the position operator. Recall, from Sect. 2.10, that the probability of a measurement of a dynamical variable  $\xi$  yielding the result  $\xi'$  when the system is in state  $A$  is given by  $|\langle \xi'|A\rangle|^2$ , assuming that the eigenvalues of  $\xi$  are discrete. This result is easily generalized to dynamical variables possessing continuous eigenvalues. In fact, the probability of a measurement of  $x$  yielding a result lying in the range  $x'$  to  $x' + dx'$  when the system is in a state  $|A\rangle$  is  $|\langle x'|A\rangle|^2 dx'$ . In other words, the probability of a measurement of position yielding a result in the range  $x'$  to  $x' + dx'$  when the wave-function of the system is  $\psi_A(x')$  is

$$P(x', dx') = |\psi_A(x')|^2 dx'. \quad (3.30)$$

This formula is only valid if the state ket  $|A\rangle$  is properly normalized: *i.e.*, if  $\langle A|A\rangle = 1$ . The corresponding normalization for the wave-function is

$$\int_{-\infty}^{+\infty} |\psi_A(x')|^2 dx' = 1. \quad (3.31)$$

Consider a second state  $B$  represented by a state ket  $|B\rangle$  and a wave-function  $\psi_B(x')$ . The inner product  $\langle B|A\rangle$  can be written

$$\langle B|A\rangle = \int_{-\infty}^{+\infty} dx' \langle B|x'\rangle \langle x'|A\rangle = \int_{-\infty}^{+\infty} \psi_B^*(x') \psi_A(x') dx', \quad (3.32)$$

where use has been made of Eqs. (3.27) and (3.29). Thus, the inner product of two states is related to the overlap integral of their wave-functions.

Consider a general function  $f(x)$  of the observable  $x$  [e.g.,  $f(x) = x^2$ ]. If  $|B\rangle = f(x)|A\rangle$  then it follows that

$$\begin{aligned}\psi_B(x') &= \langle x'|f(x) \int_{-\infty}^{+\infty} dx'' \psi_A(x'')|x''\rangle \\ &= \int_{-\infty}^{+\infty} dx'' f(x'') \psi_A(x'') \langle x'|x''\rangle,\end{aligned}\quad (3.33)$$

giving

$$\psi_B(x') = f(x') \psi_A(x'), \quad (3.34)$$

where use has been made of Eq. (3.26). Here,  $f(x')$  is the same function of the position eigenvalue  $x'$  that  $f(x)$  is of the position operator  $x$ : *i.e.*, if  $f(x) = x^2$  then  $f(x') = x'^2$ . It follows, from the above result, that a general state ket  $|A\rangle$  can be written

$$|A\rangle = \psi_A(x)\rangle, \quad (3.35)$$

where  $\psi_A(x)$  is the same function of the operator  $x$  that the wave-function  $\psi_A(x')$  is of the position eigenvalue  $x'$ , and the ket  $\rangle$  has the wave-function  $\psi(x') = 1$ . The ket  $\rangle$  is termed the *standard ket*. The dual of the standard ket is termed the *standard bra*, and is denoted  $\langle$ . It is easily seen that

$$\langle \psi_A^*(x) \leftarrow \overset{DC}{\rightarrow} \psi_A(x) \rangle. \quad (3.36)$$

Note, finally, that  $\psi_A(x)\rangle$  is often shortened to  $\psi_A\rangle$ , leaving the dependence on the position operator  $x$  tacitly understood.

### 3.4 Schrödinger's representation - I

Consider the simple system described in the previous section. A general state ket can be written  $\psi(x)\rangle$ , where  $\psi(x)$  is a general function of the position operator  $x$ , and  $\psi(x')$  is the associated wave-function. Consider the ket whose wave-function

is  $d\psi(x')/dx'$ . This ket is denoted  $d\psi/dx\rangle$ . The new ket is clearly a linear function of the original ket, so we can think of it as the result of some linear operator acting on  $\psi\rangle$ . Let us denote this operator  $d/dx$ . It follows that

$$\frac{d}{dx}\psi\rangle = \frac{d\psi}{dx}\rangle. \quad (3.37)$$

Any linear operator which acts on ket vectors can also act on bra vectors. Consider  $d/dx$  acting on a general bra  $\langle\phi(x)$ . According to Eq. (2.34), the bra  $\langle\phi d/dx$  satisfies

$$\left(\langle\phi\frac{d}{dx}\right)\psi\rangle = \langle\phi\left(\frac{d}{dx}\psi\rangle\right). \quad (3.38)$$

Making use of Eqs. (3.27) and (3.29), we can write

$$\int_{-\infty}^{+\infty} \langle\phi\frac{d}{dx}|x'\rangle dx' \psi(x') = \int_{-\infty}^{+\infty} \phi(x') dx' \frac{d\psi(x')}{dx'}. \quad (3.39)$$

The right-hand side can be transformed via integration by parts to give

$$\int_{-\infty}^{+\infty} \langle\phi\frac{d}{dx}|x'\rangle dx' \psi(x') = - \int_{-\infty}^{+\infty} \frac{d\phi(x')}{dx'} dx' \psi(x'), \quad (3.40)$$

assuming that the contributions from the limits of integration vanish. It follows that

$$\langle\phi\frac{d}{dx}|x'\rangle = -\frac{d\phi(x')}{dx'}, \quad (3.41)$$

which implies

$$\langle\phi\frac{d}{dx} = -\langle\frac{d\phi}{dx}. \quad (3.42)$$

The neglect of contributions from the limits of integration in Eq. (3.40) is reasonable because physical wave-functions are square-integrable [see Eq. (3.31)].

Note that

$$\frac{d}{dx}\psi\rangle = \frac{d\psi}{dx}\rangle \xrightarrow{\text{DC}} \langle\frac{d\psi^*}{dx} = -\langle\psi^*\frac{d}{dx}, \quad (3.43)$$

where use has been made of Eq. (3.42). It follows, by comparison with Eqs. (2.35) and (3.36), that

$$\left(\frac{d}{dx}\right)^\dagger = -\frac{d}{dx}. \quad (3.44)$$

Thus,  $d/dx$  is an *anti-Hermitian* operator.

Let us evaluate the commutation relation between the operators  $x$  and  $d/dx$ . We have

$$\frac{d}{dx}x\psi\rangle = \frac{d(x\psi)}{dx}\rangle = x\frac{d}{dx}\psi\rangle + \psi\rangle. \quad (3.45)$$

Since this holds for any ket  $\psi\rangle$ , it follows that

$$\frac{d}{dx}x - x\frac{d}{dx} = 1. \quad (3.46)$$

Let  $p$  be the momentum conjugate to  $x$  (for the simple system under consideration  $p$  is a straight-forward linear momentum). According to Eq. (3.25),  $x$  and  $p$  satisfy the commutation relation

$$xp - px = i\hbar. \quad (3.47)$$

It can be seen, by comparison with Eq. (3.46), that the Hermitian operator  $-i\hbar d/dx$  satisfies the same commutation relation with  $x$  that  $p$  does. The most general conclusion which may be drawn from a comparison of Eqs. (3.46) and (3.47) is that

$$p = -i\hbar\frac{d}{dx} + f(x), \quad (3.48)$$

since (as is easily demonstrated) a general function  $f(x)$  of the position operator automatically commutes with  $x$ .

We have chosen to normalize the eigenkets and eigenbras of the position operator so that they satisfy the normalization condition (3.26). However, this choice of normalization does not uniquely determine the eigenkets and eigenbras. Suppose that we transform to a new set of eigenbras which are related to the old set via

$$\langle x' |_{\text{new}} = e^{i\gamma'} \langle x' |_{\text{old}}, \quad (3.49)$$

where  $\gamma' \equiv \gamma(x')$  is a real function of  $x'$ . This transformation amounts to a rearrangement of the relative phases of the eigenbras. The new normalization condition is

$$\begin{aligned} \langle x' | x'' \rangle_{\text{new}} &= \langle x' | e^{i\gamma'} e^{-i\gamma''} | x'' \rangle_{\text{old}} = e^{i(\gamma' - \gamma'')} \langle x' | x'' \rangle_{\text{old}} \\ &= e^{i(\gamma' - \gamma'')} \delta(x' - x'') = \delta(x' - x''). \end{aligned} \quad (3.50)$$

Thus, the new eigenbras satisfy the same normalization condition as the old eigenbras.

By definition, the standard ket  $|x\rangle$  satisfies  $\langle x'|x\rangle = \delta(x-x')$ . It follows from Eq. (3.49) that the new standard ket is related to the old standard ket via

$$|x\rangle_{\text{new}} = e^{-i\gamma} |x\rangle_{\text{old}}, \quad (3.51)$$

where  $\gamma \equiv \gamma(x)$  is a real function of the position operator  $x$ . The dual of the above equation yields the transformation rule for the standard bra,

$$\langle x|_{\text{new}} = \langle x|_{\text{old}} e^{i\gamma}. \quad (3.52)$$

The transformation rule for a general operator  $A$  follows from Eqs. (3.51) and (3.52), plus the requirement that the triple product  $\langle A \rangle$  remain invariant (this must be the case, otherwise the probability of a measurement yielding a certain result would depend on the choice of eigenbras). Thus,

$$A_{\text{new}} = e^{-i\gamma} A_{\text{old}} e^{i\gamma}. \quad (3.53)$$

Of course, if  $A$  commutes with  $x$  then  $A$  is invariant under the transformation. In fact,  $d/dx$  is the only operator (we know of) which does not commute with  $x$ , so Eq. (3.53) yields

$$\left(\frac{d}{dx}\right)_{\text{new}} = e^{-i\gamma} \frac{d}{dx} e^{i\gamma} = \frac{d}{dx} + i \frac{d\gamma}{dx}, \quad (3.54)$$

where the subscript “old” is taken as read. It follows, from Eq. (3.48), that the momentum operator  $p$  can be written

$$p = -i \hbar \left(\frac{d}{dx}\right)_{\text{new}} - \hbar \frac{d\gamma}{dx} + f(x). \quad (3.55)$$

Thus, the special choice

$$\hbar \gamma(x) = \int^x f(x) dx \quad (3.56)$$

yields

$$p = -i \hbar \left(\frac{d}{dx}\right)_{\text{new}}. \quad (3.57)$$

Equation (3.56) fixes  $\gamma$  to within an arbitrary additive constant: *i.e.*, the special eigenkets and eigenbras for which Eq. (3.57) is true are determined to within an arbitrary common phase-factor.

In conclusion, it is possible to find a set of basis eigenkets and eigenbras of the position operator  $x$  which satisfy the normalization condition (3.26), and for which the momentum conjugate to  $x$  can be represented as the operator

$$p = -i \hbar \frac{d}{dx}. \quad (3.58)$$

A general state ket is written  $\psi(x)\rangle$ , where the standard ket  $\rangle$  satisfies  $\langle x'|\rangle = 1$ , and where  $\psi(x') = \langle x'|\psi(x)\rangle$  is the wave-function. This scheme of things is known as *Schrödinger's representation*, and is the basis of wave mechanics.

### 3.5 Schrödinger's representation - II

In the preceding sections, we have developed Schrödinger's representation for the case of a single operator  $x$  corresponding to a classical Cartesian coordinate. However, this scheme can easily be extended. Consider a system with  $N$  generalized coordinates,  $q_1 \cdots q_N$ , which can all be *simultaneously* measured. These are represented as  $N$  *commuting* operators,  $q_1 \cdots q_N$ , each with a continuous range of eigenvalues,  $q'_1 \cdots q'_N$ . Ket space is conveniently spanned by the simultaneous eigenkets of  $q_1 \cdots q_N$ , which are denoted  $|q'_1 \cdots q'_N\rangle$ . These eigenkets must form a complete set, otherwise the  $q_1 \cdots q_N$  would not be simultaneously observable.

The orthogonality condition for the eigenkets [*i.e.*, the generalization of Eq. (3.26)] is

$$\langle q'_1 \cdots q'_N | q''_1 \cdots q''_N \rangle = \delta(q'_1 - q''_1) \delta(q'_2 - q''_2) \cdots \delta(q'_N - q''_N). \quad (3.59)$$

The completeness condition [*i.e.*, the generalization of Eq. (3.27)] is

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} dq'_1 \cdots dq'_N |q'_1 \cdots q'_N\rangle \langle q'_1 \cdots q'_N| = 1. \quad (3.60)$$

The standard ket  $\rangle$  is defined such that

$$\langle q'_1 \cdots q'_N | \rangle = 1. \quad (3.61)$$

The standard bra  $\langle$  is the dual of the standard ket. A general state ket is written

$$\psi(q_1 \cdots q_N)\rangle. \quad (3.62)$$

The associated wave-function is

$$\psi(q'_1 \cdots q'_N) = \langle q'_1 \cdots q'_N | \psi \rangle. \quad (3.63)$$

Likewise, a general state bra is written

$$\langle \phi(q_1 \cdots q_N), \quad (3.64)$$

where

$$\phi(q'_1 \cdots q'_N) = \langle \phi | q'_1 \cdots q'_N \rangle. \quad (3.65)$$

The probability of an observation of the system finding the first coordinate in the range  $q'_1$  to  $q'_1 + dq'_1$ , the second coordinate in the range  $q'_2$  to  $q'_2 + dq'_2$ , etc., is

$$P(q'_1 \cdots q'_N; dq'_1 \cdots dq'_N) = |\psi(q'_1 \cdots q'_N)|^2 dq'_1 \cdots dq'_N. \quad (3.66)$$

Finally, the normalization condition for a physical wave-function is

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} |\psi(q'_1 \cdots q'_N)|^2 dq'_1 \cdots dq'_N = 1. \quad (3.67)$$

The  $N$  linear operators  $\partial/\partial q_i$  (where  $i$  runs from 1 to  $N$ ) are defined

$$\frac{\partial}{\partial q_i} \psi \rangle = \frac{\partial \psi}{\partial q_i} \rangle. \quad (3.68)$$

These linear operators can also act on bras (provided the associated wave-functions are square integrable) in accordance with [see Eq. (3.42)]

$$\langle \phi \frac{\partial}{\partial q_i} = - \langle \frac{\partial \phi}{\partial q_i}. \quad (3.69)$$

Corresponding to Eq. (3.46), we can derive the commutation relations

$$\frac{\partial}{\partial q_i} q_j - q_j \frac{\partial}{\partial q_i} = \delta_{ij}. \quad (3.70)$$

It is also clear that

$$\frac{\partial}{\partial q_i} \frac{\partial}{\partial q_j} \psi \rangle = \frac{\partial^2 \psi}{\partial q_i \partial q_j} \rangle = \frac{\partial}{\partial q_j} \frac{\partial}{\partial q_i} \psi \rangle, \quad (3.71)$$

showing that

$$\frac{\partial}{\partial q_i} \frac{\partial}{\partial q_j} = \frac{\partial}{\partial q_j} \frac{\partial}{\partial q_i}. \quad (3.72)$$

It can be seen, by comparison with Eqs. (3.23)–(3.25), that the linear operators  $-i \hbar \partial / \partial q_i$  satisfy the same commutation relations with the  $q$ 's and with each other that the  $p$ 's do. The most general conclusion we can draw from this coincidence of commutation relations is (see Dirac)

$$p_i = -i \hbar \frac{\partial}{\partial q_i} + \frac{\partial F(q_1 \cdots q_N)}{\partial q_i}. \quad (3.73)$$

However, the function  $F$  can be transformed away via a suitable readjustment of the phases of the basis eigenkets (see Sect. 3.4, and Dirac). Thus, we can always construct a set of simultaneous eigenkets of  $q_1 \cdots q_N$  for which

$$p_i = -i \hbar \frac{\partial}{\partial q_i}. \quad (3.74)$$

This is the generalized Schrödinger representation.

It follows from Eqs. (3.61), (3.68), and (3.74) that

$$p_i \rangle = 0. \quad (3.75)$$

Thus, the standard ket in Schrödinger's representation is a simultaneous eigenket of all the momentum operators belonging to the eigenvalue zero. Note that

$$\langle q'_1 \cdots q'_N | \frac{\partial}{\partial q_i} \psi \rangle = \langle q'_1 \cdots q'_N | \frac{\partial \psi}{\partial q_i} \rangle = \frac{\partial \psi(q'_1 \cdots q'_N)}{\partial q'_i} = \frac{\partial}{\partial q'_i} \langle q'_1 \cdots q'_N | \psi \rangle. \quad (3.76)$$

Hence,

$$\langle q'_1 \cdots q'_N | \frac{\partial}{\partial q_i} = \frac{\partial}{\partial q'_i} \langle q'_1 \cdots q'_N |, \quad (3.77)$$

so that

$$\langle q'_1 \cdots q'_N | p_i = -i \hbar \frac{\partial}{\partial q'_i} \langle q'_1 \cdots q'_N |. \quad (3.78)$$

The dual of the above equation gives

$$p_i |q'_1 \cdots q'_N\rangle = i \hbar \frac{\partial}{\partial q'_i} |q'_1 \cdots q'_N\rangle. \quad (3.79)$$

### 3.6 The momentum representation

Consider a system with one degree of freedom, describable in terms of a coordinate  $x$  and its conjugate momentum  $p$ , both of which have a continuous range of eigenvalues. We have seen that it is possible to represent the system in terms of the eigenkets of  $x$ . This is termed Schrödinger's representation. However, it is also possible to represent the system in terms of the eigenkets of  $p$ .

Consider the eigenkets of  $p$  which belong to the eigenvalues  $p'$ . These are denoted  $|p'\rangle$ . The orthogonality relation for the momentum eigenkets is

$$\langle p' | p'' \rangle = \delta(p' - p''), \quad (3.80)$$

and the corresponding completeness relation is

$$\int_{-\infty}^{+\infty} dp' |p'\rangle \langle p'| = 1. \quad (3.81)$$

A general state ket can be written

$$|\phi\rangle = \int dp' |p'\rangle \phi(p') \quad (3.82)$$

where the standard ket  $|p'\rangle$  satisfies

$$\langle p' | p' \rangle = 1. \quad (3.83)$$

Note that the standard ket in this representation is quite different to that in Schrödinger's representation. The momentum space wave-function  $\phi(p')$  satisfies

$$\phi(p') = \langle p' | \phi \rangle. \quad (3.84)$$

The probability that a measurement of the momentum yields a result lying in the range  $p'$  to  $p' + dp'$  is given by

$$P(p', dp') = |\phi(p')|^2 dp'. \quad (3.85)$$

Finally, the normalization condition for a physical momentum space wave-function is

$$\int_{-\infty}^{+\infty} |\phi(p')|^2 dp' = 1. \quad (3.86)$$

The fundamental commutation relations (3.23)–(3.25) exhibit a particular symmetry between coordinates and their conjugate momenta. If all the coordinates are transformed into their conjugate momenta, and *vice versa*, and  $i$  is then replaced by  $-i$ , the commutation relations are unchanged. It follows from this symmetry that we can always choose the eigenkets of  $p$  in such a manner that the coordinate  $x$  can be represented as (see Sect. 3.4)

$$x = i \hbar \frac{d}{dp}. \quad (3.87)$$

This is termed the momentum representation.

The above result is easily generalized to a system with more than one degree of freedom. Suppose the system is specified by  $N$  coordinates,  $q_1 \cdots q_N$ , and  $N$  conjugate momenta,  $p_1 \cdots p_N$ . Then, in the momentum representation, the coordinates can be written as

$$q_i = i \hbar \frac{\partial}{\partial p_i}. \quad (3.88)$$

We also have

$$q_i | \rangle = 0, \quad (3.89)$$

and

$$\langle p'_1 \cdots p'_N | q_i = i \hbar \frac{\partial}{\partial p'_i} \langle p'_1 \cdots p'_N |. \quad (3.90)$$

The momentum representation is less useful than Schrödinger's representation for a very simple reason. The energy operator (*i.e.*, the Hamiltonian) of most simple systems takes the form of a sum of quadratic terms in the momenta (*i.e.*, the kinetic energy) plus a complicated function of the coordinates (*i.e.*, the potential energy). In Schrödinger's representation, the eigenvalue problem for the energy translates into a second-order differential equation in the coordinates, with a complicated potential function. In the momentum representation, the

problem transforms into a high-order differential equation in the momenta, with a quadratic potential. With the mathematical tools at our disposal, we are far better able to solve the former type of problem than the latter. Hence, Schrödinger's representation is generally more useful than the momentum representation.

### 3.7 The uncertainty relation

How is a momentum space wave-function related to the corresponding coordinate space wave-function? To answer this question, let us consider the representative  $\langle x'|p' \rangle$  of the momentum eigenkets  $|p' \rangle$  in Schrödinger's representation for a system with a single degree of freedom. This representative satisfies

$$p' \langle x'|p' \rangle = \langle x'|p|p' \rangle = -i \hbar \frac{d}{dx'} \langle x'|p' \rangle, \quad (3.91)$$

where use has been made of Eq. (3.78) (for the case of a system with one degree of freedom). The solution of the above differential equation is

$$\langle x'|p' \rangle = c' \exp(ip'x'/\hbar), \quad (3.92)$$

where  $c' = c'(p')$ . It is easily demonstrated that

$$\langle p'|p'' \rangle = \int_{-\infty}^{+\infty} \langle p'|x' \rangle dx' \langle x'|p'' \rangle = c'^* c'' \int_{-\infty}^{\infty} \exp[-i(p' - p'')x'/\hbar] dx'. \quad (3.93)$$

The well-known mathematical result

$$\int_{-\infty}^{+\infty} \exp(i a x) dx = 2\pi \delta(a), \quad (3.94)$$

yields

$$\langle p'|p'' \rangle = |c'|^2 \hbar \delta(p' - p''). \quad (3.95)$$

This is consistent with Eq. (3.80), provided that  $c' = \hbar^{-1/2}$ . Thus,

$$\langle x'|p' \rangle = \hbar^{-1/2} \exp(ip'x'/\hbar). \quad (3.96)$$

Consider a general state ket  $|A\rangle$  whose coordinate wave-function is  $\psi(x')$ , and whose momentum wave-function is  $\Psi(p')$ . In other words,

$$\psi(x') = \langle x'|A\rangle, \quad (3.97)$$

$$\Psi(p') = \langle p'|A\rangle. \quad (3.98)$$

It is easily demonstrated that

$$\begin{aligned} \psi(x') &= \int_{-\infty}^{+\infty} dp' \langle x'|p'\rangle \langle p'|A\rangle \\ &= \frac{1}{\hbar^{1/2}} \int_{-\infty}^{+\infty} \Psi(p') \exp(ip'x'/\hbar) dp' \end{aligned} \quad (3.99)$$

and

$$\begin{aligned} \Psi(p') &= \int_{-\infty}^{+\infty} dx' \langle p'|x'\rangle \langle x'|A\rangle \\ &= \frac{1}{\hbar^{1/2}} \int_{-\infty}^{+\infty} \psi(x') \exp(-ip'x'/\hbar) dx', \end{aligned} \quad (3.100)$$

where use has been made of Eqs. (3.27), (3.81), (3.94), and (3.96). Clearly, the momentum space wave-function is the *Fourier transform* of the coordinate space wave-function.

Consider a state whose coordinate space wave-function is a *wave-packet*. In other words, the wave-function only has non-negligible amplitude in some spatially localized region of extent  $\Delta x$ . As is well-known, the Fourier transform of a wave-packet fills up a wave-number band of approximate extent  $\delta k \sim 1/\Delta x$ . Note that in Eq. (3.99) the role of the wave-number  $k$  is played by the quantity  $p'/\hbar$ . It follows that the momentum space wave-function corresponding to a wave-packet in coordinate space extends over a range of momenta  $\Delta p \sim \hbar/\Delta x$ . Clearly, a measurement of  $x$  is almost certain to give a result lying in a range of width  $\Delta x$ . Likewise, measurement of  $p$  is almost certain to yield a result lying in a range of width  $\Delta p$ . The product of these two uncertainties is

$$\Delta x \Delta p \sim \hbar. \quad (3.101)$$

This result is called *Heisenberg's uncertainty principle*.

Actually, it is possible to write Heisenberg's uncertainty principle more exactly by making use of Eq. (2.83) and the commutation relation (3.47). We obtain

$$\langle(\Delta x)^2\rangle\langle(\Delta p)^2\rangle \geq \frac{\hbar^2}{4} \quad (3.102)$$

for any general state. It is easily demonstrated that the minimum uncertainty states, for which the equality sign holds in the above relation, correspond to Gaussian wave-packets in both coordinate and momentum space.

### 3.8 Displacement operators

Consider a system with one degree of freedom corresponding to the Cartesian coordinate  $x$ . Suppose that we displace this system some distance along the  $x$ -axis. We could imagine that the system is on wheels, and we just give it a little push. The final state of the system is completely determined by its initial state, together with the direction and magnitude of the displacement. Note that the type of displacement we are considering is one in which *everything* to do with the system is displaced. So, if the system is subject to an external potential, then the potential must be displaced.

The situation is not so clear with state kets. The final state of the system only determines the *direction* of the displaced state ket. Even if we adopt the convention that all state kets have unit norms, the final ket is still not completely determined, since it can be multiplied by a constant phase-factor. However, we know that the superposition relations between states remain invariant under the displacement. This follows because the superposition relations have a physical significance which is unaffected by a displacement of the system. Thus, if

$$|R\rangle = |A\rangle + |B\rangle \quad (3.103)$$

in the undisplaced system, and the displacement causes ket  $|R\rangle$  to transform to ket  $|Rd\rangle$ , *etc.*, then in the displaced system we have

$$|Rd\rangle = |Ad\rangle + |Bd\rangle. \quad (3.104)$$

Incidentally, this determines the displaced kets to within a single arbitrary phase-factor to be multiplied into all of them. The displaced kets cannot be multiplied by individual phase-factors, because this would wreck the superposition relations.

Since Eq. (3.104) holds in the displaced system whenever Eq. (3.103) holds in the undisplaced system, it follows that the displaced ket  $|R_d\rangle$  must be the result of some linear operator acting on the undisplaced ket  $|R\rangle$ . In other words,

$$|R_d\rangle = D|R\rangle, \quad (3.105)$$

where  $D$  an operator which depends only on the nature of the displacement. The arbitrary phase-factor by which all displaced kets may be multiplied results in  $D$  being undetermined to an arbitrary multiplicative constant of modulus unity.

We now adopt the ansatz that any combination of bras, kets, and dynamical variables which possesses a physical significance is invariant under a displacement of the system. The normalization condition

$$\langle A|A\rangle = 1 \quad (3.106)$$

for a state ket  $|A\rangle$  certainly has a physical significance. Thus, we must have

$$\langle Ad|Ad\rangle = 1. \quad (3.107)$$

Now,  $|Ad\rangle = D|A\rangle$  and  $\langle Ad| = \langle A|D^\dagger$ , so

$$\langle A|D^\dagger D|A\rangle = 1. \quad (3.108)$$

Since this must hold for any state ket  $|A\rangle$ , it follows that

$$D^\dagger D = 1. \quad (3.109)$$

Hence, the displacement operator is *unitary*. Note that the above relation implies that

$$|A\rangle = D^\dagger|Ad\rangle. \quad (3.110)$$

The equation

$$v|A\rangle = |B\rangle, \quad (3.111)$$

where the operator  $v$  represents a dynamical variable, has some physical significance. Thus, we require that

$$v_d|Ad\rangle = |Bd\rangle, \quad (3.112)$$

where  $v_d$  is the displaced operator. It follows that

$$v_d|Ad\rangle = D|B\rangle = Dv|A\rangle = DvD^\dagger|Ad\rangle. \quad (3.113)$$

Since this is true for any ket  $|Ad\rangle$ , we have

$$v_d = DvD^\dagger. \quad (3.114)$$

Note that the arbitrary numerical factor in  $D$  does not affect either of the results (3.109) and (3.114).

Suppose, now, that the system is displaced an *infinitesimal* distance  $\delta x$  along the  $x$ -axis. We expect that the displaced ket  $|Ad\rangle$  should approach the undisplaced ket  $|A\rangle$  in the limit as  $\delta x \rightarrow 0$ . Thus, we expect the limit

$$\lim_{\delta x \rightarrow 0} \frac{|Ad\rangle - |A\rangle}{\delta x} = \lim_{\delta x \rightarrow 0} \frac{D - 1}{\delta x} |A\rangle \quad (3.115)$$

to exist. Let

$$d_x = \lim_{\delta x \rightarrow 0} \frac{D - 1}{\delta x}, \quad (3.116)$$

where  $d_x$  is denoted the *displacement operator* along the  $x$ -axis. The fact that  $D$  can be replaced by  $D \exp(i\gamma)$ , where  $\gamma$  is a real phase-angle, implies that  $d_x$  can be replaced by

$$\lim_{\delta x \rightarrow 0} \frac{D \exp(i\gamma) - 1}{\delta x} = \lim_{\delta x \rightarrow 0} \frac{D - 1 + i\gamma}{\delta x} = d_x + i a_x, \quad (3.117)$$

where  $a_x$  is the limit of  $\gamma/\delta x$ . We have assumed, as seems reasonable, that  $\gamma$  tends to zero as  $\delta x \rightarrow 0$ . It is clear that the displacement operator is undetermined to an arbitrary imaginary additive constant.

For small  $\delta x$ , we have

$$D = 1 + \delta x d_x. \quad (3.118)$$

It follows from Eq. (3.109) that

$$(1 + \delta x d_x^\dagger)(1 + \delta x d_x) = 1. \quad (3.119)$$

Neglecting order  $(\delta x)^2$ , we obtain

$$d_x^\dagger + d_x = 0. \quad (3.120)$$

Thus, the displacement operator is *anti-Hermitian*. Substituting into Eq. (3.114), and again neglecting order  $(\delta x)^2$ , we find that

$$v_d = (1 + \delta x d_x) v (1 - \delta x d_x) = v + \delta x (d_x v - v d_x), \quad (3.121)$$

which implies

$$\lim_{\delta x \rightarrow 0} \frac{v_d - v}{\delta x} = d_x v - v d_x. \quad (3.122)$$

Let us consider a specific example. Suppose that a state has a wave-function  $\psi(x')$ . If the system is displaced a distance  $\delta x$  along the  $x$ -axis then the new wave-function is  $\psi(x' - \delta x)$  (*i.e.*, the same shape shifted in the  $x$ -direction by a distance  $\delta x$ ). Actually, the new wave-function can be multiplied by an arbitrary number of modulus unity. It can be seen that the new wave-function is obtained from the old wave-function according to the prescription  $x' \rightarrow x' - \delta x$ . Thus,

$$x_d = x - \delta x. \quad (3.123)$$

A comparison with Eq. (3.122), using  $x = v$ , yields

$$d_x x - x d_x = -1. \quad (3.124)$$

It follows that  $i \hbar d_x$  obeys the same commutation relation with  $x$  that  $p_x$ , the momentum conjugate to  $x$ , does [see Eq. (3.25)]. The most general conclusion we can draw from this observation is that

$$p_x = i \hbar d_x + f(x), \quad (3.125)$$

where  $f$  is Hermitian (since  $p_x$  is Hermitian). However, the fact that  $d_x$  is undetermined to an arbitrary additive imaginary constant (which could be a function of  $x$ ) enables us to transform the function  $f$  out of the above equation, leaving

$$p_x = i \hbar d_x. \quad (3.126)$$

Thus, the displacement operator in the  $x$ -direction is proportional to the momentum conjugate to  $x$ . We say that  $p_x$  is the *generator* of translations along the  $x$ -axis.

A finite translation along the  $x$ -axis can be constructed from a series of very many infinitesimal translations. Thus, the operator  $D(\Delta x)$  which translates the system a distance  $\Delta x$  along the  $x$ -axis is written

$$D(\Delta x) = \lim_{N \rightarrow \infty} \left( 1 - i \frac{\Delta x p_x}{N \hbar} \right)^N, \quad (3.127)$$

where use has been made of Eqs. (3.118) and (3.126). It follows that

$$D(\Delta x) = \exp(-i p_x \Delta x / \hbar). \quad (3.128)$$

The unitary nature of the operator is now clearly apparent.

We can also construct displacement operators which translate the system along the  $y$ - and  $z$ -axes. Note that a displacement a distance  $\Delta x$  along the  $x$ -axis *commutes* with a displacement a distance  $\Delta y$  along the  $y$ -axis. In other words, if the system is moved  $\Delta x$  along the  $x$ -axis, and then  $\Delta y$  along the  $y$ -axis then it ends up in the same state as if it were moved  $\Delta y$  along the  $y$ -axis, and then  $\Delta x$  along the  $x$ -axis. The fact that translations in independent directions commute is clearly associated with the fact that the conjugate momentum operators associated with these directions also commute [see Eqs. (3.24) and (3.128)].

## 4 Quantum dynamics

### 4.1 Schrödinger's equations of motion

Up to now, we have only considered systems at one particular instant of time. Let us now investigate how quantum mechanical systems evolve with time.

Consider a system in a state  $A$  which evolves in time. At time  $t$  the state of the system is represented by the ket  $|At\rangle$ . The label  $A$  is needed to distinguish the ket from any other ket ( $|Bt\rangle$ , say) which is evolving in time. The label  $t$  is needed to distinguish the different states of the system at different times.

The final state of the system at time  $t$  is completely determined by its initial state at time  $t_0$  plus the time interval  $t - t_0$  (assuming that the system is left undisturbed during this time interval). However, the final state only determines the *direction* of the final state ket. Even if we adopt the convention that all state kets have unit norms, the final ket is still not completely determined, since it can be multiplied by an arbitrary phase-factor. However, we expect that if a superposition relation holds for certain states at time  $t_0$  then the same relation should hold between the corresponding time-evolved states at time  $t$ , assuming that the system is left undisturbed between times  $t_0$  and  $t$ . In other words, if

$$|Rt_0\rangle = |At_0\rangle + |Bt_0\rangle \quad (4.1)$$

for any three kets, then we should have

$$|Rt\rangle = |At\rangle + |Bt\rangle. \quad (4.2)$$

This rule determines the time-evolved kets to within a single arbitrary phase-factor to be multiplied into all of them. The evolved kets cannot be multiplied by individual phase-factors, since this would invalidate the superposition relation at later times.

According to Eqs. (4.1) and (4.2), the final ket  $|Rt\rangle$  depends linearly on the initial ket  $|Rt_0\rangle$ . Thus, the final ket can be regarded as the result of some linear operator acting on the initial ket: , *i.e.*,

$$|Rt\rangle = T|Rt_0\rangle, \quad (4.3)$$

where  $T$  is a linear operator which depends only on the times  $t$  and  $t_0$ . The arbitrary phase-factor by which all time evolved kets may be multiplied results in  $T(t, t_0)$  being undetermined to an arbitrary multiplicative constant of modulus unity.

Since we have adopted a convention in which the norm of any state ket is unity, it make sense to define the time evolution operator  $T$  in such a manner that it preserves the length of any ket upon which it acts (*i.e.*, if a ket is properly normalized at time  $t$  then it will remain normalized at all subsequent times  $t > t_0$ ). This is always possible, since the length of a ket possesses no physical significance. Thus, we require that

$$\langle At_0 | At_0 \rangle = \langle At | At \rangle \quad (4.4)$$

for any ket  $A$ , which immediately yields

$$T^\dagger T = 1. \quad (4.5)$$

Hence, the time evolution operator  $T$  is a *unitary* operator.

Up to now, the time evolution operator  $T$  looks very much like the spatial displacement operator  $D$  introduced in the previous section. However, there are some important differences between time evolution and spatial displacement. In general, we *do* expect the expectation value of some observable  $\xi$  to evolve with time, even if the system is left in a state of undisturbed motion (after all, time evolution has no meaning unless something *observable* changes with time). The triple product  $\langle A | \xi | A \rangle$  can evolve either because the ket  $|A\rangle$  evolves and the operator  $\xi$  stays constant, the ket  $|A\rangle$  stays constant and the operator  $\xi$  evolves, or both the ket  $|A\rangle$  and the operator  $\xi$  evolve. Since we are already committed to evolving state kets, according to Eq. (4.3), let us assume that the time evolution operator  $T$  can be chosen in such a manner that the operators representing the dynamical variables of the system *do not* evolve in time (unless they contain some specific time dependence).

We expect, from physical continuity, that as  $t \rightarrow t_0$  then  $|At\rangle \rightarrow |At_0\rangle$  for any ket  $A$ . Thus, the limit

$$\lim_{t \rightarrow t_0} \frac{|At\rangle - |At_0\rangle}{t - t_0} = \lim_{t \rightarrow t_0} \frac{T - 1}{t - t_0} |At_0\rangle \quad (4.6)$$

should exist. Note that this limit is simply the derivative of  $|At_0\rangle$  with respect to  $t_0$ . Let

$$\tau(t_0) = \lim_{t \rightarrow t_0} \frac{T(t, t_0) - 1}{t - t_0}. \quad (4.7)$$

It is easily demonstrated from Eq. (4.5) that  $\tau$  is anti-Hermitian: *i.e.*,

$$\tau^\dagger + \tau = 0. \quad (4.8)$$

The fact that  $T$  can be replaced by  $T \exp(i\gamma)$  (where  $\gamma$  is real) implies that  $\tau$  is undetermined to an arbitrary *imaginary* additive constant (see previous section). Let us define the Hermitian operator  $H(t_0) = i \hbar \tau$ . This operator is undetermined to an arbitrary *real* additive constant. It follows from Eqs. (4.6) and (4.7) that

$$i \hbar \frac{d|At_0\rangle}{dt_0} = i \hbar \lim_{t \rightarrow t_0} \frac{|At\rangle - |At_0\rangle}{t - t_0} = i \hbar \tau(t_0)|At_0\rangle = H(t_0)|At_0\rangle. \quad (4.9)$$

When written for general  $t$  this equation becomes

$$i \hbar \frac{d|At\rangle}{dt} = H(t)|At\rangle. \quad (4.10)$$

Equation (4.10) gives the general law for the time evolution of a state ket in a scheme in which the operators representing the dynamical variables remain fixed. This equation is denoted *Schrödinger's equation of motion*. It involves a Hermitian operator  $H(t)$  which is, presumably, a characteristic of the dynamical system under investigation.

We saw, in the previous section, that if the operator  $D(x, x_0)$  displaces the system along the  $x$ -axis from  $x_0$  to  $x$  then

$$p_x = i \hbar \lim_{x \rightarrow x_0} \frac{D(x, x_0) - 1}{x - x_0}, \quad (4.11)$$

where  $p_x$  is the operator representing the momentum conjugate to  $x$ . We now have that if the operator  $T(t, t_0)$  evolves the system in time from  $t_0$  to  $t$  then

$$H(t_0) = i \hbar \lim_{t \rightarrow t_0} \frac{T(t, t_0) - 1}{t - t_0}. \quad (4.12)$$

Thus, the dynamical variable corresponding to the operator  $H$  stands to time  $t$  as the momentum  $p_x$  stands to the coordinate  $x$ . By analogy with classical physics, this suggests that  $H(t)$  is the operator representing the total *energy* of the system. (Recall that, in classical physics, if the equations of motion of a system are invariant under an  $x$ -displacement of the system then this implies that the system conserves momentum in the  $x$ -direction. Likewise, if the equations of motion are invariant under a temporal displacement then this implies that the system conserves energy.) The operator  $H(t)$  is usually called the *Hamiltonian* of the system. The fact that the Hamiltonian is undetermined to an arbitrary real additive constant is related to the well-known phenomenon that energy is undetermined to an arbitrary additive constant in physics (*i.e.*, the zero of potential energy is not well-defined).

Substituting  $|At\rangle = T|At_0\rangle$  into Eq. (4.10) yields

$$i\hbar \frac{dT}{dt}|At_0\rangle = H(t) T|At_0\rangle. \quad (4.13)$$

Since this must hold for any initial state  $|At_0\rangle$  we conclude that

$$i\hbar \frac{dT}{dt} = H(t) T. \quad (4.14)$$

This equation can be integrated to give

$$T(t, t_0) = \exp\left(-i \int_{t_0}^t H(t') dt'/\hbar\right), \quad (4.15)$$

where use has been made of Eqs. (4.5) and (4.6). (Here, we assume that Hamiltonian operators evaluated at different times commute with one another). It is now clear how the fact that  $H$  is undetermined to an arbitrary real additive constant leaves  $T$  undetermined to a phase-factor. Note that, in the above analysis, time is *not* an operator (we cannot observe time, as such), it is just a parameter (or, more accurately, a continuous label). Since we are only dealing with non-relativistic quantum mechanics, the fact that position is an operator, but time is only a label, need not worry us unduly. In relativistic quantum mechanics, time and space coordinates are treated on the same footing by relegating position from being an operator to being just a label.

## 4.2 Heisenberg's equations of motion

We have seen that in Schrödinger's scheme the dynamical variables of the system remain fixed during a period of undisturbed motion, whereas the state kets evolve according to Eq. (4.10). However, this is not the only way in which to represent the time evolution of the system.

Suppose that a general state ket  $A$  is subject to the transformation

$$|A_t\rangle = T^\dagger(t, t_0)|A\rangle. \quad (4.16)$$

This is a time-dependent transformation, since the operator  $T(t, t_0)$  obviously depends on time. The subscript  $t$  is used to remind us that the transformation is time-dependent. The time evolution of the transformed state ket is given by

$$|A_t t\rangle = T^\dagger(t, t_0)|A_t\rangle = T^\dagger(t, t_0) T(t, t_0)|A_{t_0}\rangle = |A_{t_0}\rangle, \quad (4.17)$$

where use has been made of Eqs. (4.3), (4.5), and the fact that  $T(t_0, t_0) = 1$ . Clearly, the transformed state ket *does not* evolve in time. Thus, the transformation (4.16) has the effect of bringing all kets representing states of undisturbed motion of the system to rest.

The transformation must also be applied to bras. The dual of Eq. (4.16) yields

$$\langle A_t| = \langle A|T. \quad (4.18)$$

The transformation rule for a general observable  $v$  is obtained from the requirement that the expectation value  $\langle A|v|A\rangle$  should remain invariant. It is easily seen that

$$v_t = T^\dagger v T. \quad (4.19)$$

Thus, a dynamical variable, which corresponds to a fixed linear operator in Schrödinger's scheme, corresponds to a moving linear operator in this new scheme. It is clear that the transformation (4.16) leads us to a scenario in which the state of the system is represented by a fixed vector, and the dynamical variables are represented by moving linear operators. This is termed the *Heisenberg picture*, as opposed to the *Schrödinger picture*, which is outlined in Sect. 4.1.

Consider a dynamical variable  $v$  corresponding to a fixed linear operator in the Schrödinger picture. According to Eq. (4.19), we can write

$$T v_t = v T. \quad (4.20)$$

Differentiation with respect to time yields

$$\frac{dT}{dt} v_t + T \frac{dv_t}{dt} = v \frac{dT}{dt}. \quad (4.21)$$

With the help of Eq. (4.14), this reduces to

$$H T v_t + i \hbar T \frac{dv_t}{dt} = v H T, \quad (4.22)$$

or

$$i \hbar \frac{dv_t}{dt} = T^\dagger v H T - T^\dagger H T v_t = v_t H_t - H_t v_t, \quad (4.23)$$

where

$$H_t = T^\dagger H T. \quad (4.24)$$

Equation (4.23) can be written

$$i \hbar \frac{dv_t}{dt} = [v_t, H_t]. \quad (4.25)$$

Equation (4.25) shows how the dynamical variables of the system evolve in the Heisenberg picture. It is denoted *Heisenberg's equation of motion*. Note that the time-varying dynamical variables in the Heisenberg picture are usually called *Heisenberg dynamical variables* to distinguish them from *Schrödinger dynamical variables* (i.e., the corresponding variables in the Schrödinger picture), which do not evolve in time.

According to Eq. (3.22), the Heisenberg equation of motion can be written

$$\frac{dv_t}{dt} = [v_t, H_t]_{\text{quantum}}, \quad (4.26)$$

where  $[\cdot \cdot \cdot]_{\text{quantum}}$  denotes the quantum Poisson bracket. Let us compare this equation with the classical time evolution equation for a general dynamical variable  $v$ , which can be written in the form [see Eq. (3.7)]

$$\frac{dv}{dt} = [v, H]_{\text{classical}}. \quad (4.27)$$

Here,  $[\cdot \cdot \cdot]_{\text{classical}}$  is the classical Poisson bracket, and  $H$  denotes the classical Hamiltonian. The strong resemblance between Eqs. (4.26) and (4.27) provides us with further justification for our identification of the linear operator  $H$  with the energy of the system in quantum mechanics.

Note that if the Hamiltonian does not explicitly depend on time (*i.e.*, the system is not subject to some time-dependent external force) then Eq. (4.15) yields

$$T(t, t_0) = \exp[-i H (t - t_0)/\hbar]. \quad (4.28)$$

This operator manifestly commutes with  $H$ , so

$$H_t = T^\dagger H T = H. \quad (4.29)$$

Furthermore, Eq. (4.25) gives

$$i \hbar \frac{dH}{dt} = [H, H] = 0. \quad (4.30)$$

Thus, if the energy of the system has no explicit time-dependence then it is represented by the same non-time-varying operator  $H$  in both the Schrödinger and Heisenberg pictures.

Suppose that  $v$  is an observable which commutes with the Hamiltonian (and, hence, with the time evolution operator  $T$ ). It follows from Eq. (4.19) that  $v_t = v$ . Heisenberg's equation of motion yields

$$i \hbar \frac{dv}{dt} = [v, H] = 0. \quad (4.31)$$

Thus, *any observable which commutes with the Hamiltonian is a constant of the motion* (hence, it is represented by the same fixed operator in both the Schrödinger and Heisenberg pictures). Only those observables which *do not* commute with the Hamiltonian evolve in time in the Heisenberg picture.

### 4.3 Ehrenfest's theorem

We have now derived all of the basic elements of quantum mechanics. The only thing which is lacking is some rule to determine the form of the quantum mechanical Hamiltonian. For a physical system which possess a classical analogue, we

generally assume that the Hamiltonian has the same form as in classical physics (*i.e.*, we replace the classical coordinates and conjugate momenta by the corresponding quantum mechanical operators). This scheme guarantees that quantum mechanics yields the correct classical equations of motion in the classical limit. Whenever an ambiguity arises because of non-commuting observables, this can usually be resolved by requiring the Hamiltonian  $H$  to be an Hermitian operator. For instance, we would write the quantum mechanical analogue of the classical product  $x p$ , appearing in the Hamiltonian, as the Hermitian product  $(1/2)(x p + p x)$ . When the system in question has no classical analogue then we are reduced to guessing a form for  $H$  which reproduces the observed behaviour of the system.

Consider a three-dimensional system characterized by three independent Cartesian position coordinates  $x_i$  (where  $i$  runs from 1 to 3), with three corresponding conjugate momenta  $p_i$ . These are represented by three commuting position operators  $x_i$ , and three commuting momentum operators  $p_i$ , respectively. The commutation relations satisfied by the position and momentum operators are [see Eq. (3.25)]

$$[x_i, p_j] = i \hbar \delta_{ij}. \quad (4.32)$$

It is helpful to denote  $(x_1, x_2, x_3)$  as  $\mathbf{x}$  and  $(p_1, p_2, p_3)$  as  $\mathbf{p}$ . The following useful formulae,

$$[x_i, F(\mathbf{p})] = i \hbar \frac{\partial F}{\partial p_i}, \quad (4.33)$$

$$[p_i, G(\mathbf{x})] = -i \hbar \frac{\partial G}{\partial x_i}, \quad (4.34)$$

where  $F$  and  $G$  are functions which can be expanded as power series, are easily proved using the fundamental commutation relations Eq. (4.32).

Let us now consider the three-dimensional motion of a free particle of mass  $m$  in the Heisenberg picture. The Hamiltonian is assumed to have the same form as in classical physics:

$$H = \frac{\mathbf{p}^2}{2m} = \frac{1}{2m} \sum_{i=1}^3 p_i^2. \quad (4.35)$$

In the following, all dynamical variables are assumed to be Heisenberg dynamical variables, although we will omit the subscript  $t$  for the sake of clarity. The time evolution of the momentum operator  $p_i$  follows from Heisenberg's equation of motion (4.25). We find that

$$\frac{dp_i}{dt} = \frac{1}{i\hbar} [p_i, H] = 0, \quad (4.36)$$

since  $p_i$  automatically commutes with any function of the momentum operators. Thus, for a free particle the momentum operators are constants of the motion, which means that  $p_i(t) = p_i(0)$  at all times  $t$  (for  $i$  is 1 to 3). The time evolution of the position operator  $x_i$  is given by

$$\frac{dx_i}{dt} = \frac{1}{i\hbar} [x_i, H] = \frac{1}{i\hbar} \frac{1}{2m} i\hbar \frac{\partial}{\partial p_i} \left( \sum_{j=1}^3 p_j^2 \right) = \frac{p_i}{m} = \frac{p_i(0)}{m}, \quad (4.37)$$

where use has been made of Eq. (4.33). It follows that

$$x_i(t) = x_i(0) + \left[ \frac{p_i(0)}{m} \right] t, \quad (4.38)$$

which is analogous to the equation of motion of a classical free particle. Note that even though

$$[x_i(0), x_j(0)] = 0, \quad (4.39)$$

where the position operators are evaluated at equal times, the  $x_i$  *do not* commute when evaluated at different times. For instance,

$$[x_i(t), x_i(0)] = \left[ \frac{p_i(0)t}{m}, x_i(0) \right] = \frac{-i\hbar t}{m}. \quad (4.40)$$

Combining the above commutation relation with the uncertainty relation (2.83) yields

$$\langle (\Delta x_i)^2 \rangle_t \langle (\Delta x_i)^2 \rangle_{t=0} \geq \frac{\hbar^2 t^2}{4m^2}. \quad (4.41)$$

This result implies that even if a particle is well-localized at  $t = 0$ , its position becomes progressively more uncertain with time. This conclusion can also be obtained by studying the propagation of wave-packets in wave mechanics.

Let us now add a potential  $V(\mathbf{x})$  to our free particle Hamiltonian:

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}). \quad (4.42)$$

Here,  $V$  is some function of the  $x_i$  operators. Heisenberg's equation of motion gives

$$\frac{dp_i}{dt} = \frac{1}{i\hbar} [p_i, V(\mathbf{x})] = -\frac{\partial V(\mathbf{x})}{\partial x_i}, \quad (4.43)$$

where use has been made of Eq. (4.34). On the other hand, the result

$$\frac{dx_i}{dt} = \frac{p_i}{m} \quad (4.44)$$

still holds, because the  $x_i$  all commute with the new term  $V(\mathbf{x})$  in the Hamiltonian. We can use the Heisenberg equation of motion a second time to deduce that

$$\frac{d^2x_i}{dt^2} = \frac{1}{i\hbar} \left[ \frac{dx_i}{dt}, H \right] = \frac{1}{i\hbar} \left[ \frac{p_i}{m}, H \right] = \frac{1}{m} \frac{dp_i}{dt} = -\frac{1}{m} \frac{\partial V(\mathbf{x})}{\partial x_i}. \quad (4.45)$$

In vectorial form, this equation becomes

$$m \frac{d^2\mathbf{x}}{dt^2} = \frac{d\mathbf{p}}{dt} = -\nabla V(\mathbf{x}). \quad (4.46)$$

This is the quantum mechanical equivalent of Newton's second law of motion. Taking the expectation values of both sides with respect to a Heisenberg state ket that does *not* move with time, we obtain

$$m \frac{d^2\langle \mathbf{x} \rangle}{dt^2} = \frac{d\langle \mathbf{p} \rangle}{dt} = -\langle \nabla V(\mathbf{x}) \rangle. \quad (4.47)$$

This is known as *Ehrenfest's theorem*. When written in terms of expectation values, this result is independent of whether we are using the Heisenberg or Schrödinger picture. In contrast, the operator equation (4.46) only holds if  $\mathbf{x}$  and  $\mathbf{p}$  are understood to be Heisenberg dynamical variables. Note that Eq. (4.47) has no dependence on  $\hbar$ . In fact, it guarantees to us that the centre of a wave-packet always moves like a classical particle.

#### 4.4 Schrödinger's wave-equation

Let us now consider the motion of a particle in three dimensions in the Schrödinger picture. The fixed dynamical variables of the system are the position operators  $\mathbf{x} \equiv (x_1, x_2, x_3)$ , and the momentum operators  $\mathbf{p} \equiv (p_1, p_2, p_3)$ . The state of the system is represented as some time evolving ket  $|A_t\rangle$ .

Let  $|\mathbf{x}'\rangle$  represent a simultaneous eigenket of the position operators belonging to the eigenvalues  $\mathbf{x}' \equiv (x'_1, x'_2, x'_3)$ . Note that, since the position operators are *fixed* in the Schrödinger picture, we do not expect the  $|\mathbf{x}'\rangle$  to evolve in time. The wave-function of the system at time  $t$  is defined

$$\psi(\mathbf{x}', t) = \langle \mathbf{x}' | A_t \rangle. \quad (4.48)$$

The Hamiltonian of the system is taken to be

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}). \quad (4.49)$$

Schrödinger's equation of motion (4.10) yields

$$i \hbar \frac{\partial \langle \mathbf{x}' | A_t \rangle}{\partial t} = \langle \mathbf{x}' | H | A_t \rangle, \quad (4.50)$$

where use has been made of the time independence of the  $|\mathbf{x}'\rangle$ . We adopt Schrödinger's representation in which the momentum conjugate to the position operator  $x_i$  is written [see Eq. (3.74)]

$$p_i = -i \hbar \frac{\partial}{\partial x_i}. \quad (4.51)$$

Thus,

$$\left\langle \mathbf{x}' \left| \frac{\mathbf{p}^2}{2m} \right| A_t \right\rangle = - \left( \frac{\hbar^2}{2m} \right) \nabla'^2 \langle \mathbf{x}' | A_t \rangle, \quad (4.52)$$

where use has been made of Eq. (3.78). Here,  $\nabla' \equiv (\partial/\partial x', \partial/\partial y', \partial/\partial z')$  denotes the gradient operator written in terms of the position eigenvalues. We can also write

$$\langle \mathbf{x}' | V(\mathbf{x}) = V(\mathbf{x}') \langle \mathbf{x}' |, \quad (4.53)$$

where  $V(\mathbf{x}')$  is a scalar function of the position eigenvalues. Combining Eqs. (4.49), (4.50), (4.52), and (4.53), we obtain

$$i \hbar \frac{\partial \langle \mathbf{x}' | A t \rangle}{\partial t} = - \left( \frac{\hbar^2}{2m} \right) \nabla'^2 \langle \mathbf{x}' | A t \rangle + V(\mathbf{x}') \langle \mathbf{x}' | A t \rangle, \quad (4.54)$$

which can also be written

$$i \hbar \frac{\partial \psi(\mathbf{x}', t)}{\partial t} = - \left( \frac{\hbar^2}{2m} \right) \nabla'^2 \psi(\mathbf{x}', t) + V(\mathbf{x}') \psi(\mathbf{x}', t). \quad (4.55)$$

This is Schrödinger's famous wave-equation, and is the basis of wave mechanics. Note, however, that the wave-equation is just one of many possible representations of quantum mechanics. It just happens to give a type of equation which we know how to solve. In deriving the wave-equation, we have chosen to represent the system in terms of the eigenkets of the position operators, instead of those of the momentum operators. We have also fixed the relative phases of the  $|\mathbf{x}'\rangle$  according to Schrödinger's representation, so that Eq. (4.51) is valid. Finally, we have chosen to work in the Schrödinger picture, in which state kets evolve and dynamical variables are fixed, instead of the Heisenberg picture, in which the opposite is true.

Suppose that the ket  $|A t\rangle$  is an eigenket of the Hamiltonian belonging to the eigenvalue  $H'$ :

$$H|A t\rangle = H'|A t\rangle. \quad (4.56)$$

Schrödinger's equation of motion (4.10) yields

$$i \hbar \frac{d|A t\rangle}{dt} = H'|A t\rangle. \quad (4.57)$$

This can be integrated to give

$$|A t\rangle = \exp[-i H'(t - t_0)/\hbar] |A t_0\rangle. \quad (4.58)$$

Note that  $|A t\rangle$  only differs from  $|A t_0\rangle$  by a phase-factor. The direction of the vector remains fixed in ket space. This suggests that if the system is initially in an eigenstate of the Hamiltonian then it remains in this state for ever, as long as the

system is undisturbed. Such a state is called a *stationary state*. The wave-function of a stationary state satisfies

$$\psi(\mathbf{x}', t) = \psi(\mathbf{x}', t_0) \exp[-i H' (t - t_0)/\hbar]. \quad (4.59)$$

Substituting the above relation into Schrödinger's wave equation (4.55), we obtain

$$-\left(\frac{\hbar^2}{2m}\right) \nabla'^2 \psi_0(\mathbf{x}') + (V(\mathbf{x}') - E) \psi_0(\mathbf{x}') = 0, \quad (4.60)$$

where  $\psi_0(\mathbf{x}') \equiv \psi(\mathbf{x}', t_0)$ , and  $E = H'$  is the energy of the system. This is Schrödinger's time-independent wave-equation. A *bound state* solution of the above equation, in which the particle is confined within a finite region of space, satisfies the boundary condition

$$\psi_0(\mathbf{x}') \rightarrow 0 \quad \text{as } |\mathbf{x}'| \rightarrow \infty. \quad (4.61)$$

Such a solution is only possible if

$$E < \lim_{|\mathbf{x}'| \rightarrow \infty} V(\mathbf{x}'). \quad (4.62)$$

Since it is conventional to set the potential at infinity equal to zero, the above relation implies that bound states are equivalent to negative energy states. The boundary condition (4.61) is sufficient to uniquely specify the solution of Eq. (4.60).

The quantity  $\rho(\mathbf{x}', t)$ , defined by

$$\rho(\mathbf{x}', t) = |\psi(\mathbf{x}', t)|^2, \quad (4.63)$$

is termed the *probability density*. Recall, from Eq. (3.30), that the probability of observing the particle in some volume element  $d^3x'$  around position  $\mathbf{x}'$  is proportional to  $\rho(\mathbf{x}', t) d^3x'$ . The probability is *equal* to  $\rho(\mathbf{x}', t) d^3x'$  if the wave-function is properly normalized, so that

$$\int \rho(\mathbf{x}', t) d^3x' = 1. \quad (4.64)$$

Schrödinger's time-dependent wave-equation, (4.55), can easily be written in the form of a conservation equation for the probability density:

$$\frac{\partial \rho}{\partial t} + \nabla' \cdot \mathbf{j} = 0. \quad (4.65)$$

The *probability current*  $\mathbf{j}$  takes the form

$$\mathbf{j}(\mathbf{x}', t) = - \left( \frac{i \hbar}{2m} \right) [\psi^* \nabla' \psi - (\nabla' \psi^*) \psi] = \left( \frac{\hbar}{m} \right) \text{Im}(\psi^* \nabla' \psi). \quad (4.66)$$

We can integrate Eq. (4.65) over all space, using the divergence theorem, and the boundary condition  $\rho \rightarrow 0$  as  $|\mathbf{x}'| \rightarrow \infty$ , to obtain

$$\frac{\partial}{\partial t} \int \rho(\mathbf{x}', t) d^3x' = 0. \quad (4.67)$$

Thus, Schrödinger's wave-equation *conserves* probability. In particular, if the wave-function starts off properly normalized, according to Eq. (4.64), then it remains properly normalized at all subsequent times. It is easily demonstrated that

$$\int \mathbf{j}(\mathbf{x}', t) d^3x' = \frac{\langle \mathbf{p} \rangle_t}{m}, \quad (4.68)$$

where  $\langle \mathbf{p} \rangle_t$  denotes the expectation value of the momentum evaluated at time  $t$ . Clearly, the probability current is indirectly related to the particle momentum.

In deriving Eq. (4.65) we have, naturally, assumed that the potential  $V(\mathbf{x}')$  is real. Suppose, however, that the potential has an imaginary component. In this case, Eq. (4.65) generalizes to

$$\frac{\partial \rho}{\partial t} + \nabla' \cdot \mathbf{j} = \frac{2 \text{Im}(V)}{\hbar} \rho, \quad (4.69)$$

giving

$$\frac{\partial}{\partial t} \int \rho(\mathbf{x}', t) d^3x' = \frac{2}{\hbar} \text{Im} \int V(\mathbf{x}') \rho(\mathbf{x}', t) d^3x'. \quad (4.70)$$

Thus, if  $\text{Im}(V) < 0$  then the total probability of observing the particle anywhere in space decreases monotonically with time. Thus, an imaginary potential can be used to account for the disappearance of a particle. Such a potential is often employed to model nuclear reactions in which incident particles can be absorbed by nuclei.

The wave-function can always be written in the form

$$\psi(\mathbf{x}', t) = \sqrt{\rho(\mathbf{x}', t)} \exp \left[ \frac{i S(\mathbf{x}', t)}{\hbar} \right], \quad (4.71)$$

where  $\rho$  and  $S$  are both real functions. The interpretation of  $\rho$  as a probability density has already been given. What is the interpretation of  $S$ ? Note that

$$\psi^* \nabla' \psi = \sqrt{\rho} \nabla'(\sqrt{\rho}) + \left(\frac{i}{\hbar}\right) \rho \nabla' S. \quad (4.72)$$

It follows from Eq. (4.66) that

$$\mathbf{j} = \frac{\rho \nabla' S}{m}. \quad (4.73)$$

Thus, the gradient of the phase of the wave-function determines the direction of the probability current. In particular, the probability current is locally normal to the contours of the phase-function  $S$ .

Let us substitute Eq. (4.71) into Schrödinger's time-dependent wave-equation. We obtain

$$\begin{aligned} -\frac{1}{2m} [\hbar^2 \nabla'^2 \sqrt{\rho} + 2i \hbar \nabla'(\sqrt{\rho}) \cdot \nabla' S - \sqrt{\rho} |\nabla' S|^2 + i \hbar \sqrt{\rho} \nabla'^2 S] + \sqrt{\rho} V \\ = \left[ i \hbar \frac{\partial \sqrt{\rho}}{\partial t} - \sqrt{\rho} \frac{\partial S}{\partial t} \right]. \end{aligned} \quad (4.74)$$

Let us treat  $\hbar$  as a small quantity. To lowest order, Eq. (4.74) yields

$$-\frac{\partial S(\mathbf{x}', t)}{\partial t} = \frac{1}{2m} |\nabla' S(\mathbf{x}', t)|^2 + V(\mathbf{x}', t) = H(\mathbf{x}', \nabla' S, t), \quad (4.75)$$

where  $H(\mathbf{x}, \mathbf{p}, t)$  is the Hamiltonian operator. The above equation is known as the *Hamilton-Jacobi* equation, and is one of the many forms in which we can write the equations of classical mechanics. In classical mechanics,  $S$  is the *action* (*i.e.*, the path-integral of the Lagrangian). Thus, in the limit  $\hbar \rightarrow 0$ , wave mechanics reduces to classical mechanics. It is a good approximation to neglect the terms involving  $\hbar$  in Eq. (4.74) provided that

$$\hbar |\nabla'^2 S| \ll |\nabla' S|^2. \quad (4.76)$$

Note that, according to Eq. (4.71),

$$\lambda = \frac{\hbar}{|\nabla' S|}, \quad (4.77)$$

where  $\lambda$  is the de Broglie wave-length divided by  $2\pi$ . The inequality (4.76) is equivalent to

$$|\nabla'\lambda| \ll 1. \quad (4.78)$$

In other words, quantum mechanics reduces to classical mechanics whenever the de Broglie wave-length is small compared to the characteristic distance over which things (other than the quantum phase) vary. This distance is usually set by the variation scale-length of the potential.

## 5 Angular momentum

### 5.1 Orbital angular momentum

Consider a particle described by the Cartesian coordinates  $(x, y, z) \equiv \mathbf{r}$  and their conjugate momenta  $(p_x, p_y, p_z) \equiv \mathbf{p}$ . The classical definition of the orbital angular momentum of such a particle about the origin is  $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ , giving

$$L_x = y p_z - z p_y, \quad (5.1)$$

$$L_y = z p_x - x p_z, \quad (5.2)$$

$$L_z = x p_y - y p_x. \quad (5.3)$$

Let us assume that the operators  $(L_x, L_y, L_z) \equiv \mathbf{L}$  which represent the components of orbital angular momentum in quantum mechanics can be defined in an analogous manner to the corresponding components of classical angular momentum. In other words, we are going to assume that the above equations specify the angular momentum operators in terms of the position and linear momentum operators. Note that  $L_x$ ,  $L_y$ , and  $L_z$  are Hermitian, so they represent things which can, in principle, be measured. Note, also, that there is no ambiguity regarding the order in which operators appear in products on the right-hand sides of Eqs. (5.1)–(5.3), since all of the products consist of operators which commute.

The fundamental commutation relations satisfied by the position and linear momentum operators are [see Eqs. (3.23)–(3.25)]

$$[x_i, x_j] = 0, \quad (5.4)$$

$$[p_i, p_j] = 0, \quad (5.5)$$

$$[x_i, p_j] = i \hbar \delta_{ij}, \quad (5.6)$$

where  $i$  and  $j$  stand for either  $x$ ,  $y$ , or  $z$ . Consider the commutator of the operators  $L_x$  and  $L_z$ :

$$\begin{aligned} [L_x, L_z] &= [(y p_z - z p_y), (z p_x - x p_z)] = y [p_z, z] p_x + x p_y [z, p_z] \\ &= i \hbar (-y p_x + x p_y) = i \hbar L_z. \end{aligned} \quad (5.7)$$

The cyclic permutations of the above result yield the fundamental commutation relations satisfied by the components of an angular momentum:

$$[L_x, L_y] = i \hbar L_z, \quad (5.8)$$

$$[L_y, L_z] = i \hbar L_x, \quad (5.9)$$

$$[L_z, L_x] = i \hbar L_y. \quad (5.10)$$

These can be summed up more succinctly by writing

$$\mathbf{L} \times \mathbf{L} = i \hbar \mathbf{L}. \quad (5.11)$$

The three commutation relations (5.8)–(5.10) are the foundation for the whole theory of angular momentum in quantum mechanics. Whenever we encounter three operators having these commutation relations, we know that the dynamical variables which they represent have identical properties to those of the components of an angular momentum (which we are about to derive). In fact, we shall assume that *any three operators which satisfy the commutation relations (5.8)–(5.10) represent the components of an angular momentum.*

Suppose that there are  $N$  particles in the system, with angular momentum vectors  $\mathbf{L}_i$  (where  $i$  runs from 1 to  $N$ ). Each of these vectors satisfies Eq. (5.11), so that

$$\mathbf{L}_i \times \mathbf{L}_i = i \hbar \mathbf{L}_i. \quad (5.12)$$

However, we expect the angular momentum operators belonging to different particles to commute, since they represent different degrees of freedom of the system. So, we can write

$$\mathbf{L}_i \times \mathbf{L}_j + \mathbf{L}_j \times \mathbf{L}_i = 0, \quad (5.13)$$

for  $i \neq j$ . Consider the total angular momentum of the system,  $\mathbf{L} = \sum_{i=1}^N \mathbf{L}_i$ . It is clear from Eqs. (5.12) and (5.13) that

$$\begin{aligned} \mathbf{L} \times \mathbf{L} &= \sum_{i=1}^N \mathbf{L}_i \times \sum_{j=1}^N \mathbf{L}_j = \sum_{i=1}^N \mathbf{L}_i \times \mathbf{L}_i + \frac{1}{2} \sum_{i,j=1}^N (\mathbf{L}_i \times \mathbf{L}_j + \mathbf{L}_j \times \mathbf{L}_i) \\ &= i \hbar \sum_{i=1}^N \mathbf{L}_i = i \hbar \mathbf{L}. \end{aligned} \quad (5.14)$$

Thus, the sum of two or more angular momentum vectors satisfies the same commutation relation as a primitive angular momentum vector. In particular, the total angular momentum of the system satisfies the commutation relation (5.11).

The immediate conclusion which can be drawn from the commutation relations (5.8)–(5.10) is that the three components of an angular momentum vector cannot be specified (or measured) simultaneously. In fact, once we have specified one component, the values of other two components become uncertain. It is conventional to specify the z-component,  $L_z$ .

Consider the magnitude squared of the angular momentum vector,  $L^2 \equiv L_x^2 + L_y^2 + L_z^2$ . The commutator of  $L^2$  and  $L_z$  is written

$$[L^2, L_z] = [L_x^2, L_z] + [L_y^2, L_z] + [L_z^2, L_z]. \quad (5.15)$$

It is easily demonstrated that

$$[L_x^2, L_z] = -i\hbar(L_x L_y + L_y L_x), \quad (5.16)$$

$$[L_y^2, L_z] = +i\hbar(L_x L_y + L_y L_x), \quad (5.17)$$

$$[L_z^2, L_z] = 0, \quad (5.18)$$

so

$$[L^2, L_z] = 0. \quad (5.19)$$

Since there is nothing special about the z-axis, we conclude that  $L^2$  also commutes with  $L_x$  and  $L_y$ . It is clear from Eqs. (5.8)–(5.10) and (5.19) that the best we can do in quantum mechanics is to specify the magnitude of an angular momentum vector along with *one* of its components (by convention, the z-component).

It is convenient to define the *shift operators*  $L^+$  and  $L^-$ :

$$L^+ = L_x + iL_y, \quad (5.20)$$

$$L^- = L_x - iL_y. \quad (5.21)$$

Note that

$$[L^+, L_z] = -\hbar L^+, \quad (5.22)$$

$$[L^-, L_z] = +\hbar L^-, \quad (5.23)$$

$$[L^+, L^-] = 2\hbar L_z. \quad (5.24)$$

Note, also, that both shift operators commute with  $L^2$ .

## 5.2 Eigenvalues of angular momentum

Suppose that the simultaneous eigenkets of  $L^2$  and  $L_z$  are completely specified by two quantum numbers,  $l$  and  $m$ . These kets are denoted  $|l, m\rangle$ . The quantum number  $m$  is defined by

$$L_z |l, m\rangle = m \hbar |l, m\rangle. \quad (5.25)$$

Thus,  $m$  is the eigenvalue of  $L_z$  divided by  $\hbar$ . It is possible to write such an equation because  $\hbar$  has the dimensions of angular momentum. Note that  $m$  is a real number, since  $L_z$  is an Hermitian operator.

We can write

$$L^2 |l, m\rangle = f(l, m) \hbar^2 |l, m\rangle, \quad (5.26)$$

without loss of generality, where  $f(l, m)$  is some real dimensionless function of  $l$  and  $m$ . Later on, we will show that  $f(l, m) = l(l+1)$ . Now,

$$\langle l, m | L^2 - L_z^2 |l, m\rangle = \langle l, m | f(l, m) \hbar^2 - m^2 \hbar^2 |l, m\rangle = [f(l, m) - m^2] \hbar^2, \quad (5.27)$$

assuming that the  $|l, m\rangle$  have unit norms. However,

$$\begin{aligned} \langle l, m | L^2 - L_z^2 |l, m\rangle &= \langle l, m | L_x^2 + L_y^2 |l, m\rangle \\ &= \langle l, m | L_x^2 |l, m\rangle + \langle l, m | L_y^2 |l, m\rangle. \end{aligned} \quad (5.28)$$

It is easily demonstrated that

$$\langle A | \xi^2 | A \rangle \geq 0, \quad (5.29)$$

where  $|A\rangle$  is a general ket, and  $\xi$  is an Hermitian operator. The proof follows from the observation that

$$\langle A | \xi^2 | A \rangle = \langle A | \xi^\dagger \xi | A \rangle = \langle B | B \rangle, \quad (5.30)$$

where  $|B\rangle = \xi|A\rangle$ , plus the fact that  $\langle B|B\rangle \geq 0$  for a general ket  $|B\rangle$  [see Eq. (2.21)]. It follows from Eqs. (5.27)–(5.29) that

$$m^2 \leq f(l, m). \quad (5.31)$$

Consider the effect of the shift operator  $L^+$  on the eigenket  $|l, m\rangle$ . It is easily demonstrated that

$$L^2(L^+|l, m\rangle) = \hbar^2 f(l, m) (L^+|l, m\rangle), \quad (5.32)$$

where use has been made of Eq. (5.26), plus the fact that  $L^2$  and  $L_z$  commute. It follows that the ket  $L^+|l, m\rangle$  has the same eigenvalue of  $L^2$  as the ket  $|l, m\rangle$ . Thus, the shift operator  $L^+$  does not affect the magnitude of the angular momentum of any eigenket it acts upon. Note that

$$\begin{aligned} L_z L^+|l, m\rangle &= (L^+L_z + [L_z, L^+])|l, m\rangle = (L^+L_z + \hbar L^+)|l, m\rangle \\ &= (m + 1) \hbar L^+|l, m\rangle, \end{aligned} \quad (5.33)$$

where use has been made of Eq. (5.22). The above equation implies that  $L^+|l, m\rangle$  is proportional to  $|l, m + 1\rangle$ . We can write

$$L^+|l, m\rangle = c_{l,m}^+ \hbar |l, m + 1\rangle, \quad (5.34)$$

where  $c_{l,m}^+$  is a number. It is clear that when the operator  $L^+$  acts on a simultaneous eigenstate of  $L^2$  and  $L_z$ , the eigenvalue of  $L^2$  remains unchanged, but the eigenvalue of  $L_z$  is increased by  $\hbar$ . For this reason,  $L^+$  is called a *raising operator*.

Using similar arguments to those given above, it is possible to demonstrate that

$$L^-|l, m\rangle = c_{l,m}^- \hbar |l, m - 1\rangle. \quad (5.35)$$

Hence,  $L^-$  is called a *lowering operator*.

The shift operators step the value of  $m$  up and down by unity each time they operate on one of the simultaneous eigenkets of  $L^2$  and  $L_z$ . It would appear, at first sight, that any value of  $m$  can be obtained by applying the shift operators a sufficient number of times. However, according to Eq. (5.31), there is a definite upper bound to the values that  $m^2$  can take. This bound is determined by the

eigenvalue of  $L^2$  [see Eq. (5.26)]. It follows that there is a maximum and a minimum possible value which  $m$  can take. Suppose that we attempt to raise the value of  $m$  above its maximum value  $m_{\max}$ . Since there is no state with  $m > m_{\max}$ , we must have

$$L^+|l, m_{\max}\rangle = |0\rangle. \quad (5.36)$$

This implies that

$$L^- L^+|l, m_{\max}\rangle = |0\rangle. \quad (5.37)$$

However,

$$L^- L^+ = L_x^2 + L_y^2 + i[L_x, L_y] = L^2 - L_z^2 - \hbar L_z, \quad (5.38)$$

so Eq. (5.37) yields

$$(L^2 - L_z^2 - \hbar L_z)|l, m_{\max}\rangle = |0\rangle. \quad (5.39)$$

The above equation can be rearranged to give

$$L^2|l, m_{\max}\rangle = (L_z^2 + \hbar L_z)|l, m_{\max}\rangle = m_{\max}(m_{\max} + 1) \hbar^2|l, m_{\max}\rangle. \quad (5.40)$$

Comparison of this equation with Eq. (5.26) yields the result

$$f(l, m_{\max}) = m_{\max}(m_{\max} + 1). \quad (5.41)$$

But, when  $L^-$  operates on  $|n, m_{\max}\rangle$  it generates  $|n, m_{\max} - 1\rangle$ ,  $|n, m_{\max} - 2\rangle$ , etc. Since the lowering operator does not change the eigenvalue of  $L^2$ , all of these states must correspond to the same value of  $f$ , namely  $m_{\max}(m_{\max} + 1)$ . Thus,

$$L^2|l, m\rangle = m_{\max}(m_{\max} + 1) \hbar^2|l, m\rangle. \quad (5.42)$$

At this stage, we can give the unknown quantum number  $l$  the value  $m_{\max}$ , without loss of generality. We can also write the above equation in the form

$$L^2|l, m\rangle = l(l + 1) \hbar^2|l, m\rangle. \quad (5.43)$$

It is easily seen that

$$L^- L^+|l, m\rangle = (L^2 - L_z^2 - \hbar L_z)|l, m\rangle = \hbar^2[l(l + 1) - m(m + 1)]|l, m\rangle. \quad (5.44)$$

Thus,

$$\langle l, m|L^- L^+|l, m\rangle = \hbar^2[l(l + 1) - m(m + 1)]. \quad (5.45)$$

However, we also know that

$$\langle l, m | L^- L^+ | l, m \rangle = \langle l, m | L^- \hbar c_{l,m}^+ | l, m+1 \rangle = \hbar^2 c_{l,m}^+ c_{l,m+1}^-, \quad (5.46)$$

where use has been made of Eqs. (5.34) and (5.35). It follows that

$$c_{l,m}^+ c_{l,m+1}^- = [l(l+1) - m(m+1)]. \quad (5.47)$$

Consider the following:

$$\begin{aligned} \langle l, m | L^- | l, m+1 \rangle &= \langle l, m | L_x | l, m+1 \rangle - i \langle l, m | L_y | l, m+1 \rangle \\ &= \langle l, m+1 | L_x | l, m \rangle^* - i \langle l, m+1 | L_y | l, m \rangle^* \\ &= (\langle l, m+1 | L_x | l, m \rangle + i \langle l, m+1 | L_y | l, m \rangle)^* \\ &= \langle l, m+1 | L^+ | l, m \rangle^*, \end{aligned} \quad (5.48)$$

where use has been made of the fact that  $L_x$  and  $L_y$  are Hermitian. The above equation reduces to

$$c_{l,m+1}^- = (c_{l,m}^+)^* \quad (5.49)$$

with the aid of Eqs. (5.34) and (5.35).

Equations (5.47) and (5.49) can be combined to give

$$|c_{l,m}^+|^2 = [l(l+1) - m(m+1)]. \quad (5.50)$$

The solution of the above equation is

$$c_{l,m}^+ = \sqrt{l(l+1) - m(m+1)}. \quad (5.51)$$

Note that  $c_{l,m}^+$  is undetermined to an arbitrary phase-factor [*i.e.*, we can replace  $c_{l,m}^+$ , given above, by  $c_{l,m}^+ \exp(i\gamma)$ , where  $\gamma$  is real, and we still satisfy Eq. (5.50)]. We have made the arbitrary, but convenient, choice that  $c_{l,m}^+$  is real and positive. This is equivalent to choosing the relative phases of the eigenkets  $|l, m\rangle$ . According to Eq. (5.49),

$$c_{l,m}^- = (c_{l,m-1}^+)^* = \sqrt{l(l+1) - m(m-1)}. \quad (5.52)$$

We have already seen that the inequality (5.31) implies that there is a maximum and a minimum possible value of  $m$ . The maximum value of  $m$  is denoted

l. What is the minimum value? Suppose that we try to lower the value of  $m$  below its minimum value  $m_{\min}$ . Since there is no state with  $m < m_{\min}$ , we must have

$$L^-|l, m_{\min}\rangle = 0. \quad (5.53)$$

According to Eq. (5.35), this implies that

$$c_{l, m_{\min}}^- = 0. \quad (5.54)$$

It can be seen from Eq. (5.52) that  $m_{\min} = -l$ . We conclude that  $m$  can take a “ladder” of discrete values, each rung differing from its immediate neighbours by unity. The top rung is  $l$ , and the bottom rung is  $-l$ . There are only two possible choices for  $l$ . Either it is an integer (e.g.,  $l = 2$ , which allows  $m$  to take the values  $-2, -1, 0, 1, 2$ ), or it is a half-integer (e.g.,  $l = 3/2$ , which allows  $m$  to take the values  $-3/2, -1/2, 1/2, 3/2$ ). We will prove in the next section that an orbital angular momentum can only take integer values of  $l$ .

In summary, using just the fundamental commutation relations (5.8)–(5.10), plus the fact that  $L_x$ ,  $L_y$ , and  $L_z$  are Hermitian operators, we have shown that the eigenvalues of  $L^2 \equiv L_x^2 + L_y^2 + L_z^2$  can be written  $l(l+1)\hbar^2$ , where  $l$  is an integer, or a half-integer. We have also demonstrated that the eigenvalues of  $L_z$  can only take the values  $m\hbar$ , where  $m$  lies in the range  $-l, -l+1, \dots, l-1, l$ . Let  $|l, m\rangle$  denote a properly normalized simultaneous eigenket of  $L^2$  and  $L_z$ , belonging to the eigenvalues  $l(l+1)\hbar^2$  and  $m\hbar$ , respectively. We have shown that

$$L^+|l, m\rangle = \sqrt{l(l+1) - m(m+1)}\hbar|l, m+1\rangle \quad (5.55)$$

$$L^-|l, m\rangle = \sqrt{l(l+1) - m(m-1)}\hbar|l, m-1\rangle, \quad (5.56)$$

where  $L^\pm = L_x \pm iL_y$  are the so-called shift operators.

### 5.3 Rotation operators

Consider a particle described by the spherical polar coordinates  $(r, \theta, \varphi)$ . The classical momentum conjugate to the azimuthal angle  $\varphi$  is the  $z$ -component of angular momentum,  $L_z$ . According to Sect. 3.5, in quantum mechanics we can

always adopt Schrödinger's representation, for which ket space is spanned by the simultaneous eigenkets of the position operators  $r$ ,  $\theta$ , and  $\phi$ , and  $L_z$  takes the form

$$L_z = -i \hbar \frac{\partial}{\partial \phi}. \quad (5.57)$$

We can do this because there is nothing in Sect. 3.5 which specifies that we have to use Cartesian coordinates—the representation (3.74) works for any well-defined set of coordinates.

Consider an operator  $R(\Delta\phi)$  which rotates the system an angle  $\Delta\phi$  about the  $z$ -axis. This operator is very similar to the operator  $D(\Delta x)$ , introduced in Sect. 3.8, which translates the system a distance  $\Delta x$  along the  $x$  axis. We were able to demonstrate in Sect. 3.8 that

$$p_x = i \hbar \lim_{\delta x \rightarrow 0} \frac{D(\delta x) - 1}{\delta x}, \quad (5.58)$$

where  $p_x$  is the linear momentum conjugate to  $x$ . There is nothing in our derivation of this result which specifies that  $x$  has to be a Cartesian coordinate. Thus, the result should apply just as well to an angular coordinate. We conclude that

$$L_z = i \hbar \lim_{\delta\phi \rightarrow 0} \frac{R(\delta\phi) - 1}{\delta\phi}. \quad (5.59)$$

According to Eq. (5.59), we can write

$$R(\delta\phi) = 1 - i L_z \delta\phi / \hbar \quad (5.60)$$

in the limit  $\delta\phi \rightarrow 0$ . In other words, the angular momentum operator  $L_z$  can be used to rotate the system about the  $z$ -axis by an infinitesimal amount. We say that  $L_z$  is the *generator* of rotations about the  $z$ -axis. The above equation implies that

$$R(\Delta\phi) = \lim_{N \rightarrow \infty} \left( 1 - i \frac{\Delta\phi}{N} \frac{L_z}{\hbar} \right)^N, \quad (5.61)$$

which reduces to

$$R(\Delta\phi) = \exp(-i L_z \Delta\phi / \hbar). \quad (5.62)$$

Note that  $R(\Delta\varphi)$  has all of the properties we would expect of a rotation operator

$$R(0) = 1, \quad (5.63)$$

$$R(\Delta\varphi) R(-\Delta\varphi) = 1, \quad (5.64)$$

$$R(\Delta\varphi_1) R(\Delta\varphi_2) = R(\Delta\varphi_1 + \Delta\varphi_2). \quad (5.65)$$

Suppose that the system is in a simultaneous eigenstate of  $L^2$  and  $L_z$ . As before, this state is represented by the eigenket  $|l, m\rangle$ , where the eigenvalue of  $L^2$  is  $l(l+1)\hbar^2$ , and the eigenvalue of  $L_z$  is  $m\hbar$ . We expect the wave-function to remain unaltered if we rotate the system  $2\pi$  degrees about the  $z$ -axis. Thus,

$$R(2\pi)|l, m\rangle = \exp(-iL_z 2\pi/\hbar)|l, m\rangle = \exp(-i2\pi m)|l, m\rangle = |l, m\rangle. \quad (5.66)$$

We conclude that  $m$  must be an integer. This implies, from the previous section, that  $l$  must also be an integer. Thus, *orbital* angular momentum can only take on *integer* values of the quantum numbers  $l$  and  $m$ .

Consider the action of the rotation operator  $R(\Delta\varphi)$  on an eigenstate possessing zero angular momentum about the  $z$ -axis (*i.e.*, an  $m = 0$  state). We have

$$R(\Delta\varphi)|l, 0\rangle = \exp(0)|l, 0\rangle = |l, 0\rangle. \quad (5.67)$$

Thus, the eigenstate is invariant to rotations about the  $z$ -axis. Clearly, its wave-function must be symmetric about the  $z$ -axis.

There is nothing special about the  $z$ -axis, so we can write

$$R_x(\Delta\varphi_x) = \exp(-iL_x \Delta\varphi_x/\hbar), \quad (5.68)$$

$$R_y(\Delta\varphi_y) = \exp(-iL_y \Delta\varphi_y/\hbar), \quad (5.69)$$

$$R_z(\Delta\varphi_z) = \exp(-iL_z \Delta\varphi_z/\hbar), \quad (5.70)$$

by analogy with Eq. (5.62). Here,  $R_x(\Delta\varphi_x)$  denotes an operator which rotates the system by an angle  $\Delta\varphi_x$  about the  $x$ -axis, *etc.* Suppose that the system is in an eigenstate of zero overall orbital angular momentum (*i.e.*, an  $l = 0$  state). We know that the system is also in an eigenstate of zero orbital angular momentum about any particular axis. This follows because  $l = 0$  implies  $m = 0$ , according

to the previous section, and we can choose the  $z$ -axis to point in any direction. Thus,

$$\mathbf{R}_x(\Delta\varphi_x)|0,0\rangle = \exp(0)|0,0\rangle = |0,0\rangle, \quad (5.71)$$

$$\mathbf{R}_y(\Delta\varphi_y)|0,0\rangle = \exp(0)|0,0\rangle = |0,0\rangle, \quad (5.72)$$

$$\mathbf{R}_z(\Delta\varphi_z)|0,0\rangle = \exp(0)|0,0\rangle = |0,0\rangle. \quad (5.73)$$

Clearly, a zero angular momentum state is invariant to rotations about *any* axis. Such a state must possess a spherically symmetric wave-function.

Note that a rotation about the  $x$ -axis does not commute with a rotation about the  $y$ -axis. In other words, if the system is rotated an angle  $\Delta\varphi_x$  about the  $x$ -axis, and then  $\Delta\varphi_y$  about the  $y$ -axis, it ends up in a different state to that obtained by rotating an angle  $\Delta\varphi_y$  about the  $y$ -axis, and then  $\Delta\varphi_x$  about the  $x$ -axis. In quantum mechanics, this implies that  $\mathbf{R}_y(\Delta\varphi_y)\mathbf{R}_x(\Delta\varphi_x) \neq \mathbf{R}_x(\Delta\varphi_x)\mathbf{R}_y(\Delta\varphi_y)$ , or  $L_y L_x \neq L_x L_y$ , [see Eqs. (5.68)–(5.70)]. Thus, the noncommuting nature of the angular momentum operators is a direct consequence of the fact that rotations do not commute.

## 5.4 Eigenfunctions of orbital angular momentum

In Cartesian coordinates, the three components of orbital angular momentum can be written

$$L_x = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \quad (5.74)$$

$$L_y = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right), \quad (5.75)$$

$$L_z = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right), \quad (5.76)$$

using the Schrödinger representation. Transforming to standard spherical polar coordinates,

$$x = r \sin\theta \cos\varphi, \quad (5.77)$$

$$y = r \sin \theta \sin \varphi, \quad (5.78)$$

$$z = r \cos \theta, \quad (5.79)$$

we obtain

$$L_x = i \hbar \left( \sin \varphi \frac{\partial}{\partial \theta} + \cot \theta \cos \varphi \frac{\partial}{\partial \varphi} \right) \quad (5.80)$$

$$L_y = -i \hbar \left( \cos \varphi \frac{\partial}{\partial \theta} - \cot \theta \sin \varphi \frac{\partial}{\partial \varphi} \right) \quad (5.81)$$

$$L_z = -i \hbar \frac{\partial}{\partial \varphi}. \quad (5.82)$$

Note that Eq. (5.82) accords with Eq. (5.57). The shift operators  $L^\pm = L_x \pm i L_y$  become

$$L^\pm = \pm \hbar \exp(\pm i \varphi) \left( \frac{\partial}{\partial \theta} \pm i \cot \theta \frac{\partial}{\partial \varphi} \right). \quad (5.83)$$

Now,

$$L^2 = L_x^2 + L_y^2 + L_z^2 = L_z^2 + (L^+ L^- + L^- L^+)/2, \quad (5.84)$$

so

$$L^2 = -\hbar^2 \left( \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right). \quad (5.85)$$

The eigenvalue problem for  $L^2$  takes the form

$$L^2 \psi = \lambda \hbar^2 \psi, \quad (5.86)$$

where  $\psi(r, \theta, \varphi)$  is the wave-function, and  $\lambda$  is a number. Let us write

$$\psi(r, \theta, \varphi) = R(r) Y(\theta, \varphi). \quad (5.87)$$

Equation (5.86) reduces to

$$\left( \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right) Y + \lambda Y = 0, \quad (5.88)$$

where use has been made of Eq. (5.85). As is well-known, square integrable solutions to this equation only exist when  $\lambda$  takes the values  $l(l+1)$ , where  $l$  is an integer. These solutions are known as *spherical harmonics*, and can be written

$$Y_l^m(\theta, \varphi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} (-1)^m e^{im\varphi} P_l^m(\cos \theta), \quad (5.89)$$

where  $m$  is a positive integer lying in the range  $0 \leq m \leq l$ . Here,  $P_l^m(\xi)$  is an associated Legendre function satisfying the equation

$$\frac{d}{d\xi} \left[ (1 - \xi^2) \frac{dP_l^m}{d\xi} \right] - \frac{m^2}{1 - \xi^2} P_l^m + l(l+1) P_l^m = 0. \quad (5.90)$$

We define

$$Y_l^{-m} = (-1)^m (Y_l^m)^*, \quad (5.91)$$

which allows  $m$  to take the negative values  $-l \leq m < 0$ . The spherical harmonics are *orthogonal* functions, and are properly normalized with respect to integration over the entire solid angle:

$$\int_0^\pi \int_0^{2\pi} Y_l^{m*}(\theta, \varphi) Y_{l'}^{m'}(\theta, \varphi) \sin \theta \, d\theta \, d\varphi = \delta_{ll'} \delta_{mm'}. \quad (5.92)$$

The spherical harmonics also form a complete set for representing general functions of  $\theta$  and  $\varphi$ .

By definition,

$$L^2 Y_l^m = l(l+1) \hbar^2 Y_l^m, \quad (5.93)$$

where  $l$  is an integer. It follows from Eqs. (5.82) and (5.89) that

$$L_z Y_l^m = m \hbar Y_l^m, \quad (5.94)$$

where  $m$  is an integer lying in the range  $-l \leq m \leq l$ . Thus, the wave-function  $\psi(r, \theta, \varphi) = R(r) Y_l^m(\theta, \varphi)$ , where  $R$  is a general function, has all of the expected features of the wave-function of a simultaneous eigenstate of  $L^2$  and  $L_z$  belonging to the quantum numbers  $l$  and  $m$ . The well-known formula

$$\begin{aligned} \frac{dP_l^m}{d\xi} &= \frac{1}{\sqrt{1-\xi^2}} P_l^{m+1} - \frac{m\xi}{1-\xi^2} P_l^m \\ &= -\frac{(l+m)(l-m+1)}{\sqrt{1-\xi^2}} P_l^{m-1} + \frac{m\xi}{1-\xi^2} P_l^m \end{aligned} \quad (5.95)$$

can be combined with Eqs. (5.83) and (5.89) to give

$$L^+ Y_l^m = \sqrt{l(l+1) - m(m+1)} \hbar Y_l^{m+1}, \quad (5.96)$$

$$L^- Y_l^m = \sqrt{l(l+1) - m(m-1)} \hbar Y_l^{m-1}. \quad (5.97)$$

These equations are equivalent to Eqs. (5.55)–(5.56). Note that a spherical harmonic wave-function is symmetric about the  $z$ -axis (*i.e.*, independent of  $\varphi$ ) whenever  $m = 0$ , and is spherically symmetric whenever  $l = 0$  (since  $Y_0^0 = 1/\sqrt{4\pi}$ ).

In summary, by solving directly for the eigenfunctions of  $L^2$  and  $L_z$  in Schrödinger's representation, we have been able to reproduce all of the results of Sect. 5.2. Nevertheless, the results of Sect. 5.2 are more general than those obtained in this section, because they still apply when the quantum number  $l$  takes on half-integer values.

## 5.5 Motion in a central field

Consider a particle of mass  $M$  moving in a spherically symmetric potential. The Hamiltonian takes the form

$$H = \frac{\mathbf{p}^2}{2M} + V(r). \quad (5.98)$$

Adopting Schrödinger's representation, we can write  $\mathbf{p} = -(i/\hbar)\nabla$ . Hence,

$$H = -\frac{\hbar^2}{2M}\nabla^2 + V(r). \quad (5.99)$$

When written in spherical polar coordinates, the above equation becomes

$$H = -\frac{\hbar^2}{2M} \left[ \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right] + V(r). \quad (5.100)$$

Comparing this equation with Eq. (5.85), we find that

$$H = \frac{\hbar^2}{2M} \left[ -\frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{L^2}{\hbar^2 r^2} \right] + V(r). \quad (5.101)$$

Now, we know that the three components of angular momentum commute with  $L^2$  (see Sect. 5.1). We also know, from Eqs. (5.80)–(5.82), that  $L_x$ ,  $L_y$ , and  $L_z$  take the form of partial derivative operators of the *angular* coordinates,

when written in terms of spherical polar coordinates using Schrödinger's representation. It follows from Eq. (5.101) that all three components of the angular momentum commute with the Hamiltonian:

$$[\mathbf{L}, H] = 0. \quad (5.102)$$

It is also easily seen that  $L^2$  commutes with the Hamiltonian:

$$[L^2, H] = 0. \quad (5.103)$$

According to Sect. 4.2, the previous two equations ensure that the angular momentum  $\mathbf{L}$  and its magnitude squared  $L^2$  are both constants of the motion. This is as expected for a spherically symmetric potential.

Consider the energy eigenvalue problem

$$H\psi = E\psi, \quad (5.104)$$

where  $E$  is a number. Since  $L^2$  and  $L_z$  commute with each other and the Hamiltonian, it is always possible to represent the state of the system in terms of the simultaneous eigenstates of  $L^2$ ,  $L_z$ , and  $H$ . But, we already know that the most general form for the wave-function of a simultaneous eigenstate of  $L^2$  and  $L_z$  is (see previous section)

$$\psi(r, \theta, \varphi) = R(r) Y_l^m(\theta, \varphi). \quad (5.105)$$

Substituting Eq. (5.105) into Eq. (5.101), and making use of Eq. (5.93), we obtain

$$\left[ \frac{\hbar^2}{2M} \left( -\frac{1}{r^2} \frac{d}{dr} r^2 \frac{d}{dr} + \frac{l(l+1)}{r^2} \right) + V(r) - E \right] R = 0. \quad (5.106)$$

This is a Sturm-Liouville equation for the function  $R(r)$ . We know, from the general properties of this type of equation, that if  $R(r)$  is required to be well-behaved at  $r = 0$  and as  $r \rightarrow \infty$  then solutions only exist for a discrete set of values of  $E$ . These are the energy eigenvalues. In general, the energy eigenvalues depend on the quantum number  $l$ , but are independent of the quantum number  $m$ .

## 5.6 Energy levels of the hydrogen atom

Consider a hydrogen atom, for which the potential takes the specific form

$$V(r) = -\frac{e^2}{4\pi\epsilon_0 r}. \quad (5.107)$$

The radial eigenfunction  $R(r)$  satisfies Eq. (5.106), which can be written

$$\left[ \frac{\hbar^2}{2\mu} \left( -\frac{1}{r^2} \frac{d}{dr} r^2 \frac{d}{dr} + \frac{l(l+1)}{r^2} \right) - \frac{e^2}{4\pi\epsilon_0 r} - E \right] R = 0. \quad (5.108)$$

Here,  $\mu = m_e m_p / (m_e + m_p)$  is the *reduced mass*, which takes into account the fact that the electron (of mass  $m_e$ ) and the proton (of mass  $m_p$ ) both rotate about a common centre, which is equivalent to a particle of mass  $\mu$  rotating about a fixed point. Let us write the product  $r R(r)$  as the function  $P(r)$ . The above equation transforms to

$$\frac{d^2 P}{dr^2} - \frac{2\mu}{\hbar^2} \left( \frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{e^2}{4\pi\epsilon_0 r} - E \right) P = 0, \quad (5.109)$$

which is the one-dimensional Schrödinger equation for a particle of mass  $\mu$  moving in the *effective potential*

$$V_{\text{eff}}(r) = -\frac{e^2}{4\pi\epsilon_0 r} + \frac{l(l+1)\hbar^2}{2\mu r^2}. \quad (5.110)$$

The effective potential has a simple physical interpretation. The first part is the attractive Coulomb potential, and the second part corresponds to the repulsive centrifugal force.

Let

$$a = \sqrt{\frac{-\hbar^2}{2\mu E}}, \quad (5.111)$$

and  $y = r/a$ , with

$$P(r) = f(y) \exp(-y). \quad (5.112)$$

Here, it is assumed that the energy eigenvalue  $E$  is negative. Equation (5.109) transforms to

$$\left( \frac{d^2}{dy^2} - 2 \frac{d}{dy} - \frac{l(l+1)}{y^2} + \frac{2\mu e^2 a}{4\pi\epsilon_0 \hbar^2 y} \right) f = 0. \quad (5.113)$$

Let us look for a power-law solution of the form

$$f(y) = \sum_n c_n y^n. \quad (5.114)$$

Substituting this solution into Eq. (5.113), we obtain

$$\sum_n c_n \left\{ n(n-1)y^{n-2} - 2ny^{n-1} - l(l+1)y^{n-2} + \frac{2\mu e^2 a}{4\pi\epsilon_0 \hbar^2} y^{n-1} \right\} = 0. \quad (5.115)$$

Equating the coefficients of  $y^{n-2}$  gives

$$c_n[n(n-1) - l(l+1)] = c_{n-1} \left[ 2(n-1) - \frac{2\mu e^2 a}{4\pi\epsilon_0 \hbar^2} \right]. \quad (5.116)$$

Now, the power law series (5.114) must terminate at small  $n$ , at some positive value of  $n$ , otherwise  $f(y)$  behaves unphysically as  $y \rightarrow 0$ . This is only possible if  $[n_{\min}(n_{\min}-1) - l(l+1)] = 0$ , where the first term in the series is  $c_{n_{\min}} y^{n_{\min}}$ . There are two possibilities:  $n_{\min} = -l$  or  $n_{\min} = l+1$ . The former predicts unphysical behaviour of the wave-function at  $y = 0$ . Thus, we conclude that  $n_{\min} = l+1$ . Note that for an  $l = 0$  state there is a finite probability of finding the electron at the nucleus, whereas for an  $l > 0$  state there is zero probability of finding the electron at the nucleus (*i.e.*,  $|\psi|^2 = 0$  at  $r = 0$ , except when  $l = 0$ ). Note, also, that it is only possible to obtain sensible behaviour of the wave-function as  $r \rightarrow 0$  if  $l$  is an integer.

For large values of  $y$ , the ratio of successive terms in the series (5.114) is

$$\frac{c_n y}{c_{n-1}} = \frac{2y}{n}, \quad (5.117)$$

according to Eq. (5.116). This is the same as the ratio of successive terms in the series

$$\sum_n \frac{(2y)^n}{n!}, \quad (5.118)$$

which converges to  $\exp(2y)$ . We conclude that  $f(y) \rightarrow \exp(2y)$  as  $y \rightarrow \infty$ . It follows from Eq. (5.112) that  $R(r) \rightarrow \exp(r/a)/r$  as  $r \rightarrow \infty$ . This does not correspond to physically acceptable behaviour of the wave-function, since  $\int |\psi|^2 dV$  must be finite. The only way in which we can avoid this unphysical behaviour is

if the series (5.114) terminates at some maximum value of  $n$ . According to the recursion relation (5.116), this is only possible if

$$\frac{\mu e^2 a}{4\pi\epsilon_0 \hbar^2} = n, \quad (5.119)$$

where the last term in the series is  $c_n y^n$ . It follows from Eq. (5.111) that the energy eigenvalues are *quantized*, and can only take the values

$$E = -\frac{\mu e^4}{32\pi^2 \epsilon_0^2 \hbar^2 n^2}. \quad (5.120)$$

Here,  $n$  is a positive integer which must exceed the quantum number  $l$ , otherwise there would be no terms in the series (5.114).

It is clear that the wave-function for a hydrogen atom can be written

$$\psi(r, \theta, \varphi) = R(r/a) Y_l^m(\theta, \varphi), \quad (5.121)$$

where

$$a = \frac{n 4\pi\epsilon_0 \hbar^2}{\mu e^2} = 5.3 \times 10^{-11} n \text{ meters}, \quad (5.122)$$

and  $R(x)$  is a well-behaved solution of the differential equation

$$\left( \frac{1}{x^2} \frac{d}{dx} x^2 \frac{d}{dx} - \frac{l(l+1)}{x^2} + \frac{2n}{x} - 1 \right) R = 0. \quad (5.123)$$

Finally, the  $Y_l^m$  are spherical harmonics. The restrictions on the quantum numbers are  $|m| \leq l < n$ . Here,  $n$  is a positive integer,  $l$  is a non-negative integer, and  $m$  is an integer.

The ground state of hydrogen corresponds to  $n = 1$ . The only permissible values of the other quantum numbers are  $l = 0$  and  $m = 0$ . Thus, the ground state is a spherically symmetric, zero angular momentum state. The energy of the ground state is

$$E_0 = -\frac{\mu e^4}{32\pi^2 \epsilon_0^2 \hbar^2} = -13.6 \text{ electron volts}. \quad (5.124)$$

The next energy level corresponds to  $n = 2$ . The other quantum numbers are allowed to take the values  $l = 0, m = 0$  or  $l = 1, m = -1, 0, 1$ . Thus, there are  $n = 2$  states with non-zero angular momentum. Note that the energy levels given in Eq. (5.120) are independent of the quantum number  $l$ , despite the fact that  $l$  appears in the radial eigenfunction equation (5.123). This is a special property of a  $1/r$  Coulomb potential.

In addition to the quantized negative energy state of the hydrogen atom, which we have just found, there is also a continuum of unbound positive energy states.

## 5.7 Spin angular momentum

Up to now, we have tacitly assumed that the state of a particle in quantum mechanics can be completely specified by giving the wave-function  $\psi$  as a function of the spatial coordinates  $x, y$ , and  $z$ . Unfortunately, there is a wealth of experimental evidence which suggests that this simplistic approach is incomplete.

Consider an isolated system at rest, and let the eigenvalue of its total angular momentum be  $j(j+1)\hbar^2$ . According to the theory of orbital angular momentum outlined in Sects. 5.4 and 5.5, there are two possibilities. For a system consisting of a single particle,  $j = 0$ . For a system consisting of two (or more) particles,  $j$  is a non-negative integer. However, this does not agree with observations, because we often find systems which appear to be structureless, and yet have  $j \neq 0$ . Even worse, systems where  $j$  has half-integer values abound in nature. In order to explain this apparent discrepancy between theory and experiments, Goudsmit and Uhlenbeck (in 1925) introduced the concept of an internal, purely quantum mechanical, angular momentum called *spin*. For a particle with spin, the total angular momentum in the rest frame is non-vanishing.

Let us denote the three components of the spin angular momentum of a particle by the Hermitian operators  $(S_x, S_y, S_z) \equiv \mathbf{S}$ . We assume that these operators obey the fundamental commutation relations (5.8)–(5.10) for the components of an angular momentum. Thus, we can write

$$\mathbf{S} \times \mathbf{S} = i \hbar \mathbf{S}. \quad (5.125)$$

We can also define the operator

$$S^2 = S_x^2 + S_y^2 + S_z^2. \quad (5.126)$$

According to the quite general analysis of Sect. 5.1,

$$[\mathbf{S}, S^2] = 0. \quad (5.127)$$

Thus, it is possible to find simultaneous eigenstates of  $S^2$  and  $S_z$ . These are denoted  $|s, s_z\rangle$ , where

$$S_z |s, s_z\rangle = s_z \hbar |s, s_z\rangle, \quad (5.128)$$

$$S^2 |s, s_z\rangle = s(s+1) \hbar^2 |s, s_z\rangle. \quad (5.129)$$

According to the equally general analysis of Sect. 5.2, the quantum number  $s$  can, in principle, take integer or half-integer values, and the quantum number  $s_z$  can only take the values  $s, s-1, \dots, -s+1, -s$ .

Spin angular momentum clearly has many properties in common with orbital angular momentum. However, there is one vitally important difference. Spin angular momentum operators *cannot* be expressed in terms of position and momentum operators, like in Eqs. (5.1)–(5.3), since this identification depends on an analogy with classical mechanics, and the concept of spin is purely quantum mechanical: *i.e.*, it has no analogy in classical physics. Consequently, the restriction that the quantum number of the overall angular momentum must take *integer* values is lifted for spin angular momentum, since this restriction (found in Sects. 5.3 and 5.4) depends on Eqs. (5.1)–(5.3). In other words, the quantum number  $s$  is allowed to take *half-integer* values.

Consider a spin one-half particle, for which

$$S_z |\pm\rangle = \pm \frac{\hbar}{2} |\pm\rangle, \quad (5.130)$$

$$S^2 |\pm\rangle = \frac{3}{4} \hbar^2 |\pm\rangle. \quad (5.131)$$

Here, the  $|\pm\rangle$  denote eigenkets of the  $S_z$  operator corresponding to the eigenvalues  $\pm\hbar/2$ . These kets are orthonormal (since  $S_z$  is an Hermitian operator), so

$$\langle + | - \rangle = 0. \quad (5.132)$$

They are also properly normalized and complete, so that

$$\langle +|+\rangle = \langle -|-\rangle = 1, \quad (5.133)$$

and

$$|+\rangle\langle +| + |-\rangle\langle -| = 1. \quad (5.134)$$

It is easily verified that the Hermitian operators defined by

$$S_x = \frac{\hbar}{2} (|+\rangle\langle -| + |-\rangle\langle +|), \quad (5.135)$$

$$S_y = \frac{i\hbar}{2} (-|+\rangle\langle -| + |-\rangle\langle +|), \quad (5.136)$$

$$S_z = \frac{\hbar}{2} (|+\rangle\langle +| - |-\rangle\langle -|), \quad (5.137)$$

satisfy the commutation relations (5.8)–(5.10) (with the  $L_j$  replaced by the  $S_j$ ). The operator  $S^2$  takes the form

$$S^2 = \frac{3\hbar^2}{4}. \quad (5.138)$$

It is also easily demonstrated that  $S^2$  and  $S_z$ , defined in this manner, satisfy the eigenvalue relations (5.130)–(5.131). Equations (5.135)–(5.138) constitute a realization of the spin operators  $\mathbf{S}$  and  $S^2$  (for a spin one-half particle) in *spin space* (*i.e.*, that Hilbert sub-space consisting of kets which correspond to the different spin states of the particle).

## 5.8 Wave-function of a spin one-half particle

The state of a spin one-half particle is represented as a vector in ket space. Let us suppose that this space is spanned by the basis kets  $|x', y', z', \pm\rangle$ . Here,  $|x', y', z', \pm\rangle$  denotes a simultaneous eigenstate of the position operators  $x$ ,  $y$ ,  $z$ , and the spin operator  $S_z$ , corresponding to the eigenvalues  $x'$ ,  $y'$ ,  $z'$ , and  $\pm\hbar/2$ , respectively. The basis kets are assumed to satisfy the completeness relation

$$\iiint (|x', y', z', +\rangle\langle x', y', z', +| + |x', y', z', -\rangle\langle x', y', z', -|) dx' dy' dz' = 1. \quad (5.139)$$

It is helpful to think of the ket  $|x', y', z', +\rangle$  as the product of two kets—a position space ket  $|x', y', z'\rangle$ , and a spin space ket  $|+\rangle$ . We assume that such a product obeys the commutative and distributive axioms of multiplication:

$$|x', y', z'\rangle|+\rangle = |+\rangle|x', y', z'\rangle, \quad (5.140)$$

$$(c'|x', y', z'\rangle + c''|x'', y'', z''\rangle)|+\rangle = c'|x', y', z'\rangle|+\rangle + c''|x'', y'', z''\rangle|+\rangle \quad (5.141)$$

$$|x', y', z'\rangle(c_+|+\rangle + c_-|-\rangle) = c_+|x', y', z'\rangle|+\rangle + c_-|x', y', z'\rangle|-\rangle, \quad (5.142)$$

where the  $c$ 's are numbers. We can give meaning to any position space operator (such as  $L_z$ ) acting on the product  $|x', y', z'\rangle|+\rangle$  by assuming that it operates only on the  $|x', y', z'\rangle$  factor, and commutes with the  $|+\rangle$  factor. Similarly, we can give a meaning to any spin operator (such as  $S_z$ ) acting on  $|x', y', z'\rangle|+\rangle$  by assuming that it operates only on  $|+\rangle$ , and commutes with  $|x', y', z'\rangle$ . This implies that every position space operator commutes with every spin operator. In this manner, we can give meaning to the equation

$$|x', y', z', \pm\rangle = |x', y', z'\rangle|\pm\rangle = |\pm\rangle|x', y', z'\rangle. \quad (5.143)$$

The multiplication in the above equation is of quite a different type to any which we have encountered previously. The ket vectors  $|x', y', z'\rangle$  and  $|\pm\rangle$  are in two quite separate vector spaces, and their product  $|x', y', z'\rangle|\pm\rangle$  is in a third vector space. In mathematics, the latter space is termed the *product space* of the former spaces, which are termed *factor spaces*. The number of dimensions of a product space is equal to the product of the number of dimensions of each of the factor spaces. A general ket of the product space is not of the form (5.143), but is instead a sum or integral of kets of this form.

A general state  $A$  of a spin one-half particle is represented as a ket  $||A\rangle\rangle$  in the product of the spin and position spaces. This state can be completely specified by *two* wavefunctions:

$$\psi_+(x', y', z') = \langle x', y', z' | \langle + || A \rangle \rangle, \quad (5.144)$$

$$\psi_-(x', y', z') = \langle x', y', z' | \langle - || A \rangle \rangle. \quad (5.145)$$

The probability of observing the particle in the region  $x'$  to  $x' + dx'$ ,  $y'$  to  $y' + dy'$ , and  $z'$  to  $z' + dz'$ , with  $s_z = +1/2$  is  $|\psi_+(x', y', z')|^2 dx' dy' dz'$ . Likewise, the probability of observing the particle in the region  $x'$  to  $x' + dx'$ ,  $y'$  to  $y' + dy'$ , and  $z'$  to  $z' + dz'$ , with  $s_z = -1/2$  is  $|\psi_-(x', y', z')|^2 dx' dy' dz'$ . The normalization condition for the wavefunctions is

$$\iiint (|\psi_+|^2 + |\psi_-|^2) dx' dy' dz' = 1. \quad (5.146)$$

## 5.9 Rotation operators in spin space

Let us, for the moment, forget about the spatial position of the particle, and concentrate on its spin state. A general spin state  $A$  is represented by the ket

$$|A\rangle = \langle +|A\rangle|+\rangle + \langle -|A\rangle|-\rangle \quad (5.147)$$

in spin space. In Sect. 5.3, we were able to construct an operator  $R_z(\Delta\varphi)$  which rotates the system by an angle  $\Delta\varphi$  about the  $z$ -axis in position space. Can we also construct an operator  $T_z(\Delta\varphi)$  which rotates the system by an angle  $\Delta\varphi$  about the  $z$ -axis in spin space? By analogy with Eq. (5.62), we would expect such an operator to take the form

$$T_z(\Delta\varphi) = \exp(-i S_z \Delta\varphi/\hbar). \quad (5.148)$$

Thus, after rotation, the ket  $|A\rangle$  becomes

$$|A_R\rangle = T_z(\Delta\varphi)|A\rangle. \quad (5.149)$$

To demonstrate that the operator (5.148) really does rotate the spin of the system, let us consider its effect on  $\langle S_x \rangle$ . Under rotation, this expectation value changes as follows:

$$\langle S_x \rangle \rightarrow \langle A_R | S_x | A_R \rangle = \langle A | T_z^\dagger S_x T_z | A \rangle. \quad (5.150)$$

Thus, we need to compute

$$\exp(i S_z \Delta\varphi/\hbar) S_x \exp(-i S_z \Delta\varphi/\hbar). \quad (5.151)$$

This can be achieved in two different ways.

First, we can use the explicit formula for  $S_x$  given in Eq. (5.135). We find that Eq. (5.151) becomes

$$\frac{\hbar}{2} \exp(i S_z \Delta\varphi/\hbar) (|+\rangle\langle-| + |-\rangle\langle+|) \exp(-i S_z \Delta\varphi/\hbar), \quad (5.152)$$

or

$$\frac{\hbar}{2} (e^{i\Delta\varphi/2} |+\rangle\langle-| e^{i\Delta\varphi/2} + e^{-i\Delta\varphi/2} |-\rangle\langle+| e^{-i\Delta\varphi/2}), \quad (5.153)$$

which reduces to

$$S_x \cos \Delta\varphi - S_y \sin \Delta\varphi, \quad (5.154)$$

where use has been made of Eqs. (5.135)–(5.137).

A second approach is to use the so called *Baker-Hausdorff lemma*. This takes the form

$$\begin{aligned} \exp(i G \lambda) A \exp(-i G \lambda) &= A + i\lambda[G, A] + \left(\frac{i^2\lambda^2}{2!}\right) [G, [G, A]] + \quad (5.155) \\ &\dots + \left(\frac{i^n\lambda^n}{n!}\right) [G, [G, [G, \dots [G, A]] \dots], \end{aligned}$$

where  $G$  is a Hermitian operator, and  $\lambda$  is a real parameter. The proof of this lemma is left as an exercise. Applying the Baker-Hausdorff lemma to Eq. (5.151), we obtain

$$S_x + \left(\frac{i\Delta\varphi}{\hbar}\right) [S_z, S_x] + \left(\frac{1}{2!}\right) \left(\frac{i\Delta\varphi}{\hbar}\right)^2 [S_z, [S_z, S_x]] + \dots, \quad (5.156)$$

which reduces to

$$S_x \left[1 - \frac{\Delta\varphi^2}{2!} + \dots\right] - S_y \left[\varphi - \frac{\Delta\varphi^3}{3!} + \dots\right], \quad (5.157)$$

or

$$S_x \cos \Delta\varphi - S_y \sin \Delta\varphi, \quad (5.158)$$

where use has been made of Eq. (5.125). The second proof is more general than the first, since it only uses the fundamental commutation relation (5.125), and is, therefore, valid for systems with spin angular momentum higher than one-half.

For a spin one-half system, both methods imply that

$$\langle S_x \rangle \rightarrow \langle S_x \rangle \cos \Delta\varphi - \langle S_y \rangle \sin \Delta\varphi \quad (5.159)$$

under the action of the rotation operator (5.148). It is straight-forward to show that

$$\langle S_y \rangle \rightarrow \langle S_y \rangle \cos \Delta\varphi + \langle S_x \rangle \sin \Delta\varphi. \quad (5.160)$$

Furthermore,

$$\langle S_z \rangle \rightarrow \langle S_z \rangle, \quad (5.161)$$

since  $S_z$  commutes with the rotation operator. Equations (5.159)–(5.161) demonstrate that the operator (5.148) rotates the expectation value of  $\mathbf{S}$  by an angle  $\Delta\varphi$  about the  $z$ -axis. In fact, the expectation value of the spin operator behaves like a classical vector under rotation:

$$\langle S_k \rangle \rightarrow \sum_l R_{kl} \langle S_l \rangle, \quad (5.162)$$

where the  $R_{kl}$  are the elements of the conventional rotation matrix for the rotation in question. It is clear, from our second derivation of the result (5.159), that this property is not restricted to the spin operators of a spin one-half system. In fact, we have effectively demonstrated that

$$\langle J_k \rangle \rightarrow \sum_l R_{kl} \langle J_l \rangle, \quad (5.163)$$

where the  $J_k$  are the generators of rotation, satisfying the fundamental commutation relation  $\mathbf{J} \times \mathbf{J} = i \hbar \mathbf{J}$ , and the rotation operator about the  $k$ th axis is written  $R_k(\Delta\varphi) = \exp(-i J_k \Delta\varphi / \hbar)$ .

Consider the effect of the rotation operator (5.148) on the state ket (5.147). It is easily seen that

$$T_z(\Delta\varphi)|\mathbf{A}\rangle = e^{-i\Delta\varphi/2} \langle +|\mathbf{A}\rangle |+\rangle + e^{i\Delta\varphi/2} \langle -|\mathbf{A}\rangle |-\rangle. \quad (5.164)$$

Consider a rotation by  $2\pi$  radians. We find that

$$|\mathbf{A}\rangle \rightarrow T_z(2\pi)|\mathbf{A}\rangle = -|\mathbf{A}\rangle. \quad (5.165)$$

Note that a ket rotated by  $2\pi$  radians differs from the original ket by a *minus* sign. In fact, a rotation by  $4\pi$  radians is needed to transform a ket into itself. The minus sign does not affect the expectation value of  $\mathbf{S}$ , since  $\mathbf{S}$  is sandwiched between  $\langle A|$  and  $|A\rangle$ , both of which change sign. Nevertheless, the minus sign does give rise to observable consequences, as we shall see presently.

### 5.10 Magnetic moments

Consider a particle of charge  $q$  and velocity  $\mathbf{v}$  performing a circular orbit of radius  $r$  in the  $x$ - $y$  plane. The charge is equivalent to a current loop of radius  $r$  in the  $x$ - $y$  plane carrying current  $I = qv/2\pi r$ . The magnetic moment  $\boldsymbol{\mu}$  of the loop is of magnitude  $\pi r^2 I$  and is directed along the  $z$ -axis. Thus, we can write

$$\boldsymbol{\mu} = \frac{q}{2} \mathbf{r} \times \mathbf{v}, \quad (5.166)$$

where  $\mathbf{r}$  and  $\mathbf{v}$  are the vector position and velocity of the particle, respectively. However, we know that  $\mathbf{p} = \mathbf{v}/m$ , where  $\mathbf{p}$  is the vector momentum of the particle, and  $m$  is its mass. We also know that  $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ , where  $\mathbf{L}$  is the orbital angular momentum. It follows that

$$\boldsymbol{\mu} = \frac{q}{2m} \mathbf{L}. \quad (5.167)$$

Using the usual analogy between classical and quantum mechanics, we expect the above relation to also hold between the quantum mechanical operators,  $\boldsymbol{\mu}$  and  $\mathbf{L}$ , which represent magnetic moment and orbital angular momentum, respectively. This is indeed found to be the case.

Does spin angular momentum also give rise to a contribution to the magnetic moment of a charged particle? The answer is “yes”. In fact, relativistic quantum mechanics actually predicts that a charged particle possessing spin should also possess a magnetic moment (this was first demonstrated by Dirac). We can write

$$\boldsymbol{\mu} = \frac{q}{2m} (\mathbf{L} + g \mathbf{S}), \quad (5.168)$$

where  $g$  is called the *gyromagnetic ratio*. For an electron this ratio is found to be

$$g_e = 2 \left( 1 + \frac{1}{2\pi} \frac{e^2}{4\pi\epsilon_0 \hbar c} \right). \quad (5.169)$$

The factor 2 is correctly predicted by Dirac's relativistic theory of the electron. The small correction  $1/(2\pi 137)$ , derived originally by Schwinger, is due to quantum field effects. We shall ignore this correction in the following, so

$$\boldsymbol{\mu} \simeq -\frac{e}{2m_e} (\mathbf{L} + 2\mathbf{S}) \quad (5.170)$$

for an electron (here,  $e > 0$ ).

### 5.11 Spin precession

The Hamiltonian for an electron at rest in a  $z$ -directed magnetic field,  $\mathbf{B} = B \hat{\mathbf{z}}$ , is

$$H = -\boldsymbol{\mu} \cdot \mathbf{B} = \left( \frac{e}{m_e} \right) \mathbf{S} \cdot \mathbf{B} = \omega S_z, \quad (5.171)$$

where

$$\omega = \frac{eB}{m_e}. \quad (5.172)$$

According to Eq. (4.28), the time evolution operator for this system is

$$T(t, 0) = \exp(-iHt/\hbar) = \exp(-iS_z \omega t/\hbar). \quad (5.173)$$

It can be seen, by comparison with Eq. (5.148), that the time evolution operator is precisely the same as the rotation operator for spin, with  $\Delta\varphi$  set equal to  $\omega t$ . It is immediately clear that the Hamiltonian (5.171) causes the electron spin to precess about the  $z$ -axis with angular frequency  $\omega$ . In fact, Eqs. (5.159)–(5.161) imply that

$$\langle S_x \rangle_t = \langle S_x \rangle_{t=0} \cos \omega t - \langle S_y \rangle_{t=0} \sin \omega t, \quad (5.174)$$

$$\langle S_y \rangle_t = \langle S_y \rangle_{t=0} \cos \omega t + \langle S_x \rangle_{t=0} \sin \omega t, \quad (5.175)$$

$$\langle S_z \rangle_t = \langle S_z \rangle_{t=0}. \quad (5.176)$$

The time evolution of the state ket is given by analogy with Eq. (5.164):

$$|A, t\rangle = e^{-i\omega t/2} \langle +|A, 0\rangle |+\rangle + e^{i\omega t/2} \langle -|A, 0\rangle |-\rangle. \quad (5.177)$$

Note that it takes time  $t = 4\pi/\omega$  for the state ket to return to its original state. By contrast, it only takes times  $t = 2\pi/\omega$  for the spin vector to point in its original direction.

We now describe an experiment to detect the minus sign in Eq. (5.165). An almost monoenergetic beam of neutrons is split in two, sent along two different paths, A and B, and then recombined. Path A goes through a magnetic field free region. However, path B enters a small region where a static magnetic field is present. As a result, a neutron state ket going along path B acquires a phase-shift  $\exp(\mp i \omega T/2)$  (the  $\mp$  signs correspond to  $s_z = \pm 1/2$  states). Here,  $T$  is the time spent in the magnetic field, and  $\omega$  is the spin precession frequency

$$\omega = \frac{g_n e B}{m_p}. \quad (5.178)$$

This frequency is defined in an analogous manner to Eq. (5.172). The gyro-magnetic ratio for a neutron is found experimentally to be  $g_n = -1.91$ . (The magnetic moment of a neutron is entirely a quantum field effect). When neutrons from path A and path B meet they undergo interference. We expect the observed neutron intensity in the interference region to exhibit a  $\cos(\pm \omega T/2 + \delta)$  variation, where  $\delta$  is the phase difference between paths A and B in the absence of a magnetic field. In experiments, the time of flight  $T$  through the magnetic field region is kept constant, while the field-strength  $B$  is varied. It follows that the change in magnetic field required to produce successive maxima is

$$\Delta B = \frac{4\pi \hbar}{e g_n \lambda l}, \quad (5.179)$$

where  $l$  is the path-length through the magnetic field region, and  $\lambda$  is the de Broglie wavelength over  $2\pi$  of the neutrons. The above prediction has been verified experimentally to within a fraction of a percent. This prediction depends crucially on the fact that it takes a  $4\pi$  rotation to return a state ket to its original state. If it only took a  $2\pi$  rotation then  $\Delta B$  would be half of the value given above, which does not agree with the experimental data.

## 5.12 Pauli two-component formalism

We have seen, in Sect. 5.4, that the eigenstates of orbital angular momentum can be conveniently represented as spherical harmonics. In this representation, the orbital angular momentum operators take the form of differential operators involving only angular coordinates. It is conventional to represent the eigenstates of spin angular momentum as column (or row) matrices. In this representation, the spin angular momentum operators take the form of matrices.

The matrix representation of a spin one-half system was introduced by Pauli in 1926. Recall, from Sect. 5.9, that a general spin ket can be expressed as a linear combination of the two eigenkets of  $S_z$  belonging to the eigenvalues  $\pm\hbar/2$ . These are denoted  $|\pm\rangle$ . Let us represent these basis eigenkets as column matrices:

$$|+\rangle \rightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv \chi_+, \quad (5.180)$$

$$|-\rangle \rightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv \chi_-. \quad (5.181)$$

The corresponding eigenbras are represented as row matrices:

$$\langle +| \rightarrow (1, 0) \equiv \chi_+^\dagger, \quad (5.182)$$

$$\langle -| \rightarrow (0, 1) \equiv \chi_-^\dagger. \quad (5.183)$$

In this scheme, a general bra takes the form

$$|\mathcal{A}\rangle = \langle +|\mathcal{A}\rangle|+\rangle + \langle -|\mathcal{A}\rangle|-\rangle \rightarrow \begin{pmatrix} \langle +|\mathcal{A}\rangle \\ \langle -|\mathcal{A}\rangle \end{pmatrix}, \quad (5.184)$$

and a general ket becomes

$$\langle \mathcal{A}| = \langle \mathcal{A}|+\rangle\langle +| + \langle \mathcal{A}|-\rangle\langle -| \rightarrow (\langle \mathcal{A}|+, \langle \mathcal{A}|-). \quad (5.185)$$

The column matrix (5.184) is called a two-component *spinor*, and can be written

$$\chi \equiv \begin{pmatrix} \langle +|\mathcal{A}\rangle \\ \langle -|\mathcal{A}\rangle \end{pmatrix} = \begin{pmatrix} c_+ \\ c_- \end{pmatrix} = c_+ \chi_+ + c_- \chi_-, \quad (5.186)$$

where the  $c_{\pm}$  are complex numbers. The row matrix (5.185) becomes

$$\chi^{\dagger} \equiv (\langle A|+, \langle A|- \rangle) = (c_+^*, c_-^*) = c_+^* \chi_+^{\dagger} + c_-^* \chi_-^{\dagger}. \quad (5.187)$$

Consider the ket obtained by the action of a spin operator on ket  $A$ :

$$|A'\rangle = S_k |A\rangle. \quad (5.188)$$

This ket is represented as

$$|A'\rangle \rightarrow \begin{pmatrix} \langle +|A'\rangle \\ \langle -|A'\rangle \end{pmatrix} \equiv \chi'. \quad (5.189)$$

However,

$$\langle +|A'\rangle = \langle +|S_k|+\rangle \langle +|A\rangle + \langle +|S_k|-\rangle \langle -|A\rangle, \quad (5.190)$$

$$\langle -|A'\rangle = \langle -|S_k|+\rangle \langle +|A\rangle + \langle -|S_k|-\rangle \langle -|A\rangle, \quad (5.191)$$

or

$$\begin{pmatrix} \langle +|A'\rangle \\ \langle -|A'\rangle \end{pmatrix} = \begin{pmatrix} \langle +|S_k|+\rangle & \langle +|S_k|-\rangle \\ \langle -|S_k|+\rangle & \langle -|S_k|-\rangle \end{pmatrix} \begin{pmatrix} \langle +|A\rangle \\ \langle -|A\rangle \end{pmatrix}. \quad (5.192)$$

It follows that we can represent the operator/ket relation (5.188) as the matrix relation

$$\chi' = \left(\frac{\hbar}{2}\right) \sigma_k \chi, \quad (5.193)$$

where the  $\sigma_k$  are the matrices of the  $\langle \pm|S_k|\pm\rangle$  values divided by  $\hbar/2$ . These matrices, which are called the *Pauli matrices*, can easily be evaluated using the explicit forms for the spin operators given in Eqs. (5.135)–(5.137). We find that

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (5.194)$$

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad (5.195)$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (5.196)$$

Here, 1, 2, and 3 refer to  $x$ ,  $y$ , and  $z$ , respectively. Note that, in this scheme, we are effectively representing the spin operators in terms of the Pauli matrices:

$$S_k \rightarrow \left(\frac{\hbar}{2}\right) \sigma_k. \quad (5.197)$$

The expectation value of  $S_k$  can be written in terms of spinors and the Pauli matrices:

$$\langle S_k \rangle = \langle A | S_k | A \rangle = \sum_{\pm} \langle A | \pm \rangle \langle \pm | S_k | \pm \rangle \langle \pm | A \rangle = \left(\frac{\hbar}{2}\right) \chi^\dagger \sigma_k \chi. \quad (5.198)$$

The fundamental commutation relation for angular momentum, Eq. (5.125), can be combined with (5.197) to give the following commutation relation for the Pauli matrices:

$$\boldsymbol{\sigma} \times \boldsymbol{\sigma} = 2i \boldsymbol{\sigma}. \quad (5.199)$$

It is easily seen that the matrices (5.194)–(5.196) actually satisfy these relations (*i.e.*,  $\sigma_1 \sigma_2 - \sigma_2 \sigma_1 = 2i \sigma_3$ , plus all cyclic permutations). It is also easily seen that the Pauli matrices satisfy the anti-commutation relations

$$\{\sigma_i, \sigma_j\} = 2 \delta_{ij}. \quad (5.200)$$

Let us examine how the Pauli scheme can be extended to take into account the position of a spin one-half particle. Recall, from Sect. 5.8, that we can represent a general basis ket as the product of basis kets in position space and spin space:

$$|x', y', z', \pm\rangle = |x', y', z'\rangle |\pm\rangle = |\pm\rangle |x', y', z'\rangle. \quad (5.201)$$

The ket corresponding to state  $A$  is denoted  $\|A\rangle\rangle$ , and resides in the product space of the position and spin ket spaces. State  $A$  is completely specified by the two wave-functions

$$\psi_+(x', y', z') = \langle x', y', z' | \langle + \| A \rangle\rangle, \quad (5.202)$$

$$\psi_-(x', y', z') = \langle x', y', z' | \langle - \| A \rangle\rangle. \quad (5.203)$$

Consider the operator relation

$$\|A'\rangle\rangle = S_k \|A\rangle\rangle. \quad (5.204)$$

It is easily seen that

$$\begin{aligned}\langle x', y', z' | \langle + | \mathcal{A}' \rangle \rangle &= \langle + | S_k | + \rangle \langle x', y', z' | \langle + | \mathcal{A} \rangle \rangle \\ &\quad + \langle + | S_k | - \rangle \langle x', y', z' | \langle - | \mathcal{A} \rangle \rangle,\end{aligned}\quad (5.205)$$

$$\begin{aligned}\langle x', y', z' | \langle - | \mathcal{A}' \rangle \rangle &= \langle - | S_k | + \rangle \langle x', y', z' | \langle + | \mathcal{A} \rangle \rangle \\ &\quad + \langle - | S_k | - \rangle \langle x', y', z' | \langle - | \mathcal{A} \rangle \rangle,\end{aligned}\quad (5.206)$$

where use has been made of the fact that the spin operator  $S_k$  commutes with the eigenbras  $\langle x', y', z' |$ . It is fairly obvious that we can represent the operator relation (5.204) as a matrix relation if we generalize our definition of a spinor by writing

$$| \mathcal{A} \rangle \rangle \rightarrow \begin{pmatrix} \psi_+(\mathbf{r}') \\ \psi_-(\mathbf{r}') \end{pmatrix} \equiv \chi, \quad (5.207)$$

and so on. The components of a spinor are now wave-functions, instead of complex numbers. In this scheme, the operator equation (5.204) becomes simply

$$\chi' = \left( \frac{\hbar}{2} \right) \sigma_k \chi. \quad (5.208)$$

Consider the operator relation

$$| \mathcal{A}' \rangle \rangle = p_k | \mathcal{A} \rangle \rangle. \quad (5.209)$$

In the Schrödinger representation, we have

$$\begin{aligned}\langle x', y', z' | \langle + | \mathcal{A}' \rangle \rangle &= \langle x', y', z' | p_k \langle + | \mathcal{A} \rangle \rangle \\ &= -i \hbar \frac{\partial}{\partial x'_k} \langle x', y', z' | \langle + | \mathcal{A} \rangle \rangle,\end{aligned}\quad (5.210)$$

$$\begin{aligned}\langle x', y', z' | \langle - | \mathcal{A}' \rangle \rangle &= \langle x', y', z' | p_k \langle - | \mathcal{A} \rangle \rangle \\ &= -i \hbar \frac{\partial}{\partial x'_k} \langle x', y', z' | \langle - | \mathcal{A} \rangle \rangle,\end{aligned}\quad (5.211)$$

where use has been made of Eq. (3.78). The above equation reduces to

$$\begin{pmatrix} \psi'_+(\mathbf{r}') \\ \psi'_-(\mathbf{r}') \end{pmatrix} = \begin{pmatrix} -i \hbar \partial \psi_+(\mathbf{r}') / \partial x'_k \\ -i \hbar \partial \psi_-(\mathbf{r}') / \partial x'_k \end{pmatrix}. \quad (5.212)$$

Thus, the operator equation (5.209) can be written

$$\chi' = p_k \chi, \quad (5.213)$$

where

$$p_k \rightarrow -i \hbar \frac{\partial}{\partial x'_k} \mathbf{I}. \quad (5.214)$$

Here,  $\mathbf{I}$  is the  $2 \times 2$  unit matrix. In fact, any position operator (e.g.,  $p_k$  or  $L_k$ ) is represented in the Pauli scheme as some differential operator of the position eigenvalues multiplied by the  $2 \times 2$  unit matrix.

What about combinations of position and spin operators? The most commonly occurring combination is a dot product: e.g.,  $\mathbf{S} \cdot \mathbf{L} = (\hbar/2) \boldsymbol{\sigma} \cdot \mathbf{L}$ . Consider the hybrid operator  $\boldsymbol{\sigma} \cdot \mathbf{a}$ , where  $\mathbf{a} \equiv (a_x, a_y, a_z)$  is some vector position operator. This quantity is represented as a  $2 \times 2$  matrix:

$$\boldsymbol{\sigma} \cdot \mathbf{a} \equiv \sum_k a_k \sigma_k = \begin{pmatrix} +a_3 & a_1 - i a_2 \\ a_1 + i a_2 & -a_3 \end{pmatrix}. \quad (5.215)$$

Since, in the Schrödinger representation, a general position operator takes the form of a differential operator in  $x'$ ,  $y'$ , or  $z'$ , it is clear that the above quantity must be regarded as a matrix differential operator which acts on spinors of the general form (5.207). The important identity

$$(\boldsymbol{\sigma} \cdot \mathbf{a})(\boldsymbol{\sigma} \cdot \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} + i \boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}) \quad (5.216)$$

follows from the commutation and anti-commutation relations (5.199) and (5.200). Thus,

$$\begin{aligned} \sum_j \sigma_j a_j \sum_k \sigma_k b_k &= \sum_j \sum_k \left( \frac{1}{2} \{\sigma_j, \sigma_k\} + \frac{1}{2} [\sigma_j, \sigma_k] \right) a_j b_k \\ &= \sum_j \sum_k (\sigma_{jk} + i \epsilon_{jkl} \sigma_l) a_j b_k \\ &= \mathbf{a} \cdot \mathbf{b} + i \boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}). \end{aligned} \quad (5.217)$$

A general rotation operator in spin space is written

$$T(\Delta\phi) = \exp(-i \mathbf{S} \cdot \mathbf{n} \Delta\phi / \hbar), \quad (5.218)$$

by analogy with Eq. (5.148), where  $\mathbf{n}$  is a unit vector pointing along the axis of rotation, and  $\Delta\varphi$  is the angle of rotation. Here,  $\mathbf{n}$  can be regarded as a trivial position operator. The rotation operator is represented

$$\exp(-i\mathbf{S}\cdot\mathbf{n}\Delta\varphi/\hbar) \rightarrow \exp(-i\boldsymbol{\sigma}\cdot\mathbf{n}\Delta\varphi/2) \quad (5.219)$$

in the Pauli scheme. The term on the right-hand side of the above expression is the exponential of a matrix. This can easily be evaluated using the Taylor series for an exponential, plus the rules

$$(\boldsymbol{\sigma}\cdot\mathbf{n})^n = 1 \quad \text{for } n \text{ even,} \quad (5.220)$$

$$(\boldsymbol{\sigma}\cdot\mathbf{n})^n = (\boldsymbol{\sigma}\cdot\mathbf{n}) \quad \text{for } n \text{ odd.} \quad (5.221)$$

These rules follow trivially from the identity (5.216). Thus, we can write

$$\begin{aligned} \exp(-i\boldsymbol{\sigma}\cdot\mathbf{n}\Delta\varphi/2) &= \left[ 1 - \frac{(\boldsymbol{\sigma}\cdot\mathbf{n})^2}{2!} \left(\frac{\Delta\varphi}{2}\right)^2 + \frac{(\boldsymbol{\sigma}\cdot\mathbf{n})^4}{4!} \left(\frac{\Delta\varphi}{2}\right)^4 + \dots \right] \\ &\quad - i \left[ (\boldsymbol{\sigma}\cdot\mathbf{n}) \left(\frac{\Delta\varphi}{2}\right) - \frac{(\boldsymbol{\sigma}\cdot\mathbf{n})^3}{3!} \left(\frac{\Delta\varphi}{2}\right)^3 + \dots \right] \\ &= \cos(\Delta\varphi/2) \mathbf{I} - i \sin(\Delta\varphi/2) \boldsymbol{\sigma}\cdot\mathbf{n}. \end{aligned} \quad (5.222)$$

The explicit  $2 \times 2$  form of this matrix is

$$\begin{pmatrix} \cos(\Delta\varphi/2) - i n_z \sin(\Delta\varphi/2) & (-i n_x - n_y) \sin(\Delta\varphi/2) \\ (-i n_x + n_y) \sin(\Delta\varphi/2) & \cos(\Delta\varphi/2) + i n_z \sin(\Delta\varphi/2) \end{pmatrix}. \quad (5.223)$$

Rotation matrices act on spinors in much the same manner as the corresponding rotation operators act on state kets. Thus,

$$\chi' = \exp(-i\boldsymbol{\sigma}\cdot\mathbf{n}\Delta\varphi/2) \chi, \quad (5.224)$$

where  $\chi'$  denotes the spinor obtained after rotating the spinor  $\chi$  an angle  $\Delta\varphi$  about the  $\mathbf{n}$ -axis. The Pauli matrices remain unchanged under rotations. However, the quantity  $\chi^\dagger \sigma_k \chi$  is proportional to the expectation value of  $S_k$  [see Eq. (5.198)], so we would expect it to transform like a vector under rotation (see Sect. 5.9). In fact, we require

$$(\chi^\dagger \sigma_k \chi)' \equiv (\chi')^\dagger \sigma_k \chi' = \sum_l R_{kl} (\chi^\dagger \sigma_l \chi), \quad (5.225)$$

where the  $R_{kl}$  are the elements of a conventional rotation matrix. This is easily demonstrated, since

$$\exp\left(\frac{i\sigma_3\Delta\varphi}{2}\right)\sigma_1\exp\left(\frac{-i\sigma_3\Delta\varphi}{2}\right) = \sigma_1\cos\Delta\varphi - \sigma_2\sin\Delta\varphi \quad (5.226)$$

plus all cyclic permutations. The above expression is the  $2 \times 2$  matrix analogue of (see Sect. 5.9)

$$\exp\left(\frac{iS_z\Delta\varphi}{\hbar}\right)S_x\exp\left(\frac{-iS_z\Delta\varphi}{\hbar}\right) = S_x\cos\Delta\varphi - S_y\sin\Delta\varphi. \quad (5.227)$$

The previous two formulae can both be validated using the Baker-Hausdorff lemma, (5.156), which holds for Hermitian matrices, in addition to Hermitian operators.

### 5.13 Spin greater than one-half systems

In the absence of spin, the Hamiltonian can be written as some function of the position and momentum operators. Using the Schrödinger representation, in which  $\mathbf{p} \rightarrow -i\hbar\nabla$ , the energy eigenvalue problem,

$$H|E\rangle = E|E\rangle, \quad (5.228)$$

can be transformed into a partial differential equation for the wave-function  $\psi(\mathbf{r}') \equiv \langle\mathbf{r}'|E\rangle$ . This function specifies the probability density for observing the particle at a given position,  $\mathbf{r}'$ . In general, we find

$$H\psi = E\psi, \quad (5.229)$$

where  $H$  is now a partial differential operator. The boundary conditions (for a bound state) are obtained from the normalization constraint

$$\int |\psi|^2 dV = 1. \quad (5.230)$$

This is all very familiar. However, we now know how to generalize this scheme to deal with a spin one-half particle. Instead of representing the state of the particle by a single wave-function, we use *two* wave-functions. The first,  $\psi_+(\mathbf{r}')$ ,

specifies the probability density of observing the particle at position  $\mathbf{r}'$  with spin angular momentum  $+\hbar/2$  in the  $z$ -direction. The second,  $\psi_-(\mathbf{r}')$ , specifies the probability density of observing the particle at position  $\mathbf{r}'$  with spin angular momentum  $-\hbar/2$  in the  $z$ -direction. In the Pauli scheme, these wave-functions are combined into a *spinor*,  $\chi$ , which is simply the row vector of  $\psi_+$  and  $\psi_-$ . In general, the Hamiltonian is a function of the position, momentum, and spin operators. Adopting the Schrödinger representation, and the Pauli scheme, the energy eigenvalue problem reduces to

$$H\chi = E\chi, \quad (5.231)$$

where  $\chi$  is a spinor (*i.e.*, a  $1 \times 2$  matrix of wave-functions) and  $H$  is a  $2 \times 2$  matrix partial differential operator [see Eq. (5.215)]. The above spinor equation can always be written out explicitly as *two coupled partial differential equations* for  $\psi_+$  and  $\psi_-$ .

Suppose that the Hamiltonian has no dependence on the spin operators. In this case, the Hamiltonian is represented as *diagonal*  $2 \times 2$  matrix partial differential operator in the Schrödinger/Pauli scheme [see Eq. (5.214)]. In other words, the partial differential equation for  $\psi_+$  decouples from that for  $\psi_-$ . In fact, both equations have the same form, so there is only really one differential equation. In this situation, the most general solution to Eq. (5.231) can be written

$$\chi = \psi(\mathbf{r}') \begin{pmatrix} c_+ \\ c_- \end{pmatrix}. \quad (5.232)$$

Here,  $\psi(\mathbf{r}')$  is determined by the solution of the differential equation, and the  $c_{\pm}$  are arbitrary complex numbers. The physical significance of the above expression is clear. The Hamiltonian determines the relative probabilities of finding the particle at various different positions, but the direction of its spin angular momentum remains undetermined.

Suppose that the Hamiltonian depends only on the spin operators. In this case, the Hamiltonian is represented as a  $2 \times 2$  matrix of complex numbers in the Schrödinger/Pauli scheme [see Eq. (5.197)], and the spinor eigenvalue equation (5.231) reduces to a straight-forward matrix eigenvalue problem. The most

general solution can again be written

$$\chi = \psi(\mathbf{r}') \begin{pmatrix} c_+ \\ c_- \end{pmatrix}. \quad (5.233)$$

Here, the ratio  $c_+/c_-$  is determined by the matrix eigenvalue problem, and the wave-function  $\psi(\mathbf{r}')$  is arbitrary. Clearly, the Hamiltonian determines the direction of the particle's spin angular momentum, but leaves its position undetermined.

In general, of course, the Hamiltonian is a function of both position and spin operators. In this case, it is not possible to decompose the spinor as in Eqs. (5.232) and (5.233). In other words, a general Hamiltonian causes the direction of the particle's spin angular momentum to vary with position in some specified manner. This can only be represented as a spinor involving different wave-functions,  $\psi_+$  and  $\psi_-$ .

But, what happens if we have a spin one or a spin three-halves particle? It turns out that we can generalize the Pauli two-component scheme in a fairly straight-forward manner. Consider a spin- $s$  particle: *i.e.*, a particle for which the eigenvalue of  $S^2$  is  $s(s+1)\hbar^2$ . Here,  $s$  is either an integer, or a half-integer. The eigenvalues of  $S_z$  are written  $s_z\hbar$ , where  $s_z$  is allowed to take the values  $s, s-1, \dots, -s+1, -s$ . In fact, there are  $2s+1$  distinct allowed values of  $s_z$ . Not surprisingly, we can represent the state of the particle by  $2s+1$  different wave-functions, denoted  $\psi_{s_z}(\mathbf{r}')$ . Here,  $\psi_{s_z}(\mathbf{r}')$  specifies the probability density for observing the particle at position  $\mathbf{r}'$  with spin angular momentum  $s_z\hbar$  in the  $z$ -direction. More exactly,

$$\psi_{s_z}(\mathbf{r}') = \langle \mathbf{r}' | \langle s, s_z | A \rangle \rangle, \quad (5.234)$$

where  $||A\rangle\rangle$  denotes a state ket in the product space of the position and spin operators. The state of the particle can be represented more succinctly by a spinor,  $\chi$ , which is simply the  $2s+1$  component row vector of the  $\psi_{s_z}(\mathbf{r}')$ . Thus, a spin one-half particle is represented by a two-component spinor, a spin one particle by a three-component spinor, a spin three-halves particle by a four-component spinor, and so on.

In this extended Schrödinger/Pauli scheme, position space operators take the form of *diagonal*  $(2s + 1) \times (2s + 1)$  matrix differential operators. Thus, we can represent the momentum operators as [see Eq. (5.214)]

$$p_k \rightarrow -i \hbar \frac{\partial}{\partial x'_k} \mathbf{I}, \quad (5.235)$$

where  $\mathbf{I}$  is the  $(2s + 1) \times (2s + 1)$  unit matrix. We represent the spin operators as

$$S_k \rightarrow s \hbar \sigma_k, \quad (5.236)$$

where the  $(2s + 1) \times (2s + 1)$  extended Pauli matrix  $\sigma_k$  has elements

$$(\sigma_k)_{jl} = \frac{\langle s, j | S_k | s, l \rangle}{s \hbar}. \quad (5.237)$$

Here,  $j, l$  are integers, or half-integers, lying in the range  $-s$  to  $+s$ . But, how can we evaluate the brackets  $\langle s, j | S_k | s, l \rangle$  and, thereby, construct the extended Pauli matrices? In fact, it is trivial to construct the  $\sigma_z$  matrix. By definition,

$$S_z |s, j\rangle = j \hbar |s, j\rangle. \quad (5.238)$$

Hence,

$$(\sigma_z)_{jl} = \frac{\langle s, j | S_z | s, l \rangle}{s \hbar} = \frac{j}{s} \delta_{jl}, \quad (5.239)$$

where use has been made of the orthonormality property of the  $|s, j\rangle$ . Thus,  $\sigma_z$  is the suitably normalized diagonal matrix of the eigenvalues of  $S_z$ . The matrix elements of  $\sigma_x$  and  $\sigma_y$  are most easily obtained by considering the shift operators,

$$S^\pm = S_x \pm i S_y. \quad (5.240)$$

We know, from Eqs. (5.55)–(5.56), that

$$S^+ |s, j\rangle = \sqrt{s(s+1) - j(j+1)} \hbar |s, j+1\rangle, \quad (5.241)$$

$$S^- |s, j\rangle = \sqrt{s(s+1) - j(j-1)} \hbar |s, j-1\rangle. \quad (5.242)$$

It follows from Eqs. (5.237), and (5.240)–(5.242), that

$$(\sigma_x)_{jl} = \frac{\sqrt{s(s+1) - j(j-1)} \delta_{j,l+1}}{2s}$$

$$+ \frac{\sqrt{s(s+1) - j(j+1)} \delta_{j,l-1}}{2s}, \quad (5.243)$$

$$(\sigma_y)_{jl} = \frac{\sqrt{s(s+1) - j(j-1)} \delta_{j,l+1}}{2is} - \frac{\sqrt{s(s+1) - j(j+1)} \delta_{j,l-1}}{2is}. \quad (5.244)$$

According to Eqs. (5.239) and (5.243)–(5.244), the Pauli matrices for a spin one-half ( $s = 1/2$ ) particle are

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (5.245)$$

$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad (5.246)$$

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (5.247)$$

as we have seen previously. For a spin one ( $s = 1$ ) particle, we find that

$$\sigma_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad (5.248)$$

$$\sigma_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad (5.249)$$

$$\sigma_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (5.250)$$

In fact, we can now construct the Pauli matrices for a spin anything particle. This means that we can convert the general energy eigenvalue problem for a spin- $s$  particle, where the Hamiltonian is some function of position and spin operators, into  $2s + 1$  coupled partial differential equations involving the  $2s + 1$  wave-functions  $\psi_{s_z}(\mathbf{r}')$ . Unfortunately, such a system of equations is generally too complicated to solve exactly.

## 5.14 Addition of angular momentum

Consider a hydrogen atom in an  $l = 1$  state. The electron possesses orbital angular momentum of magnitude  $\hbar$ , and spin angular momentum of magnitude  $\hbar/2$ . So, what is the total angular momentum of the system?

In order to answer this question, we are going to have to learn how to add angular momentum operators. Let us consider the most general case. Suppose that we have two sets of angular momentum operators,  $\mathbf{J}_1$  and  $\mathbf{J}_2$ . By definition, these operators are Hermitian, and obey the fundamental commutation relations

$$\mathbf{J}_1 \times \mathbf{J}_1 = i \hbar \mathbf{J}_1, \quad (5.251)$$

$$\mathbf{J}_2 \times \mathbf{J}_2 = i \hbar \mathbf{J}_2. \quad (5.252)$$

We assume that the two groups of operators correspond to different degrees of freedom of the system, so that

$$[J_{1i}, J_{2j}] = 0, \quad (5.253)$$

where  $i, j$  stand for either  $x, y$ , or  $z$ . For instance,  $\mathbf{J}_1$  could be an orbital angular momentum operator, and  $\mathbf{J}_2$  a spin angular momentum operator. Alternatively,  $\mathbf{J}_1$  and  $\mathbf{J}_2$  could be the orbital angular momentum operators of two different particles in a multi-particle system. We know, from the general properties of angular momentum, that the eigenvalues of  $J_1^2$  and  $J_2^2$  can be written  $j_1(j_1 + 1)\hbar^2$  and  $j_2(j_2 + 1)\hbar^2$ , respectively, where  $j_1$  and  $j_2$  are either integers, or half-integers. We also know that the eigenvalues of  $J_{1z}$  and  $J_{2z}$  take the form  $m_1\hbar$  and  $m_2\hbar$ , respectively, where  $m_1$  and  $m_2$  are numbers lying in the ranges  $j_1, j_1 - 1, \dots, -j_1 + 1, -j_1$  and  $j_2, j_2 - 1, \dots, -j_2 + 1, -j_2$ , respectively.

Let us define the total angular momentum operator

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2. \quad (5.254)$$

Now  $\mathbf{J}$  is an Hermitian operator, since it is the sum of Hermitian operators. According to Eqs. (5.11) and (5.14),  $\mathbf{J}$  satisfies the fundamental commutation relation

$$\mathbf{J} \times \mathbf{J} = i \hbar \mathbf{J}. \quad (5.255)$$

Thus,  $\mathbf{J}$  possesses all of the expected properties of an angular momentum operator. It follows that the eigenvalue of  $J^2$  can be written  $j(j+1)\hbar^2$ , where  $j$  is an integer, or a half-integer. The eigenvalue of  $J_z$  takes the form  $m\hbar$ , where  $m$  lies in the range  $j, j-1, \dots, -j+1, -j$ . At this stage, we do not know the relationship between the quantum numbers of the total angular momentum,  $j$  and  $m$ , and those of the individual angular momenta,  $j_1, j_2, m_1$ , and  $m_2$ .

Now

$$J^2 = J_1^2 + J_2^2 + 2\mathbf{J}_1 \cdot \mathbf{J}_2. \quad (5.256)$$

We know that

$$[J_1^2, J_{1i}] = 0, \quad (5.257)$$

$$[J_2^2, J_{2i}] = 0, \quad (5.258)$$

and also that all of the  $J_{1i}$  operators commute with the  $J_{2i}$  operators. It follows from Eq. (5.256) that

$$[J^2, J_1^2] = [J^2, J_2^2] = 0. \quad (5.259)$$

This implies that the quantum numbers  $j_1, j_2$ , and  $j$  can all be measured simultaneously. In other words, we can know the magnitude of the total angular momentum together with the magnitudes of the component angular momenta. However, it is clear from Eq. (5.256) that

$$[J^2, J_{1z}] \neq 0, \quad (5.260)$$

$$[J^2, J_{2z}] \neq 0. \quad (5.261)$$

This suggests that it is not possible to measure the quantum numbers  $m_1$  and  $m_2$  simultaneously with the quantum number  $j$ . Thus, we cannot determine the projections of the individual angular momenta along the  $z$ -axis at the same time as the magnitude of the total angular momentum.

It is clear, from the preceding discussion, that we can form two alternate groups of mutually commuting operators. The first group is  $J_1^2, J_2^2, J_{1z}$ , and  $J_{2z}$ . The second group is  $J_1^2, J_2^2, J^2$ , and  $J_z$ . These two groups of operators are incompatible with one another. We can define simultaneous eigenkets of each operator group. The simultaneous eigenkets of  $J_1^2, J_2^2, J_{1z}$ , and  $J_{2z}$  are denoted

$|j_1, j_2; m_1, m_2\rangle$ , where

$$J_1^2 |j_1, j_2; m_1, m_2\rangle = j_1(j_1 + 1) \hbar^2 |j_1, j_2; m_1, m_2\rangle, \quad (5.262)$$

$$J_2^2 |j_1, j_2; m_1, m_2\rangle = j_2(j_2 + 1) \hbar^2 |j_1, j_2; m_1, m_2\rangle, \quad (5.263)$$

$$J_{1z} |j_1, j_2; m_1, m_2\rangle = m_1 \hbar |j_1, j_2; m_1, m_2\rangle, \quad (5.264)$$

$$J_{2z} |j_1, j_2; m_1, m_2\rangle = m_2 \hbar |j_1, j_2; m_1, m_2\rangle. \quad (5.265)$$

The simultaneous eigenkets of  $J_1^2, J_2^2, J^2$  and  $J_z$  are denoted  $|j_1, j_2; j, m\rangle$ , where

$$J_1^2 |j_1, j_2; j, m\rangle = j_1(j_1 + 1) \hbar^2 |j_1, j_2; j, m\rangle, \quad (5.266)$$

$$J_2^2 |j_1, j_2; j, m\rangle = j_2(j_2 + 1) \hbar^2 |j_1, j_2; j, m\rangle, \quad (5.267)$$

$$J^2 |j_1, j_2; j, m\rangle = j(j + 1) \hbar^2 |j_1, j_2; j, m\rangle, \quad (5.268)$$

$$J_z |j_1, j_2; j, m\rangle = m \hbar |j_1, j_2; j, m\rangle. \quad (5.269)$$

Each set of eigenkets are complete, mutually orthogonal (for eigenkets corresponding to different sets of eigenvalues), and have unit norms. Since the operators  $J_1^2$  and  $J_2^2$  are common to both operator groups, we can assume that the quantum numbers  $j_1$  and  $j_2$  are known. In other words, we can always determine the magnitudes of the individual angular momenta. In addition, we can either know the quantum numbers  $m_1$  and  $m_2$ , or the quantum numbers  $j$  and  $m$ , but we cannot know both pairs of quantum numbers at the same time. We can write a conventional completeness relation for both sets of eigenkets:

$$\sum_{m_1} \sum_{m_2} |j_1, j_2; m_1, m_2\rangle \langle j_1, j_2; m_1, m_2| = 1, \quad (5.270)$$

$$\sum_j \sum_m |j_1, j_2; j, m\rangle \langle j_1, j_2; j, m| = 1, \quad (5.271)$$

where the right-hand sides denote the identity operator in the ket space corresponding to states of given  $j_1$  and  $j_2$ . The summation is over all allowed values of  $m_1, m_2, j$ , and  $m$ .

The operator group  $J_1^2, J_2^2, J^2$ , and  $J_z$  is incompatible with the group  $J_1^2, J_2^2, J_{1z}$ , and  $J_{2z}$ . This means that if the system is in a simultaneous eigenstate of the former group then, in general, it is not in an eigenstate of the latter. In other

words, if the quantum numbers  $j_1$ ,  $j_2$ ,  $j$ , and  $m$  are known with certainty, then a measurement of the quantum numbers  $m_1$  and  $m_2$  will give a range of possible values. We can use the completeness relation (5.270) to write

$$|j_1, j_2; j, m\rangle = \sum_{m_1} \sum_{m_2} \langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m\rangle |j_1, j_2; m_1, m_2\rangle. \quad (5.272)$$

Thus, we can write the eigenkets of the first group of operators as a weighted sum of the eigenkets of the second set. The weights,  $\langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m\rangle$ , are called the *Clebsch-Gordon coefficients*. If the system is in a state where a measurement of  $J_1^2$ ,  $J_2^2$ ,  $J^2$ , and  $J_z$  is bound to give the results  $j_1(j_1 + 1)\hbar^2$ ,  $j_2(j_2 + 1)\hbar^2$ ,  $j(j + 1)\hbar^2$ , and  $j_z\hbar$ , respectively, then a measurement of  $J_{1z}$  and  $J_{2z}$  will give the results  $m_1\hbar$  and  $m_2\hbar$  with probability  $|\langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m\rangle|^2$ .

The Clebsch-Gordon coefficients possess a number of very important properties. First, the coefficients are zero unless

$$m = m_1 + m_2. \quad (5.273)$$

To prove this, we note that

$$(J_z - J_{1z} - J_{2z})|j_1, j_2; j, m\rangle = 0. \quad (5.274)$$

Forming the inner product with  $\langle j_1, j_2; m_1, m_2 |$ , we obtain

$$(m - m_1 - m_2)\langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m\rangle = 0, \quad (5.275)$$

which proves the assertion. Thus, the  $z$ -components of different angular momenta add algebraically. So, an electron in an  $l = 1$  state, with orbital angular momentum  $\hbar$ , and spin angular momentum  $\hbar/2$ , projected along the  $z$ -axis, constitutes a state whose total angular momentum projected along the  $z$ -axis is  $3\hbar/2$ . What is uncertain is the magnitude of the total angular momentum.

Second, the coefficients vanish unless

$$|j_1 - j_2| \leq j \leq j_1 + j_2. \quad (5.276)$$

We can assume, without loss of generality, that  $j_1 \geq j_2$ . We know, from Eq. (5.273), that for given  $j_1$  and  $j_2$  the largest possible value of  $m$  is  $j_1 + j_2$  (since  $j_1$  is the

largest possible value of  $m_1$ , etc.). This implies that the largest possible value of  $j$  is  $j_1 + j_2$  (since, by definition, the largest value of  $m$  is equal to  $j$ ). Now, there are  $(2j_1 + 1)$  allowable values of  $m_1$  and  $(2j_2 + 1)$  allowable values of  $m_2$ . Thus, there are  $(2j_1 + 1)(2j_2 + 1)$  independent eigenkets,  $|j_1, j_2; m_1, m_2\rangle$ , needed to span the ket space corresponding to fixed  $j_1$  and  $j_2$ . Since the eigenkets  $|j_1, j_2; j, m\rangle$  span the same space, they must also form a set of  $(2j_1 + 1)(2j_2 + 1)$  independent kets. In other words, there can only be  $(2j_1 + 1)(2j_2 + 1)$  distinct allowable values of the quantum numbers  $j$  and  $m$ . For each allowed value of  $j$ , there are  $2j + 1$  allowed values of  $m$ . We have already seen that the maximum allowed value of  $j$  is  $j_1 + j_2$ . It is easily seen that if the minimum allowed value of  $j$  is  $j_1 - j_2$  then the total number of allowed values of  $j$  and  $m$  is  $(2j_1 + 1)(2j_2 + 1)$ : *i.e.*,

$$\sum_{j=j_1-j_2}^{j_1+j_2} (2j + 1) \equiv (2j_1 + 1)(2j_2 + 1). \quad (5.277)$$

This proves our assertion.

Third, the sum of the modulus squared of all of the Clebsch-Gordon coefficients is unity: *i.e.*,

$$\sum_{m_1} \sum_{m_2} |\langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m \rangle|^2 = 1. \quad (5.278)$$

This assertion is proved as follows:

$$\begin{aligned} \langle j_1, j_2; j, m | j_1, j_2; j, m \rangle &= \\ \sum_{m_1} \sum_{m_2} \langle j_1, j_2; j, m | j_1, j_2; m_1, m_2 \rangle \langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m \rangle & \\ = \sum_{m_1} \sum_{m_2} |\langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m \rangle|^2 &= 1, \end{aligned} \quad (5.279)$$

where use has been made of the completeness relation (5.270).

Finally, the Clebsch-Gordon coefficients obey two recursion relations. To obtain these relations we start from

$$\begin{aligned} J^\pm |j_1, j_2; j, m\rangle &= (J_1^\pm + J_2^\pm) \\ &\times \sum_{m'_1} \sum_{m'_2} \langle j_1, j_2; m'_1, m'_2 | j_1, j_2; j, m \rangle |j_1, j_2; m'_1, m'_2\rangle. \end{aligned} \quad (5.280)$$

Making use of the well-known properties of the shift operators, which are specified by Eqs. (5.55)–(5.56), we obtain

$$\begin{aligned} & \sqrt{j(j+1) - m(m \pm 1)} |j_1, j_2; j, m \pm 1\rangle = \\ & \sum_{m'_1} \sum_{m'_2} \left( \sqrt{j_1(j_1+1) - m'_1(m'_1 \pm 1)} |j_1, j_2; m'_1 \pm 1, m'_2\rangle \right. \\ & \quad \left. + \sqrt{j_2(j_2+1) - m'_2(m'_2 \pm 1)} |j_1, j_2; m'_1, m'_2 \pm 1\rangle \right) \\ & \quad \times \langle j_1, j_2; m'_1, m'_2 | j_1, j_2; j, m \rangle. \end{aligned} \quad (5.281)$$

Taking the inner product with  $\langle j_1, j_2; m_1, m_2 |$ , and making use of the orthonormality property of the basis eigenkets, we obtain the desired recursion relations:

$$\begin{aligned} & \sqrt{j(j+1) - m(m \pm 1)} \langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m \pm 1 \rangle = \\ & \sqrt{j_1(j_1+1) - m_1(m_1 \mp 1)} \langle j_1, j_2; m_1 \mp 1, m_2 | j_1, j_2; j, m \rangle \\ & + \sqrt{j_2(j_2+1) - m_2(m_2 \mp 1)} \langle j_1, j_2; m_1, m_2 \mp 1 | j_1, j_2; j, m \rangle. \end{aligned} \quad (5.282)$$

It is clear, from the absence of complex coupling coefficients in the above relations, that we can always choose the Clebsch-Gordon coefficients to be real numbers. This is a convenient choice, since it ensures that the inverse Clebsch-Gordon coefficients,  $\langle j_1, j_2; j, m | j_1, j_2; m_1, m_2 \rangle$ , are identical to the Clebsch-Gordon coefficients. In other words,

$$\langle j_1, j_2; j, m | j_1, j_2; m_1, m_2 \rangle = \langle j_1, j_2; m_1, m_2 | j_1, j_2; j, m \rangle. \quad (5.283)$$

The inverse Clebsch-Gordon coefficients are the weights in the expansion of the  $|j_1, j_2; m_1, m_2\rangle$  in terms of the  $|j_1, j_2; j, m\rangle$ :

$$|j_1, j_2; m_1, m_2\rangle = \sum_j \sum_m \langle j_1, j_2; j, m | j_1, j_2; m_1, m_2 \rangle |j_1, j_2; j, m\rangle. \quad (5.284)$$

It turns out that the recursion relations (5.282), together with the normalization condition (5.278), are sufficient to completely determine the Clebsch-Gordon coefficients to within an arbitrary sign (multiplied into all of the coefficients). This sign is fixed by convention. The easiest way of demonstrating this assertion is by considering some specific examples.

Let us add the angular momentum of two spin one-half systems: *e.g.*, two electrons at rest. So,  $j_1 = j_2 = 1/2$ . We know, from general principles, that  $|m_1| \leq 1/2$  and  $|m_2| \leq 1/2$ . We also know, from Eq. (5.276), that  $0 \leq j \leq 1$ , where the allowed values of  $j$  differ by integer amounts. It follows that either  $j = 0$  or  $j = 1$ . Thus, two spin one-half systems can be combined to form either a spin zero system or a spin one system. It is helpful to arrange all of the possibly non-zero Clebsch-Gordon coefficients in a table:

| $m_1$     | $m_2$ |   |   |    |   |
|-----------|-------|---|---|----|---|
| 1/2       | 1/2   | ? | ? | ?  | ? |
| 1/2       | -1/2  | ? | ? | ?  | ? |
| -1/2      | 1/2   | ? | ? | ?  | ? |
| -1/2      | -1/2  | ? | ? | ?  | ? |
| $j_1=1/2$ | $j$   | 1 | 1 | 1  | 0 |
| $j_2=1/2$ | $m$   | 1 | 0 | -1 | 0 |

The box in this table corresponding to  $m_1 = 1/2, m_2 = 1/2, j = 1, m = 1$  gives the Clebsch-Gordon coefficient  $\langle 1/2, 1/2; 1/2, 1/2 | 1/2, 1/2; 1, 1 \rangle$ , or the inverse Clebsch-Gordon coefficient  $\langle 1/2, 1/2; 1, 1 | 1/2, 1/2; 1/2, 1/2 \rangle$ . All the boxes contain question marks because we do not know any Clebsch-Gordon coefficients at the moment.

A Clebsch-Gordon coefficient is automatically zero unless  $m_1 + m_2 = m$ . In other words, the  $z$ -components of angular momentum have to add algebraically. Many of the boxes in the above table correspond to  $m_1 + m_2 \neq m$ . We immediately conclude that these boxes must contain zeroes: *i.e.*,

| $m_1$     | $m_2$ |   |   |    |   |
|-----------|-------|---|---|----|---|
| 1/2       | 1/2   | ? | 0 | 0  | 0 |
| 1/2       | -1/2  | 0 | ? | 0  | ? |
| -1/2      | 1/2   | 0 | ? | 0  | ? |
| -1/2      | -1/2  | 0 | 0 | ?  | 0 |
| $j_1=1/2$ | $j$   | 1 | 1 | 1  | 0 |
| $j_2=1/2$ | $m$   | 1 | 0 | -1 | 0 |

The normalization condition (5.278) implies that the sum of the squares of all

the rows and columns of the above table must be unity. There are two rows and two columns which only contain a single non-zero entry. We conclude that these entries must be  $\pm 1$ , but we have no way of determining the signs at present. Thus,

|           |       |         |   |         |   |
|-----------|-------|---------|---|---------|---|
| $m_1$     | $m_2$ |         |   |         |   |
| 1/2       | 1/2   | $\pm 1$ | 0 | 0       | 0 |
| 1/2       | -1/2  | 0       | ? | 0       | ? |
| -1/2      | 1/2   | 0       | ? | 0       | ? |
| -1/2      | -1/2  | 0       | 0 | $\pm 1$ | 0 |
| $j_1=1/2$ | $j$   | 1       | 1 | 1       | 0 |
| $j_2=1/2$ | $m$   | 1       | 0 | -1      | 0 |

Let us evaluate the recursion relation (5.282) for  $j_1 = j_2 = 1/2$ , with  $j = 1$ ,  $m = 0$ ,  $m_1 = m_2 = \pm 1/2$ , taking the upper/lower sign. We find that

$$\langle 1/2, -1/2 | 1, 0 \rangle + \langle -1/2, 1/2 | 1, 0 \rangle = \sqrt{2} \langle 1/2, 1/2 | 1, 1 \rangle = \pm \sqrt{2}, \quad (5.285)$$

and

$$\langle 1/2, -1/2 | 1, 0 \rangle + \langle -1/2, 1/2 | 1, 0 \rangle = \sqrt{2} \langle -1/2, -1/2 | 1, -1 \rangle = \pm \sqrt{2}. \quad (5.286)$$

Here, the  $j_1$  and  $j_2$  labels have been suppressed for ease of notation. We also know that

$$\langle 1/2, -1/2 | 1, 0 \rangle^2 + \langle -1/2, 1/2 | 1, 0 \rangle^2 = 1, \quad (5.287)$$

from the normalization condition. The only real solutions to the above set of equations are

$$\begin{aligned} \sqrt{2} \langle 1/2, -1/2 | 1, 0 \rangle &= \sqrt{2} \langle -1/2, 1/2 | 1, 0 \rangle \\ &= \langle 1/2, 1/2 | 1, 1 \rangle = \langle 1/2, 1/2 | 1, -1 \rangle = \pm 1. \end{aligned} \quad (5.288)$$

The choice of sign is arbitrary—the conventional choice is a positive sign. Thus,

our table now reads

|           |       |   |              |    |   |
|-----------|-------|---|--------------|----|---|
| $m_1$     | $m_2$ |   |              |    |   |
| 1/2       | 1/2   | 1 | 0            | 0  | 0 |
| 1/2       | -1/2  | 0 | $1/\sqrt{2}$ | 0  | ? |
| -1/2      | 1/2   | 0 | $1/\sqrt{2}$ | 0  | ? |
| -1/2      | -1/2  | 0 | 0            | 1  | 0 |
| $j_1=1/2$ | $j$   | 1 | 1            | 1  | 0 |
| $j_2=1/2$ | $m$   | 1 | 0            | -1 | 0 |

We could fill in the remaining unknown entries of our table by using the recursion relation again. However, an easier method is to observe that the rows and columns of the table must all be mutually orthogonal. That is, the dot product of a row with any other row must be zero. Likewise, for the dot product of a column with any other column. This follows because the entries in the table give the expansion coefficients of one of our alternative sets of eigenkets in terms of the other set, and each set of eigenkets contains mutually orthogonal vectors with unit norms. The normalization condition tells us that the dot product of a row or column with itself must be unity. The only way that the dot product of the fourth column with the second column can be zero is if the unknown entries are equal and opposite. The requirement that the dot product of the fourth column with itself is unity tells us that the magnitudes of the unknown entries have to be  $1/\sqrt{2}$ . The unknown entries are undetermined to an arbitrary sign multiplied into them both. Thus, the final form of our table (with the conventional choice of arbitrary signs) is

|           |       |   |              |    |               |
|-----------|-------|---|--------------|----|---------------|
| $m_1$     | $m_2$ |   |              |    |               |
| 1/2       | 1/2   | 1 | 0            | 0  | 0             |
| 1/2       | -1/2  | 0 | $1/\sqrt{2}$ | 0  | $1/\sqrt{2}$  |
| -1/2      | 1/2   | 0 | $1/\sqrt{2}$ | 0  | $-1/\sqrt{2}$ |
| -1/2      | -1/2  | 0 | 0            | 1  | 0             |
| $j_1=1/2$ | $j$   | 1 | 1            | 1  | 0             |
| $j_2=1/2$ | $m$   | 1 | 0            | -1 | 0             |

The table can be read in one of two ways. The columns give the expansions of

the eigenstates of overall angular momentum in terms of the eigenstates of the individual angular momenta of the two component systems. Thus, the second column tells us that

$$|1, 0\rangle = \frac{1}{\sqrt{2}} (|1/2, -1/2\rangle + |-1/2, 1/2\rangle). \quad (5.289)$$

The ket on the left-hand side is a  $|j, m\rangle$  ket, whereas those on the right-hand side are  $|m_1, m_2\rangle$  kets. The rows give the expansions of the eigenstates of individual angular momentum in terms of those of overall angular momentum. Thus, the second row tells us that

$$|1/2, -1/2\rangle = \frac{1}{\sqrt{2}} (|1, 0\rangle + |0, 0\rangle). \quad (5.290)$$

Here, the ket on the left-hand side is a  $|m_1, m_2\rangle$  ket, whereas those on the right-hand side are  $|j, m\rangle$  kets.

Note that our table is really a combination of two sub-tables, one involving  $j = 0$  states, and one involving  $j = 1$  states. The Clebsch-Gordon coefficients corresponding to two different choices of  $j$  are completely independent: *i.e.*, there is no recursion relation linking Clebsch-Gordon coefficients corresponding to different values of  $j$ . Thus, for every choice of  $j_1$ ,  $j_2$ , and  $j$  we can construct a table of Clebsch-Gordon coefficients corresponding to the different allowed values of  $m_1$ ,  $m_2$ , and  $m$  (subject to the constraint that  $m_1 + m_2 = m$ ). A complete knowledge of angular momentum addition is equivalent to a knowing all possible tables of Clebsch-Gordon coefficients. These tables are listed (for moderate values of  $j_1$ ,  $j_2$  and  $j$ ) in many standard reference books.

## 6 Approximation methods

### 6.1 Introduction

We have developed techniques by which the general energy eigenvalue problem can be reduced to a set of coupled partial differential equations involving various wave-functions. Unfortunately, the number of such problems which yield exactly soluble equations is comparatively small. Clearly, we need to develop some techniques for finding approximate solutions to otherwise intractable problems.

Consider the following problem, which is very common. The Hamiltonian of a system is written

$$H = H_0 + H_1. \quad (6.1)$$

Here,  $H_0$  is a simple Hamiltonian for which we know the *exact* eigenvalues and eigenstates.  $H_1$  introduces some interesting additional physics into the problem, but it is sufficiently complicated that when we add it to  $H_0$  we can no longer find the exact energy eigenvalues and eigenstates. However,  $H_1$  can, in some sense (which we shall specify more exactly later on), be regarded as being *small* compared to  $H_0$ . Can we find the approximate eigenvalues and eigenstates of the modified Hamiltonian,  $H_0 + H_1$ , by performing some sort of perturbation analysis about the eigenvalues and eigenstates of the original Hamiltonian,  $H_0$ ? Let us investigate.

### 6.2 The two-state system

Let us begin by considering *time-independent perturbation theory*, in which the modification to the Hamiltonian,  $H_1$ , has no explicit dependence on time. It is usually assumed that the unperturbed Hamiltonian,  $H_0$ , is also time-independent.

Consider the simplest non-trivial system, in which there are only *two* independent eigenkets of the unperturbed Hamiltonian. These are denoted

$$H_0 |1\rangle = E_1 |1\rangle, \quad (6.2)$$

$$H_0 |2\rangle = E_2 |2\rangle. \quad (6.3)$$

It is assumed that these states, and their associated eigenvalues, are known. Since  $H_0$  is, by definition, an Hermitian operator, its two eigenkets are orthonormal and form a complete set. The lengths of these eigenkets are both normalized to unity. Let us now try to solve the modified energy eigenvalue problem

$$(H_0 + H_1)|E\rangle = E |E\rangle. \quad (6.4)$$

In fact, we can solve this problem exactly. Since, the eigenkets of  $H_0$  form a complete set, we can write

$$|E\rangle = \langle 1|E\rangle |1\rangle + \langle 2|E\rangle |2\rangle. \quad (6.5)$$

Right-multiplication of Eq. (6.4) by  $\langle 1|$  and  $\langle 2|$  yields two coupled equations, which can be written in matrix form:

$$\begin{pmatrix} E_1 - E + e_{11} & e_{12} \\ e_{12}^* & E_2 - E + e_{22} \end{pmatrix} \begin{pmatrix} \langle 1|E\rangle \\ \langle 2|E\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (6.6)$$

Here,

$$e_{11} = \langle 1|H_1|1\rangle, \quad (6.7)$$

$$e_{22} = \langle 2|H_1|2\rangle, \quad (6.8)$$

$$e_{12} = \langle 1|H_1|2\rangle. \quad (6.9)$$

In the special (but common) case of a perturbing Hamiltonian whose diagonal matrix elements (in the unperturbed eigenstates) are zero, so that

$$e_{11} = e_{22} = 0, \quad (6.10)$$

the solution of Eq. (6.6) (obtained by setting the determinant of the matrix equal to zero) is

$$E = \frac{(E_1 + E_2) \pm \sqrt{(E_1 - E_2)^2 + 4|e_{12}|^2}}{2}. \quad (6.11)$$

Let us expand in the supposedly small parameter

$$\epsilon = \frac{|e_{12}|}{|E_1 - E_2|}. \quad (6.12)$$

We obtain

$$E \simeq \frac{1}{2}(E_1 + E_2) \pm \frac{1}{2}(E_1 - E_2)(1 + 2\epsilon^2 + \dots). \quad (6.13)$$

The above expression yields the modifications to the energy eigenvalues due to the perturbing Hamiltonian:

$$E'_1 = E_1 + \frac{|e_{12}|^2}{E_1 - E_2} + \dots, \quad (6.14)$$

$$E'_2 = E_2 - \frac{|e_{12}|^2}{E_1 - E_2} + \dots. \quad (6.15)$$

Note that  $H_1$  causes the upper eigenvalue to rise, and the lower eigenvalue to fall. It is easily demonstrated that the modified eigenkets take the form

$$|1\rangle' = |1\rangle + \frac{e_{12}^*}{E_1 - E_2} |2\rangle + \dots, \quad (6.16)$$

$$|2\rangle' = |2\rangle - \frac{e_{12}}{E_1 - E_2} |1\rangle + \dots. \quad (6.17)$$

Thus, the modified energy eigenstates consist of one of the unperturbed eigenstates with a slight admixture of the other. Note that the series expansion in Eq. (6.13) only converges if  $2|\epsilon| < 1$ . This suggests that the condition for the validity of the perturbation expansion is

$$|e_{12}| < \frac{|E_1 - E_2|}{2}. \quad (6.18)$$

In other words, when we say that  $H_1$  needs to be small compared to  $H_0$ , what we really mean is that the above inequality needs to be satisfied.

### 6.3 Non-degenerate perturbation theory

Let us now generalize our perturbation analysis to deal with systems possessing more than two energy eigenstates. The energy eigenstates of the unperturbed Hamiltonian,  $H_0$ , are denoted

$$H_0 |n\rangle = E_n |n\rangle, \quad (6.19)$$

where  $n$  runs from 1 to  $N$ . The eigenkets  $|n\rangle$  are orthonormal, form a complete set, and have their lengths normalized to unity. Let us now try to solve the energy eigenvalue problem for the perturbed Hamiltonian:

$$(H_0 + H_1)|E\rangle = E|E\rangle. \quad (6.20)$$

We can express  $|E\rangle$  as a linear superposition of the unperturbed energy eigenkets,

$$|E\rangle = \sum_k \langle k|E\rangle |k\rangle, \quad (6.21)$$

where the summation is from  $k = 1$  to  $N$ . Substituting the above equation into Eq. (6.20), and right-multiplying by  $\langle m|$ , we obtain

$$(E_m + e_{mm} - E)\langle m|E\rangle + \sum_{k \neq m} e_{mk} \langle k|E\rangle = 0, \quad (6.22)$$

where

$$e_{mk} = \langle m|H_1|k\rangle. \quad (6.23)$$

Let us now develop our perturbation expansion. We assume that

$$\frac{|e_{mk}|}{E_m - E_k} \sim O(\epsilon), \quad (6.24)$$

for all  $m \neq k$ , where  $\epsilon \ll 1$  is our expansion parameter. We also assume that

$$\frac{|e_{mm}|}{E_m} \sim O(\epsilon), \quad (6.25)$$

for all  $m$ . Let us search for a modified version of the  $n$ th unperturbed energy eigenstate, for which

$$E = E_n + O(\epsilon), \quad (6.26)$$

and

$$\langle n|E\rangle = 1, \quad (6.27)$$

$$\langle m|E\rangle \sim O(\epsilon), \quad (6.28)$$

for  $m \neq n$ . Suppose that we write out Eq. (6.22) for  $m \neq n$ , neglecting terms which are  $O(\epsilon^2)$  according to our expansion scheme. We find that

$$(E_m - E_n)\langle m|E\rangle + e_{mn} \simeq 0, \quad (6.29)$$

giving

$$\langle m|E\rangle \simeq -\frac{e_{mn}}{E_m - E_n}. \quad (6.30)$$

Substituting the above expression into Eq. (6.22), evaluated for  $m = n$ , and neglecting  $O(\epsilon^3)$  terms, we obtain

$$(E_n + e_{nn} - E) - \sum_{k \neq n} \frac{|e_{nk}|^2}{E_k - E_n} = 0. \quad (6.31)$$

Thus, the modified  $n$ th energy eigenstate possesses an eigenvalue

$$E'_n = E_n + e_{nn} + \sum_{k \neq n} \frac{|e_{nk}|^2}{E_n - E_k} + O(\epsilon^3), \quad (6.32)$$

and a eigenket

$$|n\rangle' = |n\rangle + \sum_{k \neq n} \frac{e_{kn}}{E_n - E_k} |k\rangle + O(\epsilon^2). \quad (6.33)$$

Note that

$$\langle m|n\rangle' = \delta_{mn} + \frac{e_{nm}^*}{E_m - E_n} + \frac{e_{mn}}{E_n - E_m} + O(\epsilon^2) = \delta_{mn} + O(\epsilon^2). \quad (6.34)$$

Thus, the modified eigenkets remain orthonormal and properly normalized to  $O(\epsilon^2)$ .

## 6.4 The quadratic Stark effect

Suppose that a one-electron atom [*i.e.*, either a hydrogen atom, or an alkali metal atom (which possesses one valance electron orbiting outside a closed, spherically symmetric shell)] is subjected to a uniform electric field in the positive  $z$ -direction. The Hamiltonian of the system can be split into two parts. The unperturbed Hamiltonian,

$$H_0 = \frac{\mathbf{p}^2}{2m_e} + V(r), \quad (6.35)$$

and the perturbing Hamiltonian,

$$H_1 = e|\mathbf{E}|z. \quad (6.36)$$

It is assumed that the unperturbed energy eigenvalues and eigenstates are completely known. The electron spin is irrelevant in this problem (since the spin operators all commute with  $H_1$ ), so we can ignore the spin degrees of freedom of the system. This implies that the system possesses no degenerate energy eigenvalues. This is not true for the  $n \neq 1$  energy levels of the hydrogen atom, due to the special properties of a pure Coulomb potential. It is necessary to deal with this case separately, because the perturbation theory presented in Sect. 6.3 breaks down for degenerate unperturbed energy levels.

An energy eigenket of the unperturbed Hamiltonian is characterized by three quantum numbers—the radial quantum number  $n$ , and the two angular quantum numbers  $l$  and  $m$  (see Sect. 5.6). Let us denote such a ket  $|n, l, m\rangle$ , and let its energy level be  $E_{nlm}$ . According to Eq. (6.32), the change in this energy level induced by a *small* electric field is given by

$$\begin{aligned} \Delta E_{nlm} = & e|\mathbf{E}| \langle n, l, m | z | n, l, m \rangle \\ & + e^2 |\mathbf{E}|^2 \sum_{n', l', m' \neq n, l, m} \frac{|\langle n, l, m | z | n', l', m' \rangle|^2}{E_{nlm} - E_{n'l'm'}}. \end{aligned} \quad (6.37)$$

Now, since

$$L_z = x p_y - y p_x, \quad (6.38)$$

it follows that

$$[L_z, z] = 0. \quad (6.39)$$

Thus,

$$\langle n, l, m | [L_z, z] | n', l', m' \rangle = 0, \quad (6.40)$$

giving

$$(m - m') \langle n, l, m | z | n', l', m' \rangle = 0, \quad (6.41)$$

since  $|n, l, m\rangle$  is, by definition, an eigenstate of  $L_z$  with eigenvalue  $m\hbar$ . It is clear, from the above relation, that the matrix element  $\langle n, l, m | z | n', l', m' \rangle$  is zero unless  $m' = m$ . This is termed the *selection rule* for the quantum number  $m$ .

Let us now determine the selection rule for  $l$ . We have

$$[L^2, z] = [L_x^2, z] + [L_y^2, z]$$

$$\begin{aligned}
&= L_x[L_x, z] + [L_x, z]L_x + L_y[L_y, z] + [L_y, z]L_y \\
&= i\hbar(-L_x y - y L_x + L_y x + x L_y) \\
&= 2i\hbar(L_y x - L_x y + i\hbar z) \\
&= 2i\hbar(L_y x - y L_x) = 2i\hbar(x L_y - L_x y), \tag{6.42}
\end{aligned}$$

where use has been made of Eqs. (5.1)–(5.6). Similarly,

$$[L^2, y] = 2i\hbar(L_x z - x L_z), \tag{6.43}$$

$$[L^2, x] = 2i\hbar(y L_z - L_y z). \tag{6.44}$$

Thus,

$$\begin{aligned}
[L^2, [L^2, z]] &= 2i\hbar(L^2, L_y x - L_x y + i\hbar z) \\
&= 2i\hbar(L_y[L^2, x] - L_x[L^2, y] + i\hbar[L^2, z]), \\
&= -4\hbar^2 L_y(y L_z - L_y z) + 4\hbar^2 L_x(L_x z - x L_z) \\
&\quad - 2\hbar^2(L^2 z - z L^2). \tag{6.45}
\end{aligned}$$

This reduces to

$$\begin{aligned}
[L^2, [L^2, z]] &= -\hbar^2 [4(L_x x + L_y y + L_z z)L_z - 4(L_x^2 + L_y^2 + L_z^2)z \\
&\quad + 2(L^2 z - z L^2)]. \tag{6.46}
\end{aligned}$$

However, it is clear from Eqs. (5.1)–(5.3) that

$$L_x x + L_y y + L_z z = 0. \tag{6.47}$$

Hence, we obtain

$$[L^2, [L^2, z]] = 2\hbar^2(L^2 z + z L^2). \tag{6.48}$$

Finally, the above expression expands to give

$$L^4 z - 2L^2 z L^2 + z L^4 - 2\hbar^2(L^2 z + z L^2) = 0. \tag{6.49}$$

Equation (6.49) implies that

$$\langle n, l, m | L^4 z - 2L^2 z L^2 + z L^4 - 2\hbar^2(L^2 z + z L^2) | n', l', m' \rangle = 0. \tag{6.50}$$

This expression yields

$$[\mathfrak{l}^2 (\mathfrak{l} + 1)^2 - 2\mathfrak{l}(\mathfrak{l} + 1)\mathfrak{l}'(\mathfrak{l}' + 1) + \mathfrak{l}'^2 (\mathfrak{l}' + 1)^2 - 2\mathfrak{l}(\mathfrak{l} + 1) - 2\mathfrak{l}'(\mathfrak{l}' + 1)] \langle n, \mathfrak{l}, m | z | n', \mathfrak{l}', m' \rangle = 0, \quad (6.51)$$

which reduces to

$$(\mathfrak{l} + \mathfrak{l}' + 2)(\mathfrak{l} + \mathfrak{l}')(\mathfrak{l} - \mathfrak{l}' + 1)(\mathfrak{l} - \mathfrak{l}' - 1) \langle n, \mathfrak{l}, m | z | n', \mathfrak{l}', m' \rangle = 0. \quad (6.52)$$

According to the above formula, the matrix element  $\langle n, \mathfrak{l}, m | z | n', \mathfrak{l}', m' \rangle$  vanishes unless  $\mathfrak{l} = \mathfrak{l}' = 0$  or  $\mathfrak{l}' = \mathfrak{l} \pm 1$ . This matrix element can be written

$$\langle n, \mathfrak{l}, m | z | n', \mathfrak{l}', m' \rangle = \iiint \psi_{n\mathfrak{l}m}^*(\mathbf{r}', \theta', \varphi') r' \cos \theta' \psi_{n'\mathfrak{l}'m'}(\mathbf{r}', \theta', \varphi') dV', \quad (6.53)$$

where  $\psi_{n\mathfrak{l}m}(\mathbf{r}') = \langle \mathbf{r}' | n, \mathfrak{l}, m \rangle$ . Recall, however, that the wave-function of an  $\mathfrak{l} = 0$  state is spherically symmetric (see Sect. 5.3): *i.e.*,  $\psi_{n00}(\mathbf{r}') = \psi_{n00}(r')$ . It follows from Eq. (6.53) that the matrix element vanishes by symmetry when  $\mathfrak{l} = \mathfrak{l}' = 0$ . In conclusion, the matrix element  $\langle n, \mathfrak{l}, m | z | n', \mathfrak{l}', m' \rangle$  is zero unless  $\mathfrak{l}' = \mathfrak{l} \pm 1$ . This is the selection rule for the quantum number  $\mathfrak{l}$ .

Application of the selection rules to Eq. (6.37) yields

$$\Delta E_{n\mathfrak{l}m} = e^2 |\mathbf{E}|^2 \sum_{n'} \sum_{\mathfrak{l}'=\mathfrak{l}\pm 1} \frac{|\langle n, \mathfrak{l}, m | z | n', \mathfrak{l}', m \rangle|^2}{E_{n\mathfrak{l}m} - E_{n'\mathfrak{l}'m}}. \quad (6.54)$$

Note that all of the terms in Eq. (6.37) which vary linearly with the electric field-strength vanish by symmetry, according to the selection rules. Only those terms which vary *quadratically* with the field-strength survive. The polarizability of an atom is defined in terms of the energy-shift of the atomic state as follows:

$$\Delta E = -\frac{1}{2} \alpha |\mathbf{E}|^2. \quad (6.55)$$

Consider the ground state of a hydrogen atom. (Recall, that we cannot address the  $n > 1$  excited states because they are degenerate, and our theory cannot handle this at present). The polarizability of this state is given by

$$\alpha = 2e^2 \sum_{n>1} \frac{|\langle 1, 0, 0 | z | n, 1, 0 \rangle|^2}{E_{n00} - E_{100}}. \quad (6.56)$$

Here, we have made use of the fact that  $E_{n10} = E_{n00}$  for a hydrogen atom.

The sum in the above expression can be evaluated approximately by noting that [see Eq. (5.120)]

$$E_{n00} = -\frac{e^2}{8\pi\epsilon_0 a_0 n^2} \quad (6.57)$$

for a hydrogen atom, where

$$a_0 = \frac{4\pi\epsilon_0 \hbar^2}{\mu e^2} = 5.3 \times 10^{-11} \text{ meters} \quad (6.58)$$

is the *Bohr radius*. We can write

$$E_{n00} - E_{100} \geq E_{200} - E_{100} = \frac{3}{4} \frac{e^2}{8\pi\epsilon_0 a_0}. \quad (6.59)$$

Thus,

$$\alpha < \frac{16}{3} 4\pi\epsilon_0 a_0 \sum_{n>1} |\langle 1, 0, 0 | z | n, 1, 0 \rangle|^2. \quad (6.60)$$

However,

$$\begin{aligned} \sum_{n>1} |\langle 1, 0, 0 | z | n, 1, 0 \rangle|^2 &= \sum_{n', l', m'} \langle 1, 0, 0 | z | n', l', m' \rangle \langle n', m', l' | z | 1, 0, 0 \rangle \\ &= \langle 1, 0, 0 | z^2 | 1, 0, 0 \rangle, \end{aligned} \quad (6.61)$$

where we have made use of the fact that the wave-functions of a hydrogen atom form a complete set. It is easily demonstrated from the actual form of the ground state wave-function that

$$\langle 1, 0, 0 | z^2 | 1, 0, 0 \rangle = a_0^2. \quad (6.62)$$

Thus, we conclude that

$$\alpha < \frac{16}{3} 4\pi\epsilon_0 a_0^3 \simeq 5.3 4\pi\epsilon_0 a_0^3. \quad (6.63)$$

The true result is

$$\alpha = \frac{9}{2} 4\pi\epsilon_0 a_0^3 = 4.5 4\pi\epsilon_0 a_0^3. \quad (6.64)$$

It is actually possible to obtain this answer, without recourse to perturbation theory, by solving Schrödinger's equation exactly in parabolic coordinates.

## 6.5 Degenerate perturbation theory

Let us now consider systems in which the eigenstates of the unperturbed Hamiltonian,  $H_0$ , possess *degenerate* energy levels. It is always possible to represent degenerate energy eigenstates as the simultaneous eigenstates of the Hamiltonian and some other Hermitian operator (or group of operators). Let us denote this operator (or group of operators)  $L$ . We can write

$$H_0 |n, l\rangle = E_n |n, l\rangle, \quad (6.65)$$

and

$$L |n, l\rangle = L_{nl} |n, l\rangle, \quad (6.66)$$

where  $[H_0, L] = 0$ . Here, the  $E_n$  and the  $L_{nl}$  are real numbers which depend on the quantum numbers  $n$ , and  $n$  and  $l$ , respectively. It is always possible to find a sufficient number of operators which commute with the Hamiltonian in order to ensure that the  $L_{nl}$  are all different. In other words, we can choose  $L$  such that the quantum numbers  $n$  and  $l$  *uniquely* specify each eigenstate. Suppose that for each value of  $n$  there are  $N_n$  different values of  $l$ : *i.e.*, the  $n$ th energy eigenstate is  $N_n$ -fold degenerate.

In general,  $L$  *does not* commute with the perturbing Hamiltonian,  $H_1$ . This implies that the modified energy eigenstates are *not* eigenstates of  $L$ . In this situation, we expect the perturbation to split the degeneracy of the energy levels, so that each modified eigenstate  $|n, l\rangle'$  acquires a unique energy eigenvalue  $E'_{nl}$ . Let us naively attempt to use the standard perturbation theory of Sect. 6.3 to evaluate the modified energy eigenstates and energy levels. A direct generalization of Eqs. (6.32) and (6.33) yields

$$E'_{nl} = E_n + e_{nl} + \sum_{n', l' \neq n, l} \frac{|e_{n'l'nl}|^2}{E_n - E_{n'}} + O(\epsilon^3), \quad (6.67)$$

and

$$|n, l\rangle' = |n, l\rangle + \sum_{n', l' \neq n, l} \frac{e_{n'l'nl}}{E_n - E_{n'}} |n', l'\rangle + O(\epsilon^2), \quad (6.68)$$

where

$$e_{n'l'nl} = \langle n', l' | H_1 | n, l \rangle. \quad (6.69)$$

It is fairly obvious that the summations in Eqs. (6.67) and (6.68) are not well-behaved if the  $n$ th energy level is degenerate. The problem terms are those involving unperturbed eigenstates labeled by the same value of  $n$ , but different values of  $l$ : *i.e.*, those states whose unperturbed energies are  $E_n$ . These terms give rise to singular factors  $1/(E_n - E_n)$  in the summations. Note, however, that this problem would not exist if the matrix elements,  $e_{nl'l}$ , of the perturbing Hamiltonian between distinct, degenerate, unperturbed energy eigenstates corresponding to the eigenvalue  $E_n$  were zero. In other words, if

$$\langle n, l' | H_1 | n, l \rangle = \lambda_{nl} \delta_{ll'}, \quad (6.70)$$

then all of the singular terms in Eqs. (6.67) and (6.68) would vanish.

In general, Eq. (6.70) is not satisfied. Fortunately, we can always redefine the unperturbed energy eigenstates belonging to the eigenvalue  $E_n$  in such a manner that Eq. (6.70) is satisfied. Let us define  $N_n$  new states which are linear combinations of the  $N_n$  original degenerate eigenstates corresponding to the eigenvalue  $E_n$ :

$$|n, l^{(1)}\rangle = \sum_{k=1}^{N_n} \langle n, k | n, l^{(1)} \rangle |n, k\rangle. \quad (6.71)$$

Note that these new states are also degenerate energy eigenstates of the unperturbed Hamiltonian corresponding to the eigenvalue  $E_n$ . The  $|n, l^{(1)}\rangle$  are chosen in such a manner that they are eigenstates of the perturbing Hamiltonian,  $H_1$ . Thus,

$$H_1 |n, l^{(1)}\rangle = \lambda_{nl} |n, l^{(1)}\rangle. \quad (6.72)$$

The  $|n, l^{(1)}\rangle$  are also chosen so that they are orthonormal, and have unit lengths. It follows that

$$\langle n, l'^{(1)} | H_1 | n, l^{(1)} \rangle = \lambda_{nl} \delta_{ll'}. \quad (6.73)$$

Thus, if we use the new eigenstates, instead of the old ones, then we can employ Eqs. (6.67) and (6.68) directly, since all of the singular terms vanish. The only remaining difficulty is to determine the new eigenstates in terms of the original ones.

Now

$$\sum_{l=1}^{N_n} |n, l\rangle \langle n, l| = 1, \quad (6.74)$$

where 1 denotes the identity operator in the sub-space of all unperturbed energy eigenkets corresponding to the eigenvalue  $E_n$ . Using this completeness relation, the operator eigenvalue equation (6.72) can be transformed into a straightforward matrix eigenvalue equation:

$$\sum_{l''=1}^{N_n} \langle n, l'|H_1|n, l''\rangle \langle n, l''|n, l^{(1)}\rangle = \lambda_{nl} \langle n, l'|n, l^{(1)}\rangle. \quad (6.75)$$

This can be written more transparently as

$$\mathbf{U} \mathbf{x} = \lambda \mathbf{x}, \quad (6.76)$$

where the elements of the  $N_n \times N_n$  Hermitian matrix  $\mathbf{U}$  are

$$U_{jk} = \langle n, j|H_1|n, k\rangle. \quad (6.77)$$

Provided that the determinant of  $\mathbf{U}$  is non-zero, Eq. (6.76) can always be solved to give  $N_n$  eigenvalues  $\lambda_{nl}$  (for  $l = 1$  to  $N_n$ ), with  $N_n$  corresponding eigenvectors  $\mathbf{x}_{nl}$ . The eigenvectors specify the weights of the new eigenstates in terms of the original eigenstates: *i.e.*,

$$(\mathbf{x}_{nl})_k = \langle n, k|n, l^{(1)}\rangle, \quad (6.78)$$

for  $k = 1$  to  $N_n$ . In our new scheme, Eqs. (6.67) and (6.68) yield

$$E'_{nl} = E_n + \lambda_{nl} + \sum_{n' \neq n, l'} \frac{|e_{n'l'nl}|^2}{E_n - E_{n'}} + O(\epsilon^3), \quad (6.79)$$

and

$$|n, l^{(1)}\rangle' = |n, l^{(1)}\rangle + \sum_{n' \neq n, l'} \frac{e_{n'l'nl}}{E_n - E_{n'}} |n', l'\rangle + O(\epsilon^2). \quad (6.80)$$

There are no singular terms in these expressions, since the summations are over  $n' \neq n$ : *i.e.*, they specifically exclude the problematic, degenerate, unperturbed energy eigenstates corresponding to the eigenvalue  $E_n$ . Note that the first-order energy shifts are equivalent to the eigenvalues of the matrix equation (6.76).

## 6.6 The linear Stark effect

Let us examine the effect of an electric field on the excited energy levels of a hydrogen atom. For instance, consider the  $n = 2$  states. There is a single  $l = 0$  state, usually referred to as  $2s$ , and three  $l = 1$  states (with  $m = -1, 0, 1$ ), usually referred to as  $2p$ . All of these states possess the same energy,  $E_{200} = -e^2/(32\pi\epsilon_0 a_0)$ . As in Sect. 6.4, the perturbing Hamiltonian is

$$H_1 = e |\mathbf{E}| z. \quad (6.81)$$

In order to apply perturbation theory, we have to solve the matrix eigenvalue equation

$$\mathbf{U} \mathbf{x} = \lambda \mathbf{x}, \quad (6.82)$$

where  $\mathbf{U}$  is the array of the matrix elements of  $H_1$  between the degenerate  $2s$  and  $2p$  states. Thus,

$$\mathbf{U} = e |\mathbf{E}| \begin{pmatrix} 0 & \langle 2, 0, 0 | z | 2, 1, 0 \rangle & 0 & 0 \\ \langle 2, 1, 0 | z | 2, 0, 0 \rangle & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (6.83)$$

where the rows and columns correspond to the  $|2, 0, 0\rangle$ ,  $|2, 1, 0\rangle$ ,  $|2, 1, 1\rangle$ , and  $|2, 1, -1\rangle$  states, respectively. Here, we have made use of the selection rules, which tell us that the matrix element of  $z$  between two hydrogen atom states is zero unless the states possess the same  $m$  quantum number, and  $l$  quantum numbers which differ by unity. It is easily demonstrated, from the exact forms of the  $2s$  and  $2p$  wave-functions, that

$$\langle 2, 0, 0 | z | 2, 1, 0 \rangle = \langle 2, 1, 0 | z | 2, 0, 0 \rangle = 3 a_0. \quad (6.84)$$

It can be seen, by inspection, that the eigenvalues of  $\mathbf{U}$  are  $\lambda_1 = 3 e a_0 |\mathbf{E}|$ ,  $\lambda_2 = -3 e a_0 |\mathbf{E}|$ ,  $\lambda_3 = 0$ , and  $\lambda_4 = 0$ . The corresponding eigenvectors are

$$\mathbf{x}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \\ 0 \end{pmatrix}, \quad (6.85)$$

$$\mathbf{x}_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \\ 0 \end{pmatrix}, \quad (6.86)$$

$$\mathbf{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad (6.87)$$

$$\mathbf{x}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (6.88)$$

It follows from Sect. 6.5 that the simultaneous eigenstates of the unperturbed Hamiltonian and the perturbing Hamiltonian take the form

$$|1\rangle = \frac{|2, 0, 0\rangle + |2, 1, 0\rangle}{\sqrt{2}}, \quad (6.89)$$

$$|2\rangle = \frac{|2, 0, 0\rangle - |2, 1, 0\rangle}{\sqrt{2}}, \quad (6.90)$$

$$|3\rangle = |2, 1, 1\rangle, \quad (6.91)$$

$$|4\rangle = |2, 1, -1\rangle. \quad (6.92)$$

In the absence of an electric field, all of these states possess the same energy,  $E_{200}$ . The first-order energy shifts induced by an electric field are given by

$$\Delta E_1 = +3 e a_0 |\mathbf{E}|, \quad (6.93)$$

$$\Delta E_2 = -3 e a_0 |\mathbf{E}|, \quad (6.94)$$

$$\Delta E_3 = 0, \quad (6.95)$$

$$\Delta E_4 = 0. \quad (6.96)$$

Thus, the energies of states 1 and 2 are shifted upwards and downwards, respectively, by an amount  $3 e a_0 |\mathbf{E}|$  in the presence of an electric field. States 1 and 2 are orthogonal linear combinations of the original 2s and 2p( $m = 0$ ) states.

Note that the energy shifts are *linear* in the electric field-strength, so this is a much larger effect than the quadratic effect described in Sect. 6.4. The energies of states 3 and 4 (which are equivalent to the original  $2p(m = 1)$  and  $2p(m = -1)$  states, respectively) are not affected to first-order. Of course, to second-order the energies of these states are shifted by an amount which depends on the square of the electric field-strength.

Note that the linear Stark effect depends crucially on the degeneracy of the  $2s$  and  $2p$  states. This degeneracy is a special property of a pure Coulomb potential, and, therefore, only applies to a hydrogen atom. Thus, alkali metal atoms do not exhibit the linear Stark effect.

## 6.7 Fine structure

Let us now consider the energy levels of hydrogen-like atoms (*i.e.*, alkali metal atoms) in more detail. The outermost electron moves in a spherically symmetric potential  $V(r)$  due to the nuclear charge and the charges of the other electrons (which occupy spherically symmetric closed shells). The shielding effect of the inner electrons causes  $V(r)$  to depart from the pure Coulomb form. This splits the degeneracy of states characterized by the same value of  $n$ , but different values of  $l$ . In fact, higher  $l$  states have higher energies.

Let us examine a phenomenon known as *fine structure*, which is due to interaction between the spin and orbital angular momenta of the outermost electron. This electron experiences an electric field

$$\mathbf{E} = \frac{\nabla V}{e}. \quad (6.97)$$

However, a charge moving in an electric field also experiences an effective magnetic field

$$\mathbf{B} = -\mathbf{v} \times \mathbf{E}. \quad (6.98)$$

Now, an electron possesses a spin magnetic moment [see Eq. (5.170)]

$$\boldsymbol{\mu} = -\frac{e\mathbf{S}}{m_e}. \quad (6.99)$$

We, therefore, expect a spin-orbit contribution to the Hamiltonian of the form

$$\begin{aligned}
 H_{LS} &= -\boldsymbol{\mu} \cdot \mathbf{B} \\
 &= -\frac{e \mathbf{S}}{m_e} \cdot \mathbf{v} \times \left( \frac{1}{e r} \frac{dV}{dr} \right) \\
 &= \frac{1}{m_e^2 r} \frac{dV}{dr} \mathbf{L} \cdot \mathbf{S},
 \end{aligned} \tag{6.100}$$

where  $\mathbf{L} = m_e \mathbf{r} \times \mathbf{v}$  is the orbital angular momentum. When the above expression is compared to the observed spin-orbit interaction, it is found to be too large by a factor of two. There is a classical explanation for this, due to spin precession, which we need not go into. The correct quantum mechanical explanation requires a relativistically covariant treatment of electron dynamics (this is achieved using the so-called *Dirac equation*).

Let us now apply perturbation theory to a hydrogen-like atom, using  $H_{LS}$  as the perturbation (with  $H_{LS}$  taking one half of the value given above), and

$$H_0 = \frac{\mathbf{p}^2}{2 m_e} + V(r) \tag{6.101}$$

as the unperturbed Hamiltonian. We have two choices for the energy eigenstates of  $H_0$ . We can adopt the simultaneous eigenstates of  $H_0, L^2, S^2, L_z$  and  $S_z$ , or the simultaneous eigenstates of  $H_0, L^2, S^2, J^2$ , and  $J_z$ , where  $\mathbf{J} = \mathbf{L} + \mathbf{S}$  is the total angular momentum. Although the departure of  $V(r)$  from a pure  $1/r$  form splits the degeneracy of same  $n$ , different  $l$ , states, those states characterized by the same values of  $n$  and  $l$ , but different values of  $m_l$ , are still degenerate. (Here,  $m_l, m_s$ , and  $m_j$  are the quantum numbers corresponding to  $L_z, S_z$ , and  $J_z$ , respectively.) Moreover, with the addition of spin degrees of freedom, each state is doubly degenerate due to the two possible orientations of the electron spin (*i.e.*,  $m_s = \pm 1/2$ ). Thus, we are still dealing with a highly degenerate system. We know, from Sect. 6.6, that the application of perturbation theory to a degenerate system is greatly simplified if the basis eigenstates of the unperturbed Hamiltonian are also eigenstates of the perturbing Hamiltonian. Now, the perturbing Hamiltonian,  $H_{LS}$ , is proportional to  $\mathbf{L} \cdot \mathbf{S}$ , where

$$\mathbf{L} \cdot \mathbf{S} = \frac{J^2 - L^2 - S^2}{2}. \tag{6.102}$$

It is fairly obvious that the first group of operators ( $H_0, L^2, S^2, L_z$  and  $S_z$ ) does not commute with  $H_{LS}$ , whereas the second group ( $H_0, L^2, S^2, J^2$ , and  $J_z$ ) does. In fact,  $\mathbf{L} \cdot \mathbf{S}$  is just a combination of operators appearing in the second group. Thus, it is advantageous to work in terms of the eigenstates of the second group of operators, rather than those of the first group.

We now need to find the simultaneous eigenstates of  $H_0, L^2, S^2, J^2$ , and  $J_z$ . This is equivalent to finding the eigenstates of the total angular momentum resulting from the addition of two angular momenta:  $j_1 = l$ , and  $j_2 = s = 1/2$ . According to Eq. (5.276), the allowed values of the total angular momentum are  $j = l + 1/2$  and  $j = l - 1/2$ . We can write

$$|l + 1/2, m\rangle = \cos \alpha |m - 1/2, 1/2\rangle + \sin \alpha |m + 1/2, -1/2\rangle, \quad (6.103)$$

$$|l - 1/2, m\rangle = -\sin \alpha |m - 1/2, 1/2\rangle + \cos \alpha |m + 1/2, -1/2\rangle. \quad (6.104)$$

Here, the kets on the left-hand side are  $|j, m_j\rangle$  kets, whereas those on the right-hand side are  $|m_l, m_s\rangle$  kets (the  $j_1, j_2$  labels have been dropped, for the sake of clarity). We have made use of the fact that the Clebsch-Gordon coefficients are automatically zero unless  $m_j = m_l + m_s$ . We have also made use of the fact that both the  $|j, m_j\rangle$  and  $|m_l, m_s\rangle$  kets are orthonormal, and have unit lengths. We now need to determine

$$\cos \alpha = \langle m - 1/2, 1/2 | l + 1/2, m \rangle, \quad (6.105)$$

where the Clebsch-Gordon coefficient is written in  $\langle m_l, m_s | j, m_j \rangle$  form.

Let us now employ the recursion relation for Clebsch-Gordon coefficients, Eq. (5.282), with  $j_1 = l, j_2 = 1/2, j = l + 1/2, m_1 = m - 1/2, m_2 = 1/2$  (lower sign). We obtain

$$\begin{aligned} & \sqrt{(l + 1/2)(l + 3/2) - m(m + 1)} \langle m - 1/2, 1/2 | l + 1/2, m \rangle \\ &= \sqrt{l(l + 1) - (m - 1/2)(m + 1/2)} \langle m + 1/2, 1/2 | l + 1/2, m + 1 \rangle, \end{aligned} \quad (6.106)$$

which reduces to

$$\langle m - 1/2, 1/2 | l + 1/2, m \rangle = \sqrt{\frac{l + m + 1/2}{l + m + 3/2}} \langle m + 1/2, 1/2 | l + 1/2, m + 1 \rangle. \quad (6.107)$$

We can use this formula to successively increase the value of  $m_l$ . For instance,

$$\begin{aligned} \langle m - 1/2, 1/2 | l + 1/2, m \rangle &= \sqrt{\frac{l + m + 1/2}{l + m + 3/2}} \sqrt{\frac{l + m + 3/2}{l + m + 5/2}} \\ &\times \langle m + 3/2, 1/2 | l + 1/2, m + 2 \rangle. \end{aligned} \quad (6.108)$$

This procedure can be continued until  $m_l$  attains its maximum possible value,  $l$ . Thus,

$$\langle m - 1/2, 1/2 | l + 1/2, m \rangle = \sqrt{\frac{l + m + 1/2}{2l + 1}} \langle l, 1/2 | l + 1/2, l + 1/2 \rangle. \quad (6.109)$$

Consider the situation in which  $m_l$  and  $m$  both take their maximum values,  $l$  and  $1/2$ , respectively. The corresponding value of  $m_j$  is  $l + 1/2$ . This value is possible when  $j = l + 1/2$ , but not when  $j = l - 1/2$ . Thus, the  $|m_l, m_s\rangle$  ket  $|l, 1/2\rangle$  must be equal to the  $|j, m_j\rangle$  ket  $|l + 1/2, l + 1/2\rangle$ , up to an arbitrary phase-factor. By convention, this factor is taken to be unity, giving

$$\langle l, 1/2 | l + 1/2, l + 1/2 \rangle = 1. \quad (6.110)$$

It follows from Eq. (6.109) that

$$\cos \alpha = \langle m - 1/2, 1/2 | l + 1/2, m \rangle = \sqrt{\frac{l + m + 1/2}{2l + 1}}. \quad (6.111)$$

Now,

$$\sin^2 \alpha = 1 - \frac{l + m + 1/2}{2l + 1} = \frac{l - m + 1/2}{2l + 1}. \quad (6.112)$$

We now need to determine the sign of  $\sin \alpha$ . A careful examination of the recursion relation, Eq. (5.282), shows that the plus sign is appropriate. Thus,

$$\begin{aligned} |l + 1/2, m\rangle &= \sqrt{\frac{l + m + 1/2}{2l + 1}} |m - 1/2, 1/2\rangle \\ &+ \sqrt{\frac{l - m + 1/2}{2l + 1}} |m + 1/2, -1/2\rangle, \end{aligned} \quad (6.113)$$

$$\begin{aligned}
|l - 1/2, m\rangle &= -\sqrt{\frac{l - m + 1/2}{2l + 1}} |m - 1/2, 1/2\rangle \\
&\quad + \sqrt{\frac{l + m + 1/2}{2l + 1}} |m + 1/2, -1/2\rangle.
\end{aligned} \tag{6.114}$$

It is convenient to define so called *spin-angular functions* using the Pauli two-component formalism:

$$\begin{aligned}
\mathcal{Y}_l^{j=l\pm 1/2, m} &= \pm \sqrt{\frac{l \pm m + 1/2}{2l + 1}} Y_l^{m-1/2}(\theta, \varphi) \chi_+ \\
&\quad + \sqrt{\frac{l \mp m + 1/2}{2l + 1}} Y_l^{m+1/2}(\theta, \varphi) \chi_- \\
&= \frac{1}{\sqrt{2l + 1}} \begin{pmatrix} \pm \sqrt{l \pm m + 1/2} Y_l^{m-1/2}(\theta, \varphi) \\ \sqrt{l \mp m + 1/2} Y_l^{m+1/2}(\theta, \varphi) \end{pmatrix}.
\end{aligned} \tag{6.115}$$

These functions are eigenfunctions of the total angular momentum for spin one-half particles, just as the spherical harmonics are eigenfunctions of the orbital angular momentum. A general wave-function for an energy eigenstate in a hydrogen-like atom is written

$$\psi_{nlm\pm} = R_{nl}(r) \mathcal{Y}^{j=l\pm 1/2, m}. \tag{6.116}$$

The radial part of the wave-function,  $R_{nl}(r)$ , depends on the radial quantum number  $n$  and the angular quantum number  $l$ . The wave-function is also labeled by  $m$ , which is the quantum number associated with  $J_z$ . For a given choice of  $l$ , the quantum number  $j$  (*i.e.*, the quantum number associated with  $J^2$ ) can take the values  $l \pm 1/2$ .

The  $|l \pm 1/2, m\rangle$  kets are eigenstates of  $\mathbf{L} \cdot \mathbf{S}$ , according to Eq. (6.102). Thus,

$$\mathbf{L} \cdot \mathbf{S} |j = l \pm 1/2, m_j = m\rangle = \frac{\hbar^2}{2} [j(j + 1) - l(l + 1) - 3/4] |j, m\rangle, \tag{6.117}$$

giving

$$\mathbf{L} \cdot \mathbf{S} |l + 1/2, m\rangle = \frac{l \hbar^2}{2} |l + 1/2, m\rangle, \tag{6.118}$$

$$\mathbf{L} \cdot \mathbf{S} |l - 1/2, m\rangle = -\frac{(l + 1) \hbar^2}{2} |l - 1/2, m\rangle. \tag{6.119}$$

It follows that

$$\int (\mathcal{Y}^{l+1/2,m})^\dagger \mathbf{L} \cdot \mathbf{S} \mathcal{Y}^{l+1/2,m} d\Omega = \frac{l \hbar^2}{2}, \quad (6.120)$$

$$\int (\mathcal{Y}^{l-1/2,m})^\dagger \mathbf{L} \cdot \mathbf{S} \mathcal{Y}^{l-1/2,m} d\Omega = -\frac{(l+1) \hbar^2}{2}, \quad (6.121)$$

where the integrals are over all solid angle.

Let us now apply degenerate perturbation theory to evaluate the shift in energy of a state whose wave-function is  $\psi_{nlm\pm}$  due to the spin-orbit Hamiltonian  $H_{LS}$ . To first-order, the energy-shift is given by

$$\Delta E_{nlm\pm} = \int (\psi_{nlm\pm})^\dagger H_{LS} \psi_{nlm\pm} dV, \quad (6.122)$$

where the integral is over all space. Equations (6.100) (remember the factor of two), (6.116), and (6.120)–(6.121) yield

$$\Delta E_{nlm+} = +\frac{1}{2 m_e^2} \left\langle \frac{1}{r} \frac{dV}{dr} \right\rangle \frac{l \hbar^2}{2}, \quad (6.123)$$

$$\Delta E_{nlm-} = -\frac{1}{2 m_e^2} \left\langle \frac{1}{r} \frac{dV}{dr} \right\rangle \frac{(l+1) \hbar^2}{2}, \quad (6.124)$$

where

$$\left\langle \frac{1}{r} \frac{dV}{dr} \right\rangle = \int (R_{nl})^* \frac{1}{r} \frac{dV}{dr} R_{nl} r^2 dr. \quad (6.125)$$

Equations (6.123)–(6.124) are known as *Lande's interval rule*.

Let us now apply the above result to the case of a sodium atom. In chemist's notation, the ground state is written

$$(1s)^2(2s)^2(2p)^6(3s). \quad (6.126)$$

The inner ten electrons effectively form a spherically symmetric electron cloud. We are interested in the excitation of the eleventh electron from 3s to some higher energy state. The closest (in energy) unoccupied state is 3p. This state has a higher energy than 3s due to the deviations of the potential from the pure

Coulomb form. In the absence of spin-orbit interaction, there are six degenerate 3p states. The spin-orbit interaction breaks the degeneracy of these states. The modified states are labeled  $(3p)_{1/2}$  and  $(3p)_{3/2}$ , where the subscript refers to the value of  $j$ . The four  $(3p)_{3/2}$  states lie at a slightly higher energy level than the two  $(3p)_{1/2}$  states, because the radial integral (6.125) is positive. The splitting of the (3p) energy levels of the sodium atom can be observed using a spectroscope. The well-known sodium D line is associated with transitions between the 3p and 3s states. The fact that there are two slightly different 3p energy levels (note that spin-orbit coupling does not split the 3s energy levels) means that the sodium D line actually consists of two very closely spaced spectroscopic lines. It is easily demonstrated that the ratio of the typical spacing of Balmer lines to the splitting brought about by spin-orbit interaction is about  $1 : \alpha^2$ , where

$$\alpha = \frac{e^2}{2 \epsilon_0 h c} = \frac{1}{137} \quad (6.127)$$

is the *fine structure constant*. Note that Eqs. (6.123)–(6.124) are not entirely correct, since we have neglected an effect (namely, the relativistic mass correction of the electron) which is the same order of magnitude as spin-orbit coupling.

## 6.8 The Zeeman effect

Consider a hydrogen-like atom placed in a uniform  $z$ -directed magnetic field. The change in energy of the outermost electron is

$$H_B = -\boldsymbol{\mu} \cdot \mathbf{B}, \quad (6.128)$$

where

$$\boldsymbol{\mu} = -\frac{e}{2 m_e} (\mathbf{L} + 2 \mathbf{S}) \quad (6.129)$$

is its magnetic moment, including both the spin and orbital contributions. Thus,

$$H_B = \frac{e B}{2 m_e} (L_z + 2 S_z). \quad (6.130)$$

Suppose that the energy-shifts induced by the magnetic field are much smaller than those induced by spin-orbit interaction. In this situation, we can treat  $H_B$  as

a small perturbation acting on the eigenstates of  $H_0 + H_{LS}$ . Of course, these states are the simultaneous eigenstates of  $J^2$  and  $J_z$ . Let us consider one of these states, labeled by the quantum numbers  $j$  and  $m$ , where  $j = l \pm 1/2$ . From standard perturbation theory, the first-order energy-shift in the presence of a magnetic field is

$$\Delta E_{nlm\pm} = \langle l \pm 1/2, m | H_B | l \pm 1/2, m \rangle. \quad (6.131)$$

Since

$$L_z + 2S_z = J_z + S_z, \quad (6.132)$$

we find that

$$\Delta E_{nlm\pm} = \frac{eB}{2m_e} (m\hbar + \langle l \pm 1/2, m | S_z | l \pm 1/2, m \rangle). \quad (6.133)$$

Now, from Eqs. (6.113)–(6.114),

$$\begin{aligned} |l \pm 1/2, m\rangle &= \pm \sqrt{\frac{l \pm m + 1/2}{2l + 1}} |m - 1/2, 1/2\rangle \\ &\quad + \sqrt{\frac{l \mp m + 1/2}{2l + 1}} |m + 1/2, -1/2\rangle. \end{aligned} \quad (6.134)$$

It follows that

$$\begin{aligned} \langle l \pm 1/2, m | S_z | l \pm 1/2, m \rangle &= \frac{\hbar}{2(2l + 1)} [(l \pm m + 1/2) - (l \mp m + 1/2)] \\ &= \pm \frac{m\hbar}{2l + 1}. \end{aligned} \quad (6.135)$$

Thus, we obtain Lande's formula for the energy-shift induced by a weak magnetic field:

$$\Delta E_{nlm\pm} = \frac{e\hbar B}{2m_e} m \left[ 1 \pm \frac{1}{2l + 1} \right]. \quad (6.136)$$

Let us apply this theory to the sodium atom. We have already seen that the non-Coulomb potential splits the degeneracy of the 3s and 3p states, the latter states acquiring a higher energy. The spin-orbit interaction splits the six 3p states into two groups, with four  $j = 3/2$  states lying at a slightly higher energy than two  $j = 1/2$  states. According to Eq. (6.136), a magnetic field splits the  $(3p)_{3/2}$

quadruplet of states, each state acquiring a different energy. In fact, the energy of each state becomes dependent on the quantum number  $m$ , which measures the projection of the total angular momentum along the  $z$ -axis. States with higher  $m$  values have higher energies. A magnetic field also splits the  $(3p)_{1/2}$  doublet of states. However, it is evident from Eq. (6.136) that these states are split by a lesser amount than the  $j = 3/2$  states.

Suppose that we increase the strength of the magnetic field, so that the energy-shift due to the magnetic field becomes comparable to the energy-shift induced by spin-orbit interaction. Clearly, in this situation, it does not make much sense to think of  $H_B$  as a small interaction term operating on the eigenstates of  $H_0 + H_{LS}$ . In fact, this intermediate case is very difficult to analyze. Let us consider the extreme limit in which the energy-shift due to the magnetic field greatly exceeds that induced by spin-orbit effects. This is called the *Paschen-Back limit*.

In the Paschen-Back limit we can think of the spin-orbit Hamiltonian,  $H_{LS}$ , as a small interaction term operating on the eigenstates of  $H_0 + H_B$ . Note that the magnetic Hamiltonian,  $H_B$ , commutes with  $L^2, S^2, L_z, S_z$ , but does not commute with  $L^2, S^2, J^2, J_z$ . Thus, in an intense magnetic field, the energy eigenstates of a hydrogen-like atom are approximate eigenstates of the spin and orbital angular momenta, but are not eigenstates of the total angular momentum. We can label each state by the quantum numbers  $n$  (the energy quantum number),  $l$ ,  $m_l$ , and  $m_s$ . Thus, our energy eigenkets are written  $|n, l, m_l, m_s\rangle$ . The unperturbed Hamiltonian,  $H_0$ , causes states with different values of the quantum numbers  $n$  and  $l$  to have different energies. However, states with the same value of  $n$  and  $l$ , but different values of  $m_l$  and  $m_s$ , are degenerate. The shift in energy due to the magnetic field is simply

$$\begin{aligned}\Delta E_{nlm_lm_s} &= \langle n, l, m_l, m_s | H_B | n, l, m_l, m_s \rangle \\ &= \frac{e \hbar B}{2 m_e} (m_l + 2 m_s).\end{aligned}\tag{6.137}$$

Thus, states with different values of  $m_l + 2 m_s$  acquire different energies.

Let us apply this result to a sodium atom. In the absence of a magnetic field, the six  $3p$  states form two groups of four and two states, depending on the values

of their total angular momentum. In the presence of an intense magnetic field the 3p states are split into five groups. There is a state with  $m_l + 2 m_s = 2$ , a state with  $m_l + 2 m_s = 1$ , two states with  $m_l + 2 m_s = 0$ , a state with  $m_l + 2 m_s = -1$ , and a state with  $m_l + 2 m_s = -2$ . These groups are equally spaced in energy, the energy difference between adjacent groups being  $e \hbar B / 2 m_e$ .

The energy-shift induced by the spin-orbit Hamiltonian is given by

$$\Delta E_{nl m_l m_s} = \langle n, l, m_l, m_s | H_{LS} | n, l, m_l, m_s \rangle, \quad (6.138)$$

where

$$H_{LS} = \frac{1}{2 m_e^2} \frac{1}{r} \frac{dV}{dr} \mathbf{L} \cdot \mathbf{S}. \quad (6.139)$$

Now,

$$\begin{aligned} \langle \mathbf{L} \cdot \mathbf{S} \rangle &= \langle L_z S_z + (L^+ S^- + L^- S^+) / 2 \rangle \\ &= \hbar^2 m_l m_s, \end{aligned} \quad (6.140)$$

since

$$\langle L^\pm \rangle = \langle S^\pm \rangle = 0 \quad (6.141)$$

for expectation values taken between the simultaneous eigenkets of  $L_z$  and  $S_z$ . Thus,

$$\Delta E_{nl m_l m_s} = \frac{\hbar^2 m_l m_s}{2 m_e^2} \left\langle \frac{1}{r} \frac{dV}{dr} \right\rangle. \quad (6.142)$$

Let us apply the above result to a sodium atom. In the presence of an intense magnetic field, the 3p states are split into five groups with  $(m_l, m_s)$  quantum numbers  $(1, 1/2)$ ,  $(0, 1/2)$ ,  $(1, -1/2)$ , or  $(-1, 1/2)$ ,  $(0, -1/2)$ , and  $(-1, -1/2)$ , respectively, in order of decreasing energy. The spin-orbit term increases the energy of the highest energy state, does not affect the next highest energy state, decreases, but does not split, the energy of the doublet, does not affect the next lowest energy state, and increases the energy of the lowest energy state. The net result is that the five groups of states are no longer equally spaced in energy.

The sort of magnetic field-strength needed to get into the Paschen-Bach limit is given by

$$B_{PB} \sim \alpha^2 \frac{e m_e}{\epsilon_0 \hbar a_0} \simeq 25 \text{ tesla}. \quad (6.143)$$

Obviously, this is an extremely large field-strength.

## 6.9 Time-dependent perturbation theory

Suppose that the Hamiltonian of the system under consideration can be written

$$H = H_0 + H_1(t), \quad (6.144)$$

where  $H_0$  does not contain time explicitly, and  $H_1$  is a small time-dependent perturbation. It is assumed that we are able to calculate the eigenkets of the unperturbed Hamiltonian:

$$H_0 |n\rangle = E_n |n\rangle. \quad (6.145)$$

We know that if the system is in one of the eigenstates of  $H_0$  then, in the absence of the external perturbation, it remains in this state for ever. However, the presence of a small time-dependent perturbation can, in principle, give rise to a finite probability that a system initially in some eigenstate  $|i\rangle$  of the unperturbed Hamiltonian is found in some other eigenstate at a subsequent time (since  $|i\rangle$  is no longer an exact eigenstate of the total Hamiltonian). In other words, a time-dependent perturbation causes the system to make transitions between its unperturbed energy eigenstates. Let us investigate this effect.

Suppose that at  $t = t_0$  the state of the system is represented by

$$|A\rangle = \sum_n c_n |n\rangle, \quad (6.146)$$

where the  $c_n$  are complex numbers. Thus, the initial state is some linear superposition of the unperturbed energy eigenstates. In the absence of the time-dependent perturbation, the time evolution of the system is given by

$$|A, t_0, t\rangle = \sum_n c_n \exp([-i E_n(t - t_0)/\hbar] |n\rangle. \quad (6.147)$$

Now, the probability of finding the system in state  $|n\rangle$  at time  $t$  is

$$P_n(t) = |c_n \exp[-i E_n(t - t_0)/\hbar]|^2 = |c_n|^2 = P_n(t_0). \quad (6.148)$$

Clearly, with  $H_1 = 0$ , the probability of finding the system in state  $|n\rangle$  at time  $t$  is exactly the same as the probability of finding the system in this state at the initial time  $t_0$ . However, with  $H_1 \neq 0$ , we expect  $P_n(t)$  to vary with time. Thus, we can write

$$|A, t_0, t\rangle = \sum_n c_n(t) \exp[-i E_n(t - t_0)/\hbar] |n\rangle, \quad (6.149)$$

where  $P_n(t) = |c_n(t)|^2$ . Here, we have carefully separated the fast phase oscillation of the eigenkets, which depends on the unperturbed Hamiltonian, from the slow variation of the amplitudes  $c_n(t)$ , which depends entirely on the perturbation (*i.e.*,  $c_n$  is constant if  $H_1 = 0$ ). Note that in Eq. (6.149) the eigenkets  $|n\rangle$  are *time-independent* (they are actually the eigenkets of  $H_0$  evaluated at the time  $t_0$ ).

Schrödinger's time evolution equation yields

$$i \hbar \frac{\partial}{\partial t} |A, t_0, t\rangle = H |A, t_0, t\rangle = (H_0 + H_1) |A, t_0, t\rangle. \quad (6.150)$$

It follows from Eq. (6.149) that

$$(H_0 + H_1) |A, t_0, t\rangle = \sum_m c_m(t) \exp[-i E_m(t - t_0)/\hbar] (E_m + H_1) |m\rangle. \quad (6.151)$$

We also have

$$i \hbar \frac{\partial}{\partial t} |A, t_0, t\rangle = \sum_m \left( i \hbar \frac{dc_m}{dt} + c_m(t) E_m \right) \exp[-i E_m(t - t_0)/\hbar] |m\rangle, \quad (6.152)$$

where use has been made of the time-independence of the kets  $|m\rangle$ . According to Eq. (6.150), we can equate the right-hand sides of the previous two equations to obtain

$$\sum_m i \hbar \frac{dc_m}{dt} \exp[-i E_m(t - t_0)/\hbar] |m\rangle = \sum_m c_m(t) \exp[-i E_m(t - t_0)/\hbar] H_1 |m\rangle. \quad (6.153)$$

Left-multiplication by  $\langle n|$  yields

$$i \hbar \frac{dc_n}{dt} = \sum_m H_{nm}(t) \exp[i \omega_{nm}(t - t_0)] c_m(t), \quad (6.154)$$

where

$$H_{nm}(t) = \langle n|H_1(t)|m\rangle, \quad (6.155)$$

and

$$\omega_{nm} = \frac{E_n - E_m}{\hbar}. \quad (6.156)$$

Here, we have made use of the standard orthonormality result,  $\langle n|m \rangle = \delta_{nm}$ . Suppose that there are  $N$  linearly independent eigenkets of the unperturbed Hamiltonian. According to Eq. (6.154), the time variation of the coefficients  $c_n$ , which specify the probability of finding the system in state  $|n\rangle$  at time  $t$ , is determined by  $N$  coupled first-order differential equations. Note that Eq. (6.154) is exact—we have made no approximations at this stage. Unfortunately, we cannot generally find exact solutions to this equation, so we have to obtain approximate solutions via suitable expansions in small quantities. However, for the particularly simple case of a two-state system (*i.e.*,  $N = 2$ ), it is actually possible to solve Eq. (6.154) without approximation. This solution is of enormous practical importance.

## 6.10 The two-state system

Consider a system in which the time-independent Hamiltonian possesses two eigenstates, denoted

$$H_0 |1\rangle = E_1 |1\rangle, \quad (6.157)$$

$$H_0 |2\rangle = E_2 |2\rangle. \quad (6.158)$$

Suppose, for the sake of simplicity, that the diagonal matrix elements of the interaction Hamiltonian,  $H_1$ , are zero:

$$\langle 1|H_1|1\rangle = \langle 2|H_1|2\rangle = 0. \quad (6.159)$$

The off-diagonal matrix elements are assumed to oscillate sinusoidally at some frequency  $\omega$ :

$$\langle 1|H_1|2\rangle = \langle 2|H_1|1\rangle^* = \gamma \exp(i\omega t), \quad (6.160)$$

where  $\gamma$  and  $\omega$  are real. Note that it is only the off-diagonal matrix elements which give rise to the effect which we are interested in—namely, transitions between states 1 and 2.

For a two-state system, Eq. (6.154) reduces to

$$i \hbar \frac{dc_1}{dt} = \gamma \exp[+i(\omega - \omega_{21})t] c_2, \quad (6.161)$$

$$i \hbar \frac{dc_2}{dt} = \gamma \exp[-i(\omega - \omega_{21})t] c_1, \quad (6.162)$$

where  $\omega_{21} = (E_2 - E_1)/\hbar$ , and assuming that  $t_0 = 0$ . Equations (6.161) and (6.162) can be combined to give a second-order differential equation for the time variation of the amplitude  $c_2$ :

$$\frac{d^2c_2}{dt^2} + i(\omega - \omega_{21})\frac{dc_2}{dt} + \frac{\gamma^2}{\hbar^2}c_2 = 0. \quad (6.163)$$

Once we have solved for  $c_2$ , we can use Eq. (6.162) to obtain the amplitude  $c_1$ . Let us look for a solution in which the system is certain to be in state 1 at time  $t = 0$ . Thus, our boundary conditions are  $c_1(0) = 1$  and  $c_2(0) = 0$ . It is easily demonstrated that the appropriate solutions are

$$c_2(t) = \frac{-i\gamma/\hbar}{\sqrt{\gamma^2/\hbar^2 + (\omega - \omega_{21})^2/4}} \exp[-i(\omega - \omega_{21})t/2] \\ \times \sin\left(\sqrt{\gamma^2/\hbar^2 + (\omega - \omega_{21})^2/4} t\right), \quad (6.164)$$

$$c_1(t) = \exp[i(\omega - \omega_{21})t/2] \cos\left(\sqrt{\gamma^2/\hbar^2 + (\omega - \omega_{21})^2/4} t\right) \\ - \frac{i(\omega - \omega_{21})/2}{\sqrt{\gamma^2/\hbar^2 + (\omega - \omega_{21})^2/4}} \exp[i(\omega - \omega_{21})t/2] \\ \times \sin\left(\sqrt{\gamma^2/\hbar^2 + (\omega - \omega_{21})^2/4} t\right). \quad (6.165)$$

Now, the probability of finding the system in state 1 at time  $t$  is simply  $P_1(t) = |c_1|^2$ . Likewise, the probability of finding the system in state 2 at time  $t$  is  $P_2(t) = |c_2|^2$ . It follows that

$$P_2(t) = \frac{\gamma^2/\hbar^2}{\gamma^2/\hbar^2 + (\omega - \omega_{21})^2/4}$$

$$\times \sin^2\left(\sqrt{\gamma^2/\hbar^2 + (\omega - \omega_{21})^2/4} t\right), \quad (6.166)$$

$$P_1(t) = 1 - P_2(t). \quad (6.167)$$

This result is known as *Rabi's formula*.

Equation (6.166) exhibits all the features of a classic *resonance*. At resonance, when the oscillation frequency of the perturbation,  $\omega$ , matches the frequency  $\omega_{21}$ , we find that

$$P_1(t) = \cos^2(\gamma t/\hbar), \quad (6.168)$$

$$P_2(t) = \sin^2(\gamma t/\hbar). \quad (6.169)$$

According to the above result, the system starts off at  $t = 0$  in state 1. After a time interval  $\pi \hbar/2\gamma$  it is certain to be in state 2. After a further time interval  $\pi \hbar/2\gamma$  it is certain to be in state 1, and so on. Thus, the system periodically flip-flops between states 1 and 2 under the influence of the time-dependent perturbation. This implies that the system alternatively absorbs and emits energy from the source of the perturbation.

The absorption-emission cycle also take place away from the resonance, when  $\omega \neq \omega_{21}$ . However, the amplitude of oscillation of the coefficient  $c_2$  is reduced. This means that the maximum value of  $P_2(t)$  is no longer unity, nor is the minimum value of  $P_1(t)$  zero. In fact, if we plot the maximum value of  $P_2(t)$  as a function of the applied frequency,  $\omega$ , we obtain a resonance curve whose maximum (unity) lies at the resonance, and whose full-width half-maximum (in frequency) is  $4\gamma/\hbar$ . Thus, if the applied frequency differs from the resonant frequency by substantially more than  $2\gamma/\hbar$  then the probability of the system jumping from state 1 to state 2 is very small. In other words, the time-dependent perturbation is only effective at causing transitions between states 1 and 2 if its frequency of oscillation lies in the approximate range  $\omega_{21} \pm 2\gamma/\hbar$ . Clearly, the weaker the perturbation (*i.e.*, the smaller  $\gamma$  becomes), the narrower the resonance.

## 6.11 Spin magnetic resonance

Consider a spin one-half system (e.g., a bound electron) placed in a uniform  $z$ -directed magnetic field, and then subjected to a small time-dependent magnetic field rotating in the  $x$ - $y$  plane. Thus,

$$\mathbf{B} = B_0 \hat{\mathbf{z}} + B_1(\cos \omega t \hat{\mathbf{x}} + \sin \omega t \hat{\mathbf{y}}), \quad (6.170)$$

where  $B_0$  and  $B_1$  are constants, with  $B_1 \ll B_0$ . The rotating magnetic field usually represents the magnetic component of an electromagnetic wave propagating along the  $z$ -axis. In this system, the electric component of the wave has no effect. The Hamiltonian is written

$$H = -\boldsymbol{\mu} \cdot \mathbf{B} = H_0 + H_1, \quad (6.171)$$

where

$$H_0 = \frac{e B_0}{m_e} S_z, \quad (6.172)$$

and

$$H_1 = \frac{e B_1}{m_e} (\cos \omega t S_x + \sin \omega t S_y). \quad (6.173)$$

The eigenstates of the unperturbed Hamiltonian are the ‘spin up’ and ‘spin down’ states, denoted  $|+\rangle$  and  $|-\rangle$ , respectively. Thus,

$$H_0 |\pm\rangle = \pm \frac{e \hbar B_0}{2 m_e} |\pm\rangle. \quad (6.174)$$

The time-dependent Hamiltonian can be written

$$H_1 = \frac{e B_1}{2 m_e} [\exp(i \omega t) S^- + \exp(-i \omega t) S^+], \quad (6.175)$$

where  $S^+$  and  $S^-$  are the conventional raising and lowering operators for the spin angular momentum. It follows that

$$\langle + | H_1 | + \rangle = \langle - | H_1 | - \rangle = 0, \quad (6.176)$$

and

$$\langle - | H_1 | + \rangle = \langle + | H_1 | - \rangle^* = \frac{e \hbar B_1}{2 m_e} \exp(i \omega t). \quad (6.177)$$

It can be seen that this system is exactly the same as the two-state system discussed in the previous section, provided that we make the identifications

$$|1\rangle \rightarrow |-\rangle, \quad (6.178)$$

$$|2\rangle \rightarrow |+\rangle, \quad (6.179)$$

$$\omega_{21} \rightarrow \frac{e B_0}{m_e}, \quad (6.180)$$

$$\gamma \rightarrow \frac{e \hbar B_1}{2 m_e}. \quad (6.181)$$

The resonant frequency,  $\omega_{21}$ , is simply the spin precession frequency for an electron in a uniform magnetic field of strength  $B_0$ . In the absence of the perturbation, the expectation values of  $S_x$  and  $S_y$  oscillate because of the spin precession, but the expectation value of  $S_z$  remains invariant. If we now apply a magnetic perturbation rotating at the resonant frequency then, according to the analysis of the previous section, the system undergoes a succession of spin-flops,  $|+\rangle \rightleftharpoons |-\rangle$ , in addition to the spin precession. We also know that if the oscillation frequency of the applied field is very different from the resonant frequency then there is virtually zero probability of the field triggering a spin-flop. The width of the resonance (in frequency) is determined by the strength of the oscillating magnetic perturbation. Experimentalists are able to measure the magnetic moments of electrons, and other spin one-half particles, to a high degree of accuracy by placing the particles in a magnetic field, and subjecting them to an oscillating magnetic field whose frequency is gradually scanned. By determining the resonant frequency (*i.e.*, the frequency at which the particles absorb energy from the oscillating field), it is possible to calculate the magnetic moment.

## 6.12 The Dyson series

Let us now try to find approximate solutions of Eq. (6.154) for a general system. It is convenient to work in terms of the time evolution operator,  $U(t_0, t)$ , which is defined

$$|A, t_0, t\rangle = U(t_0, t) |A\rangle. \quad (6.182)$$

Here,  $|A, t_0, t\rangle$  is the state ket of the system at time  $t$ , given that the state ket at the initial time  $t_0$  is  $|A\rangle$ . It is easily seen that the time evolution operator satisfies the differential equation

$$i \hbar \frac{\partial U(t_0, t)}{\partial t} = (H_0 + H_1) U(t_0, t), \quad (6.183)$$

subject to the boundary condition

$$U(t_0, t_0) = 1. \quad (6.184)$$

In the absence of the external perturbation, the time evolution operator reduces to

$$U(t_0, t) = \exp[-i H_0(t - t_0)/\hbar]. \quad (6.185)$$

Let us switch on the perturbation and look for a solution of the form

$$U(t_0, t) = \exp[-i H_0(t - t_0)/\hbar] U_I(t_0, t). \quad (6.186)$$

It is readily demonstrated that  $U_I$  satisfies the differential equation

$$i \hbar \frac{\partial U_I(t_0, t)}{\partial t} = H_I(t_0, t) U_I(t_0, t), \quad (6.187)$$

where

$$H_I(t_0, t) = \exp[+i H_0(t - t_0)/\hbar] H_1 \exp[-i H_0(t - t_0)/\hbar], \quad (6.188)$$

subject to the boundary condition

$$U_I(t_0, t_0) = 1. \quad (6.189)$$

Note that  $U_I$  specifies that component of the time evolution operator which is due to the time-dependent perturbation. Thus, we would expect  $U_I$  to contain all of the information regarding transitions between different eigenstates of  $H_0$  caused by the perturbation.

Suppose that the system starts off at time  $t_0$  in the eigenstate  $|i\rangle$  of the unperturbed Hamiltonian. The subsequent evolution of the state ket is given by Eq. (6.149),

$$|i, t_0, t\rangle = \sum_m c_m(t) \exp[-i E_m(t - t_0)/\hbar] |m\rangle. \quad (6.190)$$

However, we also have

$$|i, t_0, t\rangle = \exp[-i H_0(t - t_0)/\hbar] U_I(t_0, t) |i\rangle. \quad (6.191)$$

It follows that

$$c_n(t) = \langle n | U_I(t_0, t) | i \rangle, \quad (6.192)$$

where use has been made of  $\langle n | m \rangle = \delta_{nm}$ . Thus, the probability that the system is found in state  $|n\rangle$  at time  $t$ , given that it is definitely in state  $|i\rangle$  at time  $t_0$ , is simply

$$P_{i \rightarrow n}(t_0, t) = |\langle n | U_I(t_0, t) | i \rangle|^2. \quad (6.193)$$

This quantity is usually termed the *transition probability* between states  $|i\rangle$  and  $|n\rangle$ .

Note that the differential equation (6.187), plus the boundary condition (6.189), are equivalent to the following integral equation,

$$U_I(t_0, t) = 1 - \frac{i}{\hbar} \int_{t_0}^t H_I(t_0, t') U_I(t_0, t') dt'. \quad (6.194)$$

We can obtain an approximate solution to this equation by iteration:

$$\begin{aligned} U_I(t_0, t) &\simeq 1 - \frac{i}{\hbar} \int_{t_0}^t H_I(t_0, t') \left[ 1 - \frac{i}{\hbar} \int_{t_0}^{t'} H_I(t_0, t'') U_I(t_0, t'') \right] dt' \\ &\simeq 1 - \frac{i}{\hbar} \int_{t_0}^t H_I(t_0, t') dt' \\ &\quad + \left( \frac{-i}{\hbar} \right)^2 \int_{t_0}^t dt' \int_{t_0}^{t'} H_I(t_0, t') H_I(t_0, t'') dt'' + \dots \end{aligned} \quad (6.195)$$

This expansion is known as the *Dyson series*. Let

$$c_n = c_n^{(0)} + c_n^{(1)} + c_n^{(2)} + \dots, \quad (6.196)$$

where the superscript <sup>(1)</sup> refers to a first-order term in the expansion, *etc.* It follows from Eqs. (6.192) and (6.195) that

$$c_n^{(0)}(t) = \delta_{in}, \quad (6.197)$$

$$c_n^{(1)}(t) = -\frac{i}{\hbar} \int_{t_0}^t \langle n | H_I(t_0, t') | i \rangle dt', \quad (6.198)$$

$$c_n^{(2)}(t) = \left(\frac{-i}{\hbar}\right)^2 \int_{t_0}^t dt' \int_{t_0}^{t'} \langle n | H_I(t_0, t') H_I(t_0, t'') | i \rangle dt''. \quad (6.199)$$

These expressions simplify to

$$c_n^{(0)}(t) = \delta_{in}, \quad (6.200)$$

$$c_n^{(1)}(t) = -\frac{i}{\hbar} \int_{t_0}^t \exp[i\omega_{ni}(t' - t_0)] H_{ni}(t') dt', \quad (6.201)$$

$$c_n^{(2)}(t) = \left(\frac{-i}{\hbar}\right)^2 \sum_m \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' \exp[i\omega_{nm}(t' - t_0)] \\ \times H_{nm}(t') \exp[i\omega_{mi}(t'' - t_0)] H_{mi}(t''), \quad (6.202)$$

where

$$\omega_{nm} = \frac{E_n - E_m}{\hbar}, \quad (6.203)$$

and

$$H_{nm}(t) = \langle n | H_I(t) | m \rangle. \quad (6.204)$$

The transition probability between states  $i$  and  $n$  is simply

$$P_{i \rightarrow n}(t_0, t) = |c_n^{(0)} + c_n^{(1)} + c_n^{(2)} + \dots|^2. \quad (6.205)$$

According to the above analysis, there is no chance of a transition between states  $|i\rangle$  and  $|n\rangle$  ( $i \neq n$ ) to zeroth-order (*i.e.*, in the absence of the perturbation). To first-order, the transition probability is proportional to the time integral of the matrix element  $\langle n | H_I | i \rangle$ , weighted by some oscillatory phase-factor. Thus, if the matrix element is zero, then there is no chance of a first-order transition between states  $|i\rangle$  and  $|n\rangle$ . However, to second-order, a transition between states  $|i\rangle$  and  $|n\rangle$  is possible even when the matrix element  $\langle n | H_I | i \rangle$  is zero.

### 6.13 Constant perturbations

Consider a constant perturbation which is suddenly switched on at time  $t = 0$ :

$$\begin{aligned} H_1(t) &= 0 && \text{for } t < 0 \\ H_1(t) &= H_1 && \text{for } t \geq 0, \end{aligned} \quad (6.206)$$

where  $H_1$  is time-independent, but is generally a function of the position, momentum, and spin operators. Suppose that the system is definitely in state  $|i\rangle$  at time  $t = 0$ . According to Eqs. (6.200)–(6.202) (with  $t_0 = 0$ ),

$$c_n^{(0)}(t) = \delta_{in}, \quad (6.207)$$

$$\begin{aligned} c_n^{(1)}(t) &= -\frac{i}{\hbar} H_{ni} \int_0^t \exp[i\omega_{ni}(t' - t)] dt' \\ &= \frac{H_{ni}}{E_n - E_i} [1 - \exp(i\omega_{ni}t)], \end{aligned} \quad (6.208)$$

giving

$$P_{i \rightarrow n}(t) \simeq |c_n^{(1)}|^2 = \frac{4|H_{ni}|^2}{|E_n - E_i|^2} \sin^2 \left[ \frac{(E_n - E_i)t}{2\hbar} \right], \quad (6.209)$$

for  $i \neq n$ . The transition probability between states  $|i\rangle$  and  $|n\rangle$  can be written

$$P_{i \rightarrow n}(t) = \frac{|H_{ni}|^2 t^2}{\hbar^2} \operatorname{sinc}^2 \left[ \frac{(E_n - E_i)t}{2\hbar} \right], \quad (6.210)$$

where

$$\operatorname{sinc}(x) = \frac{\sin x}{x}. \quad (6.211)$$

The sinc function is highly oscillatory, and decays like  $1/|x|$  at large  $|x|$ . It is a good approximation to say that  $\operatorname{sinc}(x)$  is small except when  $|x| \lesssim \pi$ . It follows that the transition probability,  $P_{i \rightarrow n}$ , is small except when

$$|E_n - E_i| \lesssim \frac{2\pi\hbar}{t}. \quad (6.212)$$

Note that in the limit  $t \rightarrow \infty$  only those transitions which conserve energy (*i.e.*,  $E_n = E_i$ ) have an appreciable probability of occurrence. At finite  $t$ , it is possible to have transitions which do not exactly conserve energy, provided that

$$\Delta E \Delta t \lesssim \hbar, \quad (6.213)$$

where  $\Delta E = |E_n - E_i|$  is change in energy of the system associated with the transition, and  $\Delta t = t$  is the time elapsed since the perturbation was switched on. Clearly, this result is just a manifestation of the well-known uncertainty relation for energy and time. This uncertainty relation is fundamentally different to the position-momentum uncertainty relation, since in non-relativistic quantum mechanics position and momentum are operators, whereas time is merely a parameter.

The probability of a transition which conserves energy (*i.e.*,  $E_n = E_i$ ) is

$$P_{i \rightarrow n}(t) = \frac{|H_{in}|^2 t^2}{\hbar^2}, \quad (6.214)$$

where use has been made of  $\text{sinc}(0) = 1$ . Note that this probability grows *quadratically* with time. This result is somewhat surprising, since it implies that the probability of a transition occurring in a fixed time interval,  $t$  to  $t + dt$ , grows linearly with  $t$ , despite the fact that  $H_1$  is constant for  $t > 0$ . In practice, there is usually a group of final states, all possessing nearly the same energy as the energy of the initial state  $|i\rangle$ . It is helpful to define the density of states,  $\rho(E)$ , where the number of final states lying in the energy range  $E$  to  $E + dE$  is given by  $\rho(E) dE$ . Thus, the probability of a transition from the initial state  $i$  to any of the continuum of possible final states is

$$P_{i \rightarrow}(t) = \int P_{i \rightarrow n}(t) \rho(E_n) dE_n, \quad (6.215)$$

giving

$$P_{i \rightarrow}(t) = \frac{2t}{\hbar} \int |H_{ni}|^2 \rho(E_n) \text{sinc}^2(x) dx, \quad (6.216)$$

where

$$x = (E_n - E_i) t / 2 \hbar, \quad (6.217)$$

and use has been made of Eq. (6.210). We know that in the limit  $t \rightarrow \infty$  the function  $\text{sinc}(x)$  is only non-zero in an infinitesimally narrow range of final energies centred on  $E_n = E_i$ . It follows that, in this limit, we can take  $\rho(E_n)$  and  $|H_{ni}|^2$  out of the integral in the above formula to obtain

$$P_{i \rightarrow [n]}(t) = \frac{2\pi}{\hbar} \overline{|H_{ni}|^2} \rho(E_n) t \Big|_{E_n \simeq E_i}, \quad (6.218)$$

where  $P_{i \rightarrow [n]}$  denotes the transition probability between the initial state  $|i\rangle$  and all final states  $|n\rangle$  which have approximately the same energy as the initial state. Here,  $\overline{|H_{ni}|^2}$  is the average of  $|H_{ni}|^2$  over all final states with approximately the same energy as the initial state. In deriving the above formula, we have made use of the result

$$\int_{-\infty}^{\infty} \text{sinc}^2(x) dx = \pi. \quad (6.219)$$

Note that the transition probability,  $P_{i \rightarrow [n]}$ , is now proportional to  $t$ , instead of  $t^2$ .

It is convenient to define the *transition rate*, which is simply the transition probability per unit time. Thus,

$$w_{i \rightarrow [n]} = \frac{dP_{i \rightarrow [n]}}{dt}, \quad (6.220)$$

giving

$$w_{i \rightarrow [n]} = \frac{2\pi}{\hbar} \overline{|H_{ni}|^2} \rho(E_n) \Big|_{E_n \simeq E_i}. \quad (6.221)$$

This appealingly simple result is known as *Fermi's golden rule*. Note that the transition rate is constant in time (for  $t > 0$ ): *i.e.*, the probability of a transition occurring in the time interval  $t$  to  $t + dt$  is independent of  $t$  for fixed  $dt$ . Fermi's golden rule is sometimes written

$$w_{i \rightarrow n} = \frac{2\pi}{\hbar} |H_{ni}|^2 \delta(E_n - E), \quad (6.222)$$

where it is understood that this formula must be integrated with  $\int \rho(E_n) dE_n$  to obtain the actual transition rate.

Let us now calculate the second-order term in the Dyson series, using the constant perturbation (6.206). From Eq. (6.202) we find that

$$\begin{aligned} c_n^{(2)}(t) &= \left(\frac{-i}{\hbar}\right)^2 \sum_m H_{nm} H_{mi} \int_0^t dt' \exp(i\omega_{nm} t') \int_0^{t'} dt'' \exp(i\omega_{mi} t) \\ &= \frac{i}{\hbar} \sum_m \frac{H_{nm} H_{mi}}{E_m - E_i} \int_0^t [\exp(i\omega_{ni} t') - \exp(i\omega_{nm} t')] dt' \\ &= \frac{it}{\hbar} \sum_m \frac{H_{nm} H_{mi}}{E_m - E_i} [\exp(i\omega_{ni} t/2) \text{sinc}(\omega_{ni} t/2)] \end{aligned}$$

$$- \exp(i \omega_{nm} t/2) \operatorname{sinc}(\omega_{nm} t/2)]. \quad (6.223)$$

Thus,

$$\begin{aligned} c_n(t) = c_n^{(1)} + c_n^{(2)} &= \frac{it}{\hbar} \exp(i \omega_{ni} t/2) \left[ \left( H_{ni} + \sum_m \frac{H_{nm} H_{mi}}{E_m - E_i} \right) \operatorname{sinc}(\omega_{ni} t/2) \right. \\ &\quad \left. - \sum_m \frac{H_{nm} H_{mi}}{E_m - E_i} \exp(i \omega_{im} t/2) \operatorname{sinc}(\omega_{nm} t/2) \right], \quad (6.224) \end{aligned}$$

where use has been made of Eq. (6.208). It follows, by analogy with the previous analysis, that

$$w_{i \rightarrow [n]} = \frac{2\pi}{\hbar} \left| H_{ni} + \sum_m \frac{H_{nm} H_{mi}}{E_m - E_i} \right|^2 \rho(E_n) \Big|_{E_n \simeq E_i}, \quad (6.225)$$

where the transition rate is calculated for all final states,  $|n\rangle$ , with approximately the same energy as the initial state,  $|i\rangle$ , and for intermediate states,  $|m\rangle$  whose energies differ from that of the initial state. The fact that  $E_m \neq E_i$  causes the last term on the right-hand side of Eq. (6.224) to average to zero (due to the oscillatory phase-factor) during the evaluation of the transition probability.

According to Eq. (6.225), a second-order transition takes place in two steps. First, the system makes a non-energy-conserving transition to some intermediate state  $|m\rangle$ . Subsequently, the system makes another non-energy-conserving transition to the final state  $|n\rangle$ . The net transition, from  $|i\rangle$  to  $|n\rangle$ , conserves energy. The non-energy-conserving transitions are generally termed *virtual transitions*, whereas the energy conserving first-order transition is termed a *real transition*. The above formula clearly breaks down if  $H_{nm} H_{mi} \neq 0$  when  $E_m = E_i$ . This problem can be avoided by gradually turning on the perturbation: *i.e.*,  $H_1 \rightarrow \exp(\eta t) H_1$  (where  $\eta$  is very small). The net result is to change the energy denominator in Eq. (6.225) from  $E_i - E_m$  to  $E_i - E_m + i \hbar \eta$ .

## 6.14 Harmonic perturbations

Consider a perturbation which oscillates sinusoidally in time. This is usually called a *harmonic perturbation*. Thus,

$$H_1(t) = V \exp(i\omega t) + V^\dagger \exp(-i\omega t), \quad (6.226)$$

where  $V$  is, in general, a function of position, momentum, and spin operators.

Let us initiate the system in the eigenstate  $|i\rangle$  of the unperturbed Hamiltonian,  $H_0$ , and switch on the harmonic perturbation at  $t = 0$ . It follows from Eq. (6.201) that

$$\begin{aligned} c_n^{(1)} &= \frac{-i}{\hbar} \int_0^t [V_{ni} \exp(i\omega t') + V_{ni}^\dagger \exp(-i\omega t')] \exp(i\omega_{ni}t') dt', \quad (6.227) \\ &= \frac{1}{\hbar} \left( \frac{1 - \exp[i(\omega_{ni} + \omega)t]}{\omega_{ni} + \omega} V_{ni} + \frac{1 - \exp[i(\omega_{ni} - \omega)t]}{\omega_{ni} - \omega} V_{ni}^\dagger \right), \end{aligned}$$

where

$$V_{ni} = \langle n|V|i\rangle, \quad (6.228)$$

$$V_{ni}^\dagger = \langle n|V^\dagger|i\rangle = \langle i|V|n\rangle^*. \quad (6.229)$$

This formula is analogous to Eq. (6.208), provided that

$$\omega_{ni} = \frac{E_n - E_i}{\hbar} \rightarrow \omega_{ni} \pm \omega. \quad (6.230)$$

Thus, it follows from the previous analysis that the transition probability  $P_{i \rightarrow n}(t) = |c_n^{(1)}|^2$  is only appreciable in the limit  $t \rightarrow \infty$  if

$$\omega_{ni} + \omega \simeq 0 \quad \text{or} \quad E_n \simeq E_i - \hbar\omega, \quad (6.231)$$

$$\omega_{ni} - \omega \simeq 0 \quad \text{or} \quad E_n \simeq E_i + \hbar\omega. \quad (6.232)$$

Clearly, (6.231) corresponds to the first term on the right-hand side of Eq. (6.227), and (6.232) corresponds to the second term. The former term describes a process by which the system gives up energy  $\hbar\omega$  to the perturbing field, whilst making a transition to a final state whose energy level is less than that of the initial state

by  $\hbar\omega$ . This process is known as *stimulated emission*. The latter term describes a process by which the system gains energy  $\hbar\omega$  from the perturbing field, whilst making a transition to a final state whose energy level exceeds that of the initial state by  $\hbar\omega$ . This process is known as *absorption*. In both cases, the total energy (*i.e.*, that of the system *plus* the perturbing field) is conserved.

By analogy with Eq. (6.221),

$$w_{i \rightarrow [n]} = \frac{2\pi}{\hbar} \overline{|V_{ni}|^2} \rho(E_n) \Big|_{E_n = E_i - \hbar\omega}, \quad (6.233)$$

$$w_{i \rightarrow [n]} = \frac{2\pi}{\hbar} \overline{|V_{ni}^\dagger|^2} \rho(E_n) \Big|_{E_n = E_i + \hbar\omega}. \quad (6.234)$$

Equation (6.233) specifies the transition rate for stimulated emission, whereas Eq. (6.234) gives the transition rate for absorption. These equations are more usually written

$$w_{i \rightarrow n} = \frac{2\pi}{\hbar} |V_{ni}|^2 \delta(E_n - E_i + \hbar\omega), \quad (6.235)$$

$$w_{i \rightarrow n} = \frac{2\pi}{\hbar} |V_{ni}^\dagger|^2 \delta(E_n - E_i - \hbar\omega). \quad (6.236)$$

It is clear from Eqs. (6.228)–(6.229) that  $|V_{ni}^\dagger|^2 = |V_{ni}|^2$ . It follows from Eqs. (6.233)–(6.234) that

$$\frac{w_{i \rightarrow [n]}}{\rho(E_n)} = \frac{w_{n \rightarrow [i]}}{\rho(E_i)}. \quad (6.237)$$

In other words, the rate of stimulated emission, divided by the density of final states for stimulated emission, equals the rate of absorption, divided by the density of final states for absorption. This result, which expresses a fundamental symmetry between absorption and stimulated emission, is known as *detailed balancing*, and is very important in statistical mechanics.

## 6.15 Absorption and stimulated emission of radiation

Let us use some of the results of time-dependent perturbation theory to investigate the interaction of an atomic electron with classical (*i.e.*, non-quantized)

electromagnetic radiation.

The unperturbed Hamiltonian is

$$H_0 = \frac{p^2}{2 m_e} + V_0(\mathbf{r}). \quad (6.238)$$

The standard classical prescription for obtaining the Hamiltonian of a particle of charge  $q$  in the presence of an electromagnetic field is

$$\mathbf{p} \rightarrow \mathbf{p} + q \mathbf{A}, \quad (6.239)$$

$$H \rightarrow H - q \phi, \quad (6.240)$$

where  $\mathbf{A}(\mathbf{r})$  is the vector potential and  $\phi(\mathbf{r})$  is the scalar potential. Note that

$$\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t}, \quad (6.241)$$

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (6.242)$$

This prescription also works in quantum mechanics. Thus, the Hamiltonian of an atomic electron placed in an electromagnetic field is

$$H = \frac{(\mathbf{p} - e\mathbf{A})^2}{2 m_e} + e\phi + V_0(\mathbf{r}), \quad (6.243)$$

where  $\mathbf{A}$  and  $\phi$  are functions of the position operators. The above equation can be written

$$H = \frac{(p^2 - e\mathbf{A}\cdot\mathbf{p} - e\mathbf{p}\cdot\mathbf{A} + e^2A^2)}{2 m_e} + e\phi + V_0(\mathbf{r}). \quad (6.244)$$

Now,

$$\mathbf{p}\cdot\mathbf{A} = \mathbf{A}\cdot\mathbf{p}, \quad (6.245)$$

provided that we adopt the gauge  $\nabla\cdot\mathbf{A} = 0$ . Hence,

$$H = \frac{p^2}{2 m_e} - \frac{e\mathbf{A}\cdot\mathbf{p}}{m_e} + \frac{e^2A^2}{2 m_e} + e\phi + V_0(\mathbf{r}). \quad (6.246)$$

Suppose that the perturbation corresponds to a monochromatic plane-wave, for which

$$\phi = 0, \quad (6.247)$$

$$\mathbf{A} = 2A_0 \boldsymbol{\epsilon} \cos\left(\frac{\omega}{c} \mathbf{n}\cdot\mathbf{r} - \omega t\right), \quad (6.248)$$

where  $\boldsymbol{\epsilon}$  and  $\mathbf{n}$  are unit vectors which specify the direction of polarization and the direction of propagation, respectively. Note that  $\boldsymbol{\epsilon} \cdot \mathbf{n} = 0$ . The Hamiltonian becomes

$$H = H_0 + H_1(t), \quad (6.249)$$

with

$$H_0 = \frac{p^2}{2m_e} + V(r), \quad (6.250)$$

and

$$H_1 \simeq -\frac{e\mathbf{A} \cdot \mathbf{p}}{m_e}, \quad (6.251)$$

where the  $A^2$  term, which is second order in  $A_0$ , has been neglected.

The perturbing Hamiltonian can be written

$$H_1 = -\frac{eA_0 \boldsymbol{\epsilon} \cdot \mathbf{p}}{m_e} (\exp[i(\omega/c) \mathbf{n} \cdot \mathbf{r} - i\omega t] + \exp[-i(\omega/c) \mathbf{n} \cdot \mathbf{r} + i\omega t]). \quad (6.252)$$

This has the same form as Eq. (6.226), provided that

$$V = -\frac{eA_0 \boldsymbol{\epsilon} \cdot \mathbf{p}}{m_e} \exp[-i(\omega/c) \mathbf{n} \cdot \mathbf{r}] \quad (6.253)$$

It is clear, by analogy with the previous analysis, that the first term on the right-hand side of Eq. (6.252) describes the absorption of a photon of energy  $\hbar\omega$ , whereas the second term describes the stimulated emission of a photon of energy  $\hbar\omega$ . It follows from Eq. (6.236) that the rate of absorption is

$$w_{i \rightarrow n} = \frac{2\pi}{\hbar} \frac{e^2}{m_e^2} |A_0|^2 |\langle n | \exp[i(\omega/c) \mathbf{n} \cdot \mathbf{r}] \boldsymbol{\epsilon} \cdot \mathbf{p} | i \rangle|^2 \delta(E_n - E_i - \hbar\omega). \quad (6.254)$$

The absorption cross-section is defined as the ratio of the power absorbed by the atom to the incident power per unit area in the electromagnetic field. Now the energy density of an electromagnetic field is

$$u = \frac{1}{2} \left( \frac{\epsilon_0 E_0^2}{2} + \frac{B_0^2}{2\mu_0} \right), \quad (6.255)$$

where  $E_0$  and  $B_0 = E_0/c = 2A_0\omega/c$  are the peak electric and magnetic field-strengths, respectively. The incident power per unit area of the electromagnetic field is

$$cU = 2\epsilon_0 c \omega^2 |A_0|^2. \quad (6.256)$$

Now,

$$\sigma_{\text{abs}} = \frac{\hbar \omega w_{i \rightarrow n}}{cU}, \quad (6.257)$$

so

$$\sigma_{\text{abs}} = \frac{\pi e^2}{\epsilon_0 m_e^2 \omega c} |\langle n | \exp[i(\omega/c) \mathbf{n} \cdot \mathbf{r}] \boldsymbol{\epsilon} \cdot \mathbf{p} | i \rangle|^2 \delta(E_n - E_i - \hbar\omega). \quad (6.258)$$

## 6.16 The electric dipole approximation

In general, the wave-length of the type of electromagnetic radiation which induces, or is emitted during, transitions between different atomic energy levels is much larger than the typical size of a light atom. Thus,

$$\exp[i(\omega/c) \mathbf{n} \cdot \mathbf{r}] = 1 + i \frac{\omega}{c} \mathbf{n} \cdot \mathbf{r} + \dots, \quad (6.259)$$

can be approximated by its first term, unity (remember that  $\omega/c = 2\pi/\lambda$ ). This approximation is known as the *electric dipole approximation*. It follows that

$$\langle n | \exp[i(\omega/c) \mathbf{n} \cdot \mathbf{r}] \boldsymbol{\epsilon} \cdot \mathbf{p} | i \rangle \simeq \boldsymbol{\epsilon} \cdot \langle n | \mathbf{p} | i \rangle. \quad (6.260)$$

It is readily demonstrated that

$$[\mathbf{r}, H_0] = \frac{i \hbar \mathbf{p}}{m_e}, \quad (6.261)$$

so

$$\langle n | \mathbf{p} | i \rangle = -i \frac{m_e}{\hbar} \langle n | [\mathbf{r}, H_0] | i \rangle = i m_e \omega_{ni} \langle n | \mathbf{r} | i \rangle. \quad (6.262)$$

Using Eq. (6.258), we obtain

$$\sigma_{\text{abs}} = 4\pi^2 \alpha \omega_{ni} |\langle n | \boldsymbol{\epsilon} \cdot \mathbf{r} | i \rangle|^2 \delta(\omega - \omega_{ni}), \quad (6.263)$$

where  $\alpha = e^2/(2\epsilon_0 \hbar c) = 1/137$  is the fine structure constant. It is clear that if the absorption cross-section is regarded as a function of the applied frequency,  $\omega$ , then it exhibits a sharp maximum at  $\omega = \omega_{ni} = (E_n - E_i)/\hbar$ .

Suppose that the radiation is polarized in the  $z$ -direction, so that  $\epsilon = \hat{z}$ . We have already seen, from Sect. 6.4, that  $\langle n|z|i \rangle = 0$  unless the initial and final states satisfy

$$\Delta l = \pm 1, \quad (6.264)$$

$$\Delta m = 0. \quad (6.265)$$

Here,  $l$  is the quantum number describing the total orbital angular momentum of the electron, and  $m$  is the quantum number describing the projection of the orbital angular momentum along the  $z$ -axis. It is easily demonstrated that  $\langle n|x|i \rangle$  and  $\langle n|y|i \rangle$  are only non-zero if

$$\Delta l = \pm 1, \quad (6.266)$$

$$\Delta m = \pm 1. \quad (6.267)$$

Thus, for generally directed radiation  $\langle n|\epsilon \cdot \mathbf{r}|i \rangle$  is only non-zero if

$$\Delta l = \pm 1, \quad (6.268)$$

$$\Delta m = 0, \pm 1. \quad (6.269)$$

These are termed the *selection rules* for electric dipole transitions. It is clear, for instance, that the electric dipole approximation allows a transition from a  $2p$  state to a  $1s$  state, but disallows a transition from a  $2s$  to a  $1s$  state. The latter transition is called a *forbidden transition*.

Forbidden transitions are not strictly forbidden. Instead, they take place at a far lower rate than transitions which are allowed according to the electric dipole approximation. After electric dipole transitions, the next most likely type of transition is a *magnetic dipole transition*, which is due to the interaction between the electron spin and the oscillating magnetic field of the incident electromagnetic radiation. Magnetic dipole transitions are typically about  $10^5$  times more unlikely than similar electric dipole transitions. The first-order term in Eq. (6.259) yields

so-called *electric quadrupole transitions*. These are typically about  $10^8$  times more unlikely than electric dipole transitions. Magnetic dipole and electric quadrupole transitions satisfy different selection rules than electric dipole transitions: for instance, the selection rules for electric quadrupole transitions are  $\Delta l = 0, \pm 2$ . Thus, transitions which are forbidden as electric dipole transitions may well be allowed as magnetic dipole or electric quadrupole transitions.

Integrating Eq. (6.263) over all possible frequencies of the incident radiation yields

$$\int \sigma_{\text{abs}}(\omega) d\omega = \sum_n 4\pi^2 \alpha \omega_{ni} |\langle n | \mathbf{e} \cdot \mathbf{r} | i \rangle|^2. \quad (6.270)$$

Suppose, for the sake of definiteness, that the incident radiation is polarized in the  $x$ -direction. It is easily demonstrated that

$$[x, [x, H_0]] = -\frac{\hbar^2}{m_e}. \quad (6.271)$$

Thus,

$$\langle i | [x, [x, H_0]] | i \rangle = \langle i | x^2 H_0 + H_0 x^2 - 2x H_0 x | i \rangle = -\frac{\hbar^2}{m_e}, \quad (6.272)$$

giving

$$2 \sum_n (\langle i | x | n \rangle E_i \langle n | x | i \rangle - \langle i | x | n \rangle E_n \langle n | x | i \rangle) = -\frac{\hbar^2}{m_e}. \quad (6.273)$$

It follows that

$$\frac{2 m_e}{\hbar} \sum_n \omega_{ni} |\langle n | x | i \rangle|^2 = 1. \quad (6.274)$$

This is known as the *Thomas-Reiche-Kuhn* sum rule. According to this rule, Eq. (6.270) reduces to

$$\int \sigma_{\text{abs}}(\omega) d\omega = \frac{2\pi^2 \alpha \hbar}{m_e} = \frac{\pi e^2}{2 \epsilon_0 m_e c}. \quad (6.275)$$

Note that  $\hbar$  has dropped out of the final result. In fact, the above formula is exactly the same as that obtained classically by treating the electron as an oscillator.

## 6.17 Energy-shifts and decay-widths

We have examined how a state  $|n\rangle$ , other than the initial state  $|i\rangle$ , becomes populated as a result of some time-dependent perturbation applied to the system. Let us now consider how the initial state becomes depopulated.

It is convenient to gradually turn on the perturbation from zero at  $t = -\infty$ . Thus,

$$H_1(t) = \exp(\eta t) H_1, \quad (6.276)$$

where  $\eta$  is small and positive, and  $H_1$  is a constant.

In the remote past,  $t \rightarrow -\infty$ , the system is assumed to be in the initial state  $|i\rangle$ . Thus,  $c_i(t \rightarrow -\infty) = 1$ , and  $c_{n \neq i}(t \rightarrow -\infty) = 0$ . Basically, we want to calculate the time evolution of the coefficient  $c_i(t)$ . First, however, let us check that our previous Fermi golden rule result still applies when the perturbing potential is turned on slowly, instead of very suddenly. For  $c_{n \neq i}(t)$  we have from Eqs. (6.200)–(6.201) that

$$c_n^{(0)}(t) = 0, \quad (6.277)$$

$$\begin{aligned} c_n^{(1)}(t) &= -\frac{i}{\hbar} H_{ni} \int_{-\infty}^t \exp[(\eta + i\omega_{ni})t'] dt' \\ &= -\frac{i}{\hbar} H_{ni} \frac{\exp[(\eta + i\omega_{ni})t]}{\eta + i\omega_{ni}}, \end{aligned} \quad (6.278)$$

where  $H_{ni} = \langle n|H_1|i\rangle$ . It follows that, to first-order, the transition probability from state  $|i\rangle$  to state  $|n\rangle$  is

$$P_{i \rightarrow n}(t) = |c_n^{(1)}|^2 = \frac{|H_{ni}|^2 \exp(2\eta t)}{\hbar^2 (\eta^2 + \omega_{ni}^2)}. \quad (6.279)$$

The transition rate is given by

$$w_{i \rightarrow n}(t) = \frac{dP_{i \rightarrow n}}{dt} = \frac{2|H_{ni}|^2 \eta \exp(2\eta t)}{\hbar^2 (\eta^2 + \omega_{ni}^2)}. \quad (6.280)$$

Consider the limit  $\eta \rightarrow 0$ . In this limit,  $\exp(\eta t) \rightarrow 1$ , but

$$\lim_{\eta \rightarrow 0} \frac{\eta}{\eta^2 + \omega_{ni}^2} = \pi \delta(\omega_{ni}) = \pi \hbar \delta(E_n - E_i). \quad (6.281)$$

Thus, Eq. (6.280) yields the standard Fermi golden rule result

$$w_{i \rightarrow n} = \frac{2\pi}{\hbar} |H_{ni}|^2 \delta(E_n - E_i). \quad (6.282)$$

It is clear that the delta-function in the above formula actually represents a function which is highly peaked at some particular energy. The width of the peak is determined by how fast the perturbation is switched on.

Let us now calculate  $c_i(t)$  using Eqs. (6.200)–(6.202). We have

$$c_i^{(0)}(t) = 1, \quad (6.283)$$

$$c_i^{(1)}(t) = -\frac{i}{\hbar} H_{ii} \int_{-\infty}^t \exp(\eta t') dt' = -\frac{i}{\hbar} H_{ii} \frac{\exp(\eta t)}{\eta}, \quad (6.284)$$

$$\begin{aligned} c_i^{(2)}(t) &= \left(\frac{-i}{\hbar}\right)^2 \sum_m |H_{mi}|^2 \int_{-\infty}^t dt' \int_{-\infty}^{t'} dt'' \\ &\quad \times \exp[(\eta + i\omega_{im})t'] \exp[(\eta + i\omega_{mi})t''], \\ &= \left(\frac{-i}{\hbar}\right)^2 \sum_m |H_{mi}|^2 \frac{\exp(2\eta t)}{2\eta(\eta + i\omega_{mi})}. \end{aligned} \quad (6.285)$$

Thus, to second-order we have

$$\begin{aligned} c_i(t) &\simeq 1 + \left(\frac{-i}{\hbar}\right) H_{ii} \frac{\exp(\eta t)}{\eta} + \left(\frac{-i}{\hbar}\right)^2 |H_{ii}|^2 \frac{\exp(2\eta t)}{2\eta^2} \\ &\quad + \left(\frac{-i}{\hbar}\right) \sum_{m \neq i} \frac{|H_{mi}|^2 \exp(2\eta t)}{2\eta(E_i - E_m + i\hbar\eta)}. \end{aligned} \quad (6.286)$$

Let us now consider the ratio  $\dot{c}_i/c_i$ , where  $\dot{c}_i \equiv dc_i/dt$ . Using Eq. (6.286), we can evaluate this ratio in the limit  $\eta \rightarrow 0$ . We obtain

$$\begin{aligned} \frac{\dot{c}_i}{c_i} &\simeq \left[ \left(\frac{-i}{\hbar}\right) H_{ii} + \left(\frac{-i}{\hbar}\right)^2 \frac{|H_{ii}|^2}{\eta} + \left(\frac{-i}{\hbar}\right) \sum_{m \neq i} \frac{|H_{mi}|^2}{E_i - E_m + i\hbar\eta} \right] \\ &\quad / \left(1 - \frac{i}{\hbar} \frac{H_{ii}}{\eta}\right) \\ &\simeq \left(\frac{-i}{\hbar}\right) H_{ii} + \lim_{\eta \rightarrow 0} \left(\frac{-i}{\hbar}\right) \sum_{m \neq i} \frac{|H_{mi}|^2}{E_i - E_m + i\hbar\eta}. \end{aligned} \quad (6.287)$$

This result is formally correct to second-order in perturbed quantities. Note that the right-hand side of Eq. (6.287) is independent of time. We can write

$$\frac{\dot{c}_i}{c_i} = \left( \frac{-i}{\hbar} \right) \Delta_i, \quad (6.288)$$

where

$$\Delta_i = H_{ii} + \lim_{\eta \rightarrow 0} \sum_{m \neq i} \frac{|H_{mi}|^2}{E_i - E_m + i \hbar \eta} \quad (6.289)$$

is a constant. According to a well-known result in pure mathematics,

$$\lim_{\epsilon \rightarrow 0} \frac{1}{x + i \epsilon} = P \frac{1}{x} - i \pi \delta(x), \quad (6.290)$$

where  $\epsilon > 0$ , and  $P$  denotes the principle part. It follows that

$$\Delta_i = H_{ii} + P \sum_{m \neq i} \frac{|H_{mi}|^2}{E_i - E_m} - i \pi \sum_{m \neq i} |H_{mi}|^2 \delta(E_i - E_m). \quad (6.291)$$

It is convenient to normalize the solution of Eq. (6.288) so that  $c_i(0) = 1$ . Thus, we obtain

$$c_i(t) = \exp\left(\frac{-i \Delta_i t}{\hbar}\right). \quad (6.292)$$

According to Eq. (6.149), the time evolution of the initial state ket  $|i\rangle$  is given by

$$|i, t\rangle = \exp[-i(\Delta_i + E_i)t/\hbar] |i\rangle. \quad (6.293)$$

We can rewrite this result as

$$|i, t\rangle = \exp(-i[E_i + \text{Re}(\Delta_i)]t/\hbar) \exp[\text{Im}(\Delta_i)t/\hbar] |i\rangle. \quad (6.294)$$

It is clear that the real part of  $\Delta_i$  gives rise to a simple shift in energy of state  $|i\rangle$ , whereas the imaginary part of  $\Delta_i$  governs the growth or decay of this state. Thus,

$$|i, t\rangle = \exp[-i(E_i + \Delta E_i)t/\hbar] \exp(-\Gamma_i t/2\hbar) |i\rangle, \quad (6.295)$$

where

$$\Delta E_i = \text{Re}(\Delta_i) = H_{ii} + P \sum_{m \neq i} \frac{|H_{mi}|^2}{E_i - E_m}, \quad (6.296)$$

and

$$\frac{\Gamma_i}{\hbar} = -\frac{2 \operatorname{Im}(\Delta_i)}{\hbar} = \frac{2\pi}{\hbar} \sum_{m \neq i} |H_{mi}|^2 \delta(E_i - E_m). \quad (6.297)$$

Note that the energy-shift  $\Delta E_i$  is the same as that predicted by standard time-independent perturbation theory.

The probability of observing the system in state  $|i\rangle$  at time  $t > 0$ , given that it is definitely in state  $|i\rangle$  at time  $t = 0$ , is given by

$$P_{i \rightarrow i}(t) = |c_i|^2 = \exp(-\Gamma_i t/\hbar), \quad (6.298)$$

where

$$\frac{\Gamma_i}{\hbar} = \sum_{m \neq i} w_{i \rightarrow m}. \quad (6.299)$$

Here, use has been made of Eq. (6.222). Clearly, the rate of decay of the initial state is a simple function of the transition rates to the other states. Note that the system conserves probability up to second-order in perturbed quantities, since

$$|c_i|^2 + \sum_{m \neq i} |c_m|^2 \simeq (1 - \Gamma_i t/\hbar) + \sum_{m \neq i} w_{i \rightarrow m} t = 1. \quad (6.300)$$

The quantity  $\Delta_i$  is called the *decay-width* of state  $|i\rangle$ . It is closely related to the mean lifetime of this state,

$$\tau_i = \frac{\hbar}{\Gamma_i}, \quad (6.301)$$

where

$$P_{i \rightarrow i} = \exp(-t/\tau_i). \quad (6.302)$$

According to Eq. (6.294), the amplitude of state  $|i\rangle$  both oscillates and decays as time progresses. Clearly, state  $|i\rangle$  is not a stationary state in the presence of the time-dependent perturbation. However, we can still represent it as a superposition of stationary states (whose amplitudes simply oscillate in time). Thus,

$$\exp[-i(E_i + \Delta E_i)t/\hbar] \exp(-\Gamma_i t/2\hbar) = \int f(E) \exp(-iEt/\hbar) dE, \quad (6.303)$$

where  $f(E)$  is the weight of the stationary state with energy  $E$  in the superposition. The Fourier inversion theorem yields

$$|f(E)|^2 \propto \frac{1}{(E - [E_i + \text{Re}(\Delta_i)])^2 + \Gamma_i^2/4}. \quad (6.304)$$

In the absence of the perturbation,  $|f(E)|^2$  is basically a delta-function centred on the unperturbed energy  $E_i$  of state  $|i\rangle$ . In other words, state  $|i\rangle$  is a stationary state whose energy is completely determined. In the presence of the perturbation, the energy of state  $|i\rangle$  is *shifted* by  $\text{Re}(\Delta_i)$ . The fact that the state is no longer stationary (*i.e.*, it decays in time) implies that its energy cannot be exactly determined. Indeed, the energy of the state is smeared over some region of width (in energy)  $\Gamma_i$  centred around the shifted energy  $E_i + \text{Re}(\Delta_i)$ . The faster the decay of the state (*i.e.*, the larger  $\Gamma_i$ ), the more its energy is spread out. This effect is clearly a manifestation of the energy-time uncertainty relation  $\Delta E \Delta t \sim \hbar$ . One consequence of this effect is the existence of a *natural width* of spectral lines associated with the decay of some excited state to the ground state (or any other lower energy state). The uncertainty in energy of the excited state, due to its propensity to decay, gives rise to a slight smearing (in wave-length) of the spectral line associated with the transition. Strong lines, which correspond to fast transitions, are smeared out more than weak lines. For this reason, spectroscopists generally favour forbidden lines for Doppler shift measurements. Such lines are not as bright as those corresponding to allowed transitions, but they are a lot sharper.

## 7 Scattering theory

### 7.1 Introduction

Historically, data regarding quantum phenomena has been obtained from two main sources—the study of spectroscopic lines, and scattering experiments. We have already developed theories which account for some aspects of the spectra of hydrogen-like atoms. Let us now examine the quantum theory of scattering.

### 7.2 The Lipmann-Schwinger equation

Consider time-independent scattering theory, for which the Hamiltonian of the system is written

$$H = H_0 + H_1, \quad (7.1)$$

where  $H_0$  is the Hamiltonian of a free particle of mass  $m$ ,

$$H_0 = \frac{p^2}{2m}, \quad (7.2)$$

and  $H_1$  represents the non-time-varying source of the scattering. Let  $|\phi\rangle$  be an energy eigenket of  $H_0$ ,

$$H_0 |\phi\rangle = E |\phi\rangle, \quad (7.3)$$

whose wave-function  $\langle \mathbf{r}' | \phi \rangle$  is  $\phi(\mathbf{r}')$ . This state is a plane-wave state or, possibly, a spherical-wave state. Schrödinger's equation for the scattering problem is

$$(H_0 + H_1)|\psi\rangle = E |\psi\rangle, \quad (7.4)$$

where  $|\psi\rangle$  is an energy eigenstate of the total Hamiltonian whose wave-function  $\langle \mathbf{r}' | \psi \rangle$  is  $\psi(\mathbf{r}')$ . In general, both  $H_0$  and  $H_0 + H_1$  have continuous energy spectra: *i.e.*, their energy eigenstates are unbound. We require a solution of Eq. (7.4) which satisfies the boundary condition  $|\psi\rangle \rightarrow |\phi\rangle$  as  $H_1 \rightarrow 0$ . Here,  $|\phi\rangle$  is a solution of the free particle Schrödinger equation, (7.3), corresponding to the same energy eigenvalue.

Formally, the desired solution can be written

$$|\psi\rangle = |\phi\rangle + \frac{1}{E - H_0} H_1 |\psi\rangle. \quad (7.5)$$

Note that we can recover Eq. (7.4) by operating on the above equation with  $E - H_0$ , and making use of Eq. (7.3). Furthermore, the solution satisfies the boundary condition  $|\psi\rangle \rightarrow |\phi\rangle$  as  $H_1 \rightarrow 0$ . Unfortunately, the operator  $(E - H_0)^{-1}$  is *singular*: i.e., it produces infinities when it operates on an eigenstate of  $H_0$  corresponding to the eigenvalue  $E$ . We need a prescription for dealing with these infinities, otherwise the above solution is useless. The standard prescription is to make the energy eigenvalue  $E$  slightly complex. Thus,

$$|\psi^\pm\rangle = |\phi\rangle + \frac{1}{E - H_0 \pm i\epsilon} H_1 |\psi^\pm\rangle, \quad (7.6)$$

where  $\epsilon$  is real, positive, and small. Equation (7.6) is called the *Lipmann-Schwinger equation*, and is non-singular as long as  $\epsilon > 0$ . The physical significance of the  $\pm$  signs will become apparent later on.

The Lipmann-Schwinger equation can be converted into an *integral equation* via left multiplication by  $\langle \mathbf{r} |$ . Thus,

$$\psi^\pm(\mathbf{r}) = \phi(\mathbf{r}) + \int \langle \mathbf{r} | \frac{1}{E - H_0 \pm i\epsilon} | \mathbf{r}' \rangle \langle \mathbf{r}' | H_1 | \psi^\pm \rangle d^3 \mathbf{r}'. \quad (7.7)$$

Adopting the Schrödinger representation, we can write the scattering problem (7.4) in the form

$$(\nabla^2 + k^2) \psi(\mathbf{r}) = \frac{2m}{\hbar^2} \langle \mathbf{r} | H_1 | \psi \rangle, \quad (7.8)$$

where

$$E = \frac{\hbar^2 k^2}{2m}. \quad (7.9)$$

This equation is called *Helmholtz's equation*, and can be inverted using standard Green's function techniques. Thus,

$$\psi(\mathbf{r}) = \phi(\mathbf{r}) + \frac{2m}{\hbar^2} \int G(\mathbf{r}, \mathbf{r}') \langle \mathbf{r}' | H_1 | \psi \rangle d^3 \mathbf{r}', \quad (7.10)$$

where

$$(\nabla^2 + k^2) G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'). \quad (7.11)$$

Note that the solution (7.10) satisfies the boundary condition  $|\psi\rangle \rightarrow |\phi\rangle$  as  $H_1 \rightarrow 0$ . As is well-known, the Green's function for the Helmholtz problem is given by

$$G(\mathbf{r}, \mathbf{r}') = -\frac{\exp(\pm i k |\mathbf{r} - \mathbf{r}'|)}{4\pi |\mathbf{r} - \mathbf{r}'|}. \quad (7.12)$$

Thus, Eq. (7.10) becomes

$$\psi^\pm(\mathbf{r}) = \phi(\mathbf{r}) - \frac{2m}{\hbar^2} \int \frac{\exp(\pm i k |\mathbf{r} - \mathbf{r}'|)}{4\pi |\mathbf{r} - \mathbf{r}'|} \langle \mathbf{r}' | H_1 | \psi \rangle d^3 \mathbf{r}'. \quad (7.13)$$

A comparison of Eqs. (7.7) and (7.13) suggests that the kernel to Eq. (7.7) takes the form

$$\left\langle \mathbf{r} \left| \frac{1}{E - H_0 \pm i \epsilon} \right| \mathbf{r}' \right\rangle = -\frac{2m}{\hbar^2} \frac{\exp(\pm i k |\mathbf{r} - \mathbf{r}'|)}{4\pi |\mathbf{r} - \mathbf{r}'|}. \quad (7.14)$$

It is not entirely clear that the  $\pm$  signs correspond on both sides of this equation. In fact, they do, as is easily proved by a more rigorous derivation of this result.

Let us suppose that the scattering Hamiltonian,  $H_1$ , is only a function of the position operators. This implies that

$$\langle \mathbf{r}' | H_1 | \mathbf{r} \rangle = V(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}'). \quad (7.15)$$

We can write

$$\begin{aligned} \langle \mathbf{r}' | H_1 | \psi^\pm \rangle &= \int \langle \mathbf{r}' | H_1 | \mathbf{r}'' \rangle \langle \mathbf{r}'' | \psi^\pm \rangle d^3 \mathbf{r}'' \\ &= V(\mathbf{r}') \psi^\pm(\mathbf{r}'). \end{aligned} \quad (7.16)$$

Thus, the integral equation (7.13) simplifies to

$$\psi^\pm(\mathbf{r}) = \phi(\mathbf{r}) - \frac{2m}{\hbar^2} \int \frac{\exp(\pm i k |\mathbf{r} - \mathbf{r}'|)}{4\pi |\mathbf{r} - \mathbf{r}'|} V(\mathbf{r}') \psi^\pm(\mathbf{r}') d^3 \mathbf{r}'. \quad (7.17)$$

Suppose that the initial state  $|\phi\rangle$  is a plane-wave with wave-vector  $\mathbf{k}$  (i.e., a stream of particles of definite momentum  $\mathbf{p} = \hbar \mathbf{k}$ ). The ket corresponding to this state is denoted  $|\mathbf{k}\rangle$ . The associated wave-function takes the form

$$\langle \mathbf{r} | \mathbf{k} \rangle = \frac{\exp(i \mathbf{k} \cdot \mathbf{r})}{(2\pi)^{3/2}}. \quad (7.18)$$

The wave-function is normalized such that

$$\begin{aligned}\langle \mathbf{k} | \mathbf{k}' \rangle &= \int \langle \mathbf{k} | \mathbf{r} \rangle \langle \mathbf{r} | \mathbf{k}' \rangle d^3 \mathbf{r} \\ &= \int \frac{\exp[-i \mathbf{r} \cdot (\mathbf{k} - \mathbf{k}')] }{(2\pi)^3} d^3 \mathbf{r} = \delta(\mathbf{k} - \mathbf{k}').\end{aligned}\quad (7.19)$$

Suppose that the scattering potential  $V(\mathbf{r})$  is only non-zero in some relatively localized region centred on the origin ( $\mathbf{r} = 0$ ). Let us calculate the wave-function  $\psi(\mathbf{r})$  a long way from the scattering region. In other words, let us adopt the ordering  $r \gg r'$ . It is easily demonstrated that

$$|\mathbf{r} - \mathbf{r}'| \simeq r - \hat{\mathbf{r}} \cdot \mathbf{r}' \quad (7.20)$$

to first-order in  $r'/r$ , where

$$\hat{\mathbf{r}} = \frac{\mathbf{r}}{r} \quad (7.21)$$

is a unit vector which points from the scattering region to the observation point. Let us define

$$\mathbf{k}' = k \hat{\mathbf{r}}. \quad (7.22)$$

Clearly,  $\mathbf{k}'$  is the wave-vector for particles which possess the same energy as the incoming particles (*i.e.*,  $k' = k$ ), but propagate from the scattering region to the observation point. Note that

$$\exp(\pm i k |\mathbf{r} - \mathbf{r}'|) \simeq \exp(\pm i k r) \exp(\mp i \mathbf{k}' \cdot \mathbf{r}'). \quad (7.23)$$

In the large- $r$  limit, Eq. (7.17) reduces to

$$\psi(\mathbf{r})^\pm \simeq \frac{\exp(i \mathbf{k} \cdot \mathbf{r})}{(2\pi)^{3/2}} - \frac{m}{2\pi \hbar^2} \frac{\exp(\pm i k r)}{r} \int \exp(\mp i \mathbf{k}' \cdot \mathbf{r}') V(\mathbf{r}') \psi^\pm(\mathbf{r}') d^3 \mathbf{r}'. \quad (7.24)$$

The first term on the right-hand side is the incident wave. The second term represents a spherical wave centred on the scattering region. The plus sign (on  $\psi^\pm$ ) corresponds to a wave propagating away from the scattering region, whereas the minus sign corresponds to a wave propagating towards the scattering region.

It is obvious that the former represents the physical solution. Thus, the wave-function a long way from the scattering region can be written

$$\psi(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \left[ \exp(i\mathbf{k}\cdot\mathbf{r}) + \frac{\exp(ikr)}{r} f(\mathbf{k}', \mathbf{k}) \right], \quad (7.25)$$

where

$$\begin{aligned} f(\mathbf{k}', \mathbf{k}) &= -\frac{(2\pi)^2 m}{\hbar^2} \int \frac{\exp(-i\mathbf{k}'\cdot\mathbf{r}')}{(2\pi)^{3/2}} V(\mathbf{r}') \psi(\mathbf{r}') d^3\mathbf{r}' \\ &= -\frac{(2\pi)^2 m}{\hbar^2} \langle \mathbf{k}' | H_1 | \psi \rangle. \end{aligned} \quad (7.26)$$

Let us define the differential cross-section  $d\sigma/d\Omega$  as the number of particles per unit time scattered into an element of solid angle  $d\Omega$ , divided by the incident flux of particles. Recall, from Sect. 4, that the probability flux (*i.e.*, the particle flux) associated with a wave-function  $\psi$  is

$$\mathbf{j} = \frac{\hbar}{m} \text{Im}(\psi^* \nabla \psi). \quad (7.27)$$

Thus, the probability flux associated with the incident wave-function,

$$\frac{\exp(i\mathbf{k}\cdot\mathbf{r})}{(2\pi)^{3/2}}, \quad (7.28)$$

is

$$\mathbf{j}_{\text{inci}} = \frac{\hbar}{(2\pi)^3 m} \mathbf{k}. \quad (7.29)$$

Likewise, the probability flux associated with the scattered wave-function,

$$\frac{\exp(ikr)}{(2\pi)^{3/2}} \frac{f(\mathbf{k}', \mathbf{k})}{r}, \quad (7.30)$$

is

$$\mathbf{j}_{\text{scat}} = \frac{\hbar}{(2\pi)^3 m} \frac{|f(\mathbf{k}', \mathbf{k})|^2}{r^2} \mathbf{k} \hat{\mathbf{r}}. \quad (7.31)$$

Now,

$$\frac{d\sigma}{d\Omega} d\Omega = \frac{r^2 d\Omega |\mathbf{j}_{\text{scat}}|}{|\mathbf{j}_{\text{inci}}|}, \quad (7.32)$$

giving

$$\frac{d\sigma}{d\Omega} = |f(\mathbf{k}', \mathbf{k})|^2. \quad (7.33)$$

Thus,  $|f(\mathbf{k}', \mathbf{k})|^2$  gives the differential cross-section for particles with incident momentum  $\hbar \mathbf{k}$  to be scattered into states whose momentum vectors are directed in a range of solid angles  $d\Omega$  about  $\hbar \mathbf{k}'$ . Note that the scattered particles possess the same energy as the incoming particles (*i.e.*,  $k' = k$ ). This is always the case for scattering Hamiltonians of the form shown in Eq. (7.15).

### 7.3 The Born approximation

Equation (7.33) is not particularly useful, as it stands, because the quantity  $f(\mathbf{k}', \mathbf{k})$  depends on the unknown ket  $|\psi\rangle$ . Recall that  $\psi(\mathbf{r}) = \langle \mathbf{r} | \psi \rangle$  is the solution of the integral equation

$$\psi(\mathbf{r}) = \phi(\mathbf{r}) - \frac{m}{2\pi \hbar^2} \frac{\exp(i \mathbf{k} \cdot \mathbf{r})}{r} \int \exp(-i \mathbf{k}' \cdot \mathbf{r}') V(\mathbf{r}') \psi(\mathbf{r}') d^3 \mathbf{r}', \quad (7.34)$$

where  $\phi(\mathbf{r})$  is the wave-function of the incident state. According to the above equation the total wave-function is a superposition of the incident wave-function and lots of spherical-waves emitted from the scattering region. The strength of the spherical-wave emitted at a given point is proportional to the local value of the scattering potential,  $V$ , as well as the local value of the wave-function,  $\psi$ .

Suppose that the scattering is not particularly strong. In this case, it is reasonable to suppose that the total wave-function,  $\psi(\mathbf{r})$ , does not differ substantially from the incident wave-function,  $\phi(\mathbf{r})$ . Thus, we can obtain an expression for  $f(\mathbf{k}', \mathbf{k})$  by making the substitution

$$\psi(\mathbf{r}) \rightarrow \phi(\mathbf{r}) = \frac{\exp(i \mathbf{k} \cdot \mathbf{r})}{(2\pi)^{3/2}}. \quad (7.35)$$

This is called the *Born approximation*.

The Born approximation yields

$$f(\mathbf{k}', \mathbf{k}) \simeq -\frac{m}{2\pi \hbar^2} \int \exp[i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}'] V(\mathbf{r}') d^3 \mathbf{r}'. \quad (7.36)$$

Thus,  $f(\mathbf{k}', \mathbf{k})$  is proportional to the Fourier transform of the scattering potential  $V(\mathbf{r})$  with respect to the wave-vector  $\mathbf{q} \equiv \mathbf{k} - \mathbf{k}'$ .

For a spherically symmetric potential,

$$f(\mathbf{k}', \mathbf{k}) \simeq -\frac{m}{2\pi \hbar^2} \iiint \exp(i \mathbf{q} \cdot \mathbf{r}') V(r') r'^2 dr' \sin \theta' d\theta' d\phi', \quad (7.37)$$

giving

$$f(\mathbf{k}', \mathbf{k}) \simeq -\frac{2m}{\hbar^2 q} \int_0^\infty r' V(r') \sin(q r') dr'. \quad (7.38)$$

Note that  $f(\mathbf{k}', \mathbf{k})$  is just a function of  $q$  for a spherically symmetric potential. It is easily demonstrated that

$$q \equiv |\mathbf{k} - \mathbf{k}'| = 2k \sin(\theta/2), \quad (7.39)$$

where  $\theta$  is the angle subtended between the vectors  $\mathbf{k}$  and  $\mathbf{k}'$ . In other words,  $\theta$  is the angle of scattering. Recall that the vectors  $\mathbf{k}$  and  $\mathbf{k}'$  have the same length by energy conservation.

Consider scattering by a Yukawa potential

$$V(r) = \frac{V_0 \exp(-\mu r)}{\mu r}, \quad (7.40)$$

where  $V_0$  is a constant and  $1/\mu$  measures the “range” of the potential. It follows from Eq. (7.38) that

$$f(\theta) = -\frac{2m V_0}{\hbar^2 \mu} \frac{1}{q^2 + \mu^2}, \quad (7.41)$$

since

$$\int_0^\infty \exp(-\mu r') \sin(q r') dr' = \frac{q}{\mu^2 + q^2}. \quad (7.42)$$

Thus, in the Born approximation, the differential cross-section for scattering by a Yukawa potential is

$$\frac{d\sigma}{d\Omega} \simeq \left( \frac{2m V_0}{\hbar^2 \mu} \right)^2 \frac{1}{[2k^2(1 - \cos \theta) + \mu^2]^2}, \quad (7.43)$$

given that

$$q^2 = 4k^2 \sin^2(\theta/2) = 2k^2(1 - \cos \theta). \quad (7.44)$$

The Yukawa potential reduces to the familiar Coulomb potential as  $\mu \rightarrow 0$ , provided that  $V_0/\mu \rightarrow ZZ'e^2/4\pi\epsilon_0$ . In this limit the Born differential cross-section becomes

$$\frac{d\sigma}{d\Omega} \simeq \left( \frac{2mZZ'e^2}{4\pi\epsilon_0\hbar^2} \right)^2 \frac{1}{16k^4 \sin^4(\theta/2)}. \quad (7.45)$$

Recall that  $\hbar k$  is equivalent to  $|\mathbf{p}|$ , so the above equation can be rewritten

$$\frac{d\sigma}{d\Omega} \simeq \left( \frac{ZZ'e^2}{16\pi\epsilon_0 E} \right)^2 \frac{1}{\sin^4(\theta/2)}, \quad (7.46)$$

where  $E = p^2/2m$  is the kinetic energy of the incident particles. Equation (7.46) is the classical Rutherford scattering cross-section formula.

The Born approximation is valid provided that  $\psi(\mathbf{r})$  is not too different from  $\phi(\mathbf{r})$  in the scattering region. It follows, from Eq. (7.17), that the condition for  $\psi(\mathbf{r}) \simeq \phi(\mathbf{r})$  in the vicinity of  $\mathbf{r} = 0$  is

$$\left| \frac{m}{2\pi\hbar^2} \int \frac{\exp(i\mathbf{k}\cdot\mathbf{r}')}{r'} V(\mathbf{r}') d^3\mathbf{r}' \right| \ll 1. \quad (7.47)$$

Consider the special case of the Yukawa potential. At low energies, (*i.e.*,  $k \ll \mu$ ) we can replace  $\exp(i\mathbf{k}\cdot\mathbf{r}')$  by unity, giving

$$\frac{2m|V_0|}{\hbar^2\mu^2} \ll 1 \quad (7.48)$$

as the condition for the validity of the Born approximation. The condition for the Yukawa potential to develop a bound state is

$$\frac{2m|V_0|}{\hbar^2\mu^2} \geq 2.7, \quad (7.49)$$

where  $V_0$  is negative. Thus, if the potential is strong enough to form a bound state then the Born approximation is likely to break down. In the high- $k$  limit, Eq. (7.47) yields

$$\frac{2m|V_0|}{\hbar^2\mu k} \ll 1. \quad (7.50)$$

This inequality becomes progressively easier to satisfy as  $k$  increases, implying that the Born approximation is more accurate at high incident particle energies.

## 7.4 Partial waves

We can assume, without loss of generality, that the incident wave-function is characterized by a wave-vector  $\mathbf{k}$  which is aligned parallel to the  $z$ -axis. The scattered wave-function is characterized by a wave-vector  $\mathbf{k}'$  which has the same magnitude as  $\mathbf{k}$ , but, in general, points in a different direction. The direction of  $\mathbf{k}'$  is specified by the polar angle  $\theta$  (*i.e.*, the angle subtended between the two wave-vectors), and an azimuthal angle  $\varphi$  about the  $z$ -axis. Equation (7.38) strongly suggests that for a spherically symmetric scattering potential [*i.e.*,  $V(\mathbf{r}) = V(r)$ ] the scattering amplitude is a function of  $\theta$  only:

$$f(\theta, \varphi) = f(\theta). \quad (7.51)$$

It follows that neither the incident wave-function,

$$\phi(\mathbf{r}) = \frac{\exp(i k z)}{(2\pi)^{3/2}} = \frac{\exp(i k r \cos \theta)}{(2\pi)^{3/2}}, \quad (7.52)$$

nor the total wave-function,

$$\psi(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \left[ \exp(i k r \cos \theta) + \frac{\exp(i k r) f(\theta)}{r} \right], \quad (7.53)$$

depend on the azimuthal angle  $\varphi$ .

Outside the range of the scattering potential, both  $\phi(\mathbf{r})$  and  $\psi(\mathbf{r})$  satisfy the free space Schrödinger equation

$$(\nabla^2 + k^2) \psi = 0. \quad (7.54)$$

What is the most general solution to this equation in spherical polar coordinates which does not depend on the azimuthal angle  $\varphi$ ? Separation of variables yields

$$\psi(r, \theta) = \sum_{\mathbf{l}} R_{\mathbf{l}}(r) P_{\mathbf{l}}(\cos \theta), \quad (7.55)$$

since the Legendre functions  $P_{\mathbf{l}}(\cos \theta)$  form a complete set in  $\theta$ -space. The Legendre functions are related to the spherical harmonics introduced in Sect. 5 via

$$P_{\mathbf{l}}(\cos \theta) = \sqrt{\frac{4\pi}{2\mathbf{l} + 1}} Y_{\mathbf{l}}^0(\theta, \varphi). \quad (7.56)$$

Equations (7.54) and (7.55) can be combined to give

$$r^2 \frac{d^2 R_l}{dr^2} + 2r \frac{dR_l}{dr} + [k^2 r^2 - l(l+1)] R_l = 0. \quad (7.57)$$

The two independent solutions to this equation are called a spherical Bessel function,  $j_l(kr)$ , and a Neumann function,  $\eta_l(kr)$ . It is easily demonstrated that

$$j_l(y) = y^l \left( -\frac{1}{y} \frac{d}{dy} \right)^l \frac{\sin y}{y}, \quad (7.58)$$

$$\eta_l(y) = -y^l \left( -\frac{1}{y} \frac{d}{dy} \right)^l \frac{\cos y}{y}. \quad (7.59)$$

Note that spherical Bessel functions are well-behaved in the limit  $y \rightarrow 0$ , whereas Neumann functions become singular. The asymptotic behaviour of these functions in the limit  $y \rightarrow \infty$  is

$$j_l(y) \rightarrow \frac{\sin(y - l\pi/2)}{y}, \quad (7.60)$$

$$\eta_l(y) \rightarrow -\frac{\cos(y - l\pi/2)}{y}. \quad (7.61)$$

We can write

$$\exp(ikr \cos \theta) = \sum_l a_l j_l(kr) P_l(\cos \theta), \quad (7.62)$$

where the  $a_l$  are constants. Note there are no Neumann functions in this expansion, because they are not well-behaved as  $r \rightarrow 0$ . The Legendre functions are orthonormal,

$$\int_{-1}^1 P_n(\mu) P_m(\mu) d\mu = \frac{\delta_{nm}}{n+1/2}, \quad (7.63)$$

so we can invert the above expansion to give

$$a_l j_l(kr) = (l+1/2) \int_{-1}^1 \exp(ikr \mu) P_l(\mu) d\mu. \quad (7.64)$$

It is well-known that

$$j_l(y) = \frac{(-i)^l}{2} \int_{-1}^1 \exp(iy \mu) P_l(\mu) d\mu, \quad (7.65)$$

where  $l = 0, 1, 2, \dots$  [see Abramowitz and Stegun (Dover, New York NY, 1965), Eq. 10.1.14]. Thus,

$$a_l = i^l (2l + 1), \quad (7.66)$$

giving

$$\exp(i k r \cos \theta) = \sum_l i^l (2l + 1) j_l(k r) P_l(\cos \theta). \quad (7.67)$$

The above expression tells us how to decompose a plane-wave into a series of spherical-waves (or “partial waves”).

The most general solution for the total wave-function outside the scattering region is

$$\psi(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \sum_l [A_l j_l(k r) + B_l \eta_l(k r)] P_l(\cos \theta), \quad (7.68)$$

where the  $A_l$  and  $B_l$  are constants. Note that the Neumann functions are allowed to appear in this expansion, because its region of validity does not include the origin. In the large- $r$  limit, the total wave-function reduces to

$$\psi(\mathbf{r}) \simeq \frac{1}{(2\pi)^{3/2}} \sum_l \left[ A_l \frac{\sin(k r - l\pi/2)}{k r} - B_l \frac{\cos(k r - l\pi/2)}{k r} \right] P_l(\cos \theta), \quad (7.69)$$

where use has been made of Eqs. (7.60)–(7.61). The above expression can also be written

$$\psi(\mathbf{r}) \simeq \frac{1}{(2\pi)^{3/2}} \sum_l C_l \frac{\sin(k r - l\pi/2 + \delta_l)}{k r} P_l(\cos \theta), \quad (7.70)$$

where the sine and cosine functions have been combined to give a sine function which is phase-shifted by  $\delta_l$ .

Equation (7.70) yields

$$\begin{aligned} \psi(\mathbf{r}) \simeq & \frac{1}{(2\pi)^{3/2}} \sum_l C_l \frac{\exp[i(k r - l\pi/2 + \delta_l)] - \exp[-i(k r - l\pi/2 + \delta_l)]}{2i k r} \\ & \times P_l(\cos \theta), \end{aligned} \quad (7.71)$$

which contains both incoming and outgoing spherical-waves. What is the source of the incoming waves? Obviously, they must be part of the large- $r$  asymptotic

expansion of the incident wave-function. In fact, it is easily seen that

$$\phi(\mathbf{r}) \simeq \frac{1}{(2\pi)^{3/2}} \sum_{\mathfrak{l}} i^{\mathfrak{l}} (2\mathfrak{l} + 1) \frac{\exp[i(\mathbf{k} \cdot \mathbf{r} - \mathfrak{l}\pi/2)] - \exp[-i(\mathbf{k} \cdot \mathbf{r} - \mathfrak{l}\pi/2)]}{2i\mathbf{k} \cdot \mathbf{r}} \times P_{\mathfrak{l}}(\cos \theta) \quad (7.72)$$

in the large- $r$  limit. Now, Eqs. (7.52) and (7.53) give

$$(2\pi)^{3/2}[\psi(\mathbf{r}) - \phi(\mathbf{r})] = \frac{\exp(i\mathbf{k} \cdot \mathbf{r})}{r} f(\theta). \quad (7.73)$$

Note that the right-hand side consists only of an outgoing spherical wave. This implies that the coefficients of the incoming spherical waves in the large- $r$  expansions of  $\psi(\mathbf{r})$  and  $\phi(\mathbf{r})$  must be equal. It follows from Eqs. (7.71) and (7.72) that

$$C_{\mathfrak{l}} = (2\mathfrak{l} + 1) \exp[i(\delta_{\mathfrak{l}} + \mathfrak{l}\pi/2)]. \quad (7.74)$$

Thus, Eqs. (7.71)–(7.73) yield

$$f(\theta) = \sum_{\mathfrak{l}=0}^{\infty} (2\mathfrak{l} + 1) \frac{\exp(i\delta_{\mathfrak{l}})}{\mathbf{k}} \sin \delta_{\mathfrak{l}} P_{\mathfrak{l}}(\cos \theta). \quad (7.75)$$

Clearly, determining the scattering amplitude  $f(\theta)$  via a decomposition into partial waves (*i.e.*, spherical-waves) is equivalent to determining the phase-shifts  $\delta_{\mathfrak{l}}$ .

## 7.5 The optical theorem

The differential scattering cross-section  $d\sigma/d\Omega$  is simply the modulus squared of the scattering amplitude  $f(\theta)$ . The total cross-section is given by

$$\begin{aligned} \sigma_{\text{total}} &= \int |f(\theta)|^2 d\Omega \\ &= \frac{1}{k^2} \oint d\varphi \int_{-1}^1 d\mu \sum_{\mathfrak{l}} \sum_{\mathfrak{l}'} (2\mathfrak{l} + 1) (2\mathfrak{l}' + 1) \exp[i(\delta_{\mathfrak{l}} - \delta_{\mathfrak{l}'})] \\ &\quad \times \sin \delta_{\mathfrak{l}} \sin \delta_{\mathfrak{l}'} P_{\mathfrak{l}}(\mu) P_{\mathfrak{l}'}(\mu), \end{aligned} \quad (7.76)$$

where  $\mu = \cos \theta$ . It follows that

$$\sigma_{\text{total}} = \frac{4\pi}{k^2} \sum_l (2l+1) \sin^2 \delta_l, \quad (7.77)$$

where use has been made of Eq. (7.63). A comparison of this result with Eq. (7.75) yields

$$\sigma_{\text{total}} = \frac{4\pi}{k} \text{Im} [f(0)], \quad (7.78)$$

since  $P_l(1) = 1$ . This result is known as the *optical theorem*. It is a reflection of the fact that the very existence of scattering requires scattering in the forward ( $\theta = 0$ ) direction in order to interfere with the incident wave, and thereby reduce the probability current in this direction.

It is usual to write

$$\sigma_{\text{total}} = \sum_{l=0}^{\infty} \sigma_l, \quad (7.79)$$

where

$$\sigma_l = \frac{4\pi}{k^2} (2l+1) \sin^2 \delta_l \quad (7.80)$$

is the  $l$ th partial cross-section: *i.e.*, the contribution to the total cross-section from the  $l$ th partial wave. Note that the maximum value for the  $l$ th partial cross-section occurs when the phase-shift  $\delta_l$  takes the value  $\pi/2$ .

## 7.6 Determination of phase-shifts

Let us now consider how the phase-shifts  $\delta_l$  can be evaluated. Consider a spherically symmetric potential  $V(r)$  which vanishes for  $r > a$ , where  $a$  is termed the range of the potential. In the region  $r > a$ , the wave-function  $\psi(\mathbf{r})$  satisfies the free-space Schrödinger equation (7.54). The most general solution which is consistent with no incoming spherical-waves is

$$\psi(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \sum_{l=0}^{\infty} i^l (2l+1) A_l(r) P_l(\cos \theta), \quad (7.81)$$

where

$$A_l(r) = \exp(i \delta_l) [\cos \delta_l j_l(kr) - \sin \delta_l \eta_l(kr)]. \quad (7.82)$$

Note that Neumann functions are allowed to appear in the above expression, because its region of validity does not include the origin (where  $V \neq 0$ ). The logarithmic derivative of the  $l$ th radial wave-function  $A_l(r)$  just outside the range of the potential is given by

$$\beta_{l+} = k a \left[ \frac{\cos \delta_l j'_l(ka) - \sin \delta_l \eta'_l(ka)}{\cos \delta_l j_l(ka) - \sin \delta_l \eta_l(ka)} \right], \quad (7.83)$$

where  $j'_l(x)$  denotes  $dj_l(x)/dx$ , etc. The above equation can be inverted to give

$$\tan \delta_l = \frac{k a j'_l(ka) - \beta_{l+} j_l(ka)}{k a \eta'_l(ka) - \beta_{l+} \eta_l(ka)}. \quad (7.84)$$

Thus, the problem of determining the phase-shift  $\delta_l$  is equivalent to that of obtaining  $\beta_{l+}$ .

The most general solution to Schrödinger's equation inside the range of the potential ( $r < a$ ) which does not depend on the azimuthal angle  $\varphi$  is

$$\psi(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \sum_{l=0}^{\infty} i^l (2l+1) R_l(r) P_l(\cos \theta), \quad (7.85)$$

where

$$R_l(r) = \frac{u_l(r)}{r}, \quad (7.86)$$

and

$$\frac{d^2 u_l}{dr^2} + \left[ k^2 - \frac{2m}{\hbar^2} V - \frac{l(l+1)}{r^2} \right] u_l = 0. \quad (7.87)$$

The boundary condition

$$u_l(0) = 0 \quad (7.88)$$

ensures that the radial wave-function is well-behaved at the origin. We can launch a well-behaved solution of the above equation from  $r = 0$ , integrate out to  $r = a$ , and form the logarithmic derivative

$$\beta_{l-} = \frac{1}{(u_l/r)} \left. \frac{d(u_l/r)}{dr} \right|_{r=a}. \quad (7.89)$$

Since  $\psi(\mathbf{r})$  and its first derivatives are necessarily continuous for physically acceptable wave-functions, it follows that

$$\beta_{l+} = \beta_{l-}. \quad (7.90)$$

The phase-shift  $\delta_l$  is obtainable from Eq. (7.84).

## 7.7 Hard sphere scattering

Let us test out this scheme using a particularly simple example. Consider scattering by a hard sphere, for which the potential is infinite for  $r < a$ , and zero for  $r > a$ . It follows that  $\psi(\mathbf{r})$  is zero in the region  $r < a$ , which implies that  $u_l = 0$  for all  $l$ . Thus,

$$\beta_{l-} = \beta_{l+} = \infty, \quad (7.91)$$

for all  $l$ . It follows from Eq. (7.84) that

$$\tan \delta_l = \frac{j_l(k a)}{\eta_l(k a)}. \quad (7.92)$$

Consider the  $l = 0$  partial wave, which is usually referred to as the  $s$ -wave. Equation (7.92) yields

$$\tan \delta_0 = \frac{\sin(k a)/k a}{-\cos(k a)/k a} = -\tan k a, \quad (7.93)$$

where use has been made of Eqs. (7.58)–(7.59). It follows that

$$\delta_0 = -k a. \quad (7.94)$$

The  $s$ -wave radial wave function is

$$\begin{aligned} A_0(r) &= \exp(-i k a) \frac{[\cos k a \sin k r - \sin k a \cos k r]}{k r} \\ &= \exp(-i k a) \frac{\sin[k(r - a)]}{k r}. \end{aligned} \quad (7.95)$$

The corresponding radial wave-function for the incident wave takes the form

$$\tilde{A}_0(r) = \frac{\sin k r}{k r}. \quad (7.96)$$

It is clear that the actual  $l = 0$  radial wave-function is similar to the incident  $l = 0$  wave-function, except that it is phase-shifted by  $k a$ .

Let us consider the low and high energy asymptotic limits of  $\tan \delta_l$ . Low energy means  $k a \ll 1$ . In this regime, the spherical Bessel functions and Neumann functions reduce to:

$$j_l(k r) \simeq \frac{(k r)^l}{(2 l + 1)!!}, \quad (7.97)$$

$$\eta_l(k r) \simeq -\frac{(2 l - 1)!!}{(k r)^{l+1}}, \quad (7.98)$$

where  $n!! = n (n - 2) (n - 4) \cdots 1$ . It follows that

$$\tan \delta_l = \frac{-(k a)^{2l+1}}{(2 l + 1) [(2 l - 1)!!]^2}. \quad (7.99)$$

It is clear that we can neglect  $\delta_l$ , with  $l > 0$ , with respect to  $\delta_0$ . In other words, at low energy only *s*-wave scattering (*i.e.*, spherically symmetric scattering) is important. It follows from Eqs. (7.33), (7.75), and (7.94) that

$$\frac{d\sigma}{d\Omega} = \frac{\sin^2 k a}{k^2} \simeq a^2 \quad (7.100)$$

for  $k a \ll 1$ . Note that the total cross-section

$$\sigma_{\text{total}} = \int \frac{d\sigma}{d\Omega} d\Omega = 4\pi a^2 \quad (7.101)$$

is *four times* the *geometric cross-section*  $\pi a^2$  (*i.e.*, the cross-section for classical particles bouncing off a hard sphere of radius  $a$ ). However, low energy scattering implies relatively long wave-lengths, so we do not expect to obtain the classical result in this limit.

Consider the high energy limit  $k a \gg 1$ . At high energies, all partial waves up to  $l_{\text{max}} = k a$  contribute significantly to the scattering cross-section. It follows from Eq. (7.77) that

$$\sigma_{\text{total}} = \frac{4\pi}{k^2} \sum_{l=0}^{l_{\text{max}}} (2 l + 1) \sin^2 \delta_l. \quad (7.102)$$

With so many  $l$  values contributing, it is legitimate to replace  $\sin^2 \delta_l$  by its average value  $1/2$ . Thus,

$$\sigma_{\text{total}} = \sum_{l=0}^{ka} \frac{2\pi}{k^2} (2l+1) \simeq 2\pi a^2. \quad (7.103)$$

This is *twice* the classical result, which is somewhat surprising, since we might expect to obtain the classical result in the short wave-length limit. For hard sphere scattering, incident waves with impact parameters less than  $a$  must be deflected. However, in order to produce a “shadow” behind the sphere, there must be scattering in the forward direction (recall the optical theorem) to produce destructive interference with the incident plane-wave. In fact, the interference is not completely destructive, and the shadow has a bright spot in the forward direction. The effective cross-section associated with this bright spot is  $\pi a^2$  which, when combined with the cross-section for classical reflection,  $\pi a^2$ , gives the actual cross-section of  $2\pi a^2$ .

## 7.8 Low energy scattering

At low energies (*i.e.*, when  $1/k$  is much larger than the range of the potential) partial waves with  $l > 0$ , in general, make a negligible contribution to the scattering cross-section. It follows that, at these energies, with a finite range potential, only  $s$ -wave scattering is important.

As a specific example, let us consider scattering by a finite potential well, characterized by  $V = V_0$  for  $r < a$ , and  $V = 0$  for  $r \geq a$ . Here,  $V_0$  is a constant. The potential is repulsive for  $V_0 > 0$ , and attractive for  $V_0 < 0$ . The outside wave-function is given by [see Eq. (7.82)]

$$A_0(r) = \exp(i\delta_0) [j_0(kr) \cos \delta_0 - \eta_0(kr) \sin \delta_0] \quad (7.104)$$

$$= \frac{\exp(i\delta_0) \sin(kr + \delta_0)}{kr}, \quad (7.105)$$

where use has been made of Eqs. (7.58)–(7.59). The inside wave-function follows

from Eq. (7.87). We obtain

$$A_0(r) = B \frac{\sin k'r}{r}, \quad (7.106)$$

where use has been made of the boundary condition (7.88). Here, B is a constant, and

$$E - V_0 = \frac{\hbar^2 k'^2}{2m}. \quad (7.107)$$

Note that Eq. (7.106) only applies when  $E > V_0$ . For  $E < V_0$ , we have

$$A_0(r) = B \frac{\sinh \kappa r}{r}, \quad (7.108)$$

where

$$V_0 - E = \frac{\hbar^2 \kappa^2}{2m}. \quad (7.109)$$

Matching  $A_0(r)$ , and its radial derivative at  $r = a$ , yields

$$\tan(k a + \delta_0) = \frac{k}{k'} \tan k' a \quad (7.110)$$

for  $E > V_0$ , and

$$\tan(k a + \delta_0) = \frac{k}{\kappa} \tanh \kappa a \quad (7.111)$$

for  $E < V_0$ .

Consider an attractive potential, for which  $E > V_0$ . Suppose that  $|V_0| \gg E$  (*i.e.*, the depth of the potential well is much larger than the energy of the incident particles), so that  $k' \gg k$ . It follows from Eq. (7.110) that, unless  $\tan k' a$  becomes extremely large, the right-hand side is much less than unity, so replacing the tangent of a small quantity with the quantity itself, we obtain

$$k a + \delta_0 \simeq \frac{k}{k'} \tan k' a. \quad (7.112)$$

This yields

$$\delta_0 \simeq k a \left( \frac{\tan k' a}{k' a} - 1 \right). \quad (7.113)$$

According to Eq. (7.102), the scattering cross-section is given by

$$\sigma_{\text{total}} \simeq \frac{4\pi}{k^2} \sin^2 \delta_0 = 4\pi a^2 \left( \frac{\tan k'a}{k'a} - 1 \right)^2. \quad (7.114)$$

Now

$$k'a = \sqrt{k^2 a^2 + \frac{2m|V_0|a^2}{\hbar^2}}, \quad (7.115)$$

so for sufficiently small values of  $ka$ ,

$$k'a \simeq \sqrt{\frac{2m|V_0|a^2}{\hbar^2}}. \quad (7.116)$$

It follows that the total (s-wave) scattering cross-section is independent of the energy of the incident particles (provided that this energy is sufficiently small).

Note that there are values of  $k'a$  (e.g.,  $k'a \simeq 4.49$ ) at which  $\delta_0 \rightarrow \pi$ , and the scattering cross-section (7.114) vanishes, despite the very strong attraction of the potential. In reality, the cross-section is not exactly zero, because of contributions from  $l > 0$  partial waves. But, at low incident energies, these contributions are small. It follows that there are certain values of  $V_0$  and  $k$  which give rise to almost perfect transmission of the incident wave. This is called the *Ramsauer-Townsend effect*, and has been observed experimentally.

## 7.9 Resonances

There is a significant exception to the independence of the cross-section on energy. Suppose that the quantity  $\sqrt{2m|V_0|a^2/\hbar^2}$  is slightly less than  $\pi/2$ . As the incident energy increases,  $k'a$ , which is given by Eq. (7.115), can reach the value  $\pi/2$ . In this case,  $\tan k'a$  becomes infinite, so we can no longer assume that the right-hand side of Eq. (7.110) is small. In fact, at the value of the incident energy when  $k'a = \pi/2$ , it follows from Eq. (7.110) that  $ka + \delta_0 = \pi/2$ , or  $\delta_0 \simeq \pi/2$  (since we are assuming that  $ka \ll 1$ ). This implies that

$$\sigma_{\text{total}} = \frac{4\pi}{k^2} \sin^2 \delta_0 = 4\pi a^2 \left( \frac{1}{k^2 a^2} \right). \quad (7.117)$$

Note that the cross-section now depends on the energy. Furthermore, the magnitude of the cross-section is much larger than that given in Eq. (7.114) for  $k'a \neq \pi/2$  (since  $ka \ll 1$ ).

The origin of this rather strange behaviour is quite simple. The condition

$$\sqrt{\frac{2m|V_0|a^2}{\hbar^2}} = \frac{\pi}{2} \quad (7.118)$$

is equivalent to the condition that a spherical well of depth  $V_0$  possesses a bound state at zero energy. Thus, for a potential well which satisfies the above equation, the energy of the scattering system is essentially the same as the energy of the bound state. In this situation, an incident particle would like to form a bound state in the potential well. However, the bound state is not stable, since the system has a small positive energy. Nevertheless, this sort of *resonance scattering* is best understood as the capture of an incident particle to form a metastable bound state, and the subsequent decay of the bound state and release of the particle. The cross-section for resonance scattering is generally far higher than that for non-resonance scattering.

We have seen that there is a resonant effect when the phase-shift of the s-wave takes the value  $\pi/2$ . There is nothing special about the  $l = 0$  partial wave, so it is reasonable to assume that there is a similar resonance when the phase-shift of the  $l$ th partial wave is  $\pi/2$ . Suppose that  $\delta_l$  attains the value  $\pi/2$  at the incident energy  $E_0$ , so that

$$\delta_l(E_0) = \frac{\pi}{2}. \quad (7.119)$$

Let us expand  $\cot \delta_l$  in the vicinity of the resonant energy:

$$\cot \delta_l(E) = \cot \delta_l(E_0) + \left( \frac{d \cot \delta_l}{dE} \right)_{E=E_0} (E - E_0) + \dots \quad (7.120)$$

$$= - \left( \frac{1}{\sin^2 \delta_l} \frac{d\delta_l}{dE} \right)_{E=E_0} (E - E_0) + \dots \quad (7.121)$$

Defining

$$\left( \frac{d\delta_l(E)}{dE} \right)_{E=E_0} = \frac{2}{\Gamma}, \quad (7.122)$$

we obtain

$$\cot \delta_l(E) = -\frac{2}{\Gamma} (E - E_0) + \dots \quad (7.123)$$

Recall, from Eq. (7.80), that the contribution of the  $l$ th partial wave to the scattering cross-section is

$$\sigma_l = \frac{4\pi}{k^2} (2l + 1) \sin^2 \delta_l = \frac{4\pi}{k^2} (2l + 1) \frac{1}{1 + \cot^2 \delta_l}. \quad (7.124)$$

Thus,

$$\sigma_l \simeq \frac{4\pi}{k^2} (2l + 1) \frac{\Gamma^2/4}{(E - E_0)^2 + \Gamma^2/4}. \quad (7.125)$$

This is the famous *Breit-Wigner* formula. The variation of the partial cross-section  $\sigma_l$  with the incident energy has the form of a classical resonance curve. The quantity  $\Gamma$  is the width of the resonance (in energy). We can interpret the Breit-Wigner formula as describing the absorption of an incident particle to form a metastable state, of energy  $E_0$ , and lifetime  $\tau = \hbar/\Gamma$  (see Sect. 6.17).