

## BASIC BOOKS IN SCIENCE

– a Series of books that start *at the beginning*

**Book 12** DRAFT VERSION 23/01/2012

# Quantum mechanics of many-particle systems: atoms, molecules – and more

**Roy McWeeny**

Professore Emerito di Chimica Teorica, Università di Pisa, Pisa (Italy)

The Series is maintained, with regular updating and improvement, at  
<http://www.learndev.org/ScienceWorkBooks.html>  
and the books may be downloaded entirely free of charge.



This book is licensed under a Creative Commons  
Attribution-ShareAlike 3.0 Unported License.

(Last updated 23 January 2012)

# Chapter 1

## The problem – and how to deal with it

### 1.1 From one particle to many

Book 11, on the principles of quantum mechanics, laid the foundations on which we hope to build a rather complete theory of the structure and properties of all the matter around us; but how can we do it? So far, the most complicated system we have studied has been one atom of Hydrogen, in which a single electron moves in the central field of a heavy nucleus (considered to be at rest). And even that was mathematically difficult: the Schrödinger equation which determines the allowed **stationary states**, in which the energy does not change with time, took the form of a partial differential equation in three **position variables**  $x, y, z$ , of the electron, relative to the nucleus. If a second electron is added and its interaction with the first is included, the corresponding Schrödinger equation cannot be solved in ‘closed form’ (i.e. in terms of known mathematical functions). But Chemistry recognizes more than a 100 atoms, in which the nucleus has a positive charge  $Ze$  and is surrounded by  $Z$  electrons each with negative charge  $-e$ .

Furthermore, matter is not composed only of free atoms: most of the atoms ‘stick together’ in more elaborate structures called **molecules**, as will be remembered from Book 5. From a few atoms of the most common **chemical elements**, an enormous number of molecules may be constructed – including the ‘molecules of life’, which may contain many thousands of atoms arranged in a way that allows them to carry the ‘genetic code’ from one generation to the next (the subject of Book 9). At first sight it would seem impossible to achieve any understanding of the material world, at the level of the particles out of which it is composed. To make any progress at all, we have to stop looking for mathematically exact solutions of the Schrödinger equation and see how far we can get with good *approximate* wave functions, often starting from simplified *models* of the systems we are studying. The next few Sections will show how this can be done, without trying to be too complete (many whole books have been written in this field) and skipping proofs whenever the mathematics becomes difficult.

The first three chapters of Book 11 introduced most of the essential ideas of Quantum

Mechanics, together with the mathematical tools for getting you started on the further applications of the theory. You'll know, for example, that a single particle moving somewhere in 3-dimensional space may be described by a wave function  $\Psi(x, y, z)$  (a function of the three coordinates of its position) and that this is just one special way of *representing* a **state vector**. If we want to talk about some observable property of the particle, such as its energy  $E$  or a momentum component  $p_x$ , which we'll denote here by  $X$  – whatever it may stand for – we first have to set up an associated **operator**<sup>1</sup>  $\mathbf{X}$ . You'll also know that an operator like  $\mathbf{X}$  works in an abstract **vector space**, simply by sending one vector into another. In Chapter 2 of Book 11 you first learnt how such operators could be defined and used to predict the average or 'expectation' value  $\bar{X}$  that would be obtained from a large number of observations on a particle in a state described by the state vector  $\Psi$ .

In Schrödinger's form of quantum mechanics (Chapter 3) the 'vectors' are replaced by *functions* but we often use the same terminology: the 'scalar product' of two functions being defined (with Dirac's 'angle-bracket' notation) as  $\langle \Psi_1 | \Psi_2 \rangle = \int \Psi_1^*(x, y, z) \Psi_2 dx dy dz$ . With this notation we often write the expectation value  $\bar{X}$  as

$$\bar{X} = \langle X \rangle = \langle \Psi | \mathbf{X} \Psi \rangle, \quad (1.1)$$

which is a Hermitian scalar product of the 'bra-vector'  $\langle \Psi |$  and the 'ket-vector'  $|\mathbf{X}\Psi\rangle$  – obtained by letting the operator  $\mathbf{X}$  work on the  $\Psi$  that stands on the right in the scalar product. Here it is assumed that the state vector is *normalized* to unity:  $\langle \Psi | \Psi \rangle = 1$ . Remember also that the same scalar product may be written with the *adjoint* operator,  $\mathbf{X}^\dagger$ , working on the left-hand  $\Psi$ . Thus

$$\bar{X} = \langle X \rangle = \langle \mathbf{X}^\dagger \Psi | \Psi \rangle. \quad (1.2)$$

This is the property of **Hermitian symmetry**. The operators associated with observables are *self*-adjoint, or 'Hermitian', so that  $\mathbf{X}^\dagger = \mathbf{X}$ .

In Schrödinger's form of quantum mechanics (Chapter 3 of Book 11)  $\mathbf{X}$  is usually represented as a partial differential operator, built up from the coordinates  $x, y, z$  and the differential operators

$$\mathbf{p}_x = \frac{\hbar}{i} \frac{\partial}{\partial x}, \quad \mathbf{p}_y = \frac{\hbar}{i} \frac{\partial}{\partial y}, \quad \mathbf{p}_z = \frac{\hbar}{i} \frac{\partial}{\partial z}, \quad (1.3)$$

which work on the wave *function*  $\Psi(x, y, z)$ .

## 1.2 The eigenvalue equation – as a variational condition

As we've given up on the idea of calculating wave functions and energy levels accurately, by directly solving Schrödinger's equation  $\mathbf{H}\Psi = E\Psi$ , we have to start thinking about

---

<sup>1</sup>Remember that a special typeface has been used for operators, vectors and other non-numerical quantities.

possible ways of getting fair approximations. To this end, let's go back to first principles – as we did in the early chapters of Book 11

The expectation value given in (1.1) would be obtained experimentally by repeating the measurement of  $X$  a large number of times, always starting from the system in state  $\Psi$ , and recording the actual results  $X_1, X_2, \dots$  etc. – which may be found  $n_1$  times,  $n_2$  times, and so on, all scattered around their average value  $\bar{X}$ . The fraction  $n_i/N$  gives the **probability**  $p_i$  of getting the result  $X_i$ ; and in terms of probabilities it follows that

$$\bar{X} = \langle X \rangle = p_1 X_1 + p_2 X_2 \dots + p_i X_i + \dots + p_N X_N = \sum_i p_i X_i. \quad (1.4)$$

Now it's much easier to calculate an expectation value, using (1.1), than it is to solve an enormous partial differential equation; so we look for some kind of condition on  $\Psi$ , involving only an expectation value, that will be satisfied when  $\Psi$  is a solution of the equation  $\mathbf{H}\Psi = E\Psi$ .

The obvious choice is to take  $\mathbf{X} = \mathbf{H} - E\mathbf{I}$ , where  $\mathbf{I}$  is the identity operator which leaves any operand unchanged, for in that case

$$\mathbf{X}\Psi = \mathbf{H}\Psi - E\Psi \quad (1.5)$$

and the state vector  $\mathbf{X}\Psi$  is zero only when the Schrödinger equation is satisfied. The test for this is simply that the vector has zero length:

$$\langle \mathbf{X}\Psi | \mathbf{X}\Psi \rangle = 0. \quad (1.6)$$

In that case,  $\Psi$  may be one of the eigenvectors of  $\mathbf{H}$ , e.g.  $\Psi_i$  with eigenvalue  $E_i$ , and the last equation gives  $\mathbf{H}\Psi_i = E_i\Psi_i$ . On taking the scalar product with  $\Psi_i$ , from the left, it follows that  $\langle \Psi_i | \mathbf{H} | \Psi_i \rangle = E_i \langle \Psi_i | \Psi_i \rangle$  and for eigenvectors normalized to unity the energy expectation value coincides with the definite eigenvalue.

Let's move on to the case where  $\Psi$  is *not* an eigenvector of  $\mathbf{H}$  but rather an arbitrary vector, which can be expressed as a mixture of a complete set of *all* the eigenvectors  $\{\Psi_i\}$  (generally infinite), with numerical 'expansion coefficients'  $c_1, c_2, \dots, c_i, \dots$ . Keeping  $\Psi$  (without subscript) to denote the arbitrary vector, we put

$$\Psi = c_1\Psi_1 + c_2\Psi_2 + \dots = \sum_i c_i\Psi_i \quad (1.7)$$

and use the general properties of eigenstates (Section 3.6 of Book 11) to obtain a general expression for the expectation value of the energy in state (1.7), which may be normalized so that  $\langle \Psi | \Psi \rangle = 1$ .

Thus, substitution of (1.7) gives

$$\bar{E} = \langle \Psi | \mathbf{H} | \Psi \rangle = \langle (\sum_i c_i\Psi_i) | \mathbf{H} | (\sum_j c_j\Psi_j) \rangle = \sum_{i,j} c_i^* c_j \langle \Psi_i | \mathbf{H} | \Psi_j \rangle$$

and since  $\mathbf{H}\Psi_i = E_i\Psi_i$ , while

$$\langle \Psi_i | \Psi_j \rangle = \delta_{ij} \quad (= 1, \text{ for } i = j; = 0 \text{ for } i \neq j),$$

this becomes

$$\bar{E}\langle\Psi|\mathbf{H}|\Psi\rangle = |c_1|^2 E_1 + |c_2|^2 E_2 + \dots = \sum_i |c_i|^2 E_i. \quad (1.8)$$

Similarly, the squared length of the normalized  $\Psi$  becomes

$$\langle\Psi|\Psi\rangle = |c_1|^2 + |c_2|^2 + \dots = \sum_i |c_i|^2 = 1. \quad (1.9)$$

Now suppose we are interested in the state of *lowest* energy, the ‘ground’ state, with  $E_1$  less than any of the others. In that case it follows from the last two equations that

$$\begin{aligned} \langle\Psi|\mathbf{H}|\Psi\rangle - E_1 &= |c_1|^2 E_1 + |c_2|^2 E_2 + \dots \\ &\quad - |c_1|^2 E_1 - |c_2|^2 E_1 + \dots \\ &= 0 + |c_2|^2 (E_2 - E_1) + \dots \end{aligned}$$

All the quantities on the right-hand side are essentially *positive*:  $|c_i|^2 > 0$  for all  $i$  and  $E_i - E_1 > 0$  because  $E_1$  is the smallest of all the eigenvalues. It follows that

Given an arbitrary state vector  $\Psi$ , which may be chosen so that  $\langle\Psi|\Psi\rangle = 1$ , the energy expectation value

$$\bar{E} = \langle\Psi|\mathbf{H}|\Psi\rangle / \langle\Psi|\Psi\rangle$$

must be greater than or equal to the lowest eigenvalue,  $E_1$ , of the Hamiltonian operator  $\mathbf{H}$

(1.10)

Here the normalization factor  $\langle\Psi|\Psi\rangle$  has been left in the denominator of  $\bar{E}$  and the result then remains valid even when  $\Psi$  is not normalized (check it!). This is a famous theorem and provides a basis for the **variation method** of calculating approximate eigenstates. In Schrödinger’s formulation of quantum mechanics, where  $\Psi$  is represented by a wave function such as  $\Psi(x, y, z)$ , one can start from any ‘trial’ function that ‘looks roughly right’ and contains adjustable parameters. By calculating a ‘variational energy’  $\langle\Psi|\mathbf{H}|\Psi\rangle$  and varying the parameters until you can’t find a lower value of this quantity you will know you have found the best approximation you can get to the ground-state energy  $E_1$  and corresponding wave function. To do better you’ll have to use a trial  $\Psi$  of different functional form.

As a first example of using the variation method we’ll get an approximate wave function for the ground state of the hydrogen atom. In Book 11 (Section 6.2) we got the energy and wave function for the ground state of an electron in a hydrogen-like atom, with nuclear charge  $Ze$ , placed at the origin. They were, using atomic units,

$$E_{1s} = -\frac{1}{2}Z^2, \quad \phi_{1s} = N_{1s}e^{-Zr},$$

where the normalizing factor is  $N_{1s} = \pi^{-1/2} Z^{3/2}$ .

We'll now try a gaussian approximation to the 1s orbital, calling it  $\phi_{1s} = N \exp -\alpha r^2$ , which correctly goes to zero for  $r \rightarrow \infty$  and to  $N$  for  $r = 0$ ; and we'll use this function (calling it  $\phi$  for short) to get an approximation to the ground state energy  $\bar{E} = \langle \phi | \mathbf{H} | \phi \rangle$ . The first step is to evaluate the new normalizing factor and this gives a useful example of the mathematics needed.

**Example 1.1** A gaussian approximation to the 1s orbital.

---

To get the normalizing factor  $N$  we must set  $\langle \phi | \phi \rangle = 1$ . Thus

$$\langle \phi | \phi \rangle = N^2 \int_0^\infty \exp(-2\alpha r^2) (4\pi r^2) dr, \quad (\text{A})$$

the volume element being that of a spherical shell of thickness  $dr$ .

To do the integration we can use the formula (very useful whenever you see a gaussian!) given in Example 5.2 of Book 11:

$$\int_{-\infty}^{+\infty} \exp(-ps^2 - qs) ds = \sqrt{\frac{\pi}{p}} \exp\left(\frac{q^2}{4p}\right),$$

which holds for any values (real or complex) of the constants  $p, q$ . Since the function we're integrating is symmetrical about  $r = 0$  and is needed only for  $q = 0$  we'll use the basic integral

$$I_0 = \int_0^\infty e^{-pr^2} dr = \frac{1}{2} \sqrt{\pi} p^{-1/2}. \quad (\text{B})$$

Now let's differentiate both sides of equation (B) with respect to the parameter  $p$ , just as if it were an ordinary variable (even though it is inside the integrand and really one should prove that this is OK). On the left we get (look back at Book 3 if you need to)

$$\frac{dI_0}{dp} = - \int_0^\infty r^2 e^{-pr^2} dr = -I_1,$$

where we've called the new integral  $I_1$  as we got it from  $I_0$  by doing *one* differentiation. On differentiating the right-hand side of (B) we get

$$\frac{d}{dp} \left( \frac{1}{2} \sqrt{\pi} p^{-1/2} \right) = \frac{1}{2} \sqrt{\pi} \left( -\frac{1}{2} p^{-3/2} \right) = -\frac{1}{4} \sqrt{\pi} / p \sqrt{p}.$$

But the two results must be equal (if two functions of  $p$  are identically equal their slopes will be equal at all points) and therefore

$$I_1 = \int_0^\infty r^2 e^{-pr^2} dr = \frac{1}{2} \sqrt{\pi} \left( \frac{1}{2} p^{-3/2} \right) = \frac{1}{4} \sqrt{\pi} / p \sqrt{p},$$

where the integral  $I_1$  on the left is the one we need as it appears in (A) above. On using this result in (A) and remembering that  $p = 2\alpha$  it follows that

$$N^2 = \left( \frac{p}{\pi} \right)^{3/2} = \left( \frac{2\alpha}{\pi} \right)^{3/2}.$$


---

Example 1.1 has given the square of the normalizing factor,

$$N^2 = \left( \frac{2\alpha}{\pi} \right)^{3/2}, \quad (1.11)$$

which will appear in all matrix elements.

Now we turn to the expectation value of the energy  $\bar{E} = \langle \phi | \mathbf{H} | \phi \rangle$ . Here the Hamiltonian will be

$$\mathbf{H} = \mathbf{T} + \mathbf{V} = -\frac{1}{2}\nabla^2 - Z/r$$

and since  $\phi$  is a function of only the radial distance  $r$  we can use the expression for  $\nabla^2$  obtained in Example 4.8 of Book 11, namely

$$\nabla^2 \equiv \frac{2}{r} \frac{d}{dr} + \frac{d^2}{dr^2}.$$

On denoting the 1-electron Hamiltonian by  $\mathbf{h}$  (we'll keep  $\mathbf{H}$  for many-electron systems) we then find  $\mathbf{h}\phi = -(Z/r)\phi - (1/r)(d\phi/dr) - \frac{1}{2}(d^2\phi/dr^2)$  and

$$\langle \phi | \mathbf{h} | \phi \rangle = -Z\langle \phi | (1/r) | \phi \rangle - \langle \phi | (1/r)(d\phi/dr) \rangle - \frac{1}{2}\langle \phi | (d^2\phi/dr^2) \rangle. \quad (1.12)$$

Let's evaluate the three terms on the right one at a time, taking the first in the next example.

### Example 1.2 Expectation value of the potential energy

---

We require  $\langle \phi | \mathbf{V} | \phi \rangle = -Z\langle \phi | (1/r) | \phi \rangle$ , where  $\phi$  is the normalized function  $\phi = Ne^{-\alpha r^2}$ . Thus

$$\langle \phi | \mathbf{V} | \phi \rangle = -ZN^2 \int_0^\infty e^{-\alpha r^2} (1/r) e^{-\alpha r^2} (4\pi r^2) dr,$$

which looks like the integral at "A" in Example 1.1 – except for the factor  $(1/r)$ . The new integral we need is  $4\pi I'_0$ , where

$$I'_0 = \int_0^\infty r e^{-pr^2} dr \quad (p = 2\alpha)$$

and the factor  $r$  spoils everything – we can no longer get  $I'_0$  from  $I_0$  by differentiating, as in Example 1.1, for that would bring down a factor  $r^2$ . However, we can use another of the tricks you learnt in Chapter 4 of Book 3. (If you've forgotten all that you'd better read it again!) It comes from 'changing the variable' by putting  $r^2 = u$  and expressing  $I'_0$  in terms of  $u$ . In that case we can use the formula you learnt long ago:

$$I'_0 = \int_0^\infty (u^{1/2} e^{-pu}) \frac{dr}{du} du.$$

To see how this works with  $u = r^2$  we note that, since  $r = u^{1/2}$ ,  $dr/du = \frac{1}{2}u^{-1/2}$ ; so in terms of the new variable  $u$

$$I'_0 = \int_0^\infty (u^{1/2} e^{-pu}) (\frac{1}{2}u^{-1/2}) du = \frac{1}{2} \int_0^\infty e^{-pu} du.$$

The integral is a simple standard integral and when the limits are put in it gives (check it!)  $I'_0 = \frac{1}{2}[-e^{-pu}/p]_0^\infty = \frac{1}{2}(1/p)$ .

---

From Example 1.2 it follows that

$$\langle \phi | \mathbf{V} | \phi \rangle = -4\pi ZN^2 \frac{1}{2} \left[ -\frac{e^{-pu}}{p} \right]_0^\infty = -2\pi ZN^2/p. \quad (1.13)$$

And now you know how to do the integrations you should be able to get the remaining terms in the expectation value of the Hamiltonian  $\mathbf{h}$ . They come from the kinetic energy operator  $\mathbf{T} = -\frac{1}{2}\nabla^2$ , as in the next example.

---

**Example 1.3** Expectation value of the kinetic energy

---

We require  $\bar{T} = \langle \phi | \mathbf{T} | \phi \rangle$  and from (1.12) this is seen to be the sum of two terms. The first one involves the first derivative of  $\phi$ , which becomes (on putting  $-\alpha r^2 = u$  in  $\phi = Ne^{-\alpha r^2}$ )

$$(d\phi/dr) = (d\phi/du)(du/dr) = N(e^{-u})(-2r\alpha) = -2N\alpha r e^{-\alpha r^2}.$$

On using this result, multiplying by  $\phi$  and integrating, it gives a contribution to  $\bar{T}$  of

$$\bar{T}_1 = \langle \phi | -\frac{1}{r} \frac{d}{dr} | \phi \rangle = N^2 p \int_0^\infty \frac{1}{r} r e^{-pr^2} (4\pi r^2) dr = 4\pi N^2 p \int_0^\infty e^{-pr^2} (r^2) dr = 4\pi N^2 p I_1$$

– the integral containing a factor  $r^2$  in the integrand (just like  $I_1$  in Example 1.1).

The second term in  $\bar{T}$  involves the second derivative of  $\phi$ ; and we already found the first derivative as  $d\phi/dr = -Npr e^{-\alpha r^2}$ . So differentiating once more (do it!) you should find

$$(d^2\phi/dr^2) = -Npe^{-\alpha r^2} - Npr(-pre^{-\alpha r^2}).$$

(check it by differentiating  $-2N\alpha r e^{-\alpha r^2}$ ).

On using this result we obtain (again with  $p = 2\alpha$ )

$$\bar{T}_2 = \langle \phi | -\frac{1}{2} \frac{d^2}{dr^2} | \phi \rangle = -\frac{1}{2} N^2 4\pi p \int_0^\infty r^2 e^{-pr^2} dr + \frac{1}{2} N^2 4\pi p^2 \int_0^\infty r^4 e^{-pr^2} dr = 2\pi N^2 (-p^2 I_2 + p I_1).$$

When the first-derivative term is added, namely  $4\pi N^2 p I_1$ , we obtain the expectation value of the kinetic energy as

$$4\pi N^2 p I_1 + 2\pi N^2 (p^2 I_2 - p I_1) = 2\pi N^2 (-p^2 I_2 + 3p I_1)$$

The two terms in the final parentheses are

$$2\pi N^2 p^2 I_2 = 2\pi N^2 \frac{3}{8} \sqrt{\frac{\pi}{2\alpha}}, \quad 2\pi N^2 p I_1 = 2\pi N^2 \frac{1}{4} \sqrt{\frac{\pi}{2\alpha}}$$

and remembering that  $p = 2\alpha$  and that  $N^2$  is given in (1.1), substitution gives the result  $\bar{T} = \bar{T}_1 + \bar{T}_2 = 2\pi N^2 (3/8) \sqrt{\pi/2\alpha}$ .

---

The expectation value of the KE is thus, noting that  $2\pi N^2 = 2p(p/\pi)^{1/2}$ ,

$$\langle \phi | \mathbf{T} | \phi \rangle = \frac{5}{8} \sqrt{\frac{\pi}{2\alpha}} \times 2\pi N^2 = \frac{3p}{4}. \quad (1.14)$$

Finally, the expectation energy with a trial wave function of the form  $\phi = Ne^{-\alpha r^2}$  becomes, on adding the PE term from (1.13),  $-2\pi Z N^2 (1/2\alpha)$

$$\bar{E} = \frac{3\alpha}{2} - 2Z \left( \frac{2}{\pi} \right)^{1/2} \alpha^{1/2}. \quad (1.15)$$

There is only one variable parameter  $\alpha$  and to get the best approximate ground state function of Gaussian form we must adjust  $\alpha$  until  $\bar{E}$  reaches a minimum value. The value of  $\bar{E}$  will be *stationary* (maximum, minimum, or turning point) when  $d\bar{E}/d\alpha = 0$ ; so we must differentiate and set the result equal to zero.

### Example 1.4 A first test of the variation method

---

Let's put  $\sqrt{\alpha} = \mu$  and write (1.15) in the form

$$\bar{E} = A\mu^2 - B\mu \quad (A = 3/2, \quad B = 2Z\sqrt{2/\pi})$$

which makes it look a bit simpler.

We can then vary  $\mu$ , finding  $d\bar{E}/d\mu = 2A\mu - B$ , and this has a stationary value when  $\mu = B/2A$ . On substituting for  $\mu$  in the energy expression, the stationary value is seen to be

$$\bar{E}_{\min} = A(B^2/4A^2) - B(B/2A),$$

where the two terms are the kinetic energy  $\bar{T} = \frac{1}{2}(B^2/2A)$  and the potential energy  $\bar{V} = (B^2/2A)$ . The total energy  $\bar{E}$  at the stationary point is thus the sum KE + PE:

$$\bar{E} = \frac{1}{2}(B^2/2A) - (B^2/2A) = -\frac{1}{2}(B^2/2A) = -\bar{T}$$

and this is an energy *minimum*, because  $d^2\bar{E}/d\mu^2 = 2A$  –which is positive.

The fact that the minimum energy is exactly  $-1 \times$  the kinetic energy is no accident: it is a consequence of the **virial theorem**, about which you'll hear more later. For the moment, we note that for a hydrogen-like atom the 1-term gaussian wave function gives a best approximate energy  $\bar{E}_{\min} = -\frac{1}{2}(2Z\sqrt{2/\pi})^2/3 = -4Z^2/3\pi$ .

---

Example 1.4 gives the result  $-0.42442 Z^2$ , where all energies are in units of  $e_H$ .

For the hydrogen atom, with  $Z = 1$ , the exact ground state energy is  $-\frac{1}{2}e_H$ , as we know from Book 11. In summary then, the conclusion from the Example is that a gaussian function gives a very poor approximation to the hydrogen atom ground state, the estimate  $-0.42442 e_H$  being in error by about 15%. The next Figure shows why:

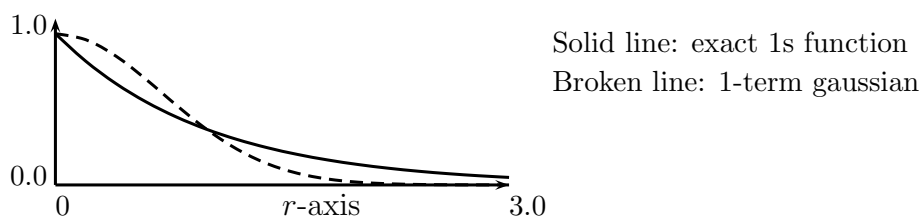


Figure 1.1 Comparison of exponential and gaussian functions

$\phi(r)$  fails to describe the sharp cusp when  $r \rightarrow 0$  and also goes to zero much too rapidly when  $r$  is large.

Of course we could get the accurate energy  $E_1 = -\frac{1}{2}e_H$  and the corresponding wave function  $\phi_1$ , by using a trial function of exponential form  $\exp -ar$  and varying the parameter  $a$  until the approximate energy reaches a minimum value. But here we'll try another approach, taking a mixture of two gaussian functions, one falling rapidly to zero as  $r$  increases and the other falling more slowly: in that way we can hope to correct the main defects in the 1-term approximation.

### Example 1.5 A 2-term gaussian approximation

---

With a trial function of the form  $\phi = A \exp -ar^2 + B \exp -br^2$  there are *three* parameters that can be independently varied,  $a, b$  and the *ratio*  $c = B/A$  – a fourth parameter not being necessary if we’re looking for a normalized function (can you say why?). So we’ll use instead a 2-term function  $\phi = \exp -ar^2 + c \exp -br^2$ .

From the previous Examples 1.1-1.3, it’s clear how you can evaluate all the integrals you need in calculating  $\langle \phi | \phi \rangle$  and the expectation values  $\langle \phi | \mathbf{V} | \phi \rangle$ ,  $\langle \phi | \mathbf{V} | \phi \rangle$ ; all you’ll need to change will be the parameter values in the integrals.

Try to work through this by yourself, without doing the variation of all three values to find the minimum value of  $\bar{E}$ . (Until you’ve learnt to use a computer that’s much too long a job! But you may like to know the result: the ‘best’ values of  $a, b, c$  are  $a = 1.32965$ ,  $b = 0.20146$ ,  $c = 0.72542$  and the best approximation to  $E_{1s}$  then comes out as  $\bar{E} = -0.4858Z^2e_H$ . This compares with the one-term approximation  $\bar{E} = -0.4244Z^2e_H$ ; the error is now reduced from about 15% to less than 3%.

---

The approximate wave function obtained in Example 1.5 is plotted in Figure 1.2 and again compared with the exact 1s function. (The functions are not normalized, being shifted vertically to show how well the cusp behaviour is corrected. Normalization improves the agreement in the middle range.)

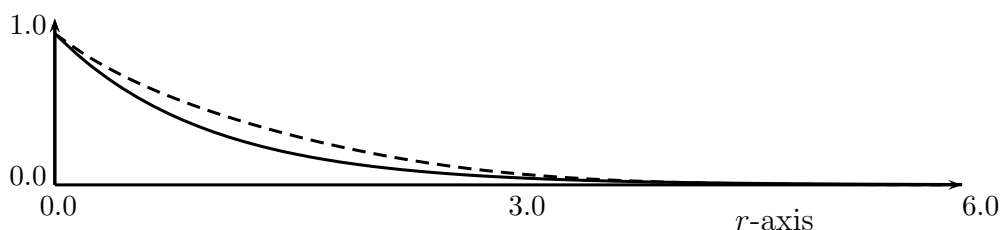


Figure 1.2 A 2-term gaussian approximation (broken line) to the hydrogen atom 1s function (solid line)

This Example suggests another form of the variation method, which is both easier to apply and much more powerful. We study it in the next Section, going back to the general case, where  $\Psi$  denotes any kind of wave function, expanded in terms of eigenfunctions  $\Psi_i$ .

## 1.3 The linear variation method

Instead of building a variational approximation to the wave function  $\Psi$  out of only two terms we may use as many as we please, taking in general

$$\Psi = c_1\Psi_1 + c_2\Psi_2 + \dots + c_N\Psi_N, \quad (1.16)$$

where (with the usual notation) the functions  $\{\Psi_i (i = 1, 2, \dots, N)\}$  are ‘fixed’ and we vary only the coefficients  $c_i$  in the linear combination: this is called a “linear variation function”

and it lies at the root of nearly all methods of constructing atomic and molecular wave functions.

From the variation theorem (1.10) we need to calculate the expectation energy  $\bar{E} = \langle \Psi | \mathbf{H} | \Psi \rangle / \langle \Psi | \Psi \rangle$ , which we know will give an upper bound to the lowest exact eigenvalue  $E_1$  of the operator  $\mathbf{H}$ . We start by putting this expression in a convenient matrix form: you used matrices a lot in Book 11, ‘representing’ the operator  $\mathbf{H}$  by a *square array* of numbers  $\mathbf{H}$  with  $H_{ij} = \langle \Psi_i | \mathbf{H} | \Psi_j \rangle$  (called a ‘matrix element’) standing at the intersection of the  $i$ th row and  $j$ th column; and collecting the coefficients  $c_i$  in a single column  $\mathbf{c}$ . (Look back at Book 11 Chapter 7 if you need reminding of the rules for using matrices.)

In matrix notation the expectation energy becomes

$$\bar{E} = \frac{\mathbf{c}^\dagger \mathbf{H} \mathbf{c}}{\mathbf{c}^\dagger \mathbf{M} \mathbf{c}}, \quad (1.17)$$

where  $\mathbf{c}^\dagger$  (the ‘Hermitian transpose’ of  $\mathbf{c}$ ) denotes the *row* of coefficients ( $c_1^* c_2^*, \dots c_N^*$ ) and  $\mathbf{M}$  (the ‘metric matrix’) looks like  $\mathbf{H}$  except that  $H_{ij}$  is replaced by  $M_{ij} = \langle \Psi_i | \Psi_j \rangle$ , the scalar product or ‘overlap’ of the two functions. This allows us to use sets of functions that are neither normalized to unity nor orthogonal – with no additional complication.

The best approximate state function (1.11) we can get is obtained by minimizing  $\bar{E}$  to make it as close as possible to the (unknown!) ground state energy  $E_1$ , and to do this we look at the effect of a small variation  $\mathbf{c} \rightarrow \mathbf{c} + \delta \mathbf{c}$ : if we have reached the minimum,  $\bar{E}$  will be *stationary*, with the corresponding change  $\delta \bar{E} = 0$ .

In the variation  $\mathbf{c} \rightarrow \mathbf{c} + \delta \mathbf{c}$ ,  $\bar{E}$  becomes

$$\bar{E} + \delta \bar{E} = \frac{\mathbf{c}^\dagger \mathbf{H} \mathbf{c} + \mathbf{c}^\dagger \mathbf{H} \delta \mathbf{c} + \delta \mathbf{c}^\dagger \mathbf{H} \mathbf{c} + \dots}{\mathbf{c}^\dagger \mathbf{M} \mathbf{c} + \mathbf{c}^\dagger \mathbf{M} \delta \mathbf{c} + \delta \mathbf{c}^\dagger \mathbf{M} \mathbf{c} + \dots},$$

where second-order terms that involve products of  $\delta$ -quantities have been dropped (vanishing in the limit  $\delta \mathbf{c} \rightarrow 0$ ).

The denominator in this expression can be re-written, since  $\mathbf{c}^\dagger \mathbf{M} \mathbf{c}$  is just a number, as

$$\mathbf{c}^\dagger \mathbf{M} \mathbf{c} [1 + (\mathbf{c}^\dagger \mathbf{M} \mathbf{c})^{-1} (\mathbf{c}^\dagger \mathbf{M} \delta \mathbf{c} + \delta \mathbf{c}^\dagger \mathbf{M} \mathbf{c})]$$

and the part in square brackets has an inverse (to first order in small quantities)

$$1 - (\mathbf{c}^\dagger \mathbf{M} \mathbf{c})^{-1} (\mathbf{c}^\dagger \mathbf{M} \delta \mathbf{c} + \delta \mathbf{c}^\dagger \mathbf{M} \mathbf{c}).$$

On putting this result in the expression for  $\bar{E} + \delta \bar{E}$  and re-arranging a bit (do it!) you’ll find

$$\bar{E} + \delta \bar{E} = \bar{E} + \mathbf{c}^\dagger \mathbf{M} \mathbf{c}^{-1} [(\mathbf{c}^\dagger \mathbf{H} \delta \mathbf{c} + \delta \mathbf{c}^\dagger \mathbf{H} \mathbf{c}) - \bar{E} (\mathbf{c}^\dagger \mathbf{M} \delta \mathbf{c} + \delta \mathbf{c}^\dagger \mathbf{M} \mathbf{c})].$$

It follows that the first-order variation is given by

$$\delta \bar{E} = \mathbf{c}^\dagger \mathbf{M} \mathbf{c}^{-1} [(\mathbf{c}^\dagger \mathbf{H} - \bar{E} \mathbf{c}^\dagger \mathbf{M}) \delta \mathbf{c} + \delta \mathbf{c}^\dagger (\mathbf{H} \mathbf{c} - \bar{E} \mathbf{M} \mathbf{c})]. \quad (1.18)$$

The two terms in (1.18) are complex conjugate, giving a real result which will vanish only when each is zero.

The condition for a stationary value thus reduces to a *matrix* eigenvalue equation

$$\mathbf{Hc} = \bar{E}\mathbf{M}\mathbf{c}. \quad (1.19)$$

To get the *minimum* value of  $\bar{E}$  we therefore take the lowest eigenvalue; and the corresponding ‘best approximation’ to the wave function  $\Psi \approx \Psi_1$  will follow on solving the simultaneous equations equivalent to (1.19), namely

$$\sum_j H_{ij}c_j = \bar{E} \sum_j M_{ij}c_j \quad (\text{all } i). \quad (1.20)$$

This is essentially what we did in Example 1.2, where the linear coefficients  $c_1, c_2$  gave a best approximation when they satisfied the two simultaneous equations

$$\begin{aligned} (H_{11} - \bar{E}M_{11})c_1 + (H_{12} - \bar{E}M_{12})c_2 &= 0, \\ (H_{21} - \bar{E}M_{21})c_1 + (H_{22} - \bar{E}M_{22})c_2 &= 0, \end{aligned}$$

the other parameters being fixed. Now we want to do the same thing generally, using a large basis of  $N$  expansion functions  $\{\Psi_i\}$ , and to make the calculation easier it’s best to use an orthonormal set. For the case  $N = 2$ ,  $M_{11} = M_{22} = 1$  and  $M_{12} = M_{21} = 0$ , the equations then become

$$\begin{aligned} (H_{11} - \bar{E})c_1 &= -H_{12}c_2, \\ H_{21}c_1 &= -(H_{22} - \bar{E})c_2. \end{aligned}$$

Here there are three unknowns,  $\bar{E}, c_1, c_2$ . However, by dividing each side of the first equation by the corresponding side of the second, we can eliminate two of them, leaving only

$$\frac{(H_{11} - \bar{E})}{H_{21}} = \frac{H_{12}}{(H_{22} - \bar{E})}.$$

This is quadratic in  $\bar{E}$  and has two possible solutions. On ‘cross-multiplying’ it follows that  $(H_{11} - \bar{E})(H_{22} - \bar{E}) = H_{12}H_{21}$  and on solving we get lower and upper values  $\bar{E}_1$  and  $\bar{E}_2$ . After substituting either value back in the original equations, we can solve to get the *ratio* of the expansion coefficients. Normalization to make  $c_1^2 + c_2^2 = 1$  then results in approximations to the first two wave functions,  $\Psi_1$  (the ground state) and  $\Psi_2$  (a state of higher energy).

### Generalization

Suppose we want a really good approximation and use a basis containing hundreds of functions  $\Psi_i$ . The set of simultaneous equations to be solved will then be enormous; but we can see how to continue by looking at the case  $N = 3$ , where they become

$$\begin{aligned} (H_{11} - \bar{E}M_{11})c_1 + (H_{12} - \bar{E}M_{12})c_2 + (H_{13} - \bar{E}M_{13})c_3 &= 0, \\ (H_{21} - \bar{E}M_{21})c_1 + (H_{22} - \bar{E}M_{22})c_2 + (H_{23} - \bar{E}M_{23})c_3 &= 0, \\ (H_{31} - \bar{E}M_{31})c_1 + (H_{32} - \bar{E}M_{32})c_2 + (H_{33} - \bar{E}M_{33})c_3 &= 0. \end{aligned}$$

We'll again take an orthonormal set, to simplify things. In that case the equations reduce to (in matrix form)

$$\begin{pmatrix} H_{11} - \bar{E} & H_{12} & H_{13} \\ H_{21} & H_{22} - \bar{E} & H_{23} \\ H_{31} & H_{32} & H_{33} - \bar{E} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

When there were only two expansion functions we had similar equations, but with only two rows and columns in the matrices:

$$\begin{pmatrix} H_{11} - \bar{E} & H_{12} \\ H_{21} & H_{22} - \bar{E} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

And we got a solution by 'cross-multiplying' in the square matrix, which gave

$$(H_{11} - \bar{E})(H_{22} - \bar{E}) - H_{21}H_{12} = 0.$$

This is called a **compatibility condition**: it determines the only values of  $\bar{E}$  for which the equations are *compatible* (i.e. can both be solved at the same time).

In the general case, there are  $N$  simultaneous equations and the condition involves the **determinant** of the square array: thus for  $N = 3$  it becomes

$$\begin{vmatrix} H_{11} - \bar{E} & H_{12} & H_{13} \\ H_{21} & H_{22} - \bar{E} & H_{23} \\ H_{31} & H_{32} & H_{33} - \bar{E} \end{vmatrix} = 0. \quad (1.21)$$

There are many books on algebra, where you can find whole chapters on the theory of determinants, but nowadays equations like (1.16) can be solved easily with the help of a small computer. All the 'theory' you really need, was explained long ago in Book 2 (Section 6.12). So here a reminder should be enough:

Given a square matrix  $\mathbf{A}$ , with three rows and columns, its determinant can be evaluated as follows. You can start from the 11-element  $A_{11}$  and then get the determinant of the  $2 \times 2$  matrix that is left when you take away the first row and first column:

$$\begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} = A_{22}A_{33} - A_{32}A_{23}.$$

– as follows from what you did just before (1.16). What you have evaluated is called the 'co-factor' of  $A_{11}$  and is denoted by  $A^{(11)}$ .

Then move to the next element in the first row, namely  $A_{12}$ , and do the same sort of thing: take away the first row and *second* column and then get the determinant of the  $2 \times 2$  matrix that is left. This would seem to be the co-factor of  $A_{12}$ ; but in fact, whenever you move from one element in the row to the next, you have to attach a minus sign; so what you have found is  $-A^{(12)}$ .

When you've finished the row you can put together the three contributions to get

$$|\mathbf{A}| = A_{11}A^{(11)} - A_{12}A^{(12)} + A_{13}A^{(13)}$$

and you've evaluated the  $3 \times 3$  determinant!

The only reason for reminding you of all that (since a small computer can do such things much better than we can) was to show that the determinant in (1.21) will give you a

polynomial of degree 3 in the energy  $\bar{E}$ . (That is clear if you take  $\mathbf{A} = \mathbf{H} - \bar{E}\mathbf{1}$ , make the expansion, and look at the terms that arise from the product of elements on the ‘principal diagonal’, namely  $(H_{11} - \bar{E}) \times (H_{22} - \bar{E}) \times (H_{33} - \bar{E})$ . These include  $-\bar{E}^3$ .) Generally, as you can see, the expansion of a determinant like (1.16), but with  $N$  rows and columns, will contain a term of highest degree in  $\bar{E}$  of the form  $(-1)^N \bar{E}^N$ . This leads to conclusions of very great importance – as you’re just about to see.

## 1.4 Going all the way! Is there a limit?

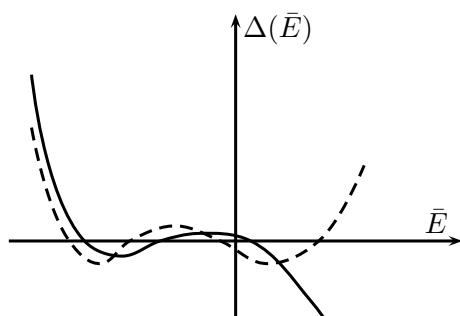
The first time you learnt anything about **eigenfunctions** and how they could be used was in Book 3 (Section 6.3). Before starting the present Section 1.4 of Book 12, you should read again what was done there. You were studying a simple differential equation, the one that describes standing waves on a vibrating string, and the solutions were sine functions (very much like the eigenfunctions coming from Schrödinger’s equation for a ‘particle in a box’, discussed in Book 11). By putting together a large number of such functions, corresponding to increasing values of the vibration frequency, you were able to get approximations to the instantaneous shape of the string for *any* kind of vibration. That was a first example of an **eigenfunction expansion**. Here we’re going to use such expansions in constructing approximate wave functions for atoms and molecules; and we’ve taken the first steps by starting from linear variation functions. What we must do now is to ask how a function of the form (1.16) can approach more and more closely an exact eigenfunction of the Hamiltonian  $\mathbf{H}$  as  $N$  is increased.

In Section 1.3 it was shown that an  $N$ -term variation function (1.16) could give an optimum approximation to the ground state wave function  $\Psi_1$ , provided the expansion coefficients  $c_i$  were chosen so as to satisfy a set of linear equations: for  $N = 3$  these took the form

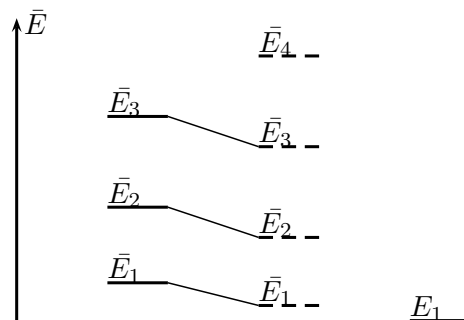
$$\begin{aligned} (H_{11} - \bar{E}M_{11})c_1 + (H_{12} - \bar{E}M_{12})c_2 + (H_{13} - \bar{E}M_{13})c_3 &= 0, \\ (H_{21} - \bar{E}M_{21})c_1 + (H_{22} - \bar{E}M_{22})c_2 + (H_{23} - \bar{E}M_{23})c_3 &= 0, \\ (H_{31} - \bar{E}M_{31})c_1 + (H_{32} - \bar{E}M_{32})c_2 + (H_{33} - \bar{E}M_{33})c_3 &= 0. \end{aligned}$$

and were compatible only when the variational energy  $\bar{E}$  satisfied the condition (1.16). There are only *three* values of  $\bar{E}$  which do so. We know that  $\bar{E}_1$  is an upper bound to the accurate lowest-energy eigenvalue  $E_1$  but what about the other two?

In general, equations of this kind are called **secular equations** and a condition like (1.16) is called a **secular determinant**. If we plot the value,  $\Delta$  say, of the determinant (having worked it out for any chosen value of  $\bar{E}$ ) against  $\bar{E}$ , we’ll get a curve something like the one in Figure 1.3; and whenever the curve crosses the horizontal axis we’ll have  $\Delta = 0$ , the compatibility condition will be satisfied and that value of  $\bar{E}$  will allow you to solve the secular equations. For *other* values you just can’t do it!



**Figure 1.3** Secular determinant  
 Solid line: for  $N = 3$   
 Broken line: for  $N = 4$



**Figure 1.4** Energy levels  
 Solid lines: for  $N = 3$   
 Broken lines: for  $N = 4$

On the far left in Fig.1.3,  $\Delta$  will become indefinitely large and positive because its expansion is a polynomial dominated by the term  $-\bar{E}^3$  and  $\bar{E}$  is negative. On the other side, where  $\bar{E}$  is positive, the curve on the far right will go off to large *negative* values. In between there will be three crossing points, showing the acceptable energy values.

Now let's look at the effect of increasing the number of basis functions by adding another,  $\Psi_4$ . The value of the secular determinant then changes and, since expansion gives a polynomial of degree 4, it will go towards  $+\infty$  for large values of  $\bar{E}$ . Figure 1.3 shows that there are now *four* crossing points on the x-axis and therefore four acceptable solutions of the secular equations. The corresponding energy levels for  $N = 3$  and  $N = 4$  are compared in Figure 1.4, where the first three are seen to go down, while one new level ( $\bar{E}_4$ ) appears at higher energy. The levels for  $N = 4$  fall in between the levels above and below for  $N = 3$  and this result is often called the “separation theorem”: it can be proved properly by studying the values of the determinant  $\Delta_N(\bar{E})$  for values of  $\bar{E}$  at the crossing points of  $\Delta_{N-1}(\bar{E})$ .

The conclusion is that, as more and more basis functions are added, the roots of the secular determinant go steadily (or ‘monotonically’) *down* and will therefore approach limiting values. The first of these,  $E_1$ , is known to be an upper bound to the exact lowest eigenvalue of  $\mathbf{H}$  (i.e. the groundstate of the system) and it now appears that the higher roots will give upper bounds to the higher ‘excited’ states. For this conclusion to be true it is necessary that the chosen basis functions form a **complete set**.

## 1.5 Complete set expansions

So far, in the last section, we've been thinking of linear variation functions in general, without saying much about the forms of the expansion functions and how they can be constructed; but for atoms and molecules they may be functions of many variables (e.g. coordinates  $x_1, y_1, z_1, x_2, y_2, z_2, x_3, \dots, z_N$  for  $N$  particles – even without including spins!). From now on we'll be dealing mainly with wave functions built up from one-particle functions, which from now on we'll denote by lower-case letters  $\{\phi_k(\mathbf{r}_i)\}$  with the index  $i$  labelling ‘Particle  $i$ ’ and  $\mathbf{r}_i$  standing for all three variables needed to indicate its position

in space (spin will be put in later); as usual the subscript on the function will just indicate which one of the whole set ( $k = 1, 2, \dots, n$ ) we mean. (It's a pity so many labels are needed, and that sometimes we have to change their names, but by now you must be getting used to the fact that you're playing a difficult game – once you're clear about what the symbols stand for the rest will be easy!)

Let's start by thinking again of the simplest case; one particle, moving in one dimension, so the particle label  $i$  is not needed and  $\mathbf{r}$  can be replaced by just one variable,  $x$ . Instead of  $\phi_k(\mathbf{r}_i)$  we can then use  $\phi_k(x)$ . We want to represent *any* function  $f(x)$  as a linear combination of these basis functions and we'll write

$$f^{(n)}(x) = c_1\phi_1(x) + c_2\phi_2(x) + \dots + c_n\phi_n(x) \quad (1.22)$$

as the ' $n$ -term approximation' to  $f(x)$ .

Our first job will be to choose the coefficients so as to get a *best* approximation to  $f(x)$  over the whole range of  $x$ -values (not just at one point). And by "the whole range" we'll mean for all  $x$  in the interval,  $(a, b)$  say, outside which the function has values that can be neglected: the range may be very small (think of the delta-function you met in Book 11) or very large (think of the interval  $(-\infty, +\infty)$  for a particle moving in free space). (When we need to show the limits of the interval we'll just use  $x = a$  and  $x = b$ .)

Generally, the curves we get on plotting  $f(x)$  and  $f^{(n)}(x)$  will differ and their difference can be measured by  $\Delta(x) = f(x) - f^{(n)}(x)$  at all points in the range. But  $\Delta(x)$  will sometimes be positive and sometimes negative. So it's no good adding these differences for all points on the curve (which will mean integrating  $\Delta(x)$ ) to get a measure of how poor the approximation is; for cancellations could lead to zero even when the curves were very different. It's really the *magnitude* of  $\Delta(x)$  that matters, or its *square* – which is always positive.

So instead let's measure the difference by  $|f(x) - f^{(n)}(x)|^2$ , at any point, and the 'total difference' by

$$D = \int_a^b \Delta(x)^2 dx = \int_a^b |f(x) - f^{(n)}(x)|^2 dx. \quad (1.23)$$

The integral gives the sum of the areas of all the strips between  $x = a$  and  $x = b$  of height  $\Delta^2$  and width  $dx$ . This quantity will measure the *error* when the whole curve is approximated by  $f^{(n)}(x)$  and we'll only get a really good fit, over the whole range of  $x$ , when  $D$  is close to zero.

The coefficients  $c_k$  should be chosen to give  $D$  its lowest possible value and you know how to do that: for a function of one variable you find a minimum value by first seeking a 'turning point' where  $(df/dx) = 0$ ; and then check that it really *is* a minimum, by verifying that  $(d^2f/dx^2)$  is positive. It's just the same here, except that we look at the variables *one at a time, keeping the others constant*. Remember too that it's the coefficients  $c_k$  that we're going to vary, not  $x$ .

Now let's put (1.17) into (1.18) and try to evaluate  $D$ . You first get (dropping the usual variable  $x$  and the limits  $a, b$  when they are obvious)

$$D = \int |f - f^{(n)}|^2 dx = \int f^2 dx + \int (f^{(n)})^2 dx - 2 \int f f^{(n)} dx. \quad (1.24)$$

So there are three terms to differentiate – only the last two really, because the first doesn't contain any  $c_k$  and so will disappear when you start differentiating. These two terms are very easy to deal with if you make use of the supposed orthonormality of the expansion functions: for real functions  $\int \phi_k^2 dx = 1$ ,  $\int \phi_k \phi_l dx = 0$  ( $k \neq l$ ). Using these two properties, we can go back to (1.19) and differentiate the last two terms, with respect to each  $c_k$  (one at a time, holding the others fixed): the first of the two terms leads to

$$\frac{\partial}{\partial c_k} \int (f^{(n)})^2 dx = \frac{\partial}{\partial c_k} c_k^2 \int \phi_k(x)^2 dx = 2c_k;$$

while the second one gives

$$-2 \frac{\partial}{\partial c_k} \int f f^{(n)} dx = -2 \frac{\partial}{\partial c_k} c_k \int f(x) \phi_k(x) dx = -2 \langle f | \phi_k \rangle,$$

where Dirac notation (see Chapter 9 of Book 11) has been used for the integral  $\int f(x) \phi_k(x) dx$ , which is the **scalar product** of the two functions  $f(x)$  and  $\phi_k(x)$ :

$$\langle f | \phi_k \rangle = \int f(x) \phi_k(x) dx.$$

We can now do the differentiation of the whole difference function  $D$  in (1.18). The result is

$$\frac{\partial D}{\partial c_k} = 2c_k - 2 \langle f | \phi_k \rangle$$

and this tells us immediately how to choose the coefficients in the  $n$ -term approximation (1.17) so as to get the best possible fit to the given function  $f(x)$ : setting all the derivatives equal to zero gives

$$c_k = \langle f | \phi_k \rangle \quad (\text{for all } k). \quad (1.25)$$

So it's really very simple: you just have to evaluate one integral to get any coefficient you want. And once you've got it, there's never any need to change it in getting a better approximation. You can make the expansion as long as you like by adding more terms, but the coefficients of the ones you've already done are *final*. Moreover, the results are quite general: if you use basis functions that are no longer real you only need change the definition of the scalar product, taking instead the *Hermitian* scalar product as in (1.1).

### Generalizations

In studying atoms and molecules we'll have to deal with functions of very many variables, not just one. But some of the examples we met in Book 11 suggest possible ways of proceeding. Thus, in going from the harmonic oscillator in one dimension (Example 4.3), with eigenfunctions  $\Psi_k(x)$ , to the 3-dimensional oscillator (Example 4.4) it was possible to find eigenfunctions of *product form*, each of the three factors being of 1-dimensional form. The same was true for a particle in a rectangular box; and also for a free particle.

To explore such possibilities more generally we first ask if a function of two variables,  $x$  and  $x'$ , defined for  $x$  in the interval  $(a, b)$  and  $x'$  in  $(a', b')$ , can be expanded in products of the form  $\phi_i(x) \phi'_j(x')$ . Suppose we write (hopefully!)

$$f(x, x') = \sum_{i,j} c_{ij} \phi_i(x) \phi'_j(x') \quad (1.26)$$

where the set  $\{\phi_i(x)\}$  is complete for functions of  $x$  defined in  $(a, b)$ , while  $\{\phi'_i(x')\}$  is complete for functions of  $x'$  defined in  $(a', b')$ . Can we justify (1.26)? A simple argument suggests that we can.

For any given value of the variable  $x'$  we may safely take (if  $\{\phi_i(x)\}$  is indeed complete)

$$f(x, x') = c_1\phi_1(x) + c_2\phi_2(x) + \dots c_i\phi_i(x) + \dots$$

where the coefficients must depend on the chosen value of  $x'$ . But then, because  $\{\phi'_i(x')\}$  is also supposed to be complete, for functions of  $x'$  in the interval  $(a', b')$ , we may express the general coefficient  $c_i$  in the previous expansion as

$$c_i = c_{i1}\phi'_1(x') + c_{i2}\phi'_2(x') + \dots c_{ij}\phi'_j(x') + \dots$$

On putting this expression for  $c_i$  in the first expansion we get the double summation postulated in (1.26) (as you should verify!). If the variables  $x, x'$  are interpreted as Cartesian coordinates the expansion may be expected to hold good within the rectangle bounded by the summation limits.

Of course, this argument would not satisfy any pure mathematician; but the further generalizations it suggests have been found satisfactory in a wide range of applications in Applied Mathematics and Physics. In the quantum mechanics of many-electron systems, for example, where the different particles are physically identical and may be described in terms of a single complete set, the many-electron wave function is commonly expanded in terms of products of 1-electron functions (or 'orbitals').

Thus, one might expect to find 2-electron wave functions constructed in the form

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i,j} c_{i,j}\phi_i(\mathbf{r}_1)\phi_j(\mathbf{r}_2), \quad (1.27)$$

where the same set of orbitals  $\{\phi_i\}$  is used for each of the identical particles, the two factors in the product being functions of the different particle variables  $\mathbf{r}_1, \mathbf{r}_2$ . Here a boldface letter  $\mathbf{r}$  stands for the set of three variables (e.g. Cartesian coordinates) defining the position of a particle (position vector  $\mathbf{r}$ ). The labels  $i$  and  $j$  run over all the orbitals of the (in principle) complete set, or (in practice) over all values 1, 2, 3, ...,  $n$ , in the finite set used in constructing an approximate wave function.

In Chapter 2 you will find applications to 2-electron atoms and molecules where the wave functions are built up from one-centre orbitals of the kind studied in Book 11. (You can find pictures of atomic orbitals there, in Chapter 3.)

# Chapter 2

## Some two-electron atoms and molecules

### 2.1 Going from one particle to two

For two electrons moving in the field provided by one or more positively charged nuclei (supposedly fixed in space), the Hamiltonian takes the form

$$H(1, 2) = h(1) + h(2) + g(1, 2) \quad (2.1)$$

where  $H(1, 2)$  operates on the variables of both particles, while  $h(i)$  operates on those of Particle  $i$  alone. (Don't get mixed up with names of the indices – here  $i = 1, 2$  label the two electrons.) The one-electron Hamiltonian  $h(i)$  has the usual form (see Book 11)

$$h(i) = -\frac{1}{2}\nabla^2(i) + V(i), \quad (2.2)$$

the first term being the kinetic energy (KE) operator and the second being the potential energy (PE) of Electron  $i$  in the given field. The operator  $g(1, 2)$  in (2.1) is simply the *interaction* potential,  $e^2/\kappa_0 r_{ij}$ , expressed in ‘atomic units’ (see Book 11)<sup>1</sup> So in (2.1) we take

$$g(1, 2) = g(1, 2) = \frac{1}{r_{12}}, \quad (2.3)$$

$r_{12}$  being the inter-electron distance. To get a very rough estimate of the total energy  $E$ , we may neglect this term altogether and use an approximate Hamiltonian

$$H_0(1, 2) = h(1) + h(2), \quad (2.4)$$

---

<sup>1</sup>A fully consistent set of units on an ‘atomic’ scale is obtained by taking the **mass** and **charge** of the electron ( $m, e$ ) to have unit values, along with the **action**  $\hbar = h/2\pi$ . Other units are  $\kappa_0 = 4\pi\epsilon_0$  ( $\epsilon_0$  being the “**electric permittivity** of free space”); **length**  $a_0 = \hbar^2\kappa_0/me^2$  and **energy**  $e_H = me^4/\kappa_0^2\hbar^2$ . These quantities may be set equal to unity wherever they appear, leading to a great simplification of all equations. If the result of an *energy* calculation is the *number*  $x$  this just means that  $E = xe_H$ ; similarly a distance calculation would give  $L = xa_0$ .

which describes an **Independent Particle ‘Model’** of the system. The resultant IPM approximation is fundamental to all that will be done in Book 12.

With a Hamiltonian of this IPM form we can look for a solution of *product* form and use the ‘separation method’ (as in Chapter 4 of Book 11). We therefore look for a wave function  $\Psi(\mathbf{r}_1, \mathbf{r}_2) = \phi_m(\mathbf{r}_1)\phi_n(\mathbf{r}_2)$ . Here each factor is a function of the position variables of only one of the two electrons, indicated by  $\mathbf{r}_1$  or  $\mathbf{r}_2$ , and (to be general!) Electron 1 is described by a wave function  $\phi_m$  while Electron 2 is described by  $\phi_n$ .

On substituting this product in the eigenvalue equation  $\mathbf{H}_0\Psi = E\Psi$  and dividing through-out by  $\Psi$  you get (do it!)

$$\frac{\mathbf{h}(1)\phi_m(\mathbf{r}_1)}{\phi_m(\mathbf{r}_1)} + \frac{\mathbf{h}(2)\phi_n(\mathbf{r}_2)}{\phi_n(\mathbf{r}_2)} = E.$$

Now the two terms on the left-hand side are quite independent, involving different sets of variables, and their sum can be a constant  $E$ , only if each term is separately a constant. Calling the two constants  $\epsilon_m$  and  $\epsilon_n$ , the product  $\Psi_{mn}(\mathbf{r}_1, \mathbf{r}_2) = \phi_m(\mathbf{r}_1)\phi_n(\mathbf{r}_2)$  will satisfy the eigenvalue equation provided

$$\begin{aligned}\mathbf{h}(1)\phi_m(\mathbf{r}_1) &= \epsilon_m\phi_m(\mathbf{r}_1), \\ \mathbf{h}(2)\phi_n(\mathbf{r}_2) &= \epsilon_n\phi_n(\mathbf{r}_2).\end{aligned}$$

The total energy will then be

$$E = \epsilon_m + \epsilon_n. \tag{2.5}$$

This means that the **orbital product** is an eigenfunction of the IPM Hamiltonian provided  $\phi_m$  and  $\phi_n$  are *any* solutions of the **one-electron** eigenvalue equation

$$\mathbf{h}\phi(\mathbf{r}) = \epsilon\phi(\mathbf{r}). \tag{2.6}$$

Note especially that the names given to the electrons, and to the corresponding variables  $\mathbf{r}_1$  and  $\mathbf{r}_2$ , don’t matter at all. The same equation applies to each electron and  $\phi = \phi(\mathbf{r})$  is a function of position for whichever electron we’re thinking of: that’s why the labels 1 and 2 have been dropped in the *one-electron* equation (2.6). Each electron has ‘its own’ **orbital energy**, depending on which solution we choose to describe it, and since  $\mathbf{H}_0$  in (2.4) does not contain any interaction energy it is not surprising that their sum gives the total energy  $E$ . We often say that the electron “is in” or “occupies” the orbital chosen to describe it. If Electron 1 is in  $\phi_m$  and Electron 2 is in  $\phi_n$ , then the two-electron function

$$\Psi_{mn}(\mathbf{r}_1, \mathbf{r}_2) = \phi_m(\mathbf{r}_1)\phi_n(\mathbf{r}_2)$$

will be an exact eigenfunction of the IPM Hamiltonian (2.4), with eigenvalue (2.5).

For example, putting both electrons in the lowest energy orbital,  $\phi_1$  say, gives a wave function  $\Psi_{11}(\mathbf{r}_1, \mathbf{r}_2) = \phi_1(\mathbf{r}_1)\phi_1(\mathbf{r}_2)$  corresponding to total energy  $E = 2\epsilon_1$ . This is the (strictly!) IPM description of the **ground state** of the system. To improve on this approximation, which is very crude, we must allow for electron interaction: the next

approximation is to use the full Hamiltonian (2.1) to calculate the energy expectation value for the IPM function (no longer an *eigen*-function of  $\mathbf{H}$ ). Thus

$$\Psi_{11}(\mathbf{r}_1, \mathbf{r}_2) = \phi_1(\mathbf{r}_1)\phi_1(\mathbf{r}_2). \quad (2.7)$$

and this gives

$$\bar{E} = \langle \Psi_{11} | \mathbf{h}(1) + \mathbf{h}(2) + \mathbf{g}(1, 2) | \Psi_{11} \rangle = 2\langle \phi_1 | \mathbf{h} | \phi_1 \rangle + \langle \phi_1 \phi_1 | g | \phi_1 \phi_1 \rangle, \quad (2.8)$$

where the first term on the right is simply twice the energy of one electron in orbital  $\phi_1$ , namely  $2\epsilon_1$ . The second term involves the *two*-electron operator given in (2.3) and has explicit form

$$\langle \phi_1 \phi_1 | g | \phi_1 \phi_1 \rangle = \int \phi_1^*(\mathbf{r}_1)\phi_1^*(\mathbf{r}_2)\frac{1}{r_{12}}\phi_1(\mathbf{r}_1)\phi_1(\mathbf{r}_2)\mathbf{d}\mathbf{r}_1\mathbf{d}\mathbf{r}_2, \quad (2.9)$$

Here the variables in the bra and the ket will always be labelled in the order 1,2 and the volume element  $\mathbf{d}\mathbf{r}_1$ , for example, will refer to integration over all particle variables (e.g. in Cartesian coordinates it is  $dx_1dy_1dz_1$ ). (Remember also that, in bra-ket notation, the functions that come from the bra should in general carry the star (complex conjugate); and even when the functions are real it is useful to keep the star.)

To evaluate the integral we need to know the form of the 1-electron wave function  $\phi_1$ , but the expression (2.9) is a valid first approximation to the electron repulsion energy in the ground state of any 2-electron system.

Let's start with the Helium atom, with just two electrons moving in the field of a nucleus of charge  $Z = 2$ .

## 2.2 The Helium atom

The function (2.7) is clearly normalized when, as we suppose, the orbitals themselves (which are now **atomic orbitals**) are normalized; for

$$\langle \phi_1 \phi_1 | \phi_1 \phi_1 \rangle = \int \phi_1^*(\mathbf{r}_1)\phi_1^*(\mathbf{r}_2)\phi_1(\mathbf{r}_1)\phi_1(\mathbf{r}_2)\mathbf{d}\mathbf{r}_1\mathbf{d}\mathbf{r}_2 = \langle \phi_1 | \phi_1 \rangle \langle \phi_1 | \phi_1 \rangle = 1 \times 1.$$

The approximate energy (2.8) is then

$$\bar{E} = 2\epsilon_1 + \langle \phi_1 \phi_1 | g | \phi_1 \phi_1 \rangle = 2\epsilon_1 + J_{11}, \quad (2.10)$$

Here  $\epsilon_1$  is the *orbital energy* of an electron, by itself, in orbital  $\phi_1$  in the field of the nucleus; the 2-electron term  $J_{11}$  is often called a 'coulomb integral' because it corresponds to the coulombic repulsion energy (see Book 10) of two distributions of electric charge, each of density  $|\phi_1(\mathbf{r})|^2$  per unit volume. For a hydrogen-like atom, with atomic number  $Z$ , we know that  $\epsilon_1 = -\frac{1}{2}Z^2e_H$ . When the coulomb integral is evaluated it turns out to be  $J_{11} = (5/8)Ze_H$  and the approximate energy thus becomes  $\bar{E} = -Z^2 + (5/8)Z$  in 'atomic' units of  $e_H$ . With  $Z = 2$  this gives a first estimate of the electronic energy of the

Helium atom in its ground state:  $\bar{E} = -2.75 e_H$ , compared with an experimental value  $-2.90374 e_H$ .

To improve the ground state wave function we may use the variation method as in Section 1.2 by choosing a new function  $\phi'_1 = N'e^{-Z'r}$ , where  $Z'$  takes the place of the actual nuclear charge and is to be treated as an adjustable parameter. This allows the electron to ‘feel’ an ‘effective nuclear charge’ a bit different from the actual  $Z = 2$ . The corresponding normalizing factor  $N'$  will have to be chosen so that

$$\langle \phi'_1 | \phi'_1 \rangle = N'^2 \int \exp(-2Z'r)(4\pi r^2) dr = 1$$

and this gives (prove it!)  $N'^2 = Z'^3/\pi$ .

The energy expectation value still has the form (2.8) and the terms can be evaluated separately

### Example 2.1 Evaluation of the one-electron term

Written in full the first 1-electron operator has an expectation value

$$\langle \Psi_{11} | \mathbf{h}(1) | \Psi_{11} \rangle = \langle \phi'_1 | \mathbf{h} | \phi'_1 \rangle \langle \phi'_1 | \phi'_1 \rangle,$$

a matrix element of the operator  $\mathbf{h}$  times the scalar product  $\langle \phi'_1 | \phi'_1 \rangle$ . In full, this is

$$N'^2 \int_0^\infty e^{-Z'r} \mathbf{h} e^{-Z'r} 4\pi r^2 dr \times N'^2 \int_0^\infty e^{-Z'r} e^{-Z'r} 4\pi r^2 dr,$$

where  $\mathbf{h}$  working on a function of  $r$  alone is equivalent to  $(-\frac{1}{2}\nabla^2 - Z'/r)$ . (Note that  $\mathbf{h}$  contains the *actual* nuclear charge ( $Z$ )!)

We can spare ourselves some work by noting that if we put  $Z' = Z$  the function  $\phi'_1 = N'e^{-Z'r}$  becomes an eigenfunction of  $(-\frac{1}{2}\nabla^2 - Z'/r)$  with eigenvalue  $\epsilon' = -\frac{1}{2}Z'^2$  (corresponding to the ‘pretend’ value  $Z = Z'$ ). So let’s write

$$\mathbf{h} = -\frac{1}{2}\nabla^2 - Z/r = (-\frac{1}{2}\nabla^2 - Z'/r) + (Z' - Z)/r,$$

where the operator in parentheses is easy to handle: when it works on  $\phi'_1$  it simply multiplies it by the eigenvalue  $-\frac{1}{2}Z'^2$ . Thus, the operator  $\mathbf{h}$ , *working on the function*  $N'e^{-Z'r}$  gives

$$\mathbf{h}(N'e^{-Z'r}) = \left(-\frac{1}{2}Z'^2 + \frac{Z'-Z}{r}\right) N'e^{-Z'r}.$$

The one-electron part of (2.8) can now be written as (two equal terms – say why!)  $2\langle \Psi_{11} | \mathbf{h}(1) | \Psi_{11} \rangle$  where

$$\begin{aligned} \langle \Psi_{11} | \mathbf{h}(1) | \Psi_{11} \rangle &= \langle \phi'_1 | \mathbf{h} | \phi'_1 \rangle \langle \phi'_1 | \phi'_1 \rangle \\ &= N'^2 \int_0^\infty e^{-Z'r} \mathbf{h} e^{-Z'r} 4\pi r^2 dr \times N'^2 \int_0^\infty e^{-2Z'r} 4\pi r^2 dr \\ &= N'^2 \int_0^\infty e^{-Z'r} \left(-\frac{1}{2}Z'^2 + \frac{Z'-Z}{r}\right) e^{-Z'r} 4\pi r^2 dr. \end{aligned}$$

(Here the last term on the second line is unity, for normalization, and leaves only the integral on the following line.)

This remaining integral is (check it out!)

$$\langle \Psi_{11} | \mathbf{h}(1) | \Psi_{11} \rangle = -\frac{1}{2}Z'^2 + 4\pi(Z' - Z)N'^2 \int_0^\infty (re^{-2Z'r}) dr$$

and from the simple definite integral  $\int_0^\infty xe^{-ax}dx = (1/a^2)$  it follows that

$$\langle \Psi_{11} | \mathbf{h}(1) | \Psi_{11} \rangle = -\frac{1}{2}Z'^2 + 4\pi(Z' - Z)N'^2(1/2Z').$$

On putting in the normalizing factor,  $N'^2 = Z'^3/\pi$ , the final result is

$$\langle \Psi_{11} | \mathbf{h}(1) | \Psi_{11} \rangle = -\frac{1}{2}Z'^2 + Z'(Z' - Z).$$

Example 2.1 has given the expectation value of the  $\mathbf{h}(1)$  term in (2.8), but  $\mathbf{h}(2)$  must give an identical result since the only difference is a change of electron label from 1 to 2; and the third term must have the value  $J'_{11} = (5/8)Z'$  since the nuclear charge  $Z$  has been given the varied value  $Z'$  only in the orbital exponent (nothing else being changed).

On putting these results together, the energy expectation value after variation of the orbital exponent will be

$$\bar{E} = -Z'^2 + 2Z'(Z' - Z) + (5/8)Z' \text{ changed } 30 \text{ jan} \quad (2.11)$$

– all, as usual, in energy units of  $e_H$ .

The variational calculation can now be completed:  $\bar{E}$  will be stationary when

$$\frac{d\bar{E}}{dZ'} = -2Z' + 4Z' - 2Z + (5/8) = 0$$

and this means that the best estimate of the total electronic energy will be found on reducing the orbital exponent from its value  $Z = 2$  for one electron *by itself* to the value  $Z' = 2 - (5/16)$  in the presence of the second electron. In other words, the central field is effectively reduced or ‘screened’ when it holds another electron: the **screening constant** (5/16) is quite large and the ground state orbital expands appreciably as a result of the screening.

The corresponding estimate of the ground state energy is

$$\bar{E} = -(27/16)^2 = -2.84765 e_H \quad (2.12)$$

– a value which compares with  $-2.75 e_H$  before the variation of  $Z$  and is much closer to the ‘exact’ value of  $-2.90374 e_H$  obtained using a very elaborate variation function.

Before moving on, we should make sure that the value used for the Coulomb integral  $J = (5/8)Z e_H$  is correct<sup>2</sup>. This is our first example of a 2-electron integral: for two electrons in the same orbital  $\phi$  it has the form (2.9), namely (dropping the orbital label ‘1’)

$$J = \int \phi^*(\mathbf{r}_1)\phi^*(\mathbf{r}_2)\frac{1}{r_{12}}\phi(\mathbf{r}_1)\phi(\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2.$$

To evaluate it, we start from Born’s interpretation of the wave function  $|\phi(\mathbf{r})|^2 = \phi^*(\mathbf{r})\phi(\mathbf{r})$  (the star allowing the function to be *complex*) as a **probability density**. It is the

<sup>2</sup>If you find the proof too difficult, just take the result on trust and keep moving!

probability *per unit volume* of finding the electron in a small element of volume  $d\mathbf{r}$  at Point  $\mathbf{r}$  and will be denoted by  $\rho(\mathbf{r}) = \phi^*(\mathbf{r})\phi(\mathbf{r})$ . As you know from Book 11, this interpretation is justified by countless experimental observations.

We now go a step further: the average value of any quantity  $f(\mathbf{r})$  that depends only on the instantaneous *position* of the moving electron will be given by  $\bar{f} = \int f(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$  where, as usual, the integration is over all space (i.e. all values of the electronic variables). Now the electron carries a charge  $-e$  and produces a potential field  $V_{\mathbf{r}'}$  at any chosen ‘field point’  $\mathbf{r}'$ .

It’s convenient to use  $\mathbf{r}_1$  for the position of the electron (instead of  $\mathbf{r}$ ) and to use  $\mathbf{r}_2$  for the second point, at which we want to get the potential  $V_{\mathbf{r}_2}$ . This will have the value  $V_{\mathbf{r}_2} = -e/\kappa_0|\mathbf{r}_{21}|$ , where  $|\mathbf{r}_{21}| = |\mathbf{r}_{12}| = r_{12}$  is the *distance* between the electron at  $\mathbf{r}_1$  and the field point  $\mathbf{r}_2$ .

When the electron moves around, its position being described by the probability density  $\rho(\mathbf{r}_1)$ , the electric potential it produces at any point  $\mathbf{r}'$  will then have an average value

$$\bar{V}(\mathbf{r}_2) = \frac{-e}{\kappa_0} \int \frac{1}{|\mathbf{r}_{21}|} d\rho(\mathbf{r}_1)\mathbf{r}_1.$$

In words, this means that

The average electric field at point  $\mathbf{r}_2$ , produced by an electron at point  $\mathbf{r}_1$  with probability density  $\rho(\mathbf{r}_1)$ , can then be calculated just as if the ‘point’ electron were ‘smeared out’ in space, with a charge density  $-e\rho(\mathbf{r}_1)$ .

(2.13)

The statement (2.13) provides the **charge cloud** picture of the probability density. It allows us to visualize very clearly, as will be seen later, the origin of many properties of atoms and molecules. As a first application let’s look at the Coulomb integral  $J$ .

**Example 2.2** Interpretation of the electron interaction.

The integral  $J$  can now be viewed as the interaction energy of two distributions of electric charge, both of density  $-e\rho(\mathbf{r})$  and of spherical form (one on top of the other). (If that seems like nonsense remember this is only a mathematical interpretation!)

The two densities are in this case  $\rho_1(\mathbf{r}_1) = N^2 \exp -2Zr_1^2$  and  $\rho_2(\mathbf{r}_2) = N^2 \exp -2Zr_2^2$ ; and the integral we need follows on putting the interaction potential  $V(\mathbf{r}_1, \mathbf{r}_2) = 1/r_{12}$  between the two and integrating over all positions of both points. Thus, giving  $e$  and  $\kappa_0$  their unit values,  $J$  becomes the double integral

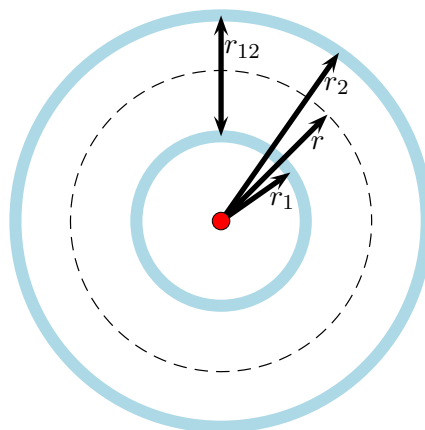
$$J = ZN^4 \int \int \exp -2Zr_1^2 \frac{1}{r_{12}} \exp -2Zr_2^2 d\mathbf{r}_1 d\mathbf{r}_2,$$

where  $r_{12}$  is simply the inverse distance between the two integration points. On the other hand,  $d\mathbf{r}_1$  and  $d\mathbf{r}_2$  are 3-dimensional elements of volume; and when the charge distributions are spherically symmetrical

functions of distance  $(r_1, r_2)$  from the origin (the nucleus), they may be divided into spherical *shells* of charge. The density is then constant within each shell, of thickness  $dr$ ; and each holds a total charge  $4\pi r^2 dr \times \rho(r)$ , the density being a function of radial distance  $(r)$  alone.

Now comes a nice connection with Electrostatics, which you should read about again in Book 10, Section 1.4. Before going on you should pause and study Figure 2.2, to have a clear picture of what we must do next.

Example 2.2 perhaps gave you an idea of how difficult it can be to deal with 2-electron integrals. The diagram below will be helpful if you want to actually evaluate  $J$ , the simplest one we've come across.



**Figure 2.2** Spherical shells of electron density (blue)

The integral  $J$  gives the electrostatic potential energy of two spherical charge distributions. Each could be built up from spherical ‘shells’ (like an onion): these are shown in blue, one for Electron 1 having radius  $r_1$  and another for Electron 2 with radius  $r_2$ . The distance between the two shells is shown with label  $r_{12}$  and this determines their potential energy as the product of the total charges they contain ( $4\pi r_1^2 dr_1$  and  $4\pi r_2^2 dr_2$ ) times the inverse distance ( $r_{12}^{-1}$ ). The total potential energy is obtained by summing (integrating) over all shells – but you need a trick! at any distance  $r$  from the nucleus, the potential due to an *inner* shell ( $r_1 < r$ ) is constant until  $r_1$  reaches  $r$  and changes form; so the first integration breaks into two parts, giving a result which depends only on where you put  $r$  (indicated by the broken line).

**Example 2.3** Evaluation of the electron interaction integral,  $J$

To summarize,  $J$  arises as the interaction energy of all pairs of spherical shells of charge, shown (blue) in Figure 2.2, and this will come from integration over all shells. We take one pair at a time.

You know that the electrostatic potential at distance  $r$  from the origin (call it  $V(r)$ ) due to a spherical shell of charge, of radius  $r_1$ , is given by

$$\begin{aligned} V(r) &= Q_{r_1} \times \frac{1}{r_1} \quad \text{for } r < r_1, \\ &= Q_{r_1} \times \frac{1}{r} \quad \text{for } r > r_1, \end{aligned}$$

where  $Q_{r_1} = 4\pi r_1^2 dr_1 \times \rho(r_1)$  is the total charge contained in the shell of radius  $r_1$  and thickness  $dr_1$ . The potential is thus constant *within* the first shell; but *outside* has a value corresponding to all the charge being put *at the origin*.

We can now do the integration over the variable  $r_1$  as it goes from 0 to  $\infty$ . For  $r_1 < r$  the sum of the contributions to  $J$  from the shells within a sphere of radius  $r$  will be

$$(1/r) \int_0^r \exp(-2Zr_1^2) 4\pi r_1^2 dr_1, \quad (A)$$

while for  $r_1 > r$  the rest of the  $r_1$  integration will give the sum of contributions from shells of radius greater than  $r$ , namely

$$\int_r^\infty \exp(-2Zr_1^2) (1/r_1) 4\pi r_1^2 dr_1. \quad (B)$$

You've met integrals a bit like these in Chapter 1, so you know how to do them and can show (do it!) that the sum of A and B is the potential function

$$V(r) = \frac{4\pi}{r} [2 - e^{-r}(2+r)].$$

This is a function of  $r$  alone, the radius of the imaginary sphere that we used to separate the integration over  $r_1$  into two parts, so now we can put  $r = r_2$  and multiply by  $(4\pi r_2^2 dr_2) N e^{-r_2}$  to obtain the energy of one shell of the second charge distribution in the field generated by the first.

After that it's all plain sailing: the integration over all the outer shells ( $r_2$ ) now goes from 0 to  $\infty$  – and you're home and dry! Integration over  $r_2$ , for all shells from  $r_2 = 0$  to  $\infty$ , will then give (check it out!)

$$J = \frac{Z}{2} \int_0^\infty [2 - e^{-r_2}(2+r_2)] e^{-r_2} r_2 dr_2 = (5/8)Z.$$

Example 2.3 gave you a small taste of how difficult it can be to actually evaluate the 2-electron integrals that are needed in describing electron interactions.

Now you know how to get a decent wave function for two electrons moving in the field of a single nucleus – the helium atom – and how the approximation can be improved as much as you wish by using the variation method with more elaborate trial functions. But following that path leads into difficult mathematics; so instead let's move on and take a quick look at some excited states.

### First excited states of He

In Book 11 we studied central-field systems, including many-electron atoms, in order to illustrate the general principles of quantum mechanics. In particular, we looked for sets of commuting operators associated with observable quantities such as angular momentum, finding that the angular momentum operators for motion in a central field commuted with the Hamiltonian  $H$  (see Chapter 6 of Book 11) and could therefore take simultaneously definite eigenvalues, along with the energy. For such a system, the energy eigenstates could be grouped into series, according to values of the angular momentum quantum numbers  $L$  and  $M$  which determine the angular momentum and *one* of its three components.

In the present chapter, we are dealing with systems of at most two electrons and the general theory is not needed: a 2-electron wave function is represented approximately as a product of 1-electron **orbitals**. And for the Helium atom we are dealing with spherically symmetrical wave functions, which involve only 's-type' orbitals, with zero angular momentum.

As a first example of an excited state we suppose one of the two electrons in the ground state, with wave function  $\Psi_{11}(\mathbf{r}_1, \mathbf{r}_2) = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)$ , is ‘promoted’ into the next higher orbital  $\phi_2$  of the s series. According to equation (6.10) of Book 11 Chapter 6, this AO corresponds to energy  $E_2 = -\frac{1}{2}(Z^2/4)$ , the whole series being depicted in Figure 13.

**Example 2.4** Excited state wave functions and energies

When one of the two electrons is promoted from the lowest-energy AO  $\phi_1$  into the next one,  $\phi_2$ , there are clearly two distinct ways of representing the state by an IPM function: it could be either

$$\Psi_{12}(\mathbf{r}_1, \mathbf{r}_2) = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2),$$

in which Electron 2 has been promoted, or

$$\Psi_{21}(\mathbf{r}_1, \mathbf{r}_2) = \phi_2(\mathbf{r}_1)\phi_1(\mathbf{r}_2),$$

in which Electron 1 (with coordinates  $\mathbf{r}_1$ ) has been put into  $\phi_2$ , the second electron staying in  $\phi_1$ . And at this point *three* product functions are available for constructing 2-electron wave functions – those we have called  $\Psi_{11}$ , the IPM ground state, and  $\Psi_{12}, \Psi_{21}$ , in which one of the electrons has been promoted. We could of course set up other products  $\Psi_{lm}$ , with both electrons promoted to higher-energy AOs, and suppose these may be used in the first few terms of a **complete set** expansion of the 2-electron wave function. The products corresponding to any particular choice of the orbitals e.g.  $\phi_1, \phi_2$  are said to belong to the same **electron configuration**.

Here, to simplify things, we’ll use a single-subscript notation to denote the first three products:  $\Psi_1 = \phi_1\phi_1$ ,  $\Psi_2 = \phi_1\phi_2$ ,  $\Psi_3 = \phi_2\phi_1$ . We can then use the linear variation method (Section 1.3) to get improved approximations to the three lowest-energy wave functions in the form

$$\Psi = c_1\Psi_1 + c_2\Psi_2 + c_3\Psi_3.$$

This involves setting up the **secular equations**

$$\begin{aligned} (H_{11} - \bar{E}M_{11})c_1 + (H_{12} - \bar{E}M_{12})c_2 + (H_{13} - \bar{E}M_{13})c_3 &= 0, \\ (H_{21} - \bar{E}M_{21})c_1 + (H_{22} - \bar{E}M_{22})c_2 + (H_{23} - \bar{E}M_{23})c_3 &= 0, \\ (H_{31} - \bar{E}M_{31})c_1 + (H_{32} - \bar{E}M_{32})c_2 + (H_{33} - \bar{E}M_{33})c_3 &= 0, \end{aligned}$$

where, as usual,  $H_{ij} = \langle \Psi_i | \mathbf{H} | \Psi_j \rangle$  and  $M_{ij} = \langle \Psi_i | \Psi_j \rangle$ . On solving them we obtain, along with the optimized mixtures, improved approximations to the energies  $E_1, E_2, E_3$  of the first three electronic states. (Read Section 1.3 again if you need to.)

Here, the approximate ground-state function  $\Psi_1$  has a very small overlap with  $\Psi_2$  and  $\Psi_3$ ; for example

$$M_{12} = \langle \Psi_1 | \Psi_2 \rangle = \langle \phi_1\phi_1 | \phi_1\phi_2 \rangle = \langle \phi_1 | \phi_1 \rangle \langle \phi_1 | \phi_2 \rangle \approx 0,$$

because  $\langle \phi_1 | \phi_1 \rangle = 1$  and  $\langle \phi_1 | \phi_2 \rangle \approx 0$  – the 1s and 2s AOs being normalized and lying mainly in different regions of space. For similar reasons, other off-diagonal terms such as  $H_{12}, H_{13}$ , which connect the IPM ground state  $\Psi_1$  with the higher-energy functions  $\Psi_2, \Psi_3$  are usually small enough to be neglected.

With such approximations (check them out!) the secular equations may be written

$$\begin{aligned} (H_{11} - \bar{E})c_1 &= 0, \\ (H_{22} - \bar{E})c_2 &= -H_{23}c_3, \\ H_{32}c_2 &= -(H_{33} - \bar{E})c_3. \end{aligned}$$

The first equation says that  $\bar{E} \approx H_{11}$  is still an approximation to the ground-state energy  $E_1$ . The other equations allow us to eliminate the expansion coefficients and to determine approximate eigenvalues for

two excited states. Thus (you've done it all before in Section 1.3!), on dividing each side of the second equation by the corresponding side of the third, the coefficients cancel and leave you with

$$\frac{(H_{22} - \bar{E})}{H_{32}} = \frac{H_{23}}{(H_{33} - \bar{E})}.$$

Now we know that  $H_{22} = H_{33}$  (say why!) and  $H_{32} = H_{23}$  (real matrix elements) and if we call these quantities  $\alpha$  and  $\beta$  the equation becomes  $(\alpha - \bar{E})^2 = \beta^2$ . The two roots are  $(\alpha - \bar{E}) = \pm\beta$  and give two approximate excited-state energies:  $\bar{E}^{(+)} = \alpha + \beta$  and  $\bar{E}^{(-)} = \alpha - \beta$ .

To end this example let's get the energies of these states, just as we did for the ground state, where we found  $\bar{E} = 2\epsilon_1 + J_{11}$  in terms of orbital energy  $\epsilon_1$  and coulomb interaction  $J_{11}$ . (You should read again, from equation (2.7) to equation (2.8), to remind yourself of how we did it.)

The excited states are linear combinations of the functions  $\Psi_2, \Psi_3$ , which belong to the configuration 1s2s. Thus  $\Psi_2^{(+)}$  for the 'plus combination', with energy  $\bar{E}_2^{(+)}$ , is obtained by putting  $\bar{E}_2^{(+)} = \alpha + \beta$  back into the second equation, which shows that  $c_3 = c_2$ . This state therefore has the (normalized) form  $\Psi_2^{(+)} = (\Psi_2 + \Psi_3)/\sqrt{2}$  and  $\Psi_2^{(-)}$  will be similar, with the plus changed to a minus.

The energy expectation value in state  $\Psi_2^{(+)}$  will be  $\langle \Psi_2^{(+)} | \mathbf{H} | \Psi_2^{(+)} \rangle = \frac{1}{2}[H_{22} + H_{33} + 2H_{23}]$ , where  $H_{22} = H_{33} = \langle \Psi_2 | \mathbf{H} | \Psi_2 \rangle$  and  $H_{23} = \langle \Psi_2 | \mathbf{H} | \Psi_3 \rangle$ . Now  $\Psi_2 = \phi_1\phi_2$  and  $\Psi_3 = \phi_2\phi_1$ , so it follows (check it, remembering that the order of the variables in an orbital product is always  $\mathbf{r}_1, \mathbf{r}_2$ ) that

$$H_{22} = H_{33} = \langle \Psi_2 | \mathbf{H} | \Psi_2 \rangle = \epsilon_1 + \epsilon_2 + J_{12} \text{ and } H_{23} = \langle \Psi_2 | \mathbf{H} | \Psi_3 \rangle = K_{12}.$$

Finally, then, the energy expectation value in state  $\Psi_2^{(+)}$  will be

$$E_2^{(+)} = \langle \Psi_2^{(+)} | \mathbf{H} | \Psi_2^{(+)} \rangle = [\epsilon_1 + \epsilon_2 + J_{12}] + K_{12},$$

while  $E_2^{(-)}$  will follow on changing the sign of the K-term.

(Note that the J and K terms are quite different:

$$J_{12} = \langle \Psi_2 | \mathbf{g} | \Psi_2 \rangle = \langle \phi_1\phi_2 | g(1,2) | \phi_1\phi_2 \rangle, \quad K_{12} = \langle \Psi_2 | \mathbf{g} | \Psi_3 \rangle = \langle \phi_1\phi_2 | g(1,2) | \phi_2\phi_1 \rangle,$$

– the 'ket' part of the matrix element  $\langle \Psi_2 | \mathbf{g} | \Psi_3 \rangle$  containing the orbitals *after* exchange of the electron labels. It's no surprise that  $K_{12}$  is called an "exchange integral"!)

Example 2.4 was tough, but was done in detail because it leads us to tremendously important conclusions, as you'll see presently. (If you didn't manage to get through it yourself, don't worry – you can move on and come back to it later.) First, however, we want to understand the general meaning of the results. In particular, we need to know how the wave functions of the two states behave under *symmetry operations* that make no apparent change to the system.

### Example 2.5 Symmetry of the wave functions

The two terms in  $\Psi_2^{(+)}$  differ only by an interchange of electronic variables  $\mathbf{r}_1, \mathbf{r}_2$  (as you can check from the definitions) and their *sum* does not change at all under such an operation: we say the wave function  $\Psi_2^{(+)}$  is **symmetric under exchange of the electrons**. On the other hand the other state, with energy  $\bar{E}_2^{(-)} = \alpha - \beta$ , has a wave function  $\Psi_2^{(-)} = (\Psi_2 - \Psi_3)/\sqrt{2}$ , which changes sign on exchanging the electrons and is said to be **antisymmetric**.

Now let's move on by adding a second *nucleus* to the 2-electron system we're considering. How can we get a wave function for two electrons in the field of *two* nuclei. When the nuclei are single protons the resultant system is the hydrogen molecule.

## 2.3 The Hydrogen molecule

It might seem that the approach used in the last Section could easily be used to describe  $H_2$ , the simplest molecule, but we meet many problems. In the first place, with two *nuclei*, we'll be dealing with a *four*-particle system, for the nuclei will also be free to move about in space. Fortunately, since the nuclei are about 2000 times as heavy as an electron, they move much more slowly and in first approximation may be taken to be at rest, with a fixed internuclear distance (as if they were 'clamped'). In this approximation, we may concentrate on the electrons alone – as if they moved in a fixed two-centre field, provided by the two nuclei. But then we meet the next big problem. For an atom we had 'ready-made' atomic orbitals, with the well-known forms (1s, 2s, 2p, 3s, 3p, 3d, etc.) discussed in Book 11, but here we know nothing about the forms of the **molecular orbitals** that will be needed in building corresponding approximations to the *molecular* wave functions. First of all, then, we need to find how to describe the one-electron system that remains when one electron is taken away. This system is experimentally well-known: it is the Hydrogen molecule ion,  $H_2^+$ .

### The Hydrogen molecule ion

For  $H_2^+$  the nuclei may be labelled *a* and *b* and placed on, say, the z-axis, with an internuclear separation *R*. The 1-electron Hamiltonian for the system is then

$$h = -\frac{1}{2}\nabla^2 - \left(\frac{1}{r_a} + \frac{1}{r_b}\right), \quad (2.14)$$

where  $r_a, r_b$  are the distances of the electron from Nucleus a and Nucleus b, respectively. The  $\nabla^2$ -operator works on the electronic coordinates, to be denoted by  $\mathbf{r}$ , and will have a form depending on the coordinate system chosen.

How can we get a reasonable first approximation to the lowest-energy molecular orbital (MO)? When the electron is close to Nucleus a, the term  $1/r_a$  will be so big that  $1/r_b$  may be neglected in (2.14). The MO will then 'shrink' into an *atomic* orbital (AO) for a single hydrogen atom. If we denote this AO by  $\phi_a(\mathbf{r})$  we can say that in this case  $\phi(\mathbf{r}) \rightarrow c_a\phi_a(\mathbf{r})$ , since this will satisfy the same single-atom eigenvalue equation for any value of a numerical factor *c*. Similarly, when  $\mathbf{r}$  is close to the second nucleus  $\phi(\mathbf{r})$  will approach a numerical multiple of the AO  $\phi_b(\mathbf{r})$ . It follows that an electron in the field of both nuclei may be fairly well represented by an MO of the form

$$\phi(\mathbf{r}) = c_a\phi_a(\mathbf{r}) + c_b\phi_b(\mathbf{r}) \quad (2.15)$$

where the constants  $c_a, c_b$  are still to be chosen (e.g. by taking them as variable parameters and using the variation method of Section 1.3) to find the MO of minimum energy. This should give at least a rough description of the ground state.

In fact, however, no calculation is needed because the molecule ion is *symmetrical* across a plane perpendicular to the molecular axis, cutting the system into two equal halves. There is no reason to expect the electron to be found with different probability on the two sides of the symmetry plane and this implies that the values of the coefficients  $c_a, c_b$  can differ, at most, in sign:  $c_b = \pm c_a$ . Two acceptable approximate MOs are thus, putting  $c_b = c_a = N_B$  in one MO and  $c_b = -c_a = N_A$  in the other

$$\phi_B(\mathbf{r}) = N_B[\phi_a(\mathbf{r}) + \phi_b(\mathbf{r})], \quad \phi_A(\mathbf{r}) = N_A[\phi_a(\mathbf{r}) - \phi_b(\mathbf{r})]. \quad (2.16)$$

This case arises only for **homonuclear** diatomic molecules – in which the two nuclei are identical. It is important because very many common diatomic molecules, such as  $\text{H}_2$ ,  $\text{N}_2$ ,  $\text{O}_2$ , are of this type.

The solutions just found are typical **Bonding and Antibonding MOs**; so called for reasons that will soon become clear. The constants  $N_A, N_B$  are normalizing factors, chosen to give unit probability of finding the electron *somewhere* in space. For normalization we require

$$N_B^2 \langle \phi_B | \phi_B \rangle = N_B^2 (2 + 2S_{ab}) = 1,$$

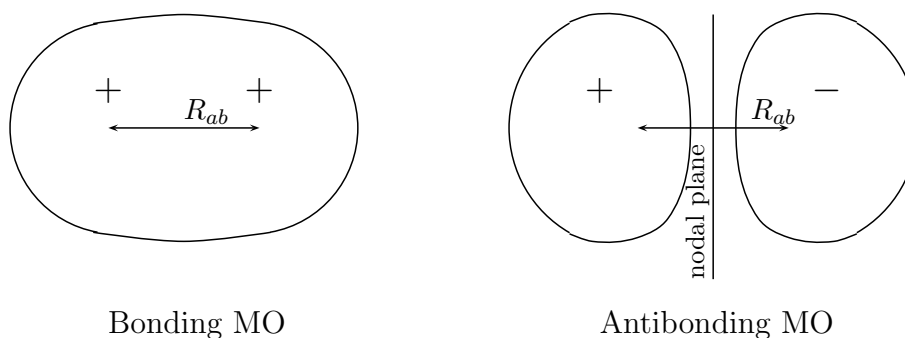
where  $S_{ab} = \langle \phi_a | \phi_b \rangle$  is the **overlap integral** between the two AOs. In this way we find MOs

$$\phi_B(\mathbf{r}) = \frac{\phi_a(\mathbf{r}) + \phi_b(\mathbf{r})}{\sqrt{2 + 2S_{ab}}} \quad (2.17)$$

for the Bonding MO, and

$$\phi_A(\mathbf{r}) = \frac{\phi_a(\mathbf{r}) - \phi_b(\mathbf{r})}{\sqrt{2 - 2S_{ab}}} \quad (2.18)$$

for the Antibonding MO. The following Figure 2.3 gives a very schematic picture of the two MOs.



**Figure 2.3** Schematic representation of the two lowest-energy MOs for  $\text{H}_2^+$ .

Here, for the ion,  $\mathbf{H} = \mathbf{h}$ , the 1-electron Hamiltonian, and the distinct quantities to be calculated are (using a common notation and supposing the AOs are normalized)

$$\alpha_a = \langle \phi_a | \mathbf{h} | \phi_a \rangle, \quad \beta_{ab} = \langle \phi_a | \mathbf{h} | \phi_b \rangle, \quad \alpha_b = \langle \phi_b | \mathbf{h} | \phi_b \rangle,$$

together with  $S_{ab} = M_{ab} = \langle \phi_a | \phi_b \rangle$ .

As in Section 1.3, the conditions for a stationary value then reduce to

$$\begin{aligned}(\alpha_a - \bar{E})c_a &= -(\beta_{ab} - \bar{E}S_{ab})c_b \\(\beta_{ab} - \bar{E}S_{ab})c_a &= -(\alpha_b - \bar{E})c_b.\end{aligned}\tag{2.19}$$

But when the system is symmetrical, as already noted, we know that  $c_b = \pm c_a$  and in that case just one equation is enough to give us both eigenvalues. Thus, putting  $\alpha_a = \alpha_b = \alpha$  and choosing  $c_b = c_a$ , the first equation reduces to  $(\alpha + \beta) - \bar{E}(1 + S) = 0$ ; while on choosing  $c_b = -c_a$  it reduces to  $(\alpha - \beta) - \bar{E}(1 - S) = 0$ . The approximate energies of the two states  $\phi_B(\mathbf{r}), \phi_A(\mathbf{r})$ , may then be written

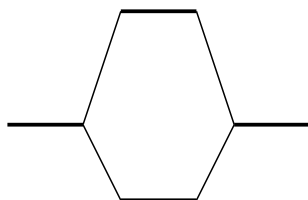
$$\bar{E}_B = \frac{\alpha + \beta}{1 + S} = \frac{\alpha(1 + S) + \beta - \alpha S}{1 + S}, \quad \bar{E}_A = \frac{\alpha - \beta}{1 - S} = \frac{\alpha(1 - S) - \beta + \alpha S}{1 - S},$$

where the numerators have been re-arranged so as to ‘separate out’ the leading terms. In this way we find

$$\bar{E}_B = \alpha + \frac{\beta - \alpha S}{1 + S}, \quad \bar{E}_A = \alpha - \frac{\beta - \alpha S}{1 - S}.\tag{2.20}$$

Since  $\alpha$  is the energy expectation value of an electron in the field of one nucleus alone and (like  $\beta$ ) has a negative value, it follows that the Bonding MO  $\phi_B$  has a *lower* energy ( $\bar{E}_B$ ) than the free-atom AO, while the Antibonding MO  $\phi_A$  has a *higher* energy ( $\bar{E}_A$ ). Note, however, that the upward displacement of the free-atom energy level in going to the antibonding level is greater than the downward displacement in going to the bonding level, owing to the overlap term. All this is shown very nicely in a **correlation diagram** which shows how the energies of the AOs on two identical atoms are related to those of the MOs which result when the atoms are combined to form a **homonuclear diatomic molecule** – a ‘homonuclear diatomic’, for short.

Such a diagram, describing the formation of  $H_2$ , is shown in Figure 2.4, energy levels for the separate atoms being indicated on the left and right with the MO energies in the centre.



**Figure 2.4** Energies of orbitals in a homonuclear diatomic.

AO levels shown left and right  
MO levels shown in the centre.

Remember that we’re still talking about a *one*-electron system, the hydrogen molecule positive ion, and that this is *homonuclear*. But before going to the neutral molecule, with two electrons, we may want to think also about other 2-electron systems such as

HeH<sup>+</sup>, with one of the Hydrogens (H) replaced by a Helium atom (He) and one of the three electrons taken away – giving you the **Helium Hydride positive ion**. In that case we'll have a **heteronuclear system** in which the two nuclei are different and the forms of the acceptable MOs must also be changed. Helium hydride is a very rare species, though it was probably important in the very early stages of the developing Universe, when there weren't many atoms around – only the very lightest ones like hydrogen and helium had already formed. But it gives us a general 'pattern' or **prototype** for the study of heteronuclear diatomic systems, which are present in great abundance in today's world. So it's worth looking at the system briefly, in the example that follows:

**Example 2.6** A heteronuclear system: HeH<sup>+</sup>.

This is a 'hydrogen-like' system, having two electrons moving in the field of two nuclei, but it differs from the hydrogen molecule in having a Helium nucleus (with charge  $Z = 2$ ) in place of one of the protons. Let's take it as 'Nucleus a' in our study of H<sub>2</sub> and ask what MOs can be formed from the AOs  $\phi_a$  and  $\phi_b$  when the different atoms come together.

Things look much the same as in the case of H<sub>2</sub>, as far as equation (2.15), except that the  $(1/r_a)$ -term in (2.14) will have  $Z = 2$  in the numerator instead of  $Z = 1$ . But this is enough to destroy the symmetry and the acceptable MOs will no longer have the simple forms (??). Instead we must go back to the stationary value conditions (2.19) to determine the mixing coefficients  $c_a, c_b$ .

Again, using  $\beta$  and  $S$  for short (in place of  $\beta_{ab}, S_{ab}$ ), the coefficients may be eliminated by division to give the single equation

$$(\alpha_a - \bar{E})(\alpha_b - \bar{E}) - (\beta - S\bar{E})^2 = 0.$$

This can be solved by the method you used in Book 1 (Section 5.3), to give two approximate eigenvalues  $\bar{E}_B$  (lower energy) and  $\bar{E}_A$  (upper energy). These correspond to the 'Bonding' and 'Antibonding' levels for a homonuclear system (see Figure 2.4), but solving the quadratic equation by the standard method doesn't give a simple result comparable with (2.19).

Instead, we use a simple approximation which shows directly how much the AO energies for the free atoms (roughly  $\alpha_a$  and  $\alpha_b$ ) are respectively 'pushed down', to give  $\bar{E}_B$ , and 'pushed up', to give  $\bar{E}_A$ . The *interaction*, which does this, is caused by the term  $(\beta - S\bar{E})^2$ . If we neglect this term,  $\bar{E} \approx \alpha_a$  – the lower of the two AO energies (corresponding to  $Z = 2$ ) – so let's use this approximation to estimate the effect of the *other terms*: the last equation is then replaced by

$$(\alpha_a - \bar{E})(\alpha_b - \alpha_a) - (\beta - \alpha_a S)^2 = 0,$$

which gives (check it!)

$$\bar{E} - \alpha_a = -\frac{(\beta - \alpha_a S)^2}{\alpha_b - \alpha_a}.$$

This is the approximation to the lowest root of the quadratic equation, which we called  $\bar{E}_B$ , the energy of the Bonding MO.

A similar argument (you should try it) shows that the higher AO energy  $\alpha_b$  is pushed *up* as a result of the mixing, giving an approximation to the energy of the Antibonding MO.

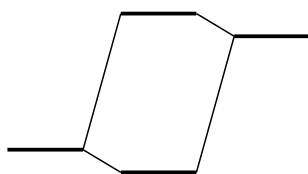
The results from Example 2.6 may be summarized as follows. The Bonding and Antibonding MOs used in describing the interaction of two different atoms to yield a heteronuclear diatomic molecule have corresponding MO energies

$$\bar{E}_B = \alpha_a - \frac{(\beta - \alpha_a S)^2}{\alpha_b - \alpha_a}, \quad \bar{E}_A = \alpha_b + \frac{(\beta - \alpha_b S)^2}{\alpha_b - \alpha_a}. \quad (2.21)$$

These results should be compared with those in (2.19), which apply to a homonuclear molecule. In particular

- the lower of the two energy levels, in this case  $\alpha_a$ , is pushed down to give the Bonding level  $\bar{E}_B$ . But whereas the shift for a homonuclear molecule was roughly  $\beta$  it is now roughly proportional to the *square* of  $\beta$  (neglecting the small overlap term  $\alpha_a S$ ), divided by the difference of the free-atom energies  $\alpha_b - \alpha_a$ ;
- the upper free-atom level is *raised* by a similar amount, to give the energy  $\bar{E}_A$  of the Antibonding MO;
- these effects are both much smaller than in the case of a homonuclear system, unless the free-atom energies are close together. They are of ‘second order’ in the interaction term  $\beta$ .

The correlation diagram in Figure 2.4 is now replaced by the one shown below:



**Figure 2.5** Energies of orbitals in a heteronuclear diatomic.

AO levels shown left and right  
MO levels shown in the centre.

It’s time to say why we’re talking about “bonding” and “antibonding” orbitals! You’ll remember from Book 5 that sometimes atoms ‘stick together’ to form molecules and other structures – gases, liquids, solids and so on. Until the early part of the last century this had to be accepted as a general ‘property of matter’ and further details had to be investigated experimentally. Only now, following the development of Quantum Mechanics, are we in a position to say *why* atoms behave like that.

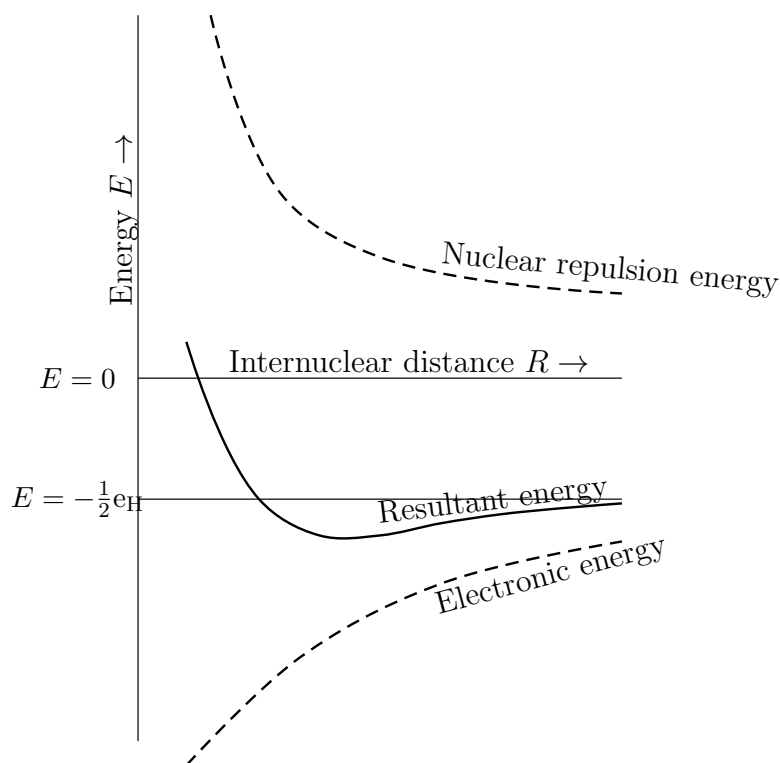
### The chemical bond

In Book 5, where you first met molecules, they were often represented in terms of ‘ball and stick models’: the ‘balls’ represented the atoms, while the ‘sticks’ that connected them, stood for the **chemical bonds** that held them together. This is still a widely used way of picturing molecules of all kinds, ranging from simple diatomics to the gigantic structures studied in the Life Sciences (see Book 9), where the molecules may contain many thousands of atoms arranged in long chains and carrying the **genetic code**.

Here, however, we are concerned with the ‘sticks’ that join the different atoms: what are they and how do they work? At bottom, they must be associated with the electrons and nuclei that carry negative and positive electric charge and with their *interaction energy*. And we have just seen how it is possible for even the single electron of a Hydrogen atom to enter a state of *lower energy* by bringing up a second proton, so that the electron is attracted to two positive nuclei instead of one. In that case we are imagining the formation of a molecular ion  $\text{H}_2^+$ , in which the electron occupies a Bonding MO.

Let's examine this case in more detail. In equation (2.19) we have an expression for the energy of an electron in the Bonding MO  $\phi_B$ , as a function of the parameters  $\alpha$  and  $\beta$  (the 'coulomb' and 'resonance' integrals). These parameters depend on the geometry of the system (i.e. the positions of the two nuclei) and are not too difficult to calculate in terms of the internuclear separation  $R$ . When this is done, the electronic energy of the system can be plotted against  $R$  and is found to increase steadily, going towards the energy of a free hydrogen atom, namely  $-\frac{1}{2}e_H$ , in the long-range limit  $R \rightarrow \infty$ . So what's wrong?

The fact is simply that we haven't yet included the energy of *repulsion* between the two nuclei: this is  $E_{\text{nuc}} \propto (1/R)$  and goes from a large positive value, when the nuclei are close together, to zero when  $R \rightarrow \infty$ . We didn't even include this term in the Hamiltonian (2.1) as it didn't depend on the *electronic* variables. Strictly it should be included (the protons are part of the system!); but then the expectation value  $\bar{E} = \langle \phi | \mathbf{H} | \phi \rangle$ , for any normalized state with wave function  $\phi(\mathbf{r})$  would contain an additive constant  $E_{\text{nuc}}$ , which can be put in at the end of the calculation. When this is done, the total energy of the system becomes the sum of two terms; the repulsion energy  $E_{\text{nuc}}$  and  $\bar{E}_B$ , the energy of the electron in the Bonding MO. The two terms are 'in competition', one falling as  $R$  increases, the other rising; and together they give a total energy  $E(R)$  which shows a shallow minimum at a certain value  $R = R_0$ . This means there is a **chemical bond** between the two atoms, with 'bond length'  $R_0$ , *say*. The variation of the three energy terms, as functions of internuclear distance, is shown in Figure 2.6 (below).



**Figure 2.6** Energy curves for the Hydrogen molecule ion  
Resultant energy  $E$  has its minimum at  $R \approx 2 a_0$

Of course, this is not for the normal hydrogen *molecule* but rather the *molecule ion* that remains when one electron is taken away. However, the 2-electron molecule  $\text{H}_2$  behaves in a very similar way: the electronic energy expression has just the same form as that for the Helium atom, given in (2.8). The big difference is that the 1-electron terms,  $\langle \Psi | \mathbf{h}(1) | \Psi \rangle$  and  $\langle \Psi | \mathbf{h}(2) | \Psi \rangle$ , and the 2-electron term  $\langle \Psi | g(1, 2) | \Psi \rangle$ , are much more difficult to evaluate. Remember that the wave function we're going to use is a product  $\Psi(\mathbf{r}_1, \mathbf{r}_2) = \phi_B(\mathbf{r}_1)\phi_B(\mathbf{r}_2)$ , where both electrons are shown in the Bonding MO  $\phi_B$ , which describes the state of lowest energy  $2E_B$  when the interelectronic repulsion energy,  $J = \langle \Psi | g(1, 2) | \Psi \rangle$ , is neglected. Since  $J$  is positive the total electronic energy will now have a lowest possible expectation value

$$\bar{E} = 2E_B + J,$$

corresponding to the molecular ground state. This has the same form as that for the 2-electron atom; but the 1-electron part,  $2E_B$ , will now depend on the attraction of the electron to *both* nuclei – and therefore on their separation  $R$ , which determines their positions in space. Apart from this weak dependence on  $R$ , the total electronic energy of the system will behave in much the same way as for the ion  $\text{H}_2^+$ , while the internuclear repulsion energy remains unchanged as  $E_{\text{nuc}}$ .

The relevant energy curves for both the normal molecule and its positive ion are therefore closely similar in form. Those for the ion are shown above. The value of  $R$  at which the energy has its minimum is usually called the **equilibrium bond length** and is denoted by  $R_e$  while the energy difference between that at the minimum and that for  $R \rightarrow \infty$  is called the **dissociation energy**, denoted by  $D_e$ . The term “equilibrium” is of course not quite correct – the nuclei are in fact moving and it is an approximation to do the calculation as if they were at rest for a series of fixed values of  $R$ . But it is usually a decent first approximation which can later be refined to take account of vibration and rotation of the system around its equilibrium configuration; and anyway we've made more serious approximations already in using such a simple form of the electronic wave function.

## 2.4 But what happened to the spin?

We started Book 11, on the basic principles of quantum mechanics, by talking about the Stern-Gerlach experiment – which showed a moving electron was not fully described by giving its *position* variables  $x, y, z$ , it needed also a **spin variable**  $s$  with only two observable values. But it seems as if we've completely forgotten about spin, using wave functions that depend only on position of the electron in space. The reason is simple: the spin (identified in Book 11 as some kind of *internal* angular momentum) has such a small effect on energy levels that it's hardly observable! You can solve the Schrödinger equation, and get meaningful results, because the usual Hamiltonian operator contains no spin operators and acts only on the position variables in the wave function. But in dealing with many-particle systems it's absolutely essential to label states according to their spin properties: as you will see presently, without spin you and I would not exist – there would be no Chemistry!

It's easy to put the spin back into our equations: just as the product function  $\Psi_{mn}(\mathbf{r}_1, \mathbf{r}_2) =$

$\phi_m(\mathbf{r}_1)\phi_n(\mathbf{r}_2)$  was used to describe two independent particles, in states  $\phi_m$  and  $\phi_n$ , so can we use a product  $\phi(\mathbf{r})\theta(s)$  to describe a particle in orbital state  $\phi$  and in spin state  $\theta$ . If  $\phi$  is an eigenstate of the *spinless* operator  $\mathbf{h}$  (with eigenvalue  $\epsilon$ ) and  $\theta$  is an eigenstate of  $\mathbf{S}_z$  (with spin component  $s = S_z$  along the z-axis), then the product is a simultaneous eigenstate of both operators:

$$\mathbf{h}[\phi\theta] = (\mathbf{h}\phi)\theta = (\epsilon\phi)\theta = \epsilon[\phi\theta]$$

since the operator  $\mathbf{h}$  doesn't touch the  $\theta$ -factor; and similarly

$$\mathbf{S}_z[\phi\theta] = \phi(\mathbf{S}_z\theta) = \phi(S_z\theta) = S_z[\phi\theta]$$

– since the operator  $\mathbf{S}_z$  doesn't touch the  $\phi$ -factor.

Now the 'spin-space' is only two-dimensional, with basis vectors denoted by  $\alpha$  and  $\beta$  corresponding to  $s = +\frac{1}{2}$  and  $s = -\frac{1}{2}$  (in units of  $\hbar$ ), respectively. So for any given orbital state  $\phi$  there will be two alternative possibilities  $\phi\alpha$  and  $\phi\beta$  when spin is taken into account. Products of this kind are called **spin-orbitals**. From now on let's agree to use Greek letters ( $\psi, \Psi$ ) for states with the spin description included, leaving  $\phi, \Phi$  for 'orbital' states (as used so far) which don't contain any spin factors. The lower-case (small) letters will be used for *one*-electron states, upper-case (capital) letters for *many*-electron states.

As long as we deal only with a *two*-electron system, the state vector (or corresponding wave function) can be expressed as a product of space and spin factors:  $\Psi(1, 2) = \Phi(1, 2)\Theta(1, 2)$ , where the electron labels are used to indicate spatial or spin variables for electrons 1 and 2. When we want to be more explicit we'll use a fuller notation, as below.

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = \Phi(\mathbf{r}_1, \mathbf{r}_2)\Theta(s_1, s_2). \quad (2.22)$$

Here  $\mathbf{x}$  stands for both space and spin variables together, so  $\mathbf{x} \equiv \mathbf{r}, s$ . This is a neat way of saying that  $\Psi(\mathbf{x}_1, \mathbf{x}_2)$  in (

In the following Example we shall be looking for a simultaneous eigenstate of all commuting operators, which will normally include  $\mathbf{H}, \mathbf{S}^2, \mathbf{S}_z$ . We suppose  $\Phi(1, 2)$  is an eigenstate (exact or approximate) of the usual spinless Hamiltonian  $\mathbf{H}(1, 2)$  and take  $\Theta(1, 2)$  as an eigenstate of *total spin* of the two particles i.e. of the operators  $\mathbf{S}^2, \mathbf{S}_z$ .

Before continuing you should turn back to Section 2.2 of Book 11 and make sure you understand the properties of the total spin operators  $\mathbf{S}_x = \mathbf{S}_x(1) + \mathbf{S}_x(2)$ ,  $\mathbf{S}_y = \mathbf{S}_y(1) + \mathbf{S}_y(2)$ ,  $\mathbf{S}_z = \mathbf{S}_z(1) + \mathbf{S}_z(2)$ . Remember, they follow the same commutation rules as for a single particle and that you can define step-up and step-down operators  $\mathbf{S}^\pm = (\mathbf{S}_x \pm i\mathbf{S}_y)$  in the same way; from them you can set up the operator  $\mathbf{S}^2$  and show that it has eigenvalues of the form  $S(S+1)$  (in units of  $\hbar^2$ ), where  $S = 1$  ('parallel-coupled' spins) or  $S = 0$  ('paired' spins). Study especially Example 2.2, which gives the spin eigenstates for a 2-electron system.

### Example 2.7 Symmetry properties of the spin eigenstates

---

In Example 2.2 of Book 11 it was shown that, for two spin-coupled electrons, the eigenstates of  $\mathbf{S}^2$  and  $\mathbf{S}_z$  with quantum numbers  $S = 0, \pm 1$  were as follows:

- (1,1)  $\Theta_{1,1} = \alpha(1)\alpha(2)$

- (1,0)  $\Theta_{1,0} = \beta(1)\alpha(2) + \alpha(1)\beta(2)$
- (0,0)  $\Theta_{0,0} = \beta(1)\alpha(2) - \alpha(1)\beta(2)$
- (1,-1)  $\Theta_{1,-1} = \beta(1)\beta(2)$

(Here the S- and M- quantum numbers are shown in parentheses and the state symbol  $\Theta$  has been used to denote a two-electron spin state)

It's important to know how these eigenstates change under a **symmetry operation** which has no observable effect on the system. In this case, all electrons are identical – we can't tell one from another – so exchanging the labels '1' and '2' (call it  $P_{12}$ ) should be a symmetry operation ( $P_{12}\alpha(1)\beta(2) = \alpha(2)\beta(1)$ ) means that Electron '1' goes into the 'down-spin' state, previously occupied by Electron '2', while Electron '2' goes into an 'up-spin' state – but the change is not observable).

If you examine all the spin states listed above you'll see at once that all the states with  $S = 1$  are unchanged, they are *symmetric* under the exchange; but the single state with  $S = 0$  *changes sign* – it is *antisymmetric* under exchange, being multiplied by  $-1$ .

We're now ready to go back and look again at the excited states of the Helium atom, but *with spin included*. The complete wave function will now be a 'space-spin' product of the form  $\Psi(1, 2) = \Phi(1, 2)\Theta(1, 2)$ , where the two factors are now re-named as agreed in the run-up to (2.22). Possible choices for the orbital factor are then  $\Phi_1$ , for the ground state, with both electrons in the first (lowest-energy) AO  $\phi_1$ ; and  $\Phi_2^{(+)}$  or  $\Phi_2^{(-)}$ , for the excited states with one electron in the AO  $\phi_1$  and the other is in the next AO  $\phi_2$  – with a 'plus' combination or a 'minus' combination of  $\Phi_2, \Phi_3$ . The available energy states for the two-electron atom, without spin, would seem to be:

- Ground state. Energy =  $E_1$ , wave function  $\Phi_1$
- 1st excited state. Energy =  $E_2^{(-)}$ , wave function  $(\Phi_2 - \Phi_3)/\sqrt{2}$  (normalized 'minus' combination),
- 2nd excited state. Energy =  $E_2^{(+)}$ , wave function  $(\phi_2 + \Phi_3)/\sqrt{2}$  (normalized 'plus' combination).

What happens when spin is taken into account? When the two electrons are interchanged, both space and spin variables change:

$$\mathbf{r}_1, \mathbf{r}_2 \rightarrow \mathbf{r}_2, \mathbf{r}_1 \quad \text{and} \quad s_1, s_2 \rightarrow s_2, s_1.$$

But the energy levels are determined essentially by the  $\Phi$  factor; so let's take the states as listed above and ask what symmetry each state will have when spin is included.

The space-spin product function  $\Psi = \Phi\Theta$  for the ground state will have  $\Phi = \Phi_1$  which is symmetric under electron exchange, but may take possible spin factors:

$$\Theta = \Theta_{1,1}, \quad \text{or} \quad \Theta_{1,0}, \quad \text{or} \quad \Theta_{1,-1},$$

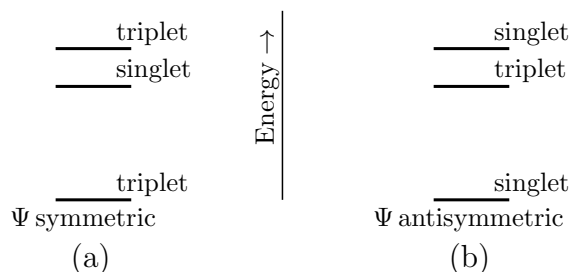
which are all symmetric under spin exchange. So three possible  $\Psi$  products can be found; all are 'totally' symmetric and correspond to the same energy – suggesting a 'three-fold degenerate triplet' ground state.

On the other hand,  $\Phi_1$  might have been combined with  $\Theta_{0,0} = \beta(1)\alpha(2) - \alpha(1)\beta(2)$  and that would have given a totally *antisymmetric* space-spin product – a ‘non-degenerate singlet’ ground state.

What about the excited state with energy  $E_2^{(-)}$ ? The antisymmetric space factor  $\Phi_2^{(-)}$  could be associated with any of the three symmetric spin factors, to give three antisymmetric space-spin products. But it could equally well be attached to the antisymmetric spin factor  $\Theta_{0,0} = \beta(1)\alpha(2) - \alpha(1)\beta(2)$  to give a single totally symmetric  $\Psi$ -product.

Finally, the excited state with energy  $E_2^{(+)}$  and symmetric space factor  $\Phi_2^{(+)}$  could be associated with the antisymmetric spin factor  $\Theta_{0,0}$  to give an antisymmetric space-spin  $\Psi$ -product; or equally well combined with any one of the three symmetric spin factors  $\Theta_{1,1}, \Theta_{1,0}, \Theta_{1,-1}$ , to give a three-fold degenerate  $\Psi$ , all products being totally antisymmetric.

That was quite a lot of work, but the results can be summarized very easily in a diagram showing the first few energy levels you might expect to find for any two-electron system. As long as there are no spin operators in the Hamiltonian the electronic energy depends only on the spatial wave function  $\Phi$ , but the nature of any state – whether it is degenerate or non-degenerate and whether or not it corresponds to definite values of the total spin – depends on the overall symmetry of the space-spin function  $\Psi$ . Remember that a state of total spin  $S$  has  $2S + 1$  degenerate components (labelled by the quantum number  $M_S$ ) and that this is the **multiplicity** of the state. The diagram below shows the two very different sets of states that result.



**Figure 2.7** Some electronic states of the He atom

Lowest level (ground state) for configuration  $1s^2$ ,  
upper levels (excited states) for configuration  $1s2s$ .  
Multiplicities of the calculated states are shown in  
(a) for *symmetric*  $\Psi$  and (b) for *antisymmetric*  $\Psi$ .

The remarkable fact is that the experimentally observed states correspond only to those shown in Figure 2.7(b), where the ground state is a singlet and the first excited state is a triplet. But wait a minute! How can we be sure the state we’re calling the “first excited state” really *is* the lowest excited state? If you look back at Example 2.4 you’ll see that the first excited state, going up in energy, was taken to be the one with wave function  $\Phi_2^{(-)}$ , namely the ‘minus’ combination of  $\Phi_2$  and  $\Phi_3$ ; and that is the one with energy

$$E_2^{(-)} = [\epsilon_1 + \epsilon_2 + J_{12}] - K_{12}.$$

On the other hand, the ‘plus’ combination gave an energy

$$E_2^{(+)} = [\epsilon_1 + \epsilon_2 + J_{12}] + K_{12}$$

and since  $K_{12}$  is an essentially *positive* quantity this energy lies *above* that of the “first excited state”. So we got it right! The energy levels on the right-hand side in Figure 2.6 are in complete agreement with experiment, while those on the left simply do not appear! Overall antisymmetry of an electronic wave function seems to be an intrinsic property of the electrons themselves – or of the ‘wave field’  $\Psi$  with which they are described. In fact this conclusion is perfectly general: it applies not just to two-electron systems but to all the electrons in the universe! – and it is confirmed by countless experiments.

## 2.5 The antisymmetry principle

This brings us to the last general principle of quantum mechanics that we’re going to need in Book 12. It wasn’t included in Book 11 because in formulating the basic principles we were thinking mainly of *one*-particle systems; but the antisymmetry of many-electron wave functions is just as important as anything we’ve discovered so far. So let’s state the **antisymmetry principle** in the general form which applies to systems of any number of electrons:

The wave function  $\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$  describing any state of an  $N$ -electron system is *antisymmetric* for any permutation  $\mathbf{P}$  of the electrons:

$$\mathbf{P}\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \epsilon_{\mathbf{P}}\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N),$$

where  $\epsilon_{\mathbf{P}} = \pm 1$  for permutations of even or odd parity, respectively.

(2.23)

Here  $\mathbf{P}$  is a general **permutation**, which acts on the numbered electronic variables  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$  and changes them into  $\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}'_N$ , where the new numbers  $1', 2', \dots, N'$  are the old ones written in a different order. This permutation can be achieved by making a series of **transpositions**  $(1, 1')(2, 2') \dots (N, N')$ , where each  $(i, i')$  interchanges one pair of numbers, one ‘old’ and one ‘new’: thus  $(1, 3)(4, 2)$  will send 1 2 3 4 into 3 4 1 2. Any permutation is equivalent to a number of transpositions: when the number is odd the **parity** of the permutation is said to be “odd”; when it is even, the parity is “even”. (Note that, in counting,  $(i, i)$  (where a number is interchanged with itself) is not included – not being a true transposition.)

Section 2.4 opened with the amazing claim that “without spin you and I would not exist – there would be no Chemistry!” To end this chapter we must ask how this can be so – and how does the Antisymmetry Principle come into the picture?

During the early development of quantum theory, before Schrödinger’s introduction of the wave function, the electrons in an atom were assigned to ‘states’ on a basis of experimental evidence. Atomic spectroscopy had shown that the emission and absorption of light could be associated with ‘quantum jumps’ of single electrons between energy levels with characteristic ‘quantum numbers’. (See Book 11 for spectral series and energy level diagrams.) A key postulate in accounting for the electronic structures of atoms, was **Pauli’s Exclusion Principle**, which stated that no two electrons could be in states with the same set of quantum numbers.

The Antisymmetry Principle is simply the modern and more general form of Pauli’s Exclusion Principle<sup>3</sup> To see how antisymmetry of the wave function contains the idea of ‘exclusion’ it’s enough to go one step beyond the two-electron systems studied in the present chapter. In an IPM description the first two spinorbitals might be  $\psi_1 = \phi_1\alpha, \psi_2 = \phi_1\beta$ , with both electrons in the same orbital  $\phi_1$ , but with opposite spins. The corresponding antisymmetric 2-electron state, found in Section 2.4, is then seen to be (before normalization)  $\psi_1\psi_2 - \psi_2\psi_1$ , which is called an “antisymmetrized spinorbital product”. It can be derived from the leading term, a ‘parent product’,  $\psi_1\psi_2$ , by subtracting the product obtained after making an electron interchange. The operator  $A = (1/2)(I - P_{12})$  ( $I$  being the usual identity operator and  $P_{12}$  being the interchange of variables for electrons 1 and 2) is called an **anti-symmetrizer**. There is a more general form of this operator, which we’ll need in Chapter 3, namely

$$A = \frac{1}{N!} \sum_{\mathbf{P}} \epsilon_{\mathbf{P}} \mathbf{P} \quad (2.24)$$

which acts on a product of  $N$  spinorbitals to produce an antisymmetric  $N$ -electron wave function. Here the summation runs over all  $N!$  permutations and the parity factor  $\epsilon_{\mathbf{P}}$  was defined after (2.23); the extra factor  $1/N!$  is included simply for convenience (you can apply the operator a second time without making any difference i.e.  $AA = A$ )

Now let’s try to get a wave function for a three-electron system, by adding another electron to orbital  $\phi_1$ . There are only two possible choices of spin factor and the third electron can therefore occupy only  $\psi_3 = \phi_1\alpha$  or  $\psi_3 = \phi_1\beta$ . The parent product will then be  $\psi_1\psi_2\psi_3$  and we want to find a function that changes sign under *any* permutation of electronic variables. To do it we use (2.24) with  $N = 3$ , noting that two spinorbitals are identical: for example,  $\psi_3 = \psi_1$ . In that case, the permutations  $\mathbf{P}$  will act on the parent product  $\psi_1\psi_2\psi_1$ , which can also be replaced by  $\psi_1\psi_1\psi_2$  (it can’t matter *which* product we antisymmetrize).

Thus

$$A[\psi_1\psi_1\psi_2] = \frac{1}{N!} \sum -P_{\epsilon_{\mathbf{P}}} \mathbf{P}[\psi_1\psi_1\psi_2].$$

---

<sup>3</sup>Over the years, starting from Pauli himself, there has been much argument about the fundamental status of the two principles, but that can be found in books on the philosophy of quantum mechanics – when you’re ready!

But now think about the effect of the ‘first’ permutation (the order doesn’t matter as the sum is over *all*  $N!$  permutations), taking it to be one that interchanges the first two spin variables. This will leave the product unchanged, and as the parity factor for a single interchange is  $-1$  the resultant term in the sum will be  $-\psi_1\psi_1\psi_2$ . But the identity permutation, included in the summation, leaves the parent product unchanged and the net result is thus exactly zero! In fact, what we have shown for three electrons is true for any number (think about it, noting that if  $P_{12}$  leaves the parent function unchanged, then any permutation can be expressed as  $P = P'P_{12}$  where  $P'$  acts on all the variables *except*  $\mathbf{x}_1, \mathbf{x}_2$ ).

To summarize,

The antisymmetrized product function  

$$A\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2) \dots \psi_N(\mathbf{x}_N),$$
representing an IPM approximation to the state of an  $N$ -electron system, can contain no repetitions of any given spinorbital: every electron must have its own distinct spinorbital. A given *spatial orbital* can hold not more than two electrons, one with spin factor  $\alpha$ , the other with  $\beta$ .

(2.25)

This is the quantum mechanical equivalent of Pauli’s Exclusion Principle: it *excludes* the possibility of finding more than two electrons in the same spatial orbital; and when two are present they must have opposite spins  $\pm\frac{1}{2}$ . It is less general than the Antisymmetry Principle, because it applies only to approximate wave functions of particular form: but is very simple to apply and leads directly to conclusions that provide the basis for all modern theories of the electronic structure of atoms and molecules. The example with which we introduced it explains at once why the 3-electron Lithium atom does not have all three electrons in the lowest-energy 1s orbital: because the Helium-like configuration  $(1s)^2$  is already ‘full’ and a third electron must then ‘overflow’ into the higher-energy 2s orbital, giving the configuration  $\text{Li}[(1s)^2(2s)]$ . Thus, there are two electrons in an **inner shell**, tightly localized around the nucleus, and one electron by itself, in a more diffuse 2s orbital. And that is the beginning of Chemistry, and of Life in all forms! Without antisymmetry and the exclusion property to which it leads, all matter would collapse – every nucleus would take all the electrons it could hold, becoming an uncharged and unreactive system, like no atom in the world we know.