

CHAPTER 1

Computational Quantum Chemistry

1.1 What Does Computational Quantum Chemistry Offer?

Computational quantum chemistry has been in development for almost nine decades. Its progress has been intimately linked to developments in computing hardware and technology. Today computational quantum chemistry provides a complementary way of investigating a wide range of chemistry. In particular it provides reliable information on molecular structures, molecular properties, reactions mechanisms and energetics. Detailed mechanistic questions can be addressed using the techniques of computational quantum chemistry. An advantage over traditional experimental techniques is that it provides a route to the study of chemical questions which may be experimentally difficult, or expensive, or dangerous. The purpose is always to answer a chemical question and in that sense computational quantum chemistry is the complement to experiment, either approach on its own is much less convincing. This complementarity of techniques is very familiar to chemists. For example, to determine a molecular structure a range of spectroscopies must be used and each provides a component of the overall picture. Now to these spectroscopies are added quantum chemical techniques that can provide further information.

Computational quantum chemistry is an elegant conjunction of chemistry, physics, mathematics and computer science. Chemistry defines the question. Physics defines the laws that are obeyed by the chemical system. Mathematics formulates a numerical representation of the problem. Computer science solves

the mathematical model, yielding numbers that encapsulate physical significance. For example, does a particular alkylation reaction proceed more efficiently with the alkyl chloride or the corresponding iodide? To answer this question, at the simplest level, we could compute the geometries of the transition structures and reactants, from which we would obtain the activation energies and so determine which reaction should be more efficient. The insight gained from such numerical answers can lead to further questions. This often results in an iterative refinement of questions, answers and models, see Figure 1.1. By such a process our understanding of a chemical question deepens.

The historical development of quantum chemistry can be categorised into a number of eras. The earliest, first age of quantum chemistry, was characterised by computational results of a qualitative nature. These did much to help develop understanding of potential energy surfaces, geometries of molecules at equilibrium, reactive transition structures, and molecular orbital concepts. These insights were able to explain the physical origins of experimentally measured properties. The second age of quantum chemistry came about through the development of computer technology and accompanying developments in numerical algorithms. This enabled much more elaborate computations to be performed. In this second era, semi-quantitative agreement with experiment was already obtained for some measured quantities. Despite this improved accuracy, quantum chemical techniques were still not able to displace experimental measurements, but had become sufficiently reliable that

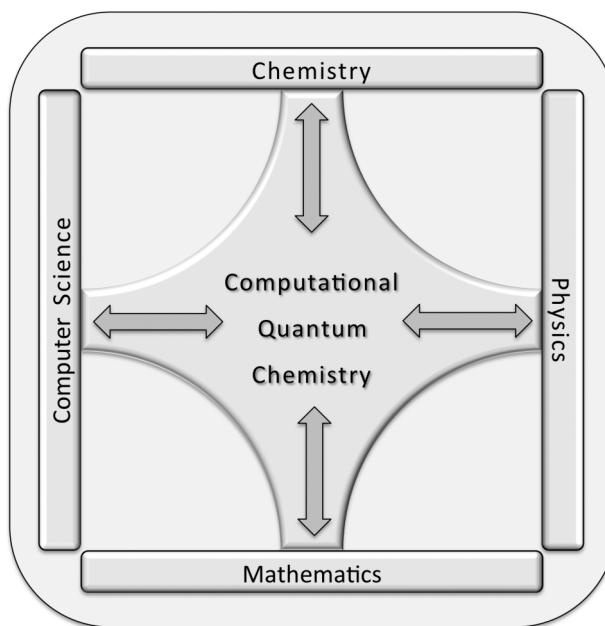


Figure 1.1 Interplay of disciplines that make up computational quantum chemistry.

they could be applied to situations in which experiments were not yet feasible. For example, the study of very short-lived molecular species, or the study of the properties of postulated molecules that had yet to be synthesised. The third age of quantum chemistry is best summarised by Graham Richards in his influential article of 1979:¹ “*The work represents perhaps a near perfect instance of theory being in harmony with experiment, each aspect vital to the other and the combination much more than the sum of the separate parts*”. Experimental measurements cannot be interpreted or understood in the absence of a reliable theoretical framework. The studies referred to by Richards showed the computational work to be an equal partner to the experiments. There have also been cases where computational studies have preceded experiments through the correct predictions of measured quantities, which have subsequently been confirmed by experiment.² Since the dawn of this new age of quantum chemistry, alluded to above, rapid developments have continued and their success has made computational quantum chemistry an essential component of many modern chemical investigations.

Historically, computational quantum chemistry was restricted to the realm of specialists who had access to high performance computing facilities, a good knowledge of software construction and numerical methods, as well as a good understanding of the underlying quantum mechanical models. There is still a strong need for this type of specialist who can push the subject forward by developing new methods, or providing very efficient computer implementations of established techniques. However, the standard models that we shall discuss in this book have been developed to the point that serious molecular questions can be tackled by any good scientist, not just the computational specialist. Today elaborate quantum chemical calculations can be carried out using fast desktop machines, and readily available software, by non-specialists. The same care and rigour must be applied to the design and execution of such calculations as would be applied to the design of any scientific investigation. A poorly thought out study, whether computational or experimental, cannot produce useful results.

1.2 The Model: Quantum Mechanics

An interesting experiment, which the author has carried out on numerous occasions, is to ask a room full of 200 undergraduate chemistry students: *What is chemistry?* Invariably, one obtains a fascinating range of answers. Many will tell you that chemistry is about “making things”, for example materials, medicines, or fuels. Others may tell you that chemistry is about understanding the physical processes that govern chemical properties, for example the rate of reaction between two molecules, or the colour of a molecule. All these answers, and many others beside, are equally valid. Yet the overriding answer is: *Chemistry is a game that electrons play!* In a sense this answer encapsulates all the other answers, since everything chemical is under the control of the

electrons that participate in the chemical process. There are no chemical phenomena that cannot be traced back to the behaviour of electrons.

So chemistry is about electrons. To understand chemistry we need to understand the behaviour of electrons. We are familiar with electrons being negatively charged particles with mass. Additionally we know from the experiments of Davisson and Germer in 1925, involving the diffraction of electrons by a crystal, that electrons can behave as waves. This wave-particle duality is quantified in the de Broglie relation

$$\lambda = \frac{h}{p} \quad (1.1)$$

where λ is the wavelength associated with a particle of mass, m , moving with velocity, v . The linear momentum is, $p=mv$, and h is Planck's constant. For slow-moving macroscopic objects the wavelength given by eqn (1.1) is undetectably small. However electrons confined within atoms and molecules are very light and fast-moving with comparatively large de Broglie wavelengths. This is the realm of quantum mechanics and the correct description of quantum mechanical particles, such as the electron, is provided by the Schrödinger equation. The electronic structure and properties of any molecule, in any of its available stationary states may be determined, in principle, by solution of Schrödinger's (time-independent) equation.

$$\hat{H}\Psi_A = E_A \Psi_A \quad (1.2)$$

In eqn (1.2), A labels the state of interest. For example, the ground state or the first electronically excited state. To begin we shall concern ourselves with the ground state only and suppress the state label. At the simplest level we want to find the energy, E , and the wavefunction, Ψ , based on the hamiltonian operator, \hat{H} , for the molecular system of interest. The Schrödinger equation can be solved exactly only for one-electron systems. Hence much of the apparatus of computational quantum chemistry is concerned with finding increasingly accurate approximations to the Schrödinger equation for many-electron molecular systems. As we shall see in due course, the accuracy of the approximations is intimately related to the computational cost of the underlying numerical algorithms.

The first chemical application of the Schrödinger equation was undertaken by Heitler and London in 1927. In their landmark paper they calculated the potential energy curve of the hydrogen molecule. Today we are able to perform calculations on much larger systems, perhaps including up to 1000 atoms, and the methods we use are very different from those used by Heitler and London. Developments in computational quantum chemistry have been closely allied to developments in computational hardware as well as algorithmic developments (Figure 1.1). This endeavour shows no sign of abating and the demand for computational studies to complement experimental work grows continually.

This is easily understood since, as we have asserted, chemistry is about the behaviour of electrons. The Schrödinger equation furnishes us, in principle, with all information about the behaviour of electrons in molecules and in turn, all information about chemistry. As we have stated already, approximations are key and it emerges that there is in practice no “best” method in computational quantum chemistry. Studies on real chemical problems always involve a trade-off between accuracy and computational cost. For certain methods we can make formal statements about their relative merits as approximations to the Schrödinger equation. However if such methods are too computationally demanding to be applicable to a problem of interest then describing them as “better” is, at best, vague.

1.2.1 The Schrödinger Equation and the Born-Oppenheimer Approximation

Before proceeding to some details, it is useful to briefly describe a very powerful notational expedience, introduced by Paul Dirac in 1939, which we shall use throughout this book. We shall write the many equations and integrals that appear using Dirac notation. For example, consider how we can obtain the energy, E , from the Schrödinger equation as shown below (the conventional notation will be shown on the left hand side and the equivalent in Dirac notation on the right).

$$\hat{H}\Psi = E\Psi \quad \equiv \quad \hat{H}|\Psi\rangle = E|\Psi\rangle \quad (1.3)$$

Now pre-multiply by Ψ^* (the complex conjugate of Ψ) and integrate over all variables, call them τ ,

$$\int \Psi^* \hat{H} \Psi d\tau = E \int \Psi^* \Psi d\tau \quad \equiv \quad \langle \Psi | H | \Psi \rangle = E \langle \Psi | \Psi \rangle \quad (1.4)$$

Now rearrange to obtain E :

$$E = \frac{\int \Psi^* \hat{H} \Psi d\tau}{\int \Psi^* \Psi d\tau} \quad \equiv \quad E = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (1.5)$$

A quantity denoted in $| \rangle$ is termed a “ket” and here represents the wavefunction. $\langle |$ is called a “bra” and represents the complex conjugate of $| \rangle$. For real quantities, $\langle |$ and $| \rangle$ are the same. When an operator is pre- and post-multiplied by a bra and ket, integration is assumed implicitly. Denoting a general operator as \hat{C} , we form

$$\langle bra | C | ket \rangle \quad e.g. \langle \Psi | H | \Psi \rangle \quad (1.6)$$

Accordingly this notation is often referred to as “bracket” notation. It is very

widely used. There are many other subtle features, related to the description of vector spaces, which are implicit in the Dirac notation, but they will not concern us here. In the rest of this text we shall use Dirac notation and conventional notations as suits the discussion.

It turns out that despite the simple form in which the Schrödinger equation can be written, its solutions are far from simple to obtain. In fact the Schrödinger equation can only be solved for one-electron systems. To deal with more complex atoms and molecules we must introduce a number of approximations. There are three key ideas which we shall adopt. To motivate the first of these, let us look in more detail at the quantities that enter the Schrödinger equation. The model of the atom that we shall use consists of a set of protons positioned at the atomic nucleus and surrounded by a number of electrons. The number of protons is given by the atomic number, Z , which tells us the number of protons carrying a unit positive charge, e , in the atomic nucleus. For neutral atoms, Z also gives the number of electrons surrounding the nucleus, each with unit negative charge, $-e$. In the absence of electric or magnetic fields, the hamiltonian operator, \hat{H} , then includes terms which specify the kinetic and potential energies of the electrons and nuclei. \hat{H} includes (i) the kinetic energy of motion for electrons and nuclei; (ii) the potential energy of attraction between electrons and nuclei; (iii) the potential energy of repulsion between electrons and similarly the potential energy of repulsion between nuclei. These terms have the following form:

$$-\frac{\hbar^2}{2m_e} \sum_i^{\text{electrons}} \nabla_i^2 - \frac{\hbar^2}{2} \sum_A^{\text{nuclei}} \frac{1}{M_A} \nabla_A^2 \quad (1.7)$$

$$-\frac{e^2}{4\pi\epsilon_0} \sum_A^{\text{nuclei}} \sum_i^{\text{electrons}} \frac{Z_A}{r_{iA}} \quad (1.8)$$

$$\frac{e^2}{4\pi\epsilon_0} \sum_{i < j}^{\text{electrons}} \frac{1}{r_{ij}} + \frac{e^2}{4\pi\epsilon_0} \sum_{A < B}^{\text{nuclei}} \frac{Z_A Z_B}{R_{AB}} \quad (1.9)$$

The quantities that enter eqns (1.7) – (1.9) are:

\hbar : Planck's constant divided by 2π ,

m_e : the rest mass of the electron

M_A : the mass of nucleus A

e : the charge on the proton

ϵ_0 : the permittivity of free-space

∇_i^2 and ∇_A^2 : are the kinetic energy operators for electron i and nucleus A , respectively. ∇^2 is known as the "laplacian operator" and has the general form

$$(in \text{ cartesian coordinates}) \nabla^2 = \left(\frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial y^2}, \frac{\partial^2}{\partial z^2} \right).$$

r_{iA} , r_{ij} and R_{AB} : are the distance vector between electron i and nucleus A ; the distance vector between electron i and electron j ; the distance vector between nucleus A and nucleus B , respectively. For example, if A and B have cartesian coordinates, (x_A, y_A, z_A) and (x_B, y_B, z_B) , respectively, the magnitude of the distance vector between A and B can be written as $R_{AB} = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2}$. All other distances are defined similarly.

Before proceeding any further we can simplify the forms of eqns (1.7) – (1.9) by introducing atomic units (au). In this system of units a number of fundamental constants take the value of unity, hence

$$\begin{aligned} e &= 1 \\ m_e &= 1 \\ \hbar &= 1 \\ 4\pi\epsilon_0 &= 1 \end{aligned}$$

Table 1.1 shows these, and other key quantities assigned a value of 1 au, with their SI equivalents. A more complete list of quantities is given in Appendix 1A.

Using these definitions we can write the kinetic energy terms in au as:

$$-\frac{1}{2} \sum_i^{\text{electrons}} \nabla_i^2 - \frac{1}{2} \sum_A^{\text{nuclei}} \frac{1}{\bar{M}_A} \nabla_A^2 \quad (1.10)$$

where \bar{M}_A is the ratio of the mass of nucleus A to the mass of the electron, $\bar{M}_A = \frac{M_A}{m_e}$. Similarly the potential energy of attraction between electrons and nuclei may be written in au as:

Table 1.1 Essential atomic units and their SI equivalents.

Quantity	Atomic Unit	Equivalent in SI units
Charge	$m_e = 1$	$1.602176 \times 10^{-19} \text{ C}$
Mass	$e = 1$	$9.109382 \times 10^{-31} \text{ kg}$
Angular Momentum	$\hbar = \frac{h}{2\pi} = 1$	$1.054571 \times 10^{-34} \text{ J s}$
Permittivity of Free-Space	$4\pi\epsilon_0 = 1$	$1.112650 \times 10^{-10} \text{ C}^2 \text{ J}^{-1} \text{ m}^{-1}$
Length	$a_0 = 1$	$5.291772 \times 10^{-11} \text{ m}$
Energy	$E_h = 1$	$4.359744 \times 10^{-18} \text{ J}$

$$-\sum_A^{\text{nuclei}} \sum_i^{\text{electrons}} \frac{Z_A}{r_{iA}} \quad (1.11)$$

and the repulsive potential energy terms as:

$$\sum_{i < j}^{\text{electrons}} \frac{1}{r_{ij}} + \sum_{A < B}^{\text{nuclei}} \frac{Z_A Z_B}{R_{AB}} \quad (1.12)$$

The full molecular hamiltonian now becomes:

$$\hat{H} = -\frac{1}{2} \sum_i^{\text{electrons}} \nabla_i^2 - \frac{1}{2} \sum_A^{\text{nuclei}} \frac{1}{M_A} \nabla_A^2 - \sum_A^{\text{nuclei}} \sum_i^{\text{electrons}} \frac{Z_A}{r_{iA}} + \sum_{i < j}^{\text{electrons}} \frac{1}{r_{ij}} + \sum_{A < B}^{\text{nuclei}} \frac{Z_A Z_B}{R_{AB}} \quad (1.13)$$

The ratio of the mass of the proton to that of the electron is $\frac{m_p}{m_e} \approx 1836$.

Therefore even the lightest nucleus, hydrogen, with its single proton is three orders of magnitude heavier than the electron. This large disparity in mass means that the relatively light electron will move much more quickly than the nucleus to which it is attached. This implies a difference in the time scales governing the motion of electrons and nuclei. The electrons will be able to execute several periods of their motion before the nuclei have moved to any significant degree. This means that we can quantise the motion of the electrons for a fixed position of the nuclei. If the nuclei are fixed then their kinetic energy

is zero. Hence we can remove the term $-\frac{1}{2} \sum_A^{\text{nuclei}} \frac{1}{M_A} \nabla_A^2$ from the molecular hamiltonian in eqn (1.13), to yield the simpler “electronic” hamiltonian

$$\hat{H}_{\text{Electronic}} = -\frac{1}{2} \sum_i^{\text{electrons}} \nabla_i^2 - \sum_A^{\text{nuclei}} \sum_i^{\text{electrons}} \frac{Z_A}{r_{iA}} + \sum_{i < j}^{\text{electrons}} \frac{1}{r_{ij}} \quad (1.14)$$

Since the kinetic energy of the nuclei are assumed to be zero, the potential energy of repulsion between nuclei assumes a constant value (for a given position of the nuclei). This is termed the “nuclear repulsion” energy and is given by

$$V_{AB} = \sum_{A > B}^{\text{nuclei}} \frac{Z_A Z_B}{R_{AB}} \quad (1.15)$$

Using the electronic hamiltonian in eqn (1.14) the electronic Schrödinger equation may be solved

$$\hat{H}_{\text{Electronic}} |\Psi_{\text{Electronic}}\rangle = E_{\text{Electronic}} |\Psi_{\text{Electronic}}\rangle \quad (1.16)$$

The solution of this equation is the electronic wavefunction, $|\Psi_{\text{Electronic}}\rangle$, which depends explicitly on the set of all electronic coordinates $\{\mathbf{r}_i\}$ and parametrically on the set of all nuclear coordinates, $\{\mathbf{R}_A\}$. Formally, we should write the electronic wavefunction as $|\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_A\})_{\text{Electronic}}\rangle$ to show this dependence. We shall omit writing the explicit dependence to simplify the notation. The parametric dependence on $\{\mathbf{R}_A\}$ implies that for every different arrangement of the nuclei, the electronic wavefunction will be different and accordingly the electronic energy obtained from eqn (1.16) will vary with the nuclear arrangement (geometry) of the system. The total molecular potential energy is given by the sum of the electronic energy, obtained from eqn (1.16), and the nuclear repulsion energy, obtained from eqn (1.15):

$$E_{\text{Total}} = E_{\text{Electronic}} + V_{AB} \quad (1.17)$$

E_{Total} is only a “potential” energy in as much as the motion of the nuclei are being considered (since $E_{\text{Electronic}}$ contains the kinetic energy of the electrons). E_{Total} depends on $\{\mathbf{R}_A\}$ and plotting E_{Total} as a function of $\{\mathbf{R}_A\}$ yields a potential energy curve or surface. This is the Born-Oppenheimer approximation. The simplest example of this is the familiar plot of potential energy against internuclear distance for a diatomic molecule, for example hydrogen chloride as shown in Figure 1.2.

We now have a form for $\hat{H}_{\text{Electronic}}$ from which we can obtain E_{Total} provided we know $|\Psi_{\text{Electronic}}\rangle$, the electronic wavefunction. The Born-Oppenheimer approximation is the first of our key approximations in dealing with the Schrödinger equation. It is the central starting point to nearly all quantum chemical methods.

We shall not consider the nuclear problem in any depth but note that if we wished to consider the nuclear motion, for example to study the effects of isotopic substitution, we would introduce the appropriate nuclear Schrödinger equation

$$\hat{H}_{\text{Nuclear}} |\Psi_{\text{Nuclear}}\rangle = E_{\text{Nuclear}} |\Psi_{\text{Nuclear}}\rangle \quad (1.18)$$

This nuclear equation governs the vibrations, rotations and translations of a molecule. These different types of nuclear motion often operate on different time and energy scales and, provided there is no strong coupling between the different types of motion, we can decompose eqn (1.18) as

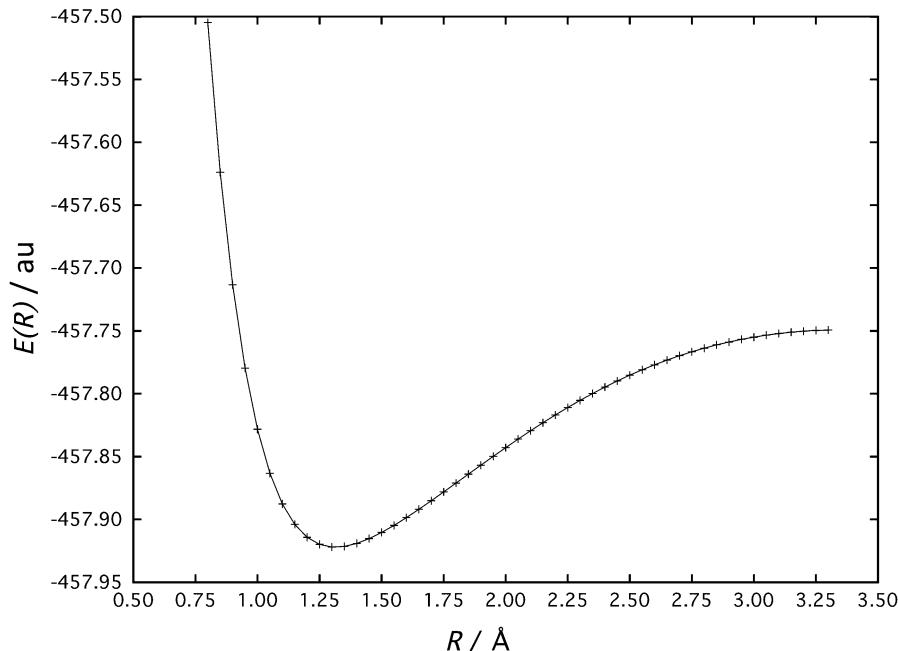


Figure 1.2 Potential energy curve for hydrogen chloride.

$$\begin{aligned}\hat{H}_{\text{Nuclear}} &= \hat{H}_{\text{Vibration}} + \hat{H}_{\text{Rotation}} + \hat{H}_{\text{Translation}} \\ |\Psi_{\text{Nuclear}}\rangle &= |\Psi_{\text{Vibration}}\rangle |\Psi_{\text{Rotation}}\rangle |\Psi_{\text{Translation}}\rangle \\ E_{\text{Nuclear}} &= E_{\text{Vibration}} + E_{\text{Rotation}} + E_{\text{Translation}}\end{aligned}\quad (1.19)$$

Since we have assumed there to be no strong coupling terms, we can proceed more efficiently by solving independent Schrödinger equations for vibration, rotation and translation:

$$\begin{aligned}\hat{H}_{\text{Vibration}} |\Psi_{\text{Vibration}}\rangle &= E_{\text{Vibration}} |\Psi_{\text{Vibration}}\rangle \\ \hat{H}_{\text{Rotation}} |\Psi_{\text{Rotation}}\rangle &= E_{\text{Rotation}} |\Psi_{\text{Rotation}}\rangle \\ \hat{H}_{\text{Translation}} |\Psi_{\text{Translation}}\rangle &= E_{\text{Translation}} |\Psi_{\text{Translation}}\rangle\end{aligned}\quad (1.20)$$

Adding in the separation of the electronic problem, as discussed above, we can summarise the outcome of applying the Born-Oppenheimer approximation to eqn (1.3) as yielding the energy, for all internal and external motions of a molecule, which can be written as

$$E = E_{\text{Electronic}} + E_{\text{Vibration}} + E_{\text{Rotation}} + E_{\text{Translation}} \quad (1.21)$$

In all that follows, unless we explicitly state otherwise, we shall be concerned with the electronic problem and the energy, E_{Total} , as given in eqn (1.17). Note that E_{Total} depends on atom type through the nuclear charge Z , but is independent of nuclear mass. In referring to the electronic hamiltonian and wavefunction we shall now omit the subscript “Electronic” to simplify the notation.

1.2.2 Electronic Wavefunctions and the Antisymmetry Principle

The next key idea necessary to develop our approach is based on experimental evidence from the electronic spectroscopy of atoms, which led Wolfgang Pauli in 1925 to introduce the concept of spin for all fundamental particles. Spin is quantified by the spin quantum number, s . For example, electrons and protons are particles with spin value, $s = \frac{1}{2}$. Non-integer values of s characterise particles known as “fermions”. Helium nuclei have, $s = 0$, and the photon, $s = 1$. Particles with integer values of s are called “bosons”. Returning to the case of electrons with $s = \frac{1}{2}$, and introducing a magnetic field, the electron can align itself either parallel or antiparallel to the field. These orientations are described by the spin magnetic quantum number, $m_s = +\frac{1}{2}$ (parallel) and $m_s = -\frac{1}{2}$ (antiparallel). To satisfy the known physical evidence of electronic states in atoms Pauli formulated, as an independent postulate of quantum mechanics, the antisymmetry principle which states that:

“The wavefunction describing any state of an N -electron system must be antisymmetric under any permutation of the electronic coordinates.”

The consequence of this principle is that no two identical fermions (in our case electrons) can occupy the same quantum state simultaneously. More familiarly, for electrons in atoms, we encounter the Pauli exclusion principle as the requirement that no two electrons can have the same four quantum numbers. If the principal quantum number, n , the angular quantum number, l , and the magnetic quantum number, m_l , are the same for two electrons, then the spin magnetic quantum number, m_s , must be different. The outcome of this is the requirement that if two electrons enter the same atomic orbital their spins must be antiparallel. The antisymmetry principle is a generalisation of this requirement that extends it to any number of electrons in any type of system.

We have referred to electrons entering orbitals in an atom and at this stage we must elaborate a little on what we mean by an orbital. An orbital is a solution to a one-electron Schrödinger equation. For example, s , p , d , f ... orbitals are all solutions of the Schrödinger equation for the hydrogen atom. The electronic hamiltonian given in eqn (1.14) and the electronic Schrödinger equation given eqn (1.16), contain no reference to spin. To accommodate the rôle of spin we must distinguish between spatial orbitals, $\phi(\mathbf{r})$, which depend solely on spatial (positional) coordinates \mathbf{r} , and spin-orbitals, $\phi(\mathbf{x})$, in which the variable \mathbf{x} includes the space (\mathbf{r}) and spin (\mathbf{s}) variables, $\mathbf{x} = \{\mathbf{r}, \mathbf{s}\}$. The spin variable has associated with it the spin operators $\hat{s}_x, \hat{s}_y, \hat{s}_z$. The hamiltonian

given in eqn (1.13) works on functions of \mathbf{r} , and the spin operators work on functions of \mathbf{s} . Writing the spin-orbital $\phi(\mathbf{x})$ as a product of space and spin-functions

$$\phi(\mathbf{x}) = \phi(\mathbf{r})\omega(\mathbf{s}) \quad (1.22)$$

we find that $\phi(\mathbf{x})$ is simultaneously a solution of (in atomic units)

$$\begin{aligned} \hat{s}_z \phi(\mathbf{x}) &= m_s \phi(\mathbf{x}) \\ \hat{h} \phi(\mathbf{x}) &= \varepsilon \phi(\mathbf{x}) \end{aligned} \quad (1.23)$$

In the second equation above, \hat{h} refers to a one-electron hamiltonian operator. Observation, in the form of the Stern-Gerlach experiment of 1921, showed that $\hat{s}_z \omega(\mathbf{s}) = m_s \omega(\mathbf{s})$ has only two solutions, which are usually denoted α and β , or \uparrow and \downarrow , respectively:

$$\begin{aligned} \hat{s}_z \alpha(\mathbf{s}) &= +\frac{1}{2} \alpha(\mathbf{s}) \\ \hat{s}_z \beta(\mathbf{s}) &= -\frac{1}{2} \beta(\mathbf{s}) \end{aligned} \quad (1.24)$$

Hence each spatial orbital $\phi(\mathbf{r})$ yields two possible spin-orbitals

$$\begin{aligned} \phi(\mathbf{x}) &= \phi(\mathbf{r})\alpha(\mathbf{s}) \\ \bar{\phi}(\mathbf{x}) &= \phi(\mathbf{r})\beta(\mathbf{s}) \end{aligned} \quad (1.25)$$

The antisymmetry principle stated above, refers to the interchange of the coordinates, \mathbf{x}_1 and \mathbf{x}_2 , of two electrons, that is the wavefunction must be antisymmetric under the permutation of both space and spin coordinates.

There are many ways to construct wavefunctions which satisfy the antisymmetry principle. Here we shall consider only one simple scheme. The electronic wavefunction, $|\Psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \dots \mathbf{x}_N)\rangle$, for an N -electron atom or molecule is written as a determinant of N spin-orbitals, called a “Slater determinant”:

$$|\Psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \dots \mathbf{x}_N)\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_i(\mathbf{x}_1) & \phi_j(\mathbf{x}_1) & \cdots & \phi_k(\mathbf{x}_1) \\ \phi_i(\mathbf{x}_2) & \phi_j(\mathbf{x}_2) & \cdots & \phi_k(\mathbf{x}_2) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_i(\mathbf{x}_N) & \phi_j(\mathbf{x}_N) & \cdots & \phi_k(\mathbf{x}_N) \end{vmatrix} \quad (1.26)$$

The $\frac{1}{\sqrt{N!}}$ term is a normalisation factor, since expanding an $N \times N$ determinant will produce $N!$ terms, see Appendix 1B. The Slater determinant is completely specified by the spin-orbitals from which it is constructed. We note the following properties of this form for $|\Psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \dots \mathbf{x}_N)\rangle$:

- (a) N electrons are distributed in N spin-orbitals, but with no specification of which electron is associated with which orbital. Electrons are physically indistinguishable particles and here they are associated equally with all orbitals in the Slater determinant.
- (b) It is a property of determinants that interchanging two rows or columns changes the sign of the determinant (Appendix 1B). Swapping two rows in eqn (1.26) corresponds to swapping the coordinates of two electrons, and this is reflected in a change of sign in $|\Psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \dots \mathbf{x}_N)\rangle$ satisfying the antisymmetry requirement.
- (c) If two electrons occupy the same spin-orbital, that is possess the same \mathbf{x} coordinates, then two rows of $|\Psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \dots \mathbf{x}_N)\rangle$ will be equivalent. Any determinant with two equivalent rows, or two equivalent columns, when evaluated will yield zero. Hence $|\Psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \dots \mathbf{x}_N)\rangle$ vanishes and is not allowed. This satisfies the Pauli exclusion principle.

For many electronically simple atoms and molecules we can adopt a single Slater determinant as a satisfactory first approximation to the electronic wavefunction. Later on we shall see that a single Slater determinant is also a satisfactory starting point for developing the exact wavefunction. It is often useful to adopt a shorthand notation for the Slater determinant in eqn (1.26) in which only the diagonal elements are specified and the electron order is implicit

$$|\Psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \dots \mathbf{x}_N)\rangle = |\phi_i(\mathbf{x}_1)\phi_j(\mathbf{x}_2) \dots \phi_k(\mathbf{x}_N)\rangle = |\phi_i\phi_j \dots \phi_k\rangle \quad (1.27)$$

1.2.3 Molecular Orbitals and Basis Set Expansions

At this stage we have a form for \hat{H} given in eqn (1.14) and a form for $|\Psi\rangle$ given in eqn (1.26). To proceed further we must add a little more detail to the specification of $|\Psi\rangle$. We have stated that the Slater determinant is completely specified by the spin-orbitals, $\phi(\mathbf{x})$, from which it is built. Some functional form must be chosen for the set of atomic or molecular orbitals $\{\phi\}$ that enter eqn (1.26). The usual choice is to approximate each $\phi(\mathbf{r})$ as a linear combination of atomic orbitals (LCAO), with the atomic orbitals being located on the nuclei. In practice, each spatial orbital, $\phi(\mathbf{r})$, is written as

$$\phi_i(\mathbf{r}) = \sum_{\mu=1}^m c_{\mu i} \chi_{\mu}(\mathbf{r}) \quad (1.28)$$

where the set $\{\chi\}$ is a set of m ‘atomic basis functions’ usually referred to as a basis set. The $c_{\mu i}$ are the mixing coefficients of the LCAO, and they are determined by minimising the energy. The process of optimising the $c_{\mu i}$ is called the “self-consistent field” (SCF) method. We shall say much more about the SCF process in chapter 2.

In principle, the set of functions $\{\chi\}$ can be of any type, provided that they are able to represent the molecular orbitals accurately and efficiently. If a great many

functions are required to obtain accuracy, then the basis set will incur a computational cost that makes it inefficient. Equally, if a particular type of basis function leads to very compact basis sets, but the resultant integrals are computationally demanding, again this will be inefficient. As is often the case in computational quantum chemistry it is necessary to find a balance between accuracy and efficiency. The use of true atomic orbitals, solutions of the Schrödinger equation for the hydrogen atom, is accurate but inefficient. In analogy with atomic orbitals it is possible to use Slater basis functions that have a radial dependence of $e^{-\zeta r}$, in which r is the distance from the nucleus and ζ is an exponent that determines the radial distribution. Large values of ζ contract the function nearer to the nucleus, whereas small values of ζ expand the function away from the nucleus. Slater orbitals have been found to be computationally less efficient than the gaussian type orbitals with a radial dependence of $e^{-\alpha r^2}$, where α is the exponent of the gaussian type orbital. This subtle difference in the radial form has a very significant effect on the speed with which integrals may be evaluated in molecular systems. While both Slater and gaussian type orbitals are used, the gaussian type orbitals are used much more widely and dominate in most quantum chemical applications and developments. These two types of function show very different radial behaviour, see Figure 1.3. The radial dependence of the Slater type orbital possesses a cusp at the nucleus: an infinite gradient. In contrast,

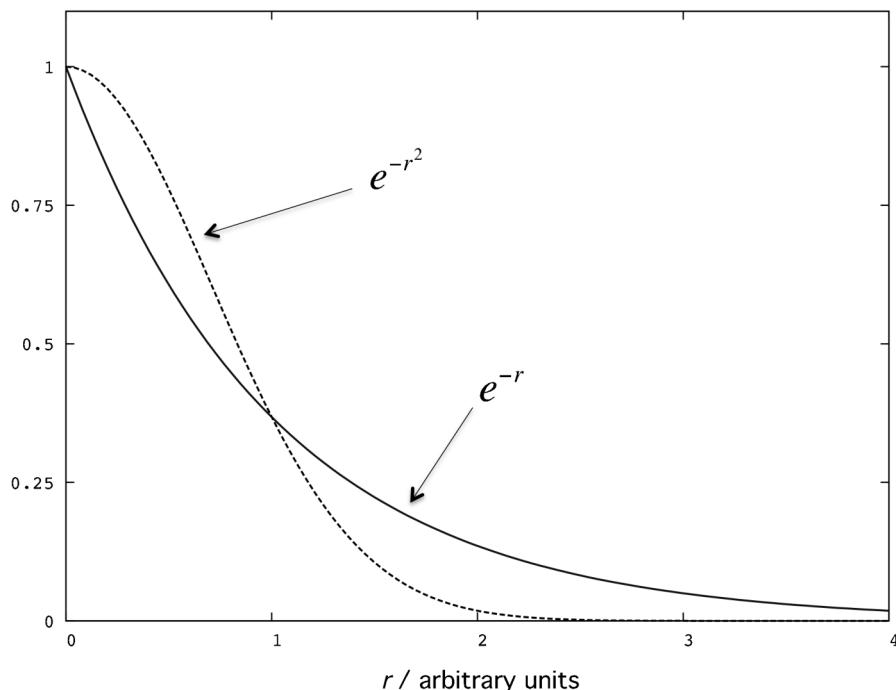


Figure 1.3 Comparison of the radial dependence of Slater type (solid line) and gaussian type orbitals (dashed line). Nucleus is located at the origin.

the gaussian type orbital has a zero gradient at the nucleus and decays more rapidly at larger distances from the nucleus. Nevertheless it turns out that the gaussian type orbitals are so much more efficient, computationally, that it is preferable to use larger sets of gaussian type orbitals than to deal with more compact Slater type orbital sets.

In all that follows we shall only concern ourselves with gaussian type orbitals. The choice of basis set, $\{\chi\}$, is critical to the success of any proposed calculation and we shall return to this subject in detail in Chapter 2.

The three approximations we have discussed so far: the Born-Oppenheimer approximation; representing the electronic wavefunction by a single Slater determinant; and the expansion of orbitals in atomic basis functions, constitute the most widely adopted approaches in computational quantum chemistry. While each approximation lends itself to improvement and refinements, a great deal of chemistry can be studied, with some reliability, using these simple ideas.

1.3 Chemistry *in Silico*: Where Do You Start?

From what has been presented so far we can assume that we are able to compute the potential energy of a molecule for any choice of the coordinates (geometry) of the constituent atoms. How does this enable us to study chemical problems? The discussion of chemical processes is often formulated in terms of the properties of reactants, products and transition states, from which information about thermochemistry, rates of reaction and molecular structure may be derived. By defining the notions of reactant, product, transition state and thermodynamic functions in terms of quantities that we can compute, we can begin to formulate a complete computational quantum chemistry.

1.3.1 Potential Energy Curves, Forces and Force Constants

Let us return to the familiar potential energy curve for a diatomic molecule, Figure 1.4, we have only one coordinate (bond distance) to consider and associate the minimum energy point on the curve with the equilibrium bond length, R_e . The minimum energy point is characterised by being a stationary point on the potential energy curve. A stationary point is defined by a zero first derivative:

$$\frac{dE(R)}{dR} = 0 \quad (1.29)$$

Physically, this quantity can be related to the force acting on the atoms

$$force = - \frac{dE(R)}{dR} \quad (1.30)$$

The negative sign shows that the force acts in the direction that will lower the

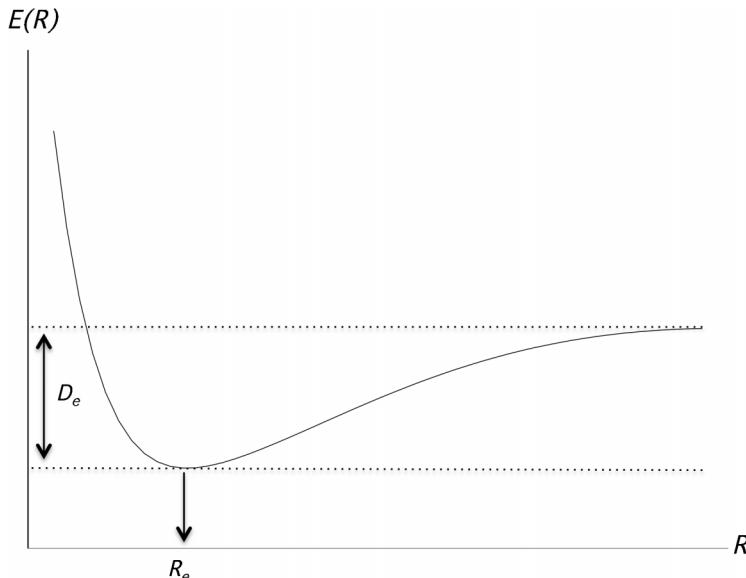


Figure 1.4 Using a potential energy curve to find the equilibrium bond length, R_e , and the bond dissociation energy, D_e .

potential energy. So any point beyond R_e has a positive slope, $dE(R)/dR$, and the force is negative. Beyond R_e , the bond distance is greater than at equilibrium and the force acts in the direction that would lower the potential energy, corresponding to contraction of the bond distance. At distances less than R_e , the slope of the curve is negative and the force acts in the direction of bond elongation.

We can also obtain the bond dissociation energy (in the absence of vibrational zero point energy) as the difference between the minimum energy point and the energy of the potential energy curve when the bond distance is large enough that any further change does not appreciably affect the potential energy (this is the plateau region on the far right hand side of the curve). This gives us two pieces of chemically useful information: the bond length and the bond strength. Another useful piece of information we can obtain is the vibrational wavenumber (frequency), which for a diatomic molecule is given by

$$\tilde{\nu} = \frac{1}{2\pi c} \sqrt{\frac{k}{\mu}} \quad (1.31)$$

where c is the speed of light, μ is the reduced mass which for a diatomic molecule A–B is written in terms of the constituent atomic masses as

$$\mu = \frac{M_A M_B}{M_A + M_B} \quad (1.32)$$

In eqn (1.31), k is the force constant which is the second derivative of the potential energy with respect to the bond distance, R , evaluated at the minimum energy point, R_e ,

$$k = \left(\frac{d^2 E(R)}{dR^2} \right)_{R_e} \quad (1.33)$$

1.3.2 Potential Energy Surfaces, Stationary Points and Reactivity

How do these ideas change when we move to molecules with more than one geometric coordinate? Consider the water molecule and its potential energy as a function of the valence angle, θ , and the symmetric stretching of the O–H bonds, R , as shown in Figure 1.5. We can now construct the potential energy as a function of both coordinates, $E(R, \theta)$. The two geometric coordinates and the energy produce a three dimensional surface shown in Figure 1.6. At the minimum energy point we can read off the equilibrium values of the two geometric coordinates. As in the case of the diatomic potential energy curve, Figure 1.4, the minimum energy point is characterised as a stationary point which implies

$$\begin{aligned} \frac{dE(R, \theta)}{dR} &= 0 \\ \frac{dE(R, \theta)}{d\theta} &= 0 \end{aligned} \quad (1.34)$$

The force is no longer a single number but a vector, $-\mathbf{g}$, containing two components:

$$\mathbf{g} = \begin{pmatrix} \frac{dE(R, \theta)}{dR} \\ \frac{dE(R, \theta)}{d\theta} \end{pmatrix} \quad (1.35)$$

Accordingly, the second derivative is no longer a single number but a 2×2

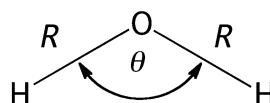


Figure 1.5 Coordinates for the symmetric stretching and bending of water.

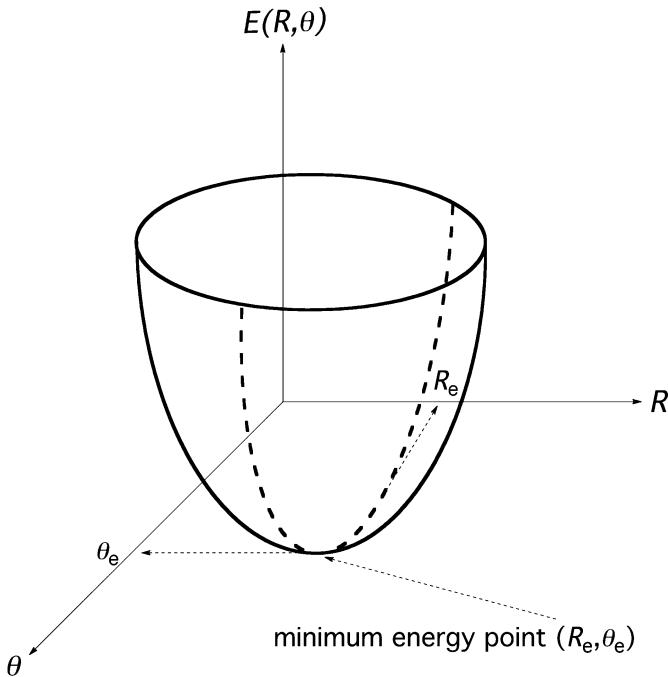


Figure 1.6 Schematic three-dimensional potential energy surface for water, the coordinates correspond to those in **Figure 1.5**.

matrix, \mathbf{F} :

$$\mathbf{F} = \begin{pmatrix} \frac{d^2E(R,\theta)}{dR^2} & \frac{d^2E(R,\theta)}{dRd\theta} \\ \frac{d^2E(R,\theta)}{d\theta dR} & \frac{d^2E(R,\theta)}{d\theta^2} \end{pmatrix} \quad (1.36)$$

If we had considered all the possible internal motions of the water molecule we would need to allow the two O–H bonds to move independently, each having its own internal coordinate. The potential energy would then depend on three coordinates, for example $E(R_1, R_2, \theta)$, yielding a four dimensional *hypersurface*. The gradient vector, \mathbf{g} , would now contain three terms and the force constant matrix, \mathbf{F} , would be a 3×3 matrix. Beyond three dimensions we should formally refer to a *hypersurface* but, in practice, potential energy functions all tend to be referred to as “*potential energy surfaces*.”

In general, for a molecule consisting of N_{atoms} atoms there will be $3N_{\text{atoms}} - 6$ independent geometric variables (bond lengths, bond angles and torsional angles). The set of independent variables is not unique. For example, the water molecule can be represented equally by either of the two sets of variables shown in Figure 1.7.

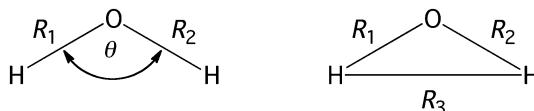


Figure 1.7 Two different sets of $3N_{\text{atoms}} - 6 = 3$ variables for describing the water molecule.

The sum, $3N_{\text{atoms}} - 6$, specifies the number of internal motions that a molecule may possess. If we consider the general set of cartesian coordinates, $\{x, y, z\}$, then each atom in a molecule may be specified by three such cartesian coordinates relative to an origin. This provides a total of $3N_{\text{atoms}}$ cartesian coordinates. This coordinate set describes not only the internal motions, one atom moving relative to another, but also the external motions. The external motions are the global translations and rotations of the molecule. We can translate an isolated molecule as a whole along any of the cartesian axes without changing the energy. To see this consider that in eqn (1.13) all the terms depend only on the relative position of one atom to another. Equally we can rotate an isolated molecule about any of the cartesian axes, again without changing the energy. Hence our set of coordinates should include $3N_{\text{atoms}} - 3$ coordinates for global translations and -3 coordinates for global rotations giving a total of $3N_{\text{atoms}} - 6$. In the case of a linear molecule there are only two axes for rotation, since rotation along the bond axis does not displace the molecule in space. Hence for linear molecules the number of internal coordinates is $3N_{\text{atoms}} - 5$.

A stationary point is defined for a general potential energy surface by

$$\frac{dE(\mathbf{R})}{dR_i} = 0 \quad i = 1, 2, \dots, 3N_{\text{atoms}} - 6 \quad (1.37)$$

where \mathbf{R} refers to the set of all nuclear coordinates and R_i to a specific member of the set. So provided we can locate a geometry that satisfies eqn (1.37), then we can find a minimum energy configuration for any molecular system. How we go about locating stationary points on surfaces of arbitrary complexity we shall discuss in chapter 3. Eqn (1.37) is a necessary condition for finding a minimum energy configuration, but it is not a sufficient condition. When we deal with chemical reactions, we are no longer just interested in minimum energy configurations. Let us continue with the simple example of water in the set of two coordinates shown in Figure 1.5 and consider the inversion of the angle as an example of a very simple reaction, as shown in Figure 1.8. We know that the minimum energy structure of water has a bond angle of $\theta_e = 104.5^\circ$, and this will be the value assumed for this variable in the reactant

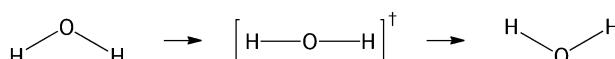


Figure 1.8 Inversion of the bond angle in water as an example of a chemical reaction.

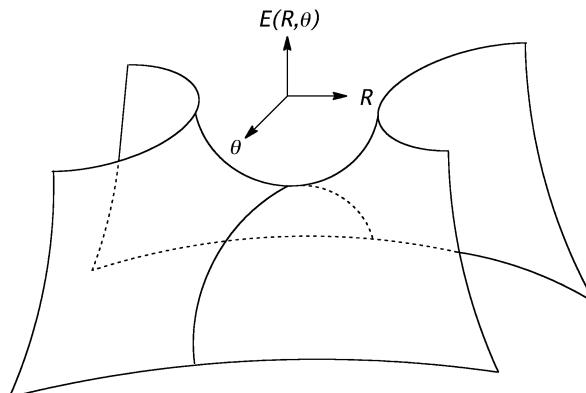


Figure 1.9 Saddle point structure around the transition state corresponding to the reaction shown in **Figure 1.8**.

and (equivalent) product shown in Figure 1.8. The linear structure will correspond to a stationary point, but it is a transition structure rather than a minimum on the potential energy surface. The potential energy surface around the linear structure is shown schematically in Figure 1.9 and corresponds to a saddle point structure. In the stationary saddle point structure one unique coordinate lies at a maximum, and all other coordinates (only one in this example) correspond to minima. Along the R coordinate the energy contours are those of a minimum, $d^2E(R,\theta)/dR^2 > 0$. Conversely, the energy contours along the θ coordinate correspond to a maximum, $d^2E(R,\theta)/d\theta^2 < 0$. If we plot the potential energy as a function of θ (at the corresponding minimum value along the R coordinate) we obtain the reaction energy profile shown in Figure 1.10. From the shape of the energy profile we can associate the coordinate, θ , with the reaction coordinate. The difference in energy between $\theta = \theta_e$ and $\theta = 180^\circ$ gives the activation barrier for this process. This is a

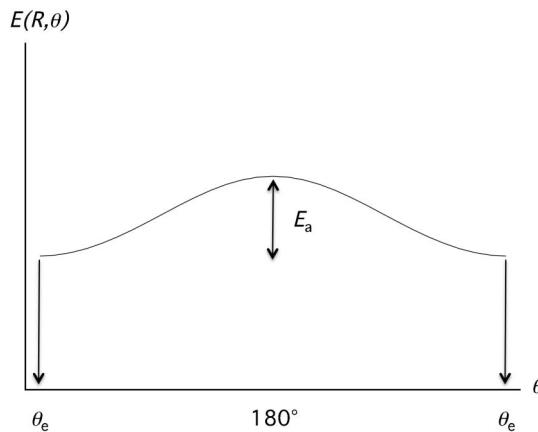


Figure 1.10 Energy profile for the reaction shown in **Figure 1.8**.

particularly simple example in which the reaction coordinate can be described in terms of just one internal coordinate, θ . In most reactions, the reaction coordinate will depend on many internal coordinates. In fact the reaction coordinate is, in principle, a function of all $3N_{\text{atoms}} - 6$ internal coordinates.

We have seen how a stationary point may correspond to a minimum or a saddle point transition structure. What distinguishes them is the curvature of the potential energy surface around the stationary point. So in general when studying chemical processes of arbitrary complexity we must first of all find stationary points that satisfy eqn (1.37) and then characterise these stationary points by computing the force constant (second derivative) matrix. The second derivatives contain the information about the curvature of the surface. Having computed the force constant matrix (eqn (1.36)), we must go one step further and transform it to diagonal form:

$$\mathbf{Q}^{-1}\mathbf{F}\mathbf{Q} = \begin{pmatrix} \frac{d^2E(R,\theta)}{dq_1^2} & 0 \\ 0 & \frac{d^2E(R,\theta)}{dq_2^2} \end{pmatrix} \quad (1.38)$$

Given the matrix \mathbf{F} , the matrix \mathbf{Q} can be found using standard numerical methods, see Appendix 1C. The columns of \mathbf{Q} give linear combinations of the initial set of variables $\{R, \theta\}$. These linear combinations are called normal coordinates, and denoted $\{q_1, q_2\}$, and can be written as

$$\begin{aligned} q_1 &= Q_{11}R + Q_{21}\theta \\ q_2 &= Q_{12}R + Q_{22}\theta \end{aligned} \quad (1.39)$$

where Q_{ij} refers to the elements of the transformation matrix \mathbf{Q} . For the simple case considered, q_1 and q_2 will correspond to symmetric (positive) and antisymmetric (negative) combinations of R and θ . Each normal coordinate will be dominated by either R or θ , so $Q_{11} >> Q_{21}$ and $Q_{22} >> Q_{12}$ in eqn (1.39), since these two coordinates do not couple very strongly. By contrast, in a reaction such as the collinear exchange shown in Figure 1.11, the reaction proceeds via the antisymmetric combination of R_1 and R_2 . The transition structure being formed from the reactants by the shortening of R_1 and the lengthening of R_2 ,

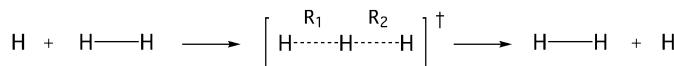


Figure 1.11 Collinear exchange reaction involving two coordinates R_1 and R_2 .

$$\begin{aligned} q_1 &= Q_{11}R_1 + Q_{21}R_2 \\ q_2 &= Q_{12}R_1 + Q_{22}R_2 \end{aligned} \quad (1.40)$$

At the transition structure geometry, the normal coordinate corresponding to the maximum in the energy profile will be an equal mixture of R_1 and R_2 ($Q_{11} = -Q_{21}$), while on the reactant side the reaction coordinate will be dominated by R_2 ($Q_{21} > Q_{11}$) and on the product side by R_1 ($Q_{11} > Q_{21}$). So we can see that the reaction coordinate is not a fixed linear combination of internal coordinates but changes as a reaction proceeds.

Chemical reactants, products and stable intermediates are characterised as minima and must have

$$\frac{d^2E(\mathbf{R})}{dq_i^2} > 0 \quad i = 1, 2, \dots, 3N_{\text{atoms}} - 6 \quad (1.41)$$

For transition structures there will be one unique normal coordinate which corresponds to the reaction coordinate. Hence a transition structure is characterised by

$$\frac{d^2E(\mathbf{R})}{dq_i^2} > 0 \quad i = 1, 2, \dots, 3N_{\text{atoms}} - 7 \quad (1.42)$$

and one unique coordinate, q_{RC} , corresponding to the reaction coordinate for which

$$\frac{d^2E(\mathbf{R})}{dq_{RC}^2} < 0 \quad (1.43)$$

Stationary points with more than one coordinate, or direction, in which the potential energy surface is a maximum are not of any chemical significance since there will always be a path of lower energy by which a reaction may proceed.

The conditions given in eqn (1.37), (1.41), (1.42) and (1.43) are sufficient to determine whether a stationary point refers to a stable structure or a transition structure, but they cannot tell us how the stationary point relates to the global structure of the potential energy surface. To illustrate this, consider the torsional conformations of 1,2-dichloroethane shown in Figure 1.12. The torsion angle, τ , is taken as the angle between the two C–Cl bonds looking along the C–C bond. The lowest energy conformation corresponds to $\tau = 180^\circ$ and is indicated in Figure 1.12 as **I**. Each 60° rotation in τ leads from conformation **I** to another stationary point. The first of these corresponds to an eclipsed conformation, **II**. **II** is a transition structure connecting conformation **I** to conformation **III**. **III** is a minimum energy structure but lies higher in energy than conformation **I**. **III** is called a *local minimum*, while **I**

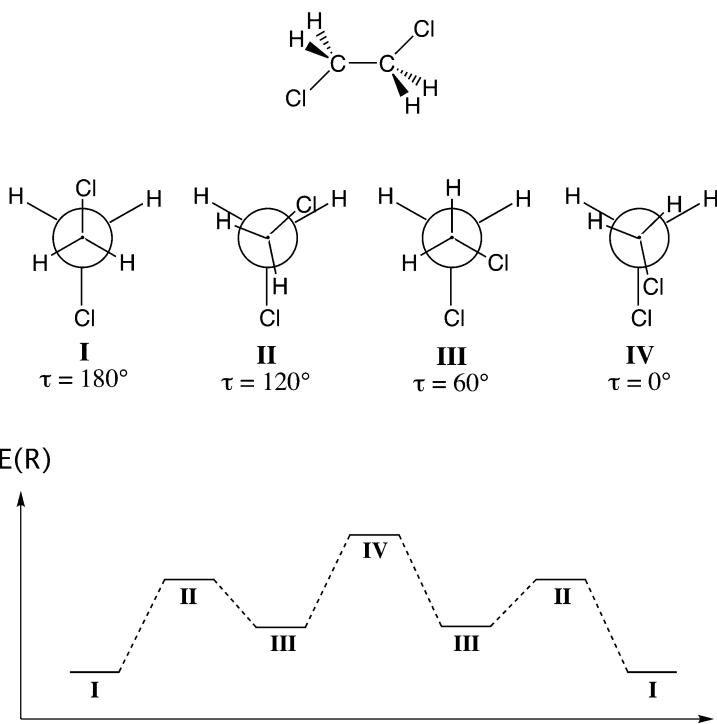


Figure 1.12 Conformations and torsional energies of 1,2-dichloroethane.

is the *global minimum*. Rotation through another 60° from **III** takes us to transition structure **IV**. Both **II** and **IV** are characterised by a maximum in the potential energy along the τ coordinate and will satisfy the conditions in eqns (1.42) and (1.43). There are two rotational barriers, the first corresponding to $E(\text{II}) - E(\text{I})$ and the second to $E(\text{IV}) - E(\text{III})$. At non-zero temperatures, the equilibrium distribution of 1,2-dichloroethane will consist of conformers **I** and **III**. The relative fractions of each component will depend on the energy difference, $E(\text{III}) - E(\text{I})$, as given by the Boltzmann distribution.

1.3.3 Linking the Electronic Energy with Thermodynamic State Functions

So now, in principle, we can study systems of arbitrary complexity by applying these simple ideas. We must use $dE(R)/dR_i$ and $d^2E(R)/dq_i^2$ to enable us to think about chemical processes on high-dimensional potential energy surfaces. The difference in energy between different stationary points provides us with energies of reaction or activation energies. Activation energies govern the kinetics of all activated processes, while energies of reaction determine reaction equilibrium and other thermochemical features of a reaction. The energies we have discussed so far refer to an isolated molecule at the absolute zero of

temperature. We must link these with the thermodynamic quantities familiar from experiments, for example enthalpies and Gibbs energies, which are usually defined for molar quantities. To do this we use statistical mechanical expressions to relate our computed internal energies at 0 K, $U(0)_{\text{Electronic}}$, to enthalpies or free energies at non-zero temperatures, $H(T)$ and $G(T)$, respectively. Note that the experimentally derived quantities are usually the changes in these state functions, ΔU , ΔH , ΔG . The computations we describe will yield the absolute quantities, U , H , G , for each reactant and product we consider, and from these ΔU , ΔH , ΔG are obtained.

Quantum chemical calculations provide us with $E(\mathbf{R}) = U(0)_{\text{Electronic}}$, we now need to relate this to the internal energy at some (non-zero) temperature, T . To proceed, using the machinery of statistical mechanics, we assume that the different energetic degrees of freedom (electronic, vibration, rotation, translation) do not couple and we can write

$$U(T) = U(T)_{\text{Translation}} + U(T)_{\text{Rotation}} + U(T)_{\text{Vibration}} + U(T)_{\text{Electronic}} \quad (1.44)$$

We now consider each term independently. In molar units we find:

$$U(T)_{\text{Translation}} = \frac{3}{2}RT \quad (1.45)$$

$$U(T)_{\text{Rotation}} = \frac{3}{2}RT \text{ (for a non-linear molecule)} \quad (1.46a)$$

$$U(T)_{\text{Rotation}} = RT \text{ (for a linear molecule)} \quad (1.46b)$$

$$U(T)_{\text{Vibration}} = U(0)_{\text{Vibration}} + N_A \sum_i^{3N_{\text{atoms}}-6} \frac{hv_i}{e^{hv_i/k_B T} - 1} \quad (1.47)$$

The vibrational term at 0 K is the quantum mechanical zero-point energy and is given by:

$$U(0)_{\text{Vibration}} = \frac{1}{2}N_A \sum_i^{3N_{\text{atoms}}-6} hv_i \quad (1.48)$$

In eqns (1.45) – (1.48), R is the gas constant, N_A is Avogadro's number, k_B is Boltzmann's constant, h is Planck's constant and v_i is the i th vibrational wavenumber (frequency).

The final term to consider is $U(T)_{\text{Electronic}}$. This is usually taken as the computed potential energy, $U(T)_{\text{Electronic}} \approx U(0)_{\text{Electronic}} = E(\mathbf{R})$, implying that there is no thermal correction to the electronic component. This is a reasonable

assumption provided that the electronic levels in a molecule are energetically well separated.

The thermodynamic relationship between internal energy, U , and enthalpy, H , is

$$H = U + pV \quad (1.49)$$

Quantum chemical calculations usually refer to the gas phase and for an ideal gas we know, $pV = nRT$ and can write (for molar quantities)

$$H(T) = U(T) + RT \quad (1.50)$$

So given a calculation of $U(0)_{\text{Electronic}}$, it is a straightforward matter to obtain $H(T)$. It is often useful to discuss chemical processes in terms of the Gibbs energy, $G(T)$. To extend our treatment to cover this state function we need to know how to calculate the entropy, $S(T)$. At this stage, the reader should be aware that there are many equivalent ways of writing the statistical mechanical expressions for the entropy. We proceed as for the internal energy and write

$$S(T) = S(T)_{\text{Translation}} + S(T)_{\text{Rotation}} + S(T)_{\text{Vibration}} + S(T)_{\text{Electronic}} \quad (1.51)$$

Taking each term in turn and again working in molar units, we obtain:

$$S(T)_{\text{Translation}} = R \left\{ \frac{5}{2} + \ln \left[\left(\frac{RT}{p} \right) \left(\frac{2\pi mk_B T}{h^2} \right)^{3/2} \right] \right\} \quad (1.52)$$

In eqn (1.52), m is the molecular mass and in the SI system of units should be expressed in kg. To evaluate the rotational entropy for a non-linear molecule we need to know the three principal moments of inertia, I_A , I_B , I_C , (obtained by diagonalising the inertia tensor, see Appendix 1D) and the rotational symmetry number, σ . With these quantities at hand we can evaluate $S(T)_{\text{Rotation}}$ for a non-linear molecule as

$$S(T)_{\text{Rotation}} = R \left\{ \frac{3}{2} + \ln \left[\left(\frac{\sqrt{\pi}}{\sigma} \right) \left(\frac{8\pi^2 k_B T}{h^2} \right)^{3/2} (I_A I_B I_C)^{1/2} \right] \right\} \quad (1.53a)$$

For a linear molecule, the moment of inertia along the bond axis will be zero and the moments of inertia in the two perpendicular directions will be equivalent. Denoting the moment of inertia as I , the rotational entropy for a linear molecule takes the form

$$S(T)_{\text{Rotation}} = R \left\{ 1 + \ln \left[\left(\frac{8\pi^2 I k_B T}{\sigma h^2} \right)^{3/2} \right] \right\} \quad (1.53b)$$

The final two contributions are given in eqns (1.54) and (1.55).

$$S(T)_{\text{Vibration}} = R \sum_i^{3N_{\text{atoms}}-6} \left[\frac{h\nu_i}{k_B T} \frac{1}{e^{h\nu_i/k_B T} - 1} - \ln \left(1 - e^{-h\nu_i/k_B T} \right) \right] \quad (1.54)$$

$$S(T)_{\text{Electronic}} = R \ln(g_0) \quad (1.55)$$

Some care must be exercised in using eqns (1.47) and (1.54). These are obtained within the vibrational harmonic-oscillator approximation. In some situations normal modes can correspond to low-frequency torsional motions, for example the torsional motion shown in Figure 1.12. The harmonic-oscillator approximation is not reliable here and simple application of eqns (1.47) and (1.54) can produce significant errors. Figure 1.13 shows a shallow anharmonic potential energy well (bold line) that deviates significantly from the harmonic form shown (light line). The first five vibrational energy levels are indicated for both potentials and it can be seen that, apart from the zero-point energy levels, the two potentials produce very different vibrational energy spacings. When the thermal energy corrections are of similar magnitude to the well depth, the molecule resides in the upper region of the potential energy well and this is the region that is poorly described by the harmonic-oscillator approximation. The general rule of thumb for the harmonic-oscillator approximation to be accurate is that the energy well depth should be $>10RT$. To put this into perspective, at 298 K, $RT = 2.5 \text{ kJ mol}^{-1}$, while the rotational barrier in ethane is only about 12 kJ mol^{-1} . In such circumstances, one should ideally obtain the potential energy as a function of the normal mode and then solve the Schrödinger equation numerically for the vibrational energy levels of this

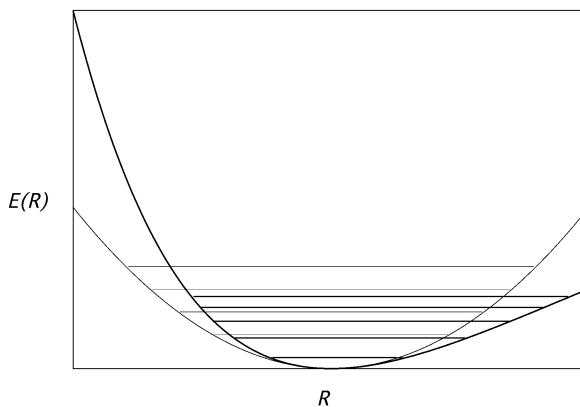


Figure 1.13 Harmonic (light line) and anharmonic (bold line) potential energy curves with vibrational energy levels superimposed.

potential. However, there are many other, simpler, approaches to this problem which provide good pragmatic accuracy.^{3,4}

Knowing how to obtain H and S from our computations, it is an easy matter to evaluate the Gibbs energy, G , using the familiar thermodynamic relationship,

$$G = H - TS \quad (1.56)$$

and hence many other chemically important quantities. For example, the equilibrium constant, K ,

$$K = e^{-\Delta H/RT} e^{\Delta S/R} \quad (1.57)$$

We now have a glimpse of how we can study geometry, vibrations and energetics using computational quantum chemistry. This already enables us to investigate a wealth of interesting chemical problems. For example, you may wish to calculate the geometric and electronic structure of a newly synthesised molecule, or predict the geometry and ground state structure of a yet to be synthesised molecule. Apart from the geometry and the electronic structure there are many other molecular properties that we shall want to be able to evaluate, for example the UV/Visible absorption spectrum, or the proton hyperfine coupling constant in a free radical. We shall discuss the calculation of molecular properties in more detail in chapter 3. Having touched on some of the molecular properties that may be obtained from quantum chemical computations we need to address the very large array of methods available for calculating $E(\mathbf{R})$. Understanding the strengths and weaknesses of commonly used quantum chemical procedures is key to successful scientific studies.

1.4 Standard Models of Electronic Structure

We have mentioned that computational quantum chemistry began in 1927 with the pioneering work of Heitler and London on the hydrogen molecule. In the intervening years a great number of techniques have emerged for dealing with the computation of electronic structure. The range of these techniques is vast and finding a suitable method for a given problem requires some careful consideration. In 1998, the Nobel prize in chemistry was awarded jointly to John Pople ("for his development of computational methods in quantum chemistry") and Walter Kohn ("for his development of the density functional theory"). Pople was one of the most important figures responsible for the widespread adoption of quantum chemical techniques. This he achieved through the development of widely available computer programs and also the development of theoretical models that can be successfully applied to chemical questions. In his Nobel lecture, Pople wrote:

"A theoretical model for any complex process is an approximate but well-defined mathematical procedure of simulation. When applied to chemistry, the task is to use input information about the number and character of component

particles (nuclei and electrons) to derive information and understanding of resultant molecular behaviour.”

Here we shall discuss a number of these models by way of providing an overview. Specifically we shall look at the form of the wavefunction used to define these models and leave much of the detail to chapter 2.

1.4.1 The Hartree-Fock Model and Electron Correlation

At the simplest level we can start by taking a molecular wavefunction to be described by a single Slater determinant of the form given in eqn (1.26). The Slater determinant is composed of a set of molecular orbitals, and each molecular orbital is expanded in terms of a set of atomic basis functions, as shown in eqn (1.28). If we vary the expansion coefficients, $c_{\mu i}$, in eqn (1.28) so as to minimise the total molecular potential energy, we obtain a set of optimal molecular orbitals. This is the Hartree-Fock model and the orbital optimisation process is the SCF method alluded to in section 1.2.3. The resultant optimal molecular orbitals are one-electron wavefunctions, and the model treats the motion of the electrons as being largely independent of each other. Since electrons are negatively charged particles they repel each other. At any instant the repulsion between all the electrons in a molecule is dependent on their positions and a repulsive energy, $\frac{1}{r_{ij}}$, is experienced between all pairs of electrons, i and j . The Hartree-Fock model *cannot* describe this instantaneous repulsion between electrons of opposite spin. There are many ways to illustrate this and here we follow an elegant exposition due to Sinanoğlu.⁵ For example, consider the two electrons of helium, denoted 1 and 2, which in the ground state reside in the 1s orbital. We familiarly write the electronic configuration as $1s^2$. The corresponding Slater determinant is

$$|\Psi(\mathbf{x}_1, \mathbf{x}_2)\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_{1s}(\mathbf{x}_1) & \bar{\phi}_{1s}(\mathbf{x}_1) \\ \phi_{1s}(\mathbf{x}_2) & \bar{\phi}_{1s}(\mathbf{x}_2) \end{vmatrix} \quad (1.58)$$

In this situation the two identically charged electrons have a coulombic repulsion inversely proportional to the distance between them. Consider the orbital, $\phi_{1s}(\mathbf{r})$, as centred at the nucleus and electron 1 located at its Bohr radius. Electron 2 now approaches electron 1 along a straight line and the inter-electronic coulombic repulsion, $\frac{1}{r_{12}}$, encloses electron 1 in a “coulomb hole” into which electron 2 cannot penetrate. The form of the coulomb hole is shown in Figure 1.14. Within the Hartree-Fock model, each electron experiences an averaged potential that arises from the electron density, $\phi_{1s}(\mathbf{r})^2$, of the other electron. This averaged potential gives rise to an inter-electronic repulsion of the form $\int \frac{\phi_{1s}(\mathbf{r}_2)^2}{r_{12}} d\mathbf{r}_2$. The orbitals and the electron

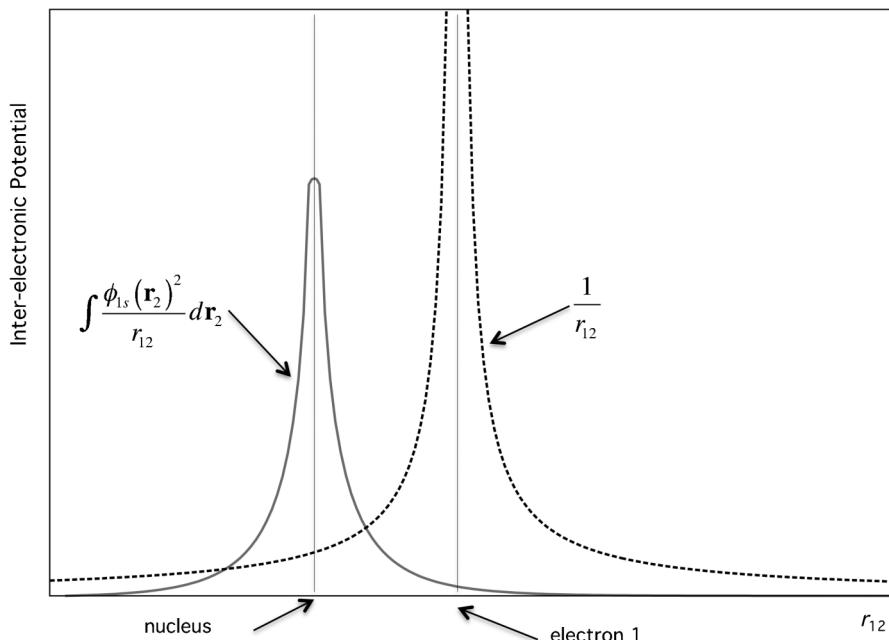


Figure 1.14 Inter-electronic coulombic potential (dashed line) around electron 1 and the averaged Hartree-Fock potential (solid line) experienced by an electron of opposite spin.

distribution are optimised in the SCF process under this averaged potential. In the averaged potential, the motion of electron 2 is not correctly correlated with that of electron 1 and the electrons are allowed to approach each other too closely. When the energy is evaluated, the repulsion energy is high and the overall energy is raised. The difference between the instantaneous and averaged forms of the inter-electronic repulsion is called the “electron correlation energy”.

Expanding the form of the wavefunction in eqn (1.58), ignoring the normalisation factor, we can separate the space and spin components and write

$$\begin{aligned} |\Psi(\mathbf{x}_1, \mathbf{x}_2)\rangle &= \phi_{1s}(\mathbf{x}_1)\bar{\phi}_{1s}(\mathbf{x}_2) - \bar{\phi}_{1s}(\mathbf{x}_1)\phi_{1s}(\mathbf{x}_2) \\ &= \phi_{1s}(\mathbf{r}_1)\phi_{1s}(\mathbf{r}_2)[\alpha(\mathbf{s}_1)\beta(\mathbf{s}_2) - \beta(\mathbf{s}_1)\alpha(\mathbf{s}_2)] \end{aligned} \quad (1.59)$$

Considering the spatial component only, the two-electron density, $\rho(\mathbf{r}_1, \mathbf{r}_2) = ||\Psi(\mathbf{r}_1, \mathbf{r}_2)\rangle|^2$, gives the probability of finding an electron at \mathbf{r}_2 when one is known to be at \mathbf{r}_1 . For the current case of helium

$$\rho(\mathbf{r}_1, \mathbf{r}_2) = ||\Psi(\mathbf{r}_1, \mathbf{r}_2)\rangle|^2 = \phi_{1s}(\mathbf{r}_1)^2 \phi_{1s}(\mathbf{r}_2)^2 \quad (1.60)$$

The two-electron density is given by the product of the density of electron 1

situated at \mathbf{r}_1 , and that of electron 2 situated at \mathbf{r}_2 . There is nothing in this form to prevent the situation of having the two electrons at the same point, $\mathbf{r}_1 = \mathbf{r}_2$. Hence a Hartree-Fock wavefunction, for electrons of opposite spin, leads to an uncorrelated (independent electron) description of the two-electron density, $\rho(\mathbf{r}_1, \mathbf{r}_2)$. The importance of this type of electron correlation cannot be overstated, since its neglect leads to errors in calculations of a number of molecular properties: incorrect prediction of molecular geometries; incorrect electron distribution leading to incorrect electrostatic properties such as dipole moments; activation barriers which are too high, and overestimation of vibrational wavenumbers and bond dissociation energies.

For electrons of parallel spin, the Hartree-Fock model *does* provide a correlated description of the electron distribution. The lowest excited state of helium, in which the electrons have parallel spin, has the configuration $1s2s$. The corresponding Slater determinant is

$$|\Psi(\mathbf{x}_1, \mathbf{x}_2)\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_{1s}(\mathbf{x}_1) & \phi_{2s}(\mathbf{x}_1) \\ \phi_{1s}(\mathbf{x}_2) & \phi_{2s}(\mathbf{x}_2) \end{vmatrix} \quad (1.61)$$

Proceeding as before, ignoring normalisation, we can separate the space and spin components and write

$$\begin{aligned} |\Psi(\mathbf{x}_1, \mathbf{x}_2)\rangle &= \phi_{1s}(\mathbf{x}_1)\phi_{2s}(\mathbf{x}_2) - \phi_{2s}(\mathbf{x}_1)\phi_{1s}(\mathbf{x}_2) \\ &= [\phi_{1s}(\mathbf{r}_1)\phi_{2s}(\mathbf{r}_2) - \phi_{2s}(\mathbf{r}_1)\phi_{1s}(\mathbf{r}_2)]\alpha(\mathbf{s}_1)\alpha(\mathbf{s}_2) \end{aligned} \quad (1.62)$$

$\rho(\mathbf{r}_1, \mathbf{r}_2)$ now takes the form

$$\rho(\mathbf{r}_1, \mathbf{r}_2) = \phi_{1s}(\mathbf{r}_1)^2\phi_{2s}(\mathbf{r}_2)^2 - 2[\phi_{1s}(\mathbf{r}_1)\phi_{2s}(\mathbf{r}_2)\phi_{2s}(\mathbf{r}_1)\phi_{1s}(\mathbf{r}_2)] + \phi_{2s}(\mathbf{r}_1)^2\phi_{1s}(\mathbf{r}_2)^2 \quad (1.63)$$

Consider again the situation in which the two electrons coalesce at the same point, $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$. In this case the first and third terms on the right hand side of eqn (1.63) are cancelled by the second term and $\rho(\mathbf{r}, \mathbf{r}) = 0$. The two electrons cannot occupy the same point in space and so the electronic motion is correlated. This type of correlation, for electrons of the same spin, arises from the antisymmetric nature of the wavefunction and is termed “Fermi correlation”. The $\rho(\mathbf{r}, \mathbf{r}) = 0$ contour in the electron density defines the “Fermi hole”.

For the reasons we have discussed, the Hartree-Fock energy, $E_{\text{H}-\text{F}}$, is always higher than the exact energy obtainable from the Schrödinger equation, E_{Exact} , and this leads us to the most common definition of the correlation energy, E_{corr} ,

$$E_{\text{corr}} = E_{\text{Exact}} - E_{\text{H}-\text{F}} \quad (1.64)$$

Since E_{Exact} is not generally within our reach, how do we proceed beyond the Hartree-Fock model?

We have introduced three principal assumptions: (i) the Born–Oppenheimer approximation (ii) the representation of the wavefunction by a single Slater determinant (iii) the expansion of the orbitals in a set of atomic functions. The orbitals, we have said, are one-electron wavefunctions arising from an approximate Schrödinger equation, namely the Hartree–Fock equation. The flexibility of the one-electron orbitals can be improved by enlarging the set of atomic functions from which they are built (eqn (1.28)). Eventually, any further expansion of the set of atomic functions will produce no useful improvement in the Hartree–Fock energy. At this point we have obtained the *Hartree-Fock limit*. Just as we can improve the one-electron orbitals by enlarging the number of functions used to expand them, we can also improve the many-electron wavefunction by mixing several Slater determinants together. The latter improves the description of the many-electron wavefunction. Hence the quality of our calculations will depend on the flexibility of the one-electron orbitals and also the form of the many-electron wavefunction. Using these two components, most computational models may be viewed in terms of a two-dimensional diagram, shown in Figure 1.15. The horizontal axis refers to the quality of the N -electron wavefunction. At the extreme left we have the case of a single Slater determinant, the Hartree-Fock model, and the accompanying lack of electron correlation. We can move to the right along the horizontal axis by combining more Slater determinants and so form a more accurate N -electron wavefunction. The vertical axis refers to the one-electron basis set used to expand the one-electron orbitals from which the Slater

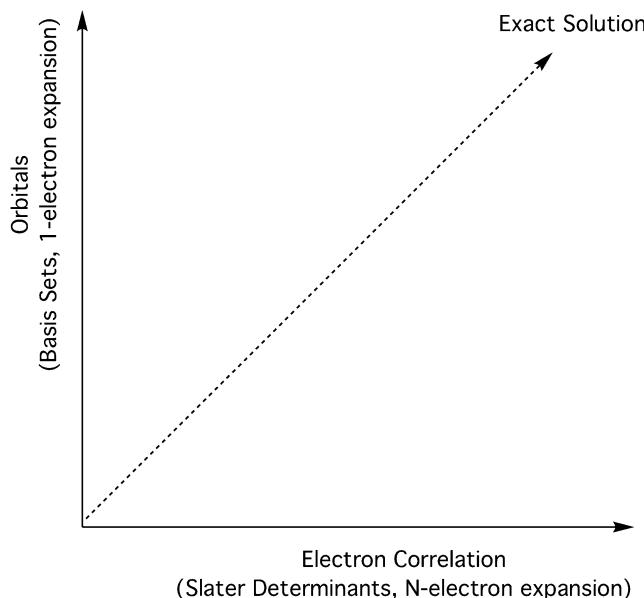


Figure 1.15 Systematic improvement of models (orbital sets and electron correlation) leads to the exact solution of the Schrödinger equation.

determinants are built. If we move along each axis until there is no useful change in the energy then we have effectively solved the electronic Schrödinger equation! Of course we are not able to do this for more than some very small molecules. However we are able to systematically improve the approximation to the Schrödinger equation, even if computational practicalities mean that we must terminate this process sooner than we would wish. In principle, any point on the diagram, specified by a coordinate along each axis, defines a computational “model”. That is, a well-defined procedure by which the energy, and other properties of a molecule may be obtained.

The Hartree–Fock model produces a set of optimal occupied molecular orbitals, from which the Slater determinant describing the ground state is built. In addition a set of unoccupied, or virtual, orbitals is produced. The number of unoccupied orbitals is determined by the size of the one-electron atomic basis set used. If there are m atomic basis functions used in the LCAO expansion of the molecular orbitals, then there will m molecular orbitals produced. For example, a calculation on water might use a basis set consisting of a single 1s type basis function on each hydrogen atom and a 1s, 2s and 2p ($2p_x$, $2p_y$, $2p_z$) set of basis functions on the oxygen atom. This gives a total of seven basis functions, and the Hartree–Fock procedure will mix these to produce seven molecular orbitals. Since water contains 10 electrons, only five of the orbitals will be doubly occupied and two will be unoccupied. Typically, the number of unoccupied orbitals is much larger than that of occupied orbitals if large flexible basis sets are used. The unoccupied orbitals do not contribute in any way to determining the Hartree–Fock energy, or any property within the Hartree–Fock model, but they are used for adding flexibility to the description of the N -electron wavefunction (moving along the horizontal axis in Figure 1.15). This is done by building additional Slater determinants in which the orbitals occupied in the Hartree–Fock ground state are now substituted by members of the set of unoccupied orbitals. For example, consider the lithium hydride molecule, with its four electrons in three molecular orbitals (ϕ_1 , ϕ_2 , ϕ_3) built from a basis set consisting of a 1s basis function on hydrogen and a 1s and 2s basis function on lithium. If we consider all possible spin states for N electrons distributed in m orbitals (including all M_S sub-levels) we can form a total of

$$\binom{2m}{N} = \frac{(2m)!}{(2m-N)!N!} \quad (1.65)$$

determinants. For the case here, $m = 3$ and $N = 4$ which gives 15 determinants. The orbital configurations corresponding to these 15 determinants, $|1\rangle - |15\rangle$, are shown in Figure 1.16. Note that determinants $|1\rangle - |9\rangle$ have the same number of α -spin electrons as β -spin electrons and consequently the total spin, $S = 0$. Determinants $|10\rangle - |12\rangle$, have three α -spin electrons and one β -spin electron, giving $S = 1$ and $M_S = 1$. Conversely, determinants $|13\rangle - |15\rangle$ have one α -spin electron and three β -spin electrons, giving $S = 1$ and $M_S = -1$. In the absence of a magnetic field, the sets of determinants $|10\rangle - |12\rangle$ and

ϕ_3	—	↑	↑	+	+	↑↑	↑↑	↑↑	↑↑
ϕ_2	↑↑	↑	↑↑	↑	↑↑	—	↑↑	↑	↑
ϕ_1	↑↑	↑↑	+	↑↑	↑	↑↑	—	↑	+
$ 1\rangle$	$ 2\rangle$	$ 3\rangle$	$ 4\rangle$	$ 5\rangle$	$ 6\rangle$	$ 7\rangle$	$ 8\rangle$	$ 9\rangle$	
ϕ_3	↑	↑	↑↑	+	+	↑↑			
ϕ_2	↑	↑↑	↑	+	↑↑	↑↑			
ϕ_1	↑↑	↑	+	↑↑	↑	+			
$ 10\rangle$	$ 11\rangle$	$ 12\rangle$	$ 13\rangle$	$ 14\rangle$	$ 15\rangle$				

Figure 1.16 Four electrons distributed in all possible ways in three orbitals $\{\phi_1, \phi_2, \phi_3\}$ give rise to 15 determinants with the orbital occupancies shown.

$|13\rangle - |15\rangle$ will be degenerate and so for the $S = 1$ states, we need only consider three determinants instead of six. We are often interested in a particular spin state only, in the present case the ground state of lithium hydride will have $S = 0$ and so we can ignore the $S = 1$ states for many purposes. If we are only interested in a spin state characterised by a fixed number of α and β electrons, N_α and N_β respectively, we can form

$$\binom{m}{N_\alpha} \binom{m}{N_\beta} = \frac{m!}{(m-N_\alpha)! N_\alpha!} \frac{m!}{(m-N_\beta)! N_\beta!} \quad (1.66)$$

determinants. So for $S = 0$, we have $m = 3$, $N_\alpha = 2$ and $N_\beta = 2$ and eqn (1.66) yields 9 determinants in accord with Figure 1.16. Configuration 1 in Figure 1.16 corresponds to the Hartree-Fock, ground state, Slater determinant. All other determinants can be described in terms of “excitations” or, more correctly, substitutions from this determinant. For example, determinants $|2\rangle - |5\rangle$ correspond to substitution of a single spin-orbital while determinants $|6\rangle - |9\rangle$ are created by two substitutions or a “double excitation”. The exact wavefunction for any state of a system can be written as a linear combination of all possible Slater determinants that can be formed within the basis set used. We shall denote the set of spin-orbitals occupied in the Hartree-Fock determinant by the letters $i, j, k, l \dots$, and the spin-orbitals in the unoccupied set as $a, b, c, d \dots$. The Hartree-Fock determinant will be denoted as $|\Psi_0\rangle$ and any determinant formed from $|\Psi_0\rangle$ by orbital substitutions, $i \rightarrow a, j \rightarrow b$ and so on as $|\Psi_{ij}^{ab}\rangle$. Using this notation the exact wavefunction (within the basis set used) can be written as

$$\begin{aligned} |\Psi\rangle = & C_0 |\Psi_0\rangle + \sum_i^{\text{occupied}} \sum_a^{\text{unoccupied}} C_i^a |\Psi_i^a\rangle + \sum_{i < j}^{\text{occupied}} \sum_{a < b}^{\text{unoccupied}} C_{ij}^{ab} |\Psi_{ij}^{ab}\rangle \\ & + \sum_{i < j < k}^{\text{occupied}} \sum_{a < b < c}^{\text{unoccupied}} C_{ijk}^{abc} |\Psi_{ijk}^{abc}\rangle + \sum_{i < j < k < l}^{\text{occupied}} \sum_{a < b < c < d}^{\text{unoccupied}} C_{ijkl}^{abcd} |\Psi_{ijkl}^{abcd}\rangle + \dots \end{aligned} \quad (1.67)$$

in which the C s are coefficients to be determined by minimising the energy. For an N -electron system there must be, up to and including, N -electron substitutions.

A useful way to write eqn (1.67), which clearly exposes its connection to the Hartree–Fock wavefunction, is in terms of a general substitution operator, \hat{T} . For an N -electron system the operator, \hat{T} , takes the form

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \dots + \hat{T}_N \quad (1.68)$$

For example \hat{T}_1 produces all single substitutions

$$\hat{T}_1 = \sum_i^{\text{occupied}} \sum_a^{\text{unoccupied}} t_i^a \hat{\tau}_i^a \quad (1.69)$$

where $\hat{\tau}_i^a$ is a specific substitution operator and t_i^a is a coefficient (related to the Cs in eqn (1.67)). Similarly \hat{T}_2 will produce all double substitutions

$$\hat{T}_2 = \sum_{i < j}^{\text{occupied}} \sum_{a < b}^{\text{unoccupied}} t_{ij}^{ab} \hat{\tau}_{ij}^{ab} \quad (1.70)$$

and so on, up to \hat{T}_N . To illustrate the function of the $\hat{\tau}$ operators, consider the Hartree–Fock determinant, $|1\rangle$, and the excited determinant, $|2\rangle$, in Figure 1.16. Determinants $|1\rangle$ and $|2\rangle$ are related by a single substitution and using the shorthand notation of eqn (1.27) we can write

$$\begin{aligned} \hat{\tau}_2^3 |\phi_1 \bar{\phi}_1 \phi_2 \bar{\phi}_2\rangle &= |\phi_1 \bar{\phi}_1 \phi_3 \bar{\phi}_2\rangle \\ \hat{\tau}_2^3 |1\rangle &= |2\rangle \end{aligned} \quad (1.71)$$

$|1\rangle$ and $|8\rangle$ are related by a double substitution

$$\begin{aligned} \hat{\tau}_{12}^{33} |\phi_1 \bar{\phi}_1 \phi_2 \bar{\phi}_2\rangle &= |\phi_1 \bar{\phi}_3 \phi_3 \bar{\phi}_2\rangle \\ \hat{\tau}_{12}^{33} |1\rangle &= |8\rangle \end{aligned} \quad (1.72)$$

and so on. This enables us to write the exact wavefunction in eqn (1.67) in a very compact notation, clearly illustrating the relationship between the Hartree–Fock determinant and the exact wavefunction, as

$$|\Psi\rangle = (1 + \hat{T}) |\Psi_0\rangle \quad (1.73)$$

It is important to note that the forms for the exact wavefunction given in eqn (1.67) and (1.73) are equivalent but that the C and t coefficients used in these two equations are subtly different. For example, the coefficient of $|\Psi_0\rangle$ in eqn (1.73) is taken as 1. We shall say more about these details in Chapter 2.

We now give a brief overview of a number of standard models which attempt to traverse the distance between the starting wavefunction, $|\Psi_0\rangle$, and the exact, $|\Psi\rangle$. These models achieve this in procedurally and theoretically distinct ways, but the aim is always to systematically improve an approximate starting wavefunction such that it approaches the exact wavefunction. This enables us to move along the horizontal axis in Figure 1.15 and, in principle, arrive at the exact solution of the electronic Schrödinger equation within a chosen basis set. The methods to be discussed are chosen because they are widely used and well understood. Additionally, each type of model has a substantial literature describing its strengths and weaknesses.

1.4.2 Configuration Interaction Methods

The first family of techniques is known as configuration interaction (CI). These methods are based closely on the ideas we introduced in the preceding section. The limiting case is known as full configuration interaction (FCI) and involves building all possible determinants, their number given by eqn (1.66), and determining the C coefficients in eqn (1.67) by minimising the energy. We can determine the C coefficients in this manner because certain types of method obey the variation principle. Such methods are termed “variational”. This property that tells us the energy obtained by the model is an upper bound to the exact energy. This means that, whatever we do to the C coefficients, the resulting energy will never descend below the exact energy. So the lower we can drive the energy, by varying the C coefficients, the closer to the exact energy we shall come. The variational property has been shown to be of slight importance in terms of ensuring the accuracy of methods, but it can provide significant computational advantages in the computation of molecular properties.

The FCI method is simply stated and provides an exact solution of the electronic Schrödinger equation (within the chosen basis set). In practice it provides many computational challenges such that FCI can only be applied to systems with very few electrons: about 10 – 12 electrons if a good basis set is used. The reason for this is that the number of determinants that can be formed grows factorially with the number of electrons and orbitals. For the very simple case of lithium hydride, all the determinants of the, $S = 0$, ground state are given by $|1\rangle - |9\rangle$ of Figure 1.16. However, this example deals with an

unrealistically small basis set. If we were to use a slightly more realistic set in which we doubled all the $1s$ and $2s$ orbitals and included a set of $2p$ orbitals on each atom we would have a total of 12 basis functions. If $N = 4$ and $m = 12$, eqn (1.66) tells us that the number of determinants in the FCI wavefunction, for the $S = 0$ states, will be 4356. Some of these determinants can be eliminated by consideration of molecular symmetry. However, the dramatic increase in the number of determinants is clear. As a further example, consider the case of benzene (42 electrons) in a minimal basis set (36 molecular orbitals) which will have a FCI expansion of $>3.1 \times 10^{19}$ determinants. Despite the inadequacy of the basis set, this calculation is not currently tractable. The FCI method is still very useful in providing benchmark calculations, on very small molecules, against which all other more approximate methods can be tested.

To develop practical CI methods it is essential to restrict the level of substitutions. Returning to the benzene example above and restricting the level of substitutions to, $\hat{T}_1 + \hat{T}_2$, results in $<144 \times 10^3$ determinants which is a very manageable problem. This is known as the “CI singles and doubles” method, usually denoted CISD. It can be applied to small to medium sized molecules. CI in the space of single substitutions alone (CIS) does not produce any lowering of the energy relative to the Hartree–Fock method. This is a consequence of Brillouin’s theorem, see section 2.6.1, which equates the orbital optimisation of the Hartree–Fock (SCF) method with the space of single substitutions. CIS can be used to study excited states. CI in the space of double substitutions (CID) is the smallest type of CI expansion that produces an improvement over the Hartree–Fock energy. The CI method is general and higher levels of substitution can be included, such as triple or quadruple substitutions, but the number of determinants again increases very rapidly with the size of the molecule. Regardless of the level of truncation, all CI methods obey the variation principle and provide upper bounds to the exact energy.

The main drawback of any form of truncated CI, with the exception of CIS, is that it is not extensive, that is the calculated energy of a system does not scale linearly with the size of the system. In quantum chemistry we often refer to a closely related, but distinct, requirement of size-consistency. The definition originates from the work of John Pople, who defined a method as *size consistent* if the calculated energy of two systems, say two helium atoms, and that of the combined atoms when separated, beyond the range of any interaction, are equal.

Taking CID as an example of a truncated CI method we can illustrate the lack of size consistency by considering the case of two helium atoms described by two s type basis functions. Helium contains only two electrons and so CID is exact (single excitations do not contribute because of Brillouin’s theorem). A CID calculation on an individual helium atom will contain two configurations as shown in Figure 1.17a. Now if the CID method is applied to two helium atoms, infinitely separated, the configurations shown in Figure 1.17(b) will be generated. Note that because the two atoms are infinitely separated there will be no transfer of electrons between atoms. The CID calculations on the

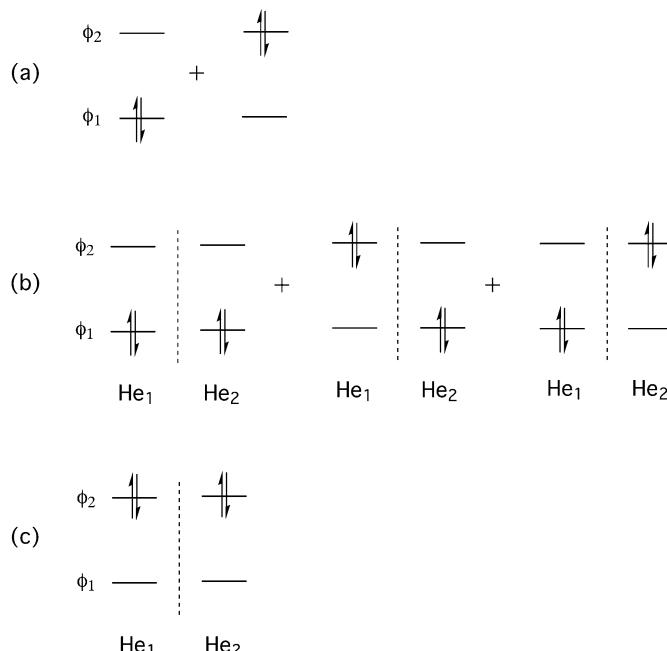


Figure 1.17 (a) CID configurations for helium in a basis set of two orbitals. (b) CID configurations for two helium atoms infinitely separated. (c) Configuration included in FCI on $(\text{helium})_2$ but missing in CID.

individual helium atoms are exact, whereas for the two atoms infinitely separated, the exact wavefunction would have to include up to four-electron excitations. The difference between the FCI wavefunction and CID is the quadruple substitution shown in Figure 1.17(c). This is the size-consistency error for two helium atoms. If more atoms are considered then the difference between the FCI wavefunction and CID will increase. This means that the size-consistency error will increase with the number of electrons.

The size consistency of a computational method is extremely important when calculating relative energies, such as activation barriers or bond dissociation energies. When comparing the energies of systems containing different numbers of electrons, if a method is not size-consistent, the relative errors in the calculated energies will be different and give unreliable energy differences.

1.4.3 Perturbation Theory Methods

We can solve the configuration mixing problem in a different way, which retains the size-consistency property, using Rayleigh–Schrödinger perturbation theory. What defines a particular form of Rayleigh–Schrödinger perturbation theory is the choice of the unperturbed or model problem. In the context of

computational quantum chemistry, the most widely used form of perturbation theory is due to Møller and Plesset. In the Møller–Plesset theory the unperturbed problem is chosen to be the Hartree–Fock model.

We can view the correlation energy as a perturbation which, when added to the model Hartree–Fock problem, should enable us to solve the correlation problem in a systematic fashion. Numerically the perturbation should yield a relatively small correction to the model problem, the latter having been solved exactly. The magnitude of the correlation energy is generally very much smaller than the Hartree–Fock energy and so the correction is relatively small. Additionally it is necessary that the Hartree–Fock wavefunction is a good approximation to the exact wavefunction, that is C_0 in eqn (1.67) should be close to 1. If these conditions are met, then a perturbed hamiltonian can be written as

$$\hat{H}(\lambda) = \hat{H}_0 + \lambda(\hat{H} - \hat{H}_0) = \hat{H}_0 + \lambda\hat{V} \quad (1.74)$$

where \hat{H}_0 is the Hartree–Fock hamiltonian and \hat{V} is the perturbation, which is the difference between the exact hamiltonian \hat{H} and the Hartree–Fock hamiltonian \hat{H}_0 . λ is a dimensionless parameter that ranges between 0 and 1. When $\lambda=0$, we have the Hartree–Fock reference or model problem

$$\hat{H}(\lambda) = \hat{H}_0 \quad (\lambda=0) \quad (1.75)$$

and when $\lambda=1$ we have the exact hamiltonian

$$\hat{H}(\lambda) = \hat{H}_0 + \hat{V} \quad (\lambda=1) \quad (1.76)$$

The exact energy and wavefunction are expanded in powers of the perturbation parameter, λ ,

$$\begin{aligned} E &= E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \lambda^3 E^{(3)} + \dots \\ \Psi &= \Psi^{(0)} + \lambda \Psi^{(1)} + \lambda^2 \Psi^{(2)} + \lambda^3 \Psi^{(3)} + \dots \end{aligned} \quad (1.77)$$

These expansions are then substituted back into the Schrödinger equation

$$\begin{aligned} (\hat{H}_0 + \lambda\hat{V}) \left(\Psi^{(0)} + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \dots \right) &= \\ \left(E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \dots \right) \left(\Psi^{(0)} + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \dots \right) & \end{aligned} \quad (1.78)$$

Terms are collected for a given order in λ yielding a series of equations that can be solved for the energy and wavefunction to any order in λ ,

$$\begin{aligned}
 \hat{H}_0\Psi^{(0)} &= E^{(0)}\Psi^{(0)} \\
 \hat{H}_0\Psi^{(1)} + \hat{V}\Psi^{(0)} &= E^{(0)}\Psi^{(1)} + E^{(1)}\Psi^{(0)} \\
 \hat{H}_0\Psi^{(2)} + \hat{V}\Psi^{(1)} &= E^{(0)}\Psi^{(2)} + E^{(1)}\Psi^{(1)} + E^{(2)}\Psi^{(0)} \\
 &\dots
 \end{aligned} \tag{1.79}$$

We shall discuss some more details of the Møller–Plesset perturbation theory in Chapter 2. The important feature being that all Møller–Plesset methods yield energies which are size-consistent, but lack the variational bound. The theory at second-order, denoted “MP2”, is very widely used. Successive higher orders of the method are available, but are very much more complicated from both a theoretical and a computational perspective. Their formulation is best discussed in terms of diagrammatic techniques, which are outside the scope of this book.

1.4.4 Coupled-Cluster Methods

In formulating the configuration mixing problem, we previously defined an excitation operator, \hat{T} , in eqn (1.68). For application to general systems, this operator is only practicable when truncated at some low order. For example, \hat{T}_2 , which leads to the CID method and its attendant lack of size-consistency. We saw for the case of two separated helium atoms, in Figure 1.17, that the size-consistency error comes about from a missing higher excitation, specifically a quadruple substitution. How might we correct for this missing term, without resorting to the inclusion of \hat{T}_4 ? In the coupled-cluster (CC) approach the exact wavefunction is generated by applying the exponential of the substitution operator to the reference Hartree–Fock wavefunction:

$$|\Psi\rangle = e^{\hat{T}}|\Psi_0\rangle \tag{1.80}$$

The expansion of the exponential function is

$$e^X = \sum_{p=0}^{\infty} \frac{X^p}{p!} = 1 + X + \frac{X^2}{2} + \frac{X^3}{6} + \frac{X^4}{24} + \dots \tag{1.81}$$

Returning to the example of two separated helium atoms, let us apply this expansion to eqn (1.80) with the substitution operator restricted to \hat{T}_2

$$e^{\hat{T}_2}|\Psi_0\rangle = \left(1 + \hat{T}_2 + \frac{1}{2}\hat{T}_2\hat{T}_2\right)|\Psi_0\rangle \tag{1.82}$$

Since there are only four electrons in our example, we can truncate the

expansion at the third term. In analogy with our previous notation, the wavefunction in eqn (1.82) is denoted CCD, coupled cluster in the space of double substitutions. The first two terms on the right hand side of eqn (1.82) produce substitutions equivalent to the CID wavefunction. The third term introduces a four-electron substitution as a product of two-electron substitutions. From our discussion of Figure 1.17, this last term will remove the size-consistency error. In fact all coupled-cluster methods, regardless of the level of truncation of \hat{T} , produce size-consistent energies.

If the complete \hat{T} operator is included, then eqn (1.80) yields the exact wavefunction and is equivalent to FCI. Any truncation of \hat{T} , while including the effects of higher substitutions, will not be equivalent to FCI. This is because in the FCI method the contribution of the quadruple substitution has its own variationally determined coefficient, see eqn (1.67). Conversely, in eqn (1.82) the form of the \hat{T}_2 operator, eqn (1.70), means that the coefficient of the quadruple substitution is given as the product of the coefficients of two double substitutions. The double substitutions, which occur linearly, are termed “connected” clusters. Higher substitutions which appear as products of lower substitution operators, $\frac{1}{2}\hat{T}_2\hat{T}_2$, are termed “disconnected” clusters.

The coefficients t_{ij}^{ab} , eqn (1.70), are obtained by projection and the resultant energy, while size-consistent, is not variationally bound. Approximating the effects of higher substitutions as products of lower substitutions is often very reliable and CC techniques provide some of the most accurate results available for general chemical systems. The most widely used level is CCSD. Including triple substitutions, CCSDT, incurs a significant rise in computational cost and consequently a model termed CCSD(T) has found very widespread application. In CCSD(T) the effects of triple substitutions are included in an approximate fashion using ideas from the Møller–Plesset perturbation theory.

We have said something about the influence of computational cost on the choice of a method. There must always be a balance between cost and required accuracy. For the types of method we have discussed so far, is it possible to establish a hierarchy of accuracy? The vast literature associated with these methods suggests that the following is a plausible hierarchy:

$$\text{HF} \ll \text{MP2} < \text{CISD} < \text{CCSD} < \text{CCSD(T)} < \text{FCI}$$

The extremes of ‘best’, FCI, and ‘worst’, HF, are irrefutable, but the intermediate methods are less clear and depend on the type of chemical problem being addressed. What about computational cost? This is not an easy thing to quantify because there are many details of implementation and approximation to be taken into account before meaningful comparisons can be made. For example, there have been significant advances in developing “linear-scaling techniques” that can be applied to large molecular systems. Linear-scaling techniques usually involve quite elaborate approximations, these may be valid for large systems but less appropriate for small to medium sized systems. However, we can look at the

Table 1.2 Formal scaling of computational cost with the size of the orbital space, m . $\text{iter}\cdot m^x$ indicates there are many iterative steps, each of computational cost m^x .

Method:	HF	MP2	CISD	CCSD	CCSD(T)	FCI
Formal Scaling	$\text{iter}\cdot m^4$	m^5	$\text{iter}\cdot m^6$	$\text{iter}\cdot m^6$	$\text{iter}\cdot m^6 + m^7$	$m!$
Variational?	Yes	No	Yes	No	No	Yes
Size-consistent?	Yes	Yes	No	Yes	Yes	Yes

formal cost of the methods we have discussed, assuming no approximations, and this does provide a useful comparison, see Table 1.2.

1.4.5 Multiconfigurational and Multireference Methods

The families of correlated methods we have described so far all depend on the Hartree-Fock determinant dominating the final wavefunction, that is $C_0 \approx 1$ in eqn (1.67), and all the many other determinants providing a relatively small, but essential, component of the total wavefunction. In certain systems the reference wavefunction cannot be represented adequately by a single Slater determinant. Such situations arise, for example, when chemical bonds are stretched or when spin-paired electrons occupy energetically degenerate but spatially distinct orbitals. Consider the familiar molecular orbital picture of an electron pair bond between two identical atoms, in which atomic orbitals on the two atoms mix to form bonding and antibonding molecular orbitals, as shown in Figure 1.18a. If the bond between atom A and atom B is stretched, the energy separation between the bonding and antibonding molecular orbitals is reduced, as in Figure 1.18b. If the atoms are moved to very large separation, as might be required when calculating a diatomic potential energy curve, the bonding and antibonding combinations will become degenerate. In the first case, Figure 1.18a, it is obvious that the two electrons should be placed in the lower lying bonding orbital. In the final case, Figure 1.18c, there is no reason to place the electrons in the bonding orbital, since now it is degenerate with the antibonding orbital. In such situations we must include two determinants: one with the bonding orbital doubly occupied and the other with the antibonding orbital doubly occupied. This is the extreme case of exact degeneracy and we would expect each Slater determinant to contribute equally to the wavefunction. This is easily illustrated with the simple case of the hydrogen molecule described with a minimal basis set consisting of a single $1s$ type function on each atom. The problem arises in the Hartree-Fock method because, even at dissociation, both electrons are assigned to a single symmetric doubly occupied molecular orbital of the form

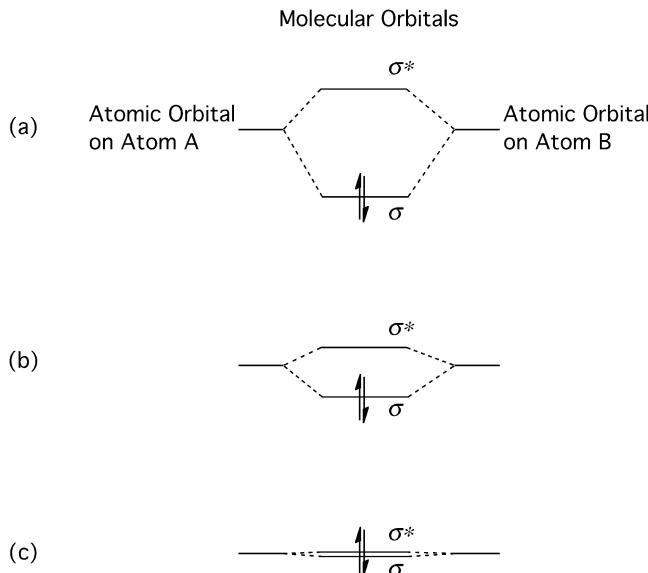


Figure 1.18 (a) Homonuclear diatomic bonding and antibonding orbital energy levels at equilibrium bond distance. (b) Bonding and antibonding energy levels at stretched bond distance. (c) Bonding and antibonding energy levels for completely separated atoms.

$$\sigma = \frac{1}{\sqrt{2}} [\chi_A + \chi_B] \quad H_A \cdots \cdots \cdots H_B \quad (1.83)$$

where χ_A is the $1s$ type function on atom A , and χ_B that on atom B . The corresponding Slater determinant is

$$|\Psi_0\rangle = |\sigma\bar{\sigma}\rangle = \frac{1}{\sqrt{2}} \sigma(\mathbf{r}_1)\sigma(\mathbf{r}_2)[\alpha(\mathbf{s}_1)\beta(\mathbf{s}_2) - \beta(\mathbf{s}_1)\alpha(\mathbf{s}_2)] \quad (1.84)$$

Regardless of the internuclear distance, the form of the bonding orbital remains that given in eqn (1.83). The spatial part of the wavefunction can be expanded to give

$$\sigma(\mathbf{r}_1)\sigma(\mathbf{r}_2) = \frac{1}{2} [\chi_A(\mathbf{r}_1)\chi_A(\mathbf{r}_2) + \chi_A(\mathbf{r}_1)\chi_B(\mathbf{r}_2) + \chi_B(\mathbf{r}_1)\chi_A(\mathbf{r}_2) + \chi_B(\mathbf{r}_1)\chi_B(\mathbf{r}_2)] \quad (1.85)$$

According to this the two electrons spend half of their time on different atoms, one on atom A and one on atom B , and the other half of their time on one atom, either atom A or atom B . In terms of structures, even at dissociation, the wavefunction contains 50% $H_A \bullet \bullet H_B$, 25% $H_A^- + H_B$ and 25% $H_A^+ - H_B$. This is

clearly incorrect, since the molecule should dissociate into $2\text{H}\bullet$. The correct form required for the wavefunction at dissociation is

$$\frac{1}{2}[\chi_A(\mathbf{r}_1)\chi_B(\mathbf{r}_2) + \chi_B(\mathbf{r}_1)\chi_A(\mathbf{r}_2)][\alpha(\mathbf{s}_1)\beta(\mathbf{s}_2) - \beta(\mathbf{s}_1)\alpha(\mathbf{s}_2)] \quad (1.86)$$

This form of the wavefunction cannot be expressed as a single determinant. However if we use the antibonding orbital

$$\sigma^* = \frac{1}{\sqrt{2}}[\chi_A - \chi_B] \quad (1.87)$$

and expand the spatial component of the corresponding determinant

$$\sigma^*(\mathbf{r}_1)\sigma^*(\mathbf{r}_2) = \frac{1}{2}[\chi_A(\mathbf{r}_1)\chi_A(\mathbf{r}_2) - \chi_A(\mathbf{r}_1)\chi_B(\mathbf{r}_2) - \chi_B(\mathbf{r}_1)\chi_A(\mathbf{r}_2) + \chi_B(\mathbf{r}_1)\chi_B(\mathbf{r}_2)] \quad (1.88)$$

We can see that the $\text{H}_A\bullet\bullet\text{H}_B$ type terms occur with the opposite sign to that in eqn (1.85). If we subtract the determinants we obtain the required form for dissociated hydrogen

$$|\sigma\bar{\sigma}\rangle - |\sigma^*\bar{\sigma}^*\rangle = \frac{1}{2}[\chi_A(\mathbf{r}_1)\chi_B(\mathbf{r}_2) + \chi_B(\mathbf{r}_1)\chi_A(\mathbf{r}_2)][\alpha(\mathbf{s}_1)\beta(\mathbf{s}_2) - \beta(\mathbf{s}_1)\alpha(\mathbf{s}_2)] \quad (1.89)$$

In the intermediate situation of stretched bond lengths, we will not have exact degeneracy, but can speak of near-degeneracy of the bonding and antibonding orbitals. The contribution of the doubly substituted determinant, eqn (1.88), will not be 50% of the wavefunction but it will be significant. We can summarise the three situations depicted in Figure 1.18 for the case of the hydrogen molecule by writing a CI type wavefunction

$$|\Psi_{\text{CI}}\rangle = C_1|\sigma\bar{\sigma}\rangle + C_2|\sigma^*\bar{\sigma}^*\rangle \quad (1.90)$$

At equilibrium, Figure 1.18a, $C_1 > C_2$ ($C_1 \approx 1, C_2 \approx 0$). At stretched bond distances, Figure 1.18b, $C_1 > C_2$ and at dissociation, Figure 1.18c, $C_1 = C_2$. In the latter two circumstances the qualitatively correct wavefunction is not given by eqn (1.84), but rather by eqn (1.90). It consists of more than one Slater determinant and is said to be multiconfigurational in character. Figure 1.19 shows the potential energy curve for molecular hydrogen calculated using eqns (1.84) and (1.90). The difference between the two curves in Figure 1.19 is a type of electron correlation energy, it is highly structure dependent and is termed non-dynamic electron correlation. The non-dynamic electron correlation comes from the inability of a single determinant to provide a qualitatively correct reference wavefunction. In Figure 1.19, if we were to adopt the definition of eqn (1.64) for the correlation energy, we can see that as the bond distance increases the correlation energy also increases, reaching a maximum

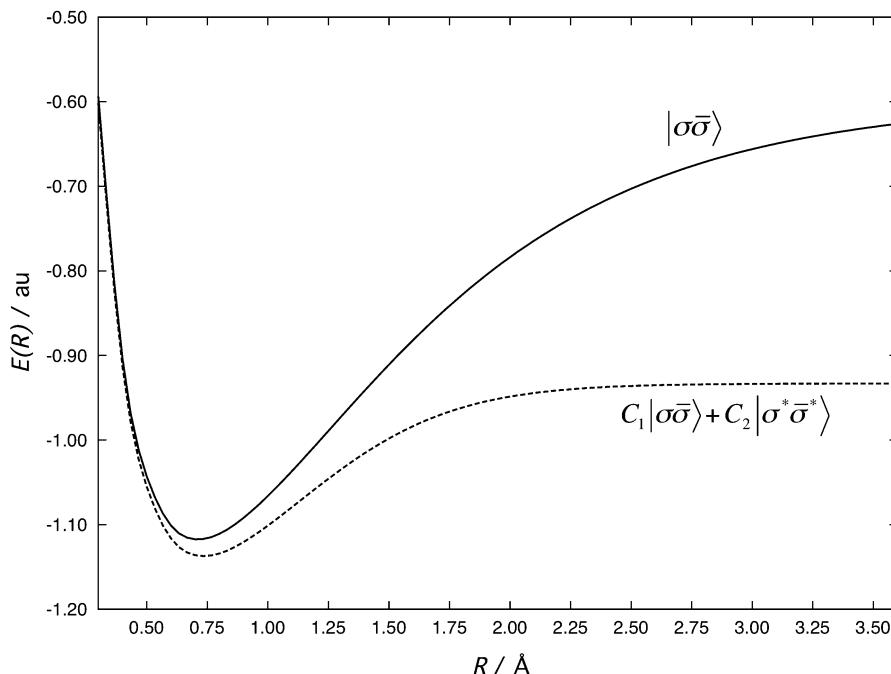


Figure 1.19 Potential energy curve for H_2 using the wavefunction forms given in eqns (1.84) (solid line) and (1.90) (dashed line).

value at dissociation! This is clearly nonsensical, since at dissociation we have two separated hydrogen atoms and there must be no electron correlation. In this case the correct reference wavefunction is given by eqn (1.90), since at dissociation eqn (1.84) describes a physically incorrect wavefunction.

The orbitals where near-degeneracy effects operate may be a very small subset of the full orbital space. The majority of the orbitals will be adequately described as doubly occupied or empty. To define an appropriate reference wavefunction we must optimise the orbitals, as in the Hartree-Fock method, but also allow a mixing of determinants, as in eqn (1.90), so that optimal mixing coefficients may be determined through minimisation of the energy. This is the realm of the multiconfigurational SCF (MCSCF) method. To proceed we partition the orbital space into three subspaces. The first subspace, termed “inactive”, consists of orbitals that remain doubly occupied in all determinants. The second subspace, termed “active”, consists of orbitals that are allowed to have variable occupancy through the mixing of determinants. The third subspace is the “virtual” space and contains orbitals that are unoccupied in all determinants. Returning to the lithium hydride example, with its four electrons in three molecular orbitals (ϕ_1, ϕ_2, ϕ_3) built from a basis set consisting of a $1s$ function on hydrogen and a $1s$ and $2s$ function on lithium, we could choose the inactive subspace to consist of the ϕ_1 orbital (essentially the $1s$ orbital on lithium) and the active subspace to be the ϕ_2 and ϕ_3 bonding

and antibonding orbitals. In this limited basis there will be no virtual orbital subspace. The active space consists of two electrons and two orbitals, if we allow all possible distributions of the active electrons in the active orbitals for an overall $S = 0$ then we shall have a complete active space (CAS). The CAS will comprise the four determinants shown in Figure 1.20.

Keeping ϕ_1 as doubly occupied in all configurations has reduced the number of determinants from nine (Figure 1.16) to four (Figure 1.20). However if we were to consider the same approach for, say the nitrogen molecule, the number of active orbitals would have to include, at least, each bonding and antibonding orbital of the triple bond. The active space would now include six electrons and six orbitals and the CAS would involve 400 determinants. This small extension of the active space results in a dramatic increase in the number of determinants involved.

Multi-configurational methods are conceptually and computationally much more difficult than the single determinant Hartree-Fock method. For diatomic molecules the choice of active space is simple. However for general molecular systems, especially in chemical reactions, great care must be exercised in the choice of active space to keep it small enough to ensure that the calculation remains computationally tractable but flexible enough to adequately describe the chemical situation being explored. Provided that a suitable active space can be chosen, the orbitals are optimised in a similar fashion to the Hartree-Fock method. This defines a suitable reference wavefunction that can then be subjected to further elaboration using CI, Møller-Plesset or CC methods to account for the dynamic correlation energy. The techniques we have discussed for correlated calculations can all be extended to the case of multiconfigurational reference wavefunctions. Such multireference techniques (MR-CI, MR-MP2, MR-CC) are very much more computationally demanding but their use is sometimes essential in order to arrive at meaningful results.

Multi-configurational and multi-reference methods are not as widely used as their single-determinant-based counterparts. The calculations do require greater computational resources but perhaps more importantly they require a high degree of chemical understanding on the part of the user if they are to produce useful studies. The choice of active space provides a great degree of flexibility in the application of multi-reference techniques. This means that they are, in principle,

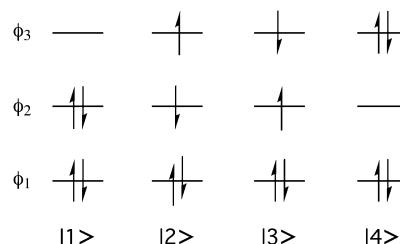


Figure 1.20 Complete active space (CAS) determinants for two electrons in two orbitals and $S = 0$.

applicable to any type of chemical structure but in practice the pragmatic choice of active space provides the opportunity for erroneous modeling. These are not automatic ‘black-box’ techniques and should be used with due care and caution.

1.4.6 Density Functional Methods

The final type of standard model we shall discuss is that based on density functional theory (DFT). These are currently the most widely used techniques for the study of general chemical problems. The ultimate aim is to deal directly with the one-electron density and circumvent consideration of the many-electron wavefunction. An N -electron wavefunction, $|\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)\rangle$, contains all information on the system and depends on $4N$ coordinates, that is three spatial and one spin coordinate per electron. A remarkable theorem due to Hohenberg and Kohn establishes that the ground state energy and all ground state properties can be obtained from functionals of the one-electron density alone. Formally, the one-electron density, $\rho(\mathbf{r})$, can be obtained from the N -electron wavefunction, $|\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)\rangle$, by integrating over all electronic spin coordinates and over all spatial coordinates, except those of electron 1.

$$\rho(\mathbf{r}_1) = N \int |\psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 d\mathbf{s}_1 d\mathbf{x}_2 \dots d\mathbf{x}_N \quad (1.91)$$

Consequently $\rho(\mathbf{r}_1)$ depends only on the three spatial coordinates of electron 1, regardless of the number of electrons in the system. If we can work with $\rho(\mathbf{r}_1)$ as a fundamental quantity, instead of $|\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)\rangle$, it should lead to less demanding computations. Additionally it provides us with a simple quantity defined in three-dimensional physical space to deal with, rather than the many-dimensional space of the N -electron wavefunction.

Unfortunately the Hohenberg-Kohn theorem does not tell us how we might find suitable functionals of $\rho(\mathbf{r}_1)$. To proceed we can make use of the known exact solution to an idealised electronic problem, that of the uniform electron gas. The “uniform” refers to the distribution of the electron density throughout the gas. The Coulomb and Fermi correlation properties of the uniform electron gas can be extracted in the form of exchange (Fermi) – correlation (Coulomb) functionals of the one-electron density. These exchange-correlation functionals can be added to a self-consistent field type formalism providing a computationally efficient technique. This approach, with suitable modifications, is capable of wide application and good accuracy for chemical problems. Modifications and refinements are necessary since the uniform electron gas is generally not a good approximation to the electron distribution in chemical systems. For example, the electron distribution around two bonded atoms of very different electronegativity will produce a very non-uniform distribution of the electron density. Many formal results in DFT are known and formal developments in the theory continue. If the exact form of the exchange and correlation terms was known for

a general electronic system, rather than the uniform electron gas, then DFT would provide an exact solution to the Schrödinger equation.

Practical forms of DFT have developed with appeal to empirical input. This can be a contentious issue, but it is clear that many of the ‘parameters’ that enter practical DFT methods have been obtained by comparison to some form of external data, for example atomic correlation energies or thermochemical databases. DFT methods have received huge efforts in their development and implementation and provide a cost-effective approach to many chemical questions. Comparing with the methods in Table 1.2, pure DFT methods have a formal scaling, with basis set size, of m^3 . At present there is no systematic way to refine the density functional approach to the exact limit in analogy with Figure 1.15. A similar construct to Figure 1.15 has been suggested for DFT methods,⁶ but it refers to an elaboration of the formal techniques used and provides no guarantee of approaching the limiting case of exactly solving the Schrödinger equation.

1.5 How Do You Select a Theoretical Model?

From the preceding section it should be clear that the choice of theoretical model on which to base the study of a chemical question is vast. A key consideration in narrowing down the method to be used is the balance of computational cost against accuracy. Computational costs include: the amount of processor time used; the quantity of disk space needed; the amount of core (random access) memory necessary to efficiently complete the task. It is invariably the case that there must be a compromise between cost, accuracy and ease of carrying out a chemical study. A study will usually involve many calculations of different types and it is not enough to be able to perform a single demanding calculation and stop. Certain methods are only feasible for small molecules and may provide very accurate results. For larger molecules, less accurate methods must be employed. Even if large computational resources are available, the skill and experience of the user in applying more complex techniques must be a consideration. As in any computational endeavour the maxim “garbage in, garbage out”, holds unassailable sway.

We have already mentioned that it is a difficult task to quantify the precise cost of a computational method. There are many considerations of computational implementation that can influence the cost and scaling of a given technique. For example, we have already stated in Table 1.2 that the Hartree-Fock technique formally scales with the basis set size as m^4 . This scaling comes from the $\sim m^4/8$ unique two-electron integrals that must be evaluated in a Hartree-Fock calculation. Various integral approximations may be introduced to reduce this dependence to between m^2 and m^3 . These approximations are typically integral screening techniques, which circumvent the evaluation of batches of integrals, thereby reducing the computational effort. For very large molecules other types of distant-dependent approximations may be employed to further reduce the scaling to m .

Many of the standard methods we have discussed in the previous section have well-known strengths and weaknesses. This knowledge has been gained through wide application of the techniques over many years to a wide range of chemical problems. For example, the Hartree-Fock method is arguably the simplest approach. Its weakness stems from the lack of correlation between the motion of electrons of opposite spin. Despite this severe approximation, the Hartree-Fock method is both variational and size-consistent and able to provide reasonable estimates of equilibrium bond lengths and conformational energy differences. Its ability to provide accurate estimates of reaction energies and activation barriers is rather poor.

To assess the accuracy of a method for a given task we can look at the performance of the method in the calculation of reaction energies; equilibrium geometries; harmonic vibrational wavenumbers; dipole moments and any other molecular properties of relevance. Many molecular databases exist which provide this type of information for a variety of methods. If the methods of interest have not been assessed for a given property then it is important that a validation is carried out by the user on a selection of molecules for which reliable data is available for comparison. For some chemical questions, for example the geometry of a transition structure for a reaction, experimental data will not be available for comparison. In such cases comparison may be made with higher-level computational techniques. Table 1.3 lists a number of databases that can be used for validation against experimental data or high-level computational results.

1.6 The Apparatus: Hardware and Software

Since the 1980s, computer hardware has developed at an extraordinary pace. In real terms this has meant that fast computing platforms have become available, at a decreasing cost, to a wide range of users. There have been a number of key developments in the structure of central processing units (CPUs) that have enabled this remarkable progress. It would take us too far from our purpose here to attempt any detailed discussion of these. However, it is to be hoped that some of these remarkable historical developments in computer design are of sufficient interest to merit the briefest mention here.

There has been an impressive increase in the operational clock speed at which modern CPUs work, amounting to some three orders of magnitude. There have also been four key changes in the way instructions are processed.

- (a) A CPU design, termed “RISC” (Reduced Instruction Set Computing) exploited the idea that it was possible to carry out a small set of simple instructions in a highly optimised manner and so achieve much faster effective speeds than when using more complex instructions. For example, a complex instruction might involve fetching from memory, executing an arithmetic operation and reloading memory with the result. In the RISC CPU this complex instruction is decomposed into constituent lower level instructions, each being performed very efficiently.

Table 1.3 Selected databases for the assessment of computational methods.**1 Atomic Reference Data for Electronic Structure Calculations**

Location: <http://www.nist.gov/physlab/data/dftdata/index.cfm>
 References: S. Kotochigova, Z. H. Levine, E. L. Shirley, M. D. Stiles and C. W. Clark, *Phys. Rev. A*, 1997, **55**, 191 and S. Kotochigova, Z. H. Levine, E. L. Shirley, M. D. Stiles and C. W. Clark, *Phys. Rev. A*, 1997, **56**, 5191.

2 Benchmark Energy and Geometry Database

Location: <http://www.begdb.com>
 Reference: J. Rezáč, P. Jurečka, K. E. Riley, J. Černý, H. Valdés, K. Pluháčková, K. Berka, T. Rezáč, M. Pitonák, J. Vondrášek and P. Hobza, *Collect. Czech. Chem. Commun.*, 2008, **73**, 1261.

3 Barrier Heights of Diverse Reaction Types

References: J. Zheng, Y. Zhao and D. G. Truhlar, *J. Chem. Theory Comput.*, 2007, **3**, 569 and J. Zheng, Y. Zhao and D. G. Truhlar, *J. Chem. Theory Comput.*, 2009, **5**, 808.

4 NIST: Computational Chemistry Comparison and Benchmark Database

Location: <http://cccbdb.nist.gov>

5 The Computational Results Database

Location: <http://tyr3.chem.wsu.edu/~feller/Site/Database.html>
 Reference: D. Feller, *J. Comput. Chem.*, 1996, **17**, 1571.

- (b) The RISC structure lends itself to pipelining, in which a CPU module is dedicated to performing each of the basic tasks of fetching from memory, decoding the instruction, executing the instruction and writing back the result. The advantage of pipelining is that, as the first instruction exits the first operational module, a second instruction can begin processing. As in a pipeline, several instructions can begin processing before the outcome of the first instruction is completed, see Figure 1.21. Pipelining may not be faster for the execution of a single instruction but will be faster when several instructions must be performed.
- (c) Multiple-issue techniques are complex procedures which effectively allow more than one instruction to be performed per clock cycle, for example by fetching a word in which several instructions are packed.
- (d) More efficiency can be gained by enabling several instructions to begin execution simultaneously. Multi-core processors are now commonplace in everyday desktop machines. Multi-core processors are single components (chips) with two (dual-core) or more (e.g. quad-core) independent processors, which read and execute instructions. Since the cores are independent of each other, parallel execution of instructions may be performed.

More recently, graphical processing units (GPUs) have been the focus of much interest in the computational chemistry community, since they allow high levels of parallelism at relatively low cost. We now also have access to chips with very large numbers of core processing units. The question of how these new hardware structures are to be exploited efficiently for computational quantum chemistry has led to new avenues of research and endeavour.

Clock Cycle:	1	2	3	4	5	6	7	8	9	10	11	12
Non-Pipelined CPU												
Instruction 1	F	D	X	W								
Instruction 2					F	D	X	W				
Instruction 3									F	D	X	W
Pipelined CPU												
Instruction 1	F	D	X	W								
Instruction 2	F	D	X	W								
Instruction 3		F	D	X	W							

F=Fetch; D=Decode; X=Execute; W=Write.

Figure 1.21 The execution of three consecutive instructions on a non-pipelining CPU compared to a pipelining CPU.

Developments in computer hardware continue unabated. The speed of these developments means that almost anything one writes about hardware will be out of date by the time it is read! For the interested reader, a recent review of high performance computing in chemistry may be found in ref. 7.

The cost of computer hardware has decreased and its efficiency has increased. Consequently, computational quantum chemistry can be carried out on a wide range of systems, from simple desktop workstations to massive machines with hundreds of thousands of CPUs. In monetary currency, this range spans some eight orders of magnitude as indicated in Table 1.4. The top supercomputers in the world are constantly changing and what was impressive in the recent past is quickly overtaken. At the time of writing, the top supercomputer in the world is deemed to be the Sequoia, which is installed at the U.S. Department of Energy's Lawrence Livermore National Laboratory in California. The Sequoia contains more than 1.5×10^6 processing cores!

The other essential component of computational quantum chemistry is the software: the sets of instructions, to be carried out by a computer to produce energies, geometries, molecular orbitals and anything else of interest. Computational quantum chemistry software is complicated because it contains many, many, different components. For example, any program must start with a procedure to read in the specification of a molecule in terms of coordinates and atom types, which is a relatively simple task. Having read in a molecule, the program must assign basis functions to the atoms, and then begin efficiently calculating the many necessary integrals involving the basis functions, which is a relatively complicated task.

Efficient quantum chemistry programs, see Table 1.5 for a partial list, have often undergone many tens, and in some cases hundreds, of man-years of development. It is possible to put together a simple program to perform quantum chemical calculations, but it should be borne in mind that a working program is not necessarily an efficient program. To perform some of the elaborate studies currently being carried out by researchers does require validated programs of extreme efficiency. The list in Table 1.5 is by no means exhaustive, but illustrates that there are many programs available offering a

Table 1.4 Relative cost of some typical computing platforms used for computational quantum chemistry.

<i>Platform</i>	<i>Relative cost</i>
Multi-core desktop PC	1
Multi-CPU, multi-core workstation	10
Workstation cluster	100
Supercomputer	$10^6\text{--}10^8$

wide range of functionality. The specialist abilities of these programs varies considerably and the interested reader should explore which programs might be most suitable for the work they wish to carry out. This should include considerations of the methods available, interfaces to other software (for example for graphical display of output), availability of source code to enable addition of functionality and the nature and limitations of the software licence.

Table 1.5 An incomplete list of widely used computational quantum chemistry programs, listed alphabetically.

1 ACES	Website: http://www.qtp.ufl.edu/ACES
2 ADF	Website: http://www.scm.com/Products/ADF.html
3 Columbus	Website: http://www.univie.ac.at/columbus
4 Dalton	Website: http://dirac.chem.sdu.dk/daltonprogram.org
5 DeMon	Website: http://www.demon-software.com/public_html/program.html
6 GAMESS	Website: http://www.msg.ameslab.gov/gamess/index.html
7 GAMESS-UK	Website: http://www.cfs.dl.ac.uk
8 Gaussian	Website: http://www.gaussian.com
9 Jaguar	Website: http://www.schrodinger.com/products/14/
10 MOLCAS	Website: http://www.molcas.org/introduction.html
11 Molpro	Website: http://www.molpro.net
12 NWChem	Website: http://www.nwchem-sw.org/index.php/Main_Page
13 ORCA	Website: http://www.mpibac.mpg.de/bac/logins/neese/description.php
14 Q-Chem	Website: http://www.q-chem.com/features.htm
15 TURBOMOLE	Website: http://www.cosmologic.de/index.php?cosId=3010&crId=3

Appendix 1A Physical Constants, Atomic Units and Conversion Factors

Quantity	Symbol or Formula	Atomic Units	SI Units
Planck's constant	h	2π	6.626069×10^{-34} J s
Planck's constant/ 2π	\hbar	1	1.054571×10^{-34} J s
Charge	e	1	1.602176×10^{-19} C
	$4\pi\epsilon_0$	1	1.112650×10^{-10} J $^{-1}$ C 2 m $^{-1}$
Unified atomic mass unit	u	1822.888	1.660538×10^{-27} kg
Electron rest mass	m_e	1	9.109382×10^{-31} kg
Proton mass	m_p	1836.152	1.672621×10^{-27} kg
Neutron mass	m_n	1838.683	1.674927×10^{-27} kg
Fine structure constant	α		7.297352×10^{-3}
Speed of light in vacuum	c $\left(= \frac{1}{\alpha} \right)$	137.035999	299792458 m s $^{-1}$
Energy	$E_h = \frac{e^2}{4\pi\epsilon_0 a_0}$	1	4.359744×10^{-18} J (27.21138 eV)
Time	$\frac{\hbar}{E_h}$	1	2.418884×10^{-17} s
Length	a_0	1	$0.5291772 \times 10^{-10}$ m
Charge density	$\frac{e}{a_0^3}$	1	1.081202×10^{12} C m $^{-3}$
Electric Field	$\frac{E_h}{ea_0}$	1	5.142206×10^{11} V m $^{-1}$
Dipole Moment	ea_0	1	8.478353×10^{-30} C m
Quadrupole moment	ea_0^2	1	4.486551×10^{-40} C m 2
Electric polarisability	$\frac{e^2 a_0^2}{E_h}$	1	1.648777×10^{-41} C 2 m 2 J $^{-1}$
Magnetic dipole moment	$\frac{he}{m_e}$	1	1.854801×10^{-23} J T $^{-1}$
Bohr Magneton	μ_B	1	9.274009×10^{-24} J T $^{-1}$
Nuclear magneton	μ_N	2.723085×10^{-4}	5.050783×10^{-27} J T $^{-1}$

It is often convenient to express energy in terms of energy equivalents according to the equations: $E = hv = \frac{hc}{\lambda} = mc^2 = kT$.

$$\begin{aligned}
 1 \text{ au of energy} &\equiv \frac{E}{h} = 6.579683 \times 10^{15} \text{ Hz} \\
 &\equiv \frac{E}{hc} = 2.194746 \times 10^7 \text{ m}^{-1} \\
 &\equiv \frac{E}{c^2} = 4.850869 \times 10^{-35} \text{ kg} = 2.921262 \times 10^{-8} \text{ u} \\
 &\equiv \frac{E}{k_B} = 3.157750 \times 10^5 \text{ K}
 \end{aligned}$$

A complete listing of currently accepted values of physical constants is available in P. J. Mohr, B. N. Taylor, and D. B. Newell, *Rev. Mod. Phys.*, 2012, **84**, 1527.

Appendix 1B Elementary Properties of Determinants

Consider the square 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (1B.1)$$

The determinant of \mathbf{A} , $\det \mathbf{A}$, which we shall denote as \mathbf{D} is defined as

$$\mathbf{D} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \quad (1B.2)$$

For determinants of larger dimension we can compute the value of the determinant by a Laplace expansion along any row or column. The idea of the Laplace expansion is to reduce the larger determinant into a combination of 2×2 determinants that can be evaluated using eqn (1B.2). To illustrate this consider the 3×3 determinant

$$\mathbf{D} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \quad (1B.3)$$

We shall expand this determinant along the top row. To begin we take the first element in row 1 and column 1, which is a . Now strike out the elements in the row and column containing a

$$\mathbf{D} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \quad (1B.4)$$

This leaves the 2×2 determinant

$$\begin{vmatrix} e & f \\ h & i \end{vmatrix} \quad (1B.5)$$

The first term in the expansion is

$$a \begin{vmatrix} e & f \\ h & i \end{vmatrix} \quad (1B.6)$$

We now proceed to the second element in row 1 column 2, b , and strike out the row and column containing b

$$\begin{vmatrix} a & b & e \\ d & e & f \\ g & h & i \end{vmatrix} \quad (1B.7)$$

leaving the determinant

$$\begin{vmatrix} d & f \\ g & i \end{vmatrix} \quad (1B.8)$$

The second term in the expansion becomes

$$b \begin{vmatrix} d & f \\ g & i \end{vmatrix} \quad (1B.9)$$

Repeating the same process for the element in row 1 column 3 gives

$$\begin{vmatrix} a & b & e \\ d & e & f \\ g & h & i \end{vmatrix} \quad (1B.10)$$

leaving as the third term in the expansion

$$c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \quad (1B.11)$$

Each of the three terms in eqn (1B.6), (1B.9) and (1B.11) are called the *minors* of the 3×3 determinant **D**. To complete the expansion we must include a phase factor, $(-1)^{i+j}$ where i is the row and j the column which has been struck out. The phase factor takes the form

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix} \quad (1B.12)$$

Now we can write \mathbf{D} in terms of the three *cofactors* (being the minors multiplied by the phase factor) as

$$\begin{aligned} \mathbf{D} &= \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = (-1)^{1+1} a \begin{vmatrix} e & f \\ h & i \end{vmatrix} + (-1)^{1+2} b \begin{vmatrix} d & f \\ g & i \end{vmatrix} (-1)^{1+3} c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \quad (1B.13) \\ &= a(ei-fh) - b(di-fg) + c(dh-eg) \end{aligned}$$

The expansion of an $N \times N$ determinant contains $N!$ terms. For \mathbf{D} , $N = 3$ and $N! = 6$. Moving to $N = 4$, will yield $N! = 24$ terms. The procedure remains the same, the $N \times N$ determinant is reduced to an expansion of cofactors containing $(N-1) \times (N-1)$ determinants. These in turn are reduced until all the cofactors contain 2×2 determinants that can be evaluated using eqn (1B.2).

For our purposes we need note only three important properties of determinants that we will illustrate with \mathbf{D} given in eqn (1B.3).

1. A determinant changes sign if two rows are interchanged

$$\mathbf{D} = a(ei-fh) - b(di-fg) + c(dh-eg)$$

swap the top two rows

$$\mathbf{D}' = \begin{vmatrix} d & e & f \\ a & b & c \\ g & h & i \end{vmatrix} = dbi - dch - eai + ecg + fah - fbg \quad (1B.14)$$

Comparing term by term we find that $\mathbf{D}' = -\mathbf{D}$.

2. A determinant changes sign if two columns are interchanged
Swapping the first two columns of \mathbf{D} gives

$$\mathbf{D}'' = \begin{vmatrix} b & a & c \\ e & d & f \\ h & g & i \end{vmatrix} = bdi - bfg - aei + afh + ceg - cdh \quad (1B.15)$$

Again comparing term by term we find that $\mathbf{D}'' = -\mathbf{D}$.

3. A determinant is zero if two rows or columns are the same
Taking the first two columns as the same

$$\mathbf{D}''' = \begin{vmatrix} a & a & c \\ d & d & f \\ g & g & i \end{vmatrix} = adi - afg - adi + afg + cdg - cdg = 0 \quad (1B.16)$$

In the context of Slater determinants, consider the LiH⁺ molecule with orbital occupancy 1σ²2σ

$$\begin{aligned} \Psi^{\text{LiH}^+} &= \frac{1}{\sqrt{6}} \begin{vmatrix} \phi_{1\sigma}(1) & \bar{\phi}_{1\sigma}(1) & \phi_{2\sigma}(1) \\ \phi_{1\sigma}(2) & \bar{\phi}_{1\sigma}(2) & \phi_{2\sigma}(2) \\ \phi_{1\sigma}(3) & \bar{\phi}_{1\sigma}(3) & \phi_{2\sigma}(3) \end{vmatrix} \\ &= \frac{1}{\sqrt{6}} \left\{ \begin{array}{l} \phi_{1\sigma}(1)[\bar{\phi}_{1\sigma}(2)\phi_{2\sigma}(3) - \phi_{2\sigma}(2)\bar{\phi}_{1\sigma}(3)] \\ -\bar{\phi}_{1\sigma}(1)[\phi_{1\sigma}(2)\phi_{2\sigma}(3) - \phi_{2\sigma}(2)\phi_{1\sigma}(3)] \\ +\phi_{2\sigma}(1)[\phi_{1\sigma}(2)\bar{\phi}_{1\sigma}(3) - \bar{\phi}_{1\sigma}(2)\phi_{1\sigma}(3)] \end{array} \right\} \end{aligned} \quad (1B.17)$$

If the 1σ² electrons were of parallel spin, then we would substitute $\bar{\phi}_{1\sigma} \rightarrow \phi_{1\sigma}$. In which case the first two terms in eqn (1B.17) would cancel each other and the part of the third term in square brackets would be zero. Consequently $\Psi^{\text{LiH}^+} = 0$, in accordance with the Pauli exclusion principle which states that such a wavefunction cannot describe the electronic structure of LiH⁺.

Appendix 1C Diagonalisation of Matrices

For symmetric matrices, such as the ones we shall deal with here, there are many techniques for obtaining the eigenvalues and eigenvectors. We shall briefly describe only one technique, known as the *Jacobi method*. It is not the most efficient method but is easy to explain and fairly robust in numerical application.

A symmetric $n \times n$ matrix, \mathbf{A} , can be reduced to diagonal form, \mathbf{a} , through a unitary transformation, \mathbf{U} . A unitary transformation is one which preserves orthonormality of vectors. Hence, \mathbf{A} obeys the eigenvalue equation

$$\mathbf{AU} = \mathbf{Ua} \quad (1C.1)$$

where

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \quad \mathbf{U} = \begin{pmatrix} U_{11} & U_{12} & \cdots & U_{1n} \\ U_{21} & U_{22} & \cdots & U_{2n} \\ \vdots & \vdots & & \vdots \\ U_{n1} & U_{n2} & \cdots & U_{nn} \end{pmatrix}$$

$$\mathbf{a} = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{11} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

The diagonal elements of the matrix \mathbf{a} contain the eigenvalues of the matrix \mathbf{A} , and the *columns* of the transformation matrix \mathbf{U} contain the eigenvectors of the matrix \mathbf{A} .

A unitary matrix such as \mathbf{U} has the special property that its inverse, \mathbf{U}^{-1} , is equal to its *conjugate transpose*, \mathbf{U}^* . The conjugate transpose of a matrix is obtained by taking the complex conjugate of all elements, U_{ij}^* , and then swapping all elements U_{ij}^* with U_{ji} . In most of this book we are concerned with real matrices, and so \mathbf{U} is orthogonal rather than unitary and the inverse matrix is simply the transpose of \mathbf{U} : $\mathbf{U}^{-1} = \mathbf{U}^T$. Thus to obtain \mathbf{a} , we pre-multiply \mathbf{A} by \mathbf{U}^T and post-multiply by \mathbf{U}

$$\mathbf{U}^T \mathbf{A} \mathbf{U} = \mathbf{a} \quad (1C.2)$$

All that remains is to specify the form of \mathbf{U} . We can specify an elementary unitary matrix, \mathbf{u} , as $\mathbf{u}(i,j,\theta)$, for example

$$\mathbf{u}(2,4,\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\theta & 0 & \cos\theta \end{pmatrix} \quad (1C.3)$$

This matrix has 1 along the diagonal and 0 everywhere else except for 4 entries in the “*ij-plane*”. In the example above the *ij* refers to 2,4. The entries in the 2,2 and 4,4 positions are replaced by $\cos\theta$ and those in the 2,4 and 4,2 positions by $\sin\theta$ and $-\sin\theta$, respectively. A matrix such as \mathbf{u} describes a rotation in the *ij* (2,4) plane

$$\mathbf{u}^T \mathbf{A} \mathbf{u} = \mathbf{A}' \quad (1C.4)$$

in which we have omitted the dependence on the plane ij and the rotation angle θ , for notational convenience. The angle of rotation, θ , may be chosen so as to eliminate the ij and ji elements of the matrix \mathbf{A}' . This is most easily illustrated by a 2×2 example. Consider the symmetric matrix,

$$\mathbf{A} = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}$$

there is only one off-diagonal element to eliminate and

$$\mathbf{u} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \text{ and } \tan 2\theta = \frac{2A_{12}}{A_{22} - A_{11}}$$

$$\Rightarrow \theta = 0.463647 \text{ radians}; \cos \theta = 0.894427; \sin \theta = 0.447213$$

$$\begin{aligned} \mathbf{u}^T \mathbf{A} \mathbf{u} &= \begin{pmatrix} 0.894427 & -0.447213 \\ 0.447213 & 0.894427 \end{pmatrix} \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 0.894427 & 0.447213 \\ -0.447213 & 0.894427 \end{pmatrix} \\ &= \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

So the matrix \mathbf{A} has eigenvalues of 6 and 1 with corresponding eigenvectors of $(0.894427, -0.447213)$ and $(0.447213, 0.894427)$.

For a matrix of dimension >2 , we proceed by searching the lower triangle (since the matrix is symmetric) for the largest off-diagonal matrix element, which we denote as A_{ij} . The angle of rotation required to eliminate this off-diagonal element is evaluated as

$$\tan 2\theta = \frac{2A_{ij}}{A_{ii} - A_{jj}} \left\{ \text{if } A_{ii} = A_{jj} \text{ then } \theta = \frac{\pi}{4} \right\} \quad (1C.5)$$

and we form $\mathbf{u}_1^T \mathbf{A} \mathbf{u}_1 = \mathbf{A}'$. We next search the lower triangle of the matrix \mathbf{A}' for the largest off-diagonal element, A'_{ij} , and repeat our procedure until no off-diagonal elements remain

$$\mathbf{u}_2^T \mathbf{A}' \mathbf{u}_2 = \mathbf{A}''$$

$$\mathbf{u}_3^T \mathbf{A}'' \mathbf{u}_3 = \mathbf{A}'''$$

$$\vdots$$

$$\mathbf{u}_m^T \mathbf{A}''' \mathbf{u}_m = \mathbf{a} \text{ (diagonal)}$$

The total rotation matrix \mathbf{U} is then given by

$$\mathbf{U} = \mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3 \cdots \mathbf{u}_m \quad (1C.6)$$

By such a procedure we can find the eigenvalues and eigenvectors of a symmetric matrix of essentially any size (provided it fits into the memory of a computer). This also provides a general scheme for computing *powers of a matrix*, e.g. $\mathbf{A}^3, \mathbf{A}^{\frac{1}{2}}, \mathbf{A}^{-\frac{3}{2}}$. In general

$$\mathbf{A}^x = \mathbf{U} \mathbf{a}^x \mathbf{U}^{-1}$$

First the eigenvalues and eigenvectors must be found. Then the eigenvalues (diagonal elements of \mathbf{a}) are raised to the appropriate power to produce \mathbf{a}^x , finally this matrix is back transformed by the eigenvector matrix to yield \mathbf{A}^x .

To implement this diagonalisation procedure as a computer program we can exploit the fact that each plane rotation matrix $\mathbf{u}(i,j,\theta)$ given by eqn (1C.3), only affects the rows and columns of the matrix \mathbf{A} that contain the indices i and j . The updated elements of $\mathbf{A}' = \mathbf{u}^T(i,j,\theta) \mathbf{A} \mathbf{u}(i,j,\theta)$ are:

$$\begin{aligned} A'_{ij} &= A'_{ji} = 0 \\ A'_{ii} &= \frac{1}{2}(A_{ii} + A_{jj}) + \frac{1}{2}(A_{ii} - A_{jj})\cos(2\theta) - A_{ij}\sin(2\theta) \\ A'_{jj} &= \frac{1}{2}(A_{ii} + A_{jj}) - \frac{1}{2}(A_{ii} - A_{jj})\cos(2\theta) + A_{ij}\sin(2\theta) \\ A'_{ik} &= A'_{ki} = A_{ik}\cos(\theta) - A_{jk}\sin(\theta) \quad [k \neq i, j] \\ A'_{jk} &= A'_{kj} = A_{ik}\sin(\theta) + A_{jk}\cos(\theta) \quad [k \neq i, j] \end{aligned} \quad (1C.7)$$

To evaluate these quantities we need to know the cosine and sine of θ and 2θ . These can be obtained directly in terms of the elements of \mathbf{A} as:

$$\begin{aligned} \cos(2\theta) &= \frac{|A_{ii} - A_{jj}|}{\sqrt{(A_{ii} - A_{jj})^2 + 4A_{ij}^2}} \\ \sin(2\theta) &= sign \frac{2A_{ij}}{\sqrt{(A_{ii} - A_{jj})^2 + 4A_{ij}^2}} \\ \cos(\theta) &= \sqrt{\frac{1}{2}(\cos(2\theta) + 1)} \\ \sin(\theta) &= \frac{\sin(2\theta)}{2\cos(\theta)} \end{aligned} \quad (1C.8)$$

In the second equation above, the *sign* factor is taken depending on whether the quantity $(A_{ii} - A_{jj})$ is positive or negative. Similarly, the elements of

$\mathbf{U}' = \mathbf{U} \mathbf{u}(i,j,\theta)$ can be accumulated as:

$$\begin{aligned} U'_{ki} &= U_{ki}\cos(\theta) - U_{kj}\sin(\theta) \\ U'_{kj} &= U_{ki}\sin(\theta) + U_{kj}\cos(\theta) \end{aligned} \quad (1C.9)$$

This process will produce the eigenvalues along the diagonal of \mathbf{A} in random order and it will be necessary to order them into ascending (or descending) order and, in doing so, remembering to swap the columns of the accumulated transformation matrix \mathbf{U} .

Appendix 1D Moments of Inertia and the Inertia Tensor

Let us assume that we have the cartesian coordinates of a general molecule relative to some origin, $X' = (x', y', z')$. The centre of mass in this system has coordinates:

$$x^{\text{CoM}} = \frac{\sum_{\text{atoms}} m_A x'_A}{\sum_{\text{atoms}} m_A} : y^{\text{CoM}} = \frac{\sum_{\text{atoms}} m_A y'_A}{\sum_{\text{atoms}} m_A} : z^{\text{CoM}} = \frac{\sum_{\text{atoms}} m_A z'_A}{\sum_{\text{atoms}} m_A} \quad (1D.1)$$

The coordinates must now be shifted so that the origin is at the centre of mass:

$$x_A = x'_A - x^{\text{CoM}} : y_A = y'_A - y^{\text{CoM}} : z_A = z'_A - z^{\text{CoM}} \quad (1D.2)$$

Relative to the centre of mass coordinate system, the inertia tensor, \mathbf{I} , has elements

$$\mathbf{I} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \quad (1D.3)$$

The tensor is symmetric: $I_{yx} = I_{xy}$; $I_{zx} = I_{xz}$; $I_{zy} = I_{yz}$. The elements of \mathbf{I} are given by:

$$\begin{aligned} I_{xx} &= \sum_{\text{atoms}} m_A (y_A^2 + z_A^2) : I_{yy} = \sum_{\text{atoms}} m_A (x_A^2 + z_A^2) : I_{zz} = \sum_{\text{atoms}} m_A (x_A^2 + y_A^2) \\ I_{xy} &= \sum_{\text{atoms}} m_A x_A y_A : I_{xz} = \sum_{\text{atoms}} m_A x_A z_A : I_{yz} = \sum_{\text{atoms}} m_A y_A z_A \end{aligned} \quad (1D.4)$$

To obtain the principal moments of inertia, I_A , I_B , I_C , we must assemble and diagonalise \mathbf{I} . The eigenvalues of \mathbf{I} give the principal moments of inertia, and by convention we choose $I_A \leq I_B \leq I_C$.

For example consider the hydrogen peroxide molecule with coordinates (au):

Coordinates (au) of HOOH in initial frame

Atom	<i>x</i>	<i>y</i>	<i>z</i>	Masses of atoms / <i>u</i>
O	-7.58339	-1.97542	0.49607	15.99491
H	-8.11665	-0.22102	0.49684	1.00783
O	-4.82324	-1.61428	0.49604	15.99491
H	-4.28999	-3.36869	0.49715	1.00783

Using eqns (1D.1) and (1D.2) we can change the origin of the coordinate system to the centre of mass to obtain

Coordinates (au) of HOOH with origin at centre of mass

Atom	<i>x</i>	<i>y</i>	<i>z</i>
O	-1.38007	-0.18057	-0.00004
H	-1.91333	1.57384	0.00073
O	1.38007	0.18057	-0.00007
H	1.91333	-1.57384	0.00103

We can now form the elements of the inertia tensor using eqn (1D.4)

Components of the inertia tensor

$$I_{xx} = 6.03576$$

$$I_{yy} = 68.30694$$

$$I_{zz} = 74.34270$$

$$I_{xy} = 1.90217$$

$$I_{xz} = -0.00016$$

$$I_{yz} = -0.00058$$

Assembling \mathbf{I} as in eqn (1D.3) and diagonalising yields the eigenvalues and eigenvectors:

Eigenvalues:	(1) 5.97768	(2) 68.36489	(3) 74.34257
--------------	-------------	--------------	--------------

Eigenvectors:

<i>x</i>	0.99566	-0.09302	-0.00001
<i>y</i>	0.09302	0.99566	0.00010
<i>z</i>	0.00000	-0.00010	1.00000

From which we determine that $I_A = 5.97768$, $I_B = 68.36489$, $I_C = 74.34257$, with the units being ua_0^2 .

References

1. W. G. Richards, *Nature*, 1979, **278**, 507.
2. H. F. Schaefer III, *Science*, 1986, **231**, 1100.
3. P. Y. Ayala and H. B. Schlegel, *J. Chem. Phys.*, 1998, **108**, 2314.
4. D. G. Truhlar, *J. Comput. Chem.*, 1991, **12**, 266.
5. O. Sinanoğlu, *J. Chem. Phys.*, 1962, **36**, 706.
6. J. P. Perdew and K. Schmidt, *AIP Conf. Proc.*, 2001, **577**, 1.
7. W. A. de Jong, E. Bylaska, N. Govind, C. L. Janssen, K. Kowalski, T. Müller, D. I. M. B. Nielsen, H. J. J. van Dam, V. Veryazovb and R. Lindh, *Phys. Chem. Chem. Phys.*, 2010, **12**, 6896.