Theoretical Techniques

University of Cambridge Part II Natural Sciences Tripos

Yue Wu

Yusuf Hamied Department of Chemistry
Lensfield Road,
Cambridge, CB2 1EW

yw628@cam.ac.uk

Acknowledgements

Nothing in these lecture notes is completely original. The structure of this notes is based on what was originally lectured in the A4 course by Prof. Angelos Michaelides and Dr. Alex Thom in 2024.

They are nowhere near accurate representations of what was actually lectured, as I have modified everything significantly.

Contents

1 The Schrödinger Equation **1**

1.1 Born–Oppenheimer Approximation 2

2 Force Fields **4**

2.1 Traditional Force Fields 4

2.2 Machine Learning Force Fields 6

3 Density Functional Theory **7**

3.1 The Hohenberg–Kohn Theorems 7

3.2 Kohn–Sham Theory 8

3.3 Exchange and Correlation 10

3.4 Dispersion Forces in DFT 13

4 Molecular Orbitals **19**

4.1 The Orbital Approximation 19

4.2 Hückel Theory 19

4.3 Is Hückel Theory Predictive? 23

4.4 Populations and Bond Orders 25

4.5 Alternant Hydrocarbons 26

4.6 Heteroconjugated Systems and Perturbation Theory 27

4.7 Degenerate Perturbation Theory 29

5 Many-electron Wavefunction **31**

5.1 Spin 31

5.2 Slater Determinants 32

5.3 Hartree–Fock Theory 36

5.4 Electron Correlation 39

6 Reaction Theory **45**

6.1 Beyond the Born–Oppenheimer Approximation (Not Lectured) 45

6.2 Avoided Crossing and Conical Intersections 45

6.3 Chemical Reactions 46

1 The Schrödinger Equation

To deal with a molecular system quantum mechanically, we need the Schrödinger equation. Let's start with the time-independent Schrödinger equation

$$\hat{H}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_n,\mathbf{R}_1\ldots,\mathbf{R}_N) = E\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_n,\mathbf{R}_1\ldots,\mathbf{R}_N)\,, \tag{1.1}$$

where $\hat{H}$ is the Hamiltonian of the system, Ψ is the total wavefunction that depends both on the nuclei coordinates $\{\mathbf{R}_I\}_{I=1}^N$ and the electron coordinates $\{\mathbf{r}_i\}_{i=1}^n$, and E is the energy of the system.

For an isolated molecular system, which is our main interest, the Hamiltonian is given by

$$\hat{H} = \hat{T}^{\text{n}} + \hat{T}^{\text{e}} + \hat{V}^{\text{n-n}} + \hat{V}^{\text{n-e}} + \hat{V}^{\text{e-e}}\,, \tag{1.2}$$

where

- $\hat{T}^{\text{n}}$ is the kinetic energy of the nuclei;
- $\hat{T}^{\text{e}}$ is the kinetic energy of the electrons;
- $\hat{V}^{\text{n-n}}$ is the potential energy from nucleus-nucleus repulsion;
- $\hat{V}^{\text{n-e}}$ is the potential energy from nucleus-electron attraction;
- $\hat{V}^{\text{e-e}}$ is the potential energy from electron-electron repulsion.

We may expand everything out and write

$$\begin{aligned}\hat{H} = &\sum_{I=1}^{N} -\frac{\hbar^2}{2M_I}\nabla^2_{\mathbf{R}_I} + \sum_{i=1}^{n} -\frac{\hbar^2}{2m_e}\nabla^2_{\mathbf{r}_i} \\ &+ \sum_{I<J}^{N} \frac{Z_I Z_J e^2}{4\pi\epsilon_0\,|\mathbf{R}_I - \mathbf{R}_J|} - \sum_{I=1}^{N}\sum_{i=1}^{n} \frac{Z_I e^2}{4\pi\epsilon_0\,|\mathbf{R}_I - \mathbf{r}_i|} + \sum_{i<j}^{n} \frac{e^2}{4\pi\epsilon_0\,|\mathbf{r}_i - \mathbf{r}_j|}\,.\end{aligned} \tag{1.3}$$

To avoid writing out so many constants, we introduce the atomic units, in which

(i) unit of length is *Bohr* a_0;

(ii) unit of charge is the magnitude of the charge on the electron e;

(iii) unit of action is *reduced Planck's constant* $\hbar$;

(iv) unit of mass is the mass of the electron m_e.

Then

$$\hat{H} = \sum_{I=1}^{N} -\frac{1}{2M_I}\nabla^2_{\mathbf{R}_I} + \sum_{i=1}^{n} -\frac{1}{2}\nabla^2_{\mathbf{r}_i} + \sum_{I<J}^{N} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} - \sum_{I=1}^{N}\sum_{i=1}^{n} \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|} + \sum_{i<j}^{n} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}\,. \tag{1.4}$$

We have a horrible quantum many-body equation to solve! The motions of electrons and nuclei are coupled in a complicated manner, making it impossible to give exact solutions of the Schrödinger equation apart from the simplest systems (hydrogen-like atoms), or the limiting case like an infinite 'jellium'. We therefore need to make approximations.

1.1 Born–Oppenheimer Approximation

We will start with a simple but powerful approximation which you have met in Part IB: the *Born–Oppenheimer approximation*.

The Born–Oppenheimer approximation is based on the fact that the nuclei have masses much greater than the electrons: the proton-to-electron mass ratio is $m_p/m_e \approx 1836$ and the neutron-to-electron mass ratio is $m_n/m_e \approx 1839$. Since nuclei are much heavier than electrons, they move much more slowly. Essentially, the Born–Oppenheimer approximation assumes that the fast-moving electrons can adjust instantaneously to the changing positions of the slow-moving nuclei. This is called an *adiabatic approximation*, where the fast electronic degrees of freedom are assumed to adjust instantaneously to the nuclear positions. In practice one often follows the electronic ground-state potential energy surface, so that the electron relaxes to the ground state for each nuclear geometry, although excited-state surfaces can also be considered. An immediate result is that this allows us to decouple the motion of the electrons and nuclei, so we can approximate the total wavefunction as a product of an electronic wavefunction and a nuclear wavefunction:

$$\Psi^{\text{B–O}}(\mathbf{r}_1, \ldots, \mathbf{r}_n, \mathbf{R}_1 \ldots, \mathbf{R}_N) = \Psi^{\text{e}}(\mathbf{r}_1, \ldots, \mathbf{r}_n; \mathbf{R}_1 \ldots, \mathbf{R}_N)\Psi^{n}(\mathbf{R}_1 \ldots, \mathbf{R}_N)\,. \tag{1.5}$$

It is important to note that the electronic wavefunction Ψ^{e} is parametrically dependent upon the nuclear coordinates, i.e. they are treated as fixed.

For convenience, we will denote $\mathsf{r} = (\mathbf{r}_1, \ldots, \mathbf{r}_n)$ and $\mathsf{R} = (\mathbf{R}_1 \ldots, \mathbf{R}_N)$.

Then, for a given set of nuclear geometries, we can write out the electronic Hamiltonian[1]

$$\hat{H}^{\text{e}}(\mathsf{r}; \mathsf{R}) = \hat{T}^{\text{e}}(\mathsf{r}) + \hat{V}^{\text{n-n}}(\mathsf{R}) + \hat{V}^{\text{n-e}}(\mathsf{r}; \mathsf{R}) + \hat{V}^{\text{e-e}}(\mathsf{r}) \tag{1.6}$$

and solve for the electronic Schrödinger equation

$$\hat{H}^{\text{e}}(\mathsf{r}; \mathsf{R})\Psi^{\text{e}}(\mathsf{r}; \mathsf{R}) = E^{\text{e}}(\mathsf{R})\Psi^{\text{e}}(\mathsf{r}; \mathsf{R})\,. \tag{1.7}$$

We may repeat this calculation for a large number of different nuclear geometries and get the electronic energy as a function of the nuclear geometry. This is seen as the average energy of a molecule at a certain nuclear configuration would have, assuming the nuclei are not moving.

The energy we obtained above can be thought of as the potential energy for nuclear motion, and can be combined with the nuclear kinetic energy operator to give the nuclear motion Hamiltonian

$$\hat{H}^{\text{n}}(\mathsf{R}) = \hat{T}^{\text{n}}(\mathsf{R}) + E^{\text{e}}(\mathsf{R})\,. \tag{1.8}$$

Solving the Schrödinger equation with this Hamiltonian gives the total energy E and the nuclear wavefunction $\Psi^{\text{n}}(\mathsf{R})$.

The Born–Oppenheimer approximation is a foundational approximation used in the vast majority of simulations in computational chemistry. It allows for a computationally tractable approach to molecular systems by significantly reducing the complexity of the coupled electron-nuclei problem. We are now left with only one term in the Hamiltonian that is difficult to deal with, which is the electron-electron repulsion.

Born–Oppenheimer approximation works well in most situations where the nuclei move much slower than the electrons. However, it breaks in situations where we have e.g. excited electronic states, conical intersections (where different potential energy surfaces become degenerate and intersect), high energy collisions, certain diffusion processes on metal surfaces.

[1] There is no clear convention whether $\hat{V}^{\text{n-n}}(\mathsf{R})$ should appear in the electronic Hamiltonian or in the nuclear Hamiltonian — but it should definitely appear once, and only once. This is just a constant energy term for each particular nuclear geometry, so it does not affect the solution. But remember to always include it in the potential energy surface (see later).

1.1.1 The Potential Energy Surfaces

As we pointed out above, since $E^{\mathrm{e}}(\mathsf{R})$ can be thought of as the potential energy for a given nuclear configuration, a plot of E^{e} as a function of the nuclear coordinates $\mathsf{R} = (\mathbf{R}_1, \ldots, \mathbf{R}_N)$ is known as a *potential energy surface* (PES).

However, note that translations and rotations do not affect the electronic energies of a molecular system, so E^{e} actually has fewer degrees of freedom than $\dim \mathsf{R} = 3N$. A molecule has 3 translational degrees of freedom and 3 rotational degrees of freedom (or 2 if the molecule is linear), hence a PES is a $3N - 6$ dimensional surface (or $3N - 5$ for linear molecules).

The PES is crucial in studying chemical reactions, as it allows us to identify transition states (the highest energy points along the reaction pathway) and compute activation energies. These features of the PES help determine how fast a reaction will occur.

Since the Born–Oppenheimer approximation underpins the concept of the PES, PESs are often called Born–Oppenheimer surfaces or Born–Oppenheimer PESs.

PESs will be covered in more detail in the Part III M4: *Energy Landscapes and Soft Matter* course. We limit our discussion here to key features that are helpful for understanding what comes later in this course.

Figure 1.1 is an example of a (simplified) potential energy surface with key features labeled. The key features include

(i) Minima: Local minima on the PES correspond to stable molecular geometries where the molecule has low energy. These are labeled as the Initial State and Final State on figure 1.1. The global minimum represents the most stable structure (the equilibrium geometry) of the system.

(ii) Transition States: Saddle points on the PES correspond to transition states, which are higher-energy configurations that the system may pass through during a chemical reaction. These are critical in determining the activation energy required for a reaction to occur.

(iii) Energy Barriers: The energy difference between a reactant minimum and a transition state is the activation energy.

(iv) The reaction pathway is the route taken from reactants to products on the PES. It shows how the nuclei rearrange during a chemical reaction. By following the lowest-energy path (the minimum energy pathway, red line in figure 1.1) between two minima, we can understand how a reaction proceeds.

One limitation of PESs is that for systems with a large number of atoms, the PES becomes a high-dimensional object that is difficult to visualize or compute. Therefore, what one often do is to 'project' the PES to some lower-dimensional object so that it can be displayed on a 2D graph.

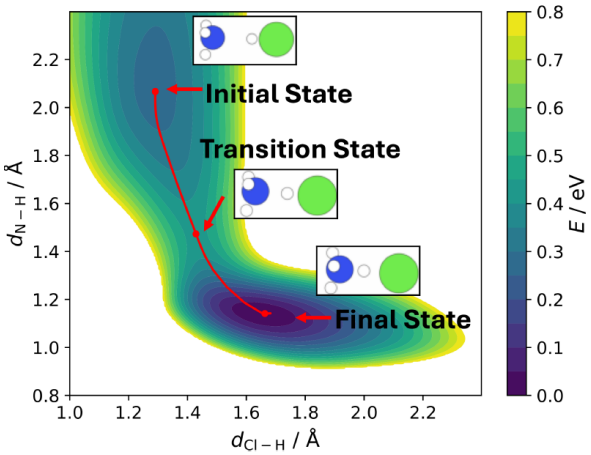


Figure 1.1: A 2-dimensional PES showing the reaction of NH_3 with HCl to produce NH_4^+ and Cl^-. Figure adapted from the official course notes by Prof. Angelos Michaelides.

2 Force Fields

2.1 Traditional Force Fields

Schrödinger equations for molecular systems are infamously hard to solve, so a genius idea would be discarding them and going back to Newton's classical mechanics.

Force field (FF) methods neglect the electrons as well as the quantum mechanical nature of molecules, and calculate the energies of chemical systems using empirical potentials.

A familiar example is the *Morse potential*, which is often used to model the interaction between a pair of atoms in a diatomic molecule. It has the form

$$V(r) = D_e \left(1 - \mathrm{e}^{-\beta(r-r_e)}\right)^2 . \tag{2.1}$$

Comments (that should be familiar from IB):

(i) D_e is the dissociation energy, i.e. $D_e = \lim_{r\to\infty} V(r)$.

(ii) The minimum of $V(r)$ occurs at $r = r_e$, the equilibrium bond length.

(iii) β controls the width of the potential, which is related to the stiffness of the bond.

(iv) The potential is anharmonic: it is steeper for compression ($r < r_e$) than for stretching ($r > r_e$). The potential is approximately harmonic near the equilibrium point (from Taylor expansion).

Another example is the *Lennard-Jones potential*, which is used to describe the intermolecular interaction between a pair of neutral atoms or molecules. It has the expression

$$V(r) = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] . \tag{2.2}$$

Comments (that should be familiar from IB):

(i) ϵ is the depth of the potential well.

(ii) σ is the zero point of the potential.

(iii) The $(\sigma/r)^6$ term represents the long-ranged attractive van der Waals interaction.

(iv) The $(\sigma/r)^{12}$ term represents the strong repulsive interaction at short distances due to Pauli repulsion. The choice of the power 12 here is arbitrary — it's just for computational convenience while making sure that it dominates over the $(\sigma/r)^6$ term at short distances. A more physically motivated form replaces this with an exponential (the exp-6 or Buckingham potential).

To describe more complicated molecules, more complicated potentials have been developed. It's common to refer to these as *molecular mechanics potentials*. Examples of these include widely used force fields such as AMBER, CHARMM, and OPLS. A generic energy expression for potentials of this type can be written as

$$E_{\text{total}} = E_{\text{stretch}} + E_{\text{bend}} + E_{\text{torsion}} + E_{\text{non-bonded}} \,. \tag{2.3}$$

Some of the simplest possible forms of these potentials might be

- Bond stretching energy: This may be modeled by a harmonic potential

$$E_{\text{stretch}} = \sum_{\text{bonds}} k_b (r - r_e)^2 \,. \tag{2.4}$$

- Angle bending energy: This is the energy due to deviation from equilibrium bond angles, which can also be modeled by a harmonic potential

$$E_{\text{bend}} = \sum_{\text{angles}} k_\theta (\theta - \theta_e)^2 \,. \tag{2.5}$$

- Torsional (dihedral) energy: This is the energy associated with rotation around bonds, typically modeled using a periodic potential function

$$E_{\text{torsion}} = \sum_{\text{dihedrals}} \frac{V_n}{2} [1 + \cos(n\phi - \gamma)] \,, \tag{2.6}$$

  where n is the periodicity of the torsion and γ is the phase offset.

- Non-bonded interaction energies: this includes the van der Waals energy, which may be modeled using the Lennard-Jones potential

$$E_{\text{vdW}} = \sum_{\text{non-bonded}} 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right] , \tag{2.7}$$

  and the electrostatic energy, which is computed by Coulomb's law

$$E_{\text{electrostatic}} = \sum_{\text{non-bonded}} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \,. \tag{2.8}$$

The advantages of the force field methods include

- Fast and has good computational scaling $O(n)$, meaning that computational time scales with n, the number of atoms in the system. For comparison, the scaling of quantum methods: DFT is $O(n^3)$ and CCSD(T) is $O(n^7)$.
- Can be accurate, if the system is inside the space the force field is parameterised.
- Tuning parameters can give insight.
- Simple and easy to use (once a model is developed).

Disadvantages include

- Not accurate outside training space.

- Generally not good for reactions.
- No electrons (most are not polarisable).
- Difficult and labour-intensive to parameterise — tons of parameters need to be fitted.
- Struggle for complex systems and interfaces.
- Functional forms can be too restrictive.

2.2 Machine Learning Force Fields

You may have noticed that machine learning (ML) has become phenomenal in society over the last few years. It has also had a huge impact in the physical sciences, including chemistry, winning two Nobel prizes in 2024.

This gives rise to a new type of force field known as the *machine learning interatomic potentials* (MLIPs). We will only give a quick overview of MLIPs. You will learn more about it in Part III L10 course.

The basic idea of MLIP is that you do a large set of high-accuracy *ab initio* calculations on a range of structures, and use machine learning to studying them (which is basically a huge network of matrices that can be parameterised to fit the input data). To use it, you simply put your structure into the machine learning model so it returns you the energy and forces on atoms.

The accuracy of MLIP depends on your input data and how well you fitting is, but usually it is able to deliver near DFT/*ab initio* accuracy with a computational cost slightly above a traditional force field.

3 Density Functional Theory

Rather than stepping back to classical mechanics, what else can we do to avoid solving the horrible looking Schrödinger equation

$$\hat{H}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_n,\mathbf{R}_1\ldots,\mathbf{R}_N) = E\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_n,\mathbf{R}_1\ldots,\mathbf{R}_N) \tag{3.1}$$

while preserving the quantum mechanical nature of molecular systems?

The answer is, rather than looking at the molecular wavefunction $\Psi : \mathbb{R}^{3(N+n)} \to \mathbb{C}$ of $3(N+n)$ variables, we look at the electron density $n : \mathbb{R}^3 \to \mathbb{R}$, an apparently lower dimensional and so simpler object. This leads to the *density functional theory* (DFT).[2]

Before introducing density functional theory, let's first introduce what a *functional* is. Abstractly, let V be a vector space over a field $\mathbb{F}$, then a *functional* is a map $F : V \to \mathbb{F}$. Within our context, we can think of a functional as special type of function whose input is a function, and the output is a number. For example, $\max_{x\in\mathbb{R}}$ is a functional, giving $\max_{x\in\mathbb{R}}[\sin x] = 1$, $\max_{x\in\mathbb{R}}[4-x^2] = 4$ etc. A particularly important class of functionals is definite integrals, such as

$$F[f] = \int_\alpha^\beta \mathrm{d}x\, f(x) \tag{3.2}$$

for some α, β. Note that it is conventional to put the argument of a functional in square brackets [ ].

3.1 The Hohenberg–Kohn Theorems

3.1.1 The First Hohenberg–Kohn Theorem

The essence of density functional theory is that the ground state electron energy is a unique functional of the ground state electron density. This means that given a ground state electron density $n(\mathbf{r})$, then we know which external potential (generated by the nuclei under Born–Oppenheimer approximation) uniques produces it, and then we know the unique electron wavefunction it corresponds to, and also the ground state electron energy, and everything else.

This is known as the First Hohenberg–Kohn theorem.

Theorem 3.1 (The first Hohenberg–Kohn theorem). The external potential $V_{\text{ext}}(\mathbf{r})$ and hence the total energy are unique functionals of the electron density $n(\mathbf{r})$ up to an additive constant.

Proof. Suppose there are two external potentials $V_1(\mathbf{r})$ and $V_2(\mathbf{r})$ both lead to the same $n(\mathbf{r})$. There will be two Hamiltonians $\hat{H}_1$ and $\hat{H}_2$ with the same ground state density, but different wavefunctions Ψ_1 and Ψ_2. Now use the variational principle,

$$\begin{aligned} E_1^0 < \left\langle \Psi_2 \middle| \hat{H}_1 \middle| \Psi_2 \right\rangle &= \left\langle \Psi_2 \middle| \hat{H}_2 \middle| \Psi_2 \right\rangle + \left\langle \Psi_2 \middle| \hat{H}_2 - \hat{H}_1 \middle| \Psi_2 \right\rangle \\ &= E_2^0 + \int \mathrm{d}^3\mathbf{r}\, n(\mathbf{r})[V_1(\mathbf{r}) - V_2(\mathbf{r})]\,, \end{aligned} \tag{3.3}$$

where E_0^1 and E_2^0 are the ground state energies for $\hat{H}_1$ and $\hat{H}_2$, respectively. The subscripts 1 and 2 may be interchanged to give a second inequality. These two inequalities may be added to give

$$E_1^0 + E_2^0 < E_2^0 + E_1^0\,, \tag{3.4}$$

which is a contradiction. □

[2] A more mathematical treatment of DFT can be found in my notes on *B10: Electronic Structure*. Pathetically, this part of the course is now removed, so you can't see it in the official course notes.

Thus, at least in principle, the ground-state density determines (within a constant) the external potential in the Schrödinger equation of which it is a solution. Therefore, the electronic energy is also uniquely determined, via

$$\begin{aligned} E^{\mathrm{e}}[n] &= \left\langle \Psi[n] \middle| \hat{T}^{\mathrm{e}} + \hat{V}^{\mathrm{n\text{-}e}}[n] + \hat{V}^{\mathrm{e\text{-}e}} \middle| \Psi[n] \right\rangle \\ &= T^{\mathrm{e}}[n] + V^{\mathrm{n\text{-}e}}[n] + V^{\mathrm{e\text{-}e}}[n] \\ &= T^{\mathrm{e}}[n] + V^{\mathrm{e\text{-}e}}[n] + \int \mathrm{d}^3\mathbf{r}\, \tilde{V}_{\mathrm{ext}}[n](\mathbf{r}) n(\mathbf{r}) \,, \end{aligned} \tag{3.5}$$

where $\Psi[n]$ and $\tilde{V}_{\mathrm{ext}}[n](\mathbf{r})$ are the wavefunction and the external potential determined by $n(\mathbf{r})$.

3.1.2 The Second Hohenberg–Kohn theorem

The First Hohenberg–Kohn theorem shows that working with electron density is theoretically possible, however, the form of $E^{\mathrm{e}}[n]$ it gives is troublesome: the $\tilde{V}_{\mathrm{ext}}[n](\mathbf{r})$ depends on the electron density. Rather, in a real problem, under Born–Oppenheimer approximation, we have a specific nuclear arrangement and hence a known potential, and we want to find a optimum electron density. This is realised by the second Hohenberg–Kohn theorem.

We introduce the shorthand notation

$$F[n] = \left\langle \Psi \middle| \hat{F} \middle| \Psi \right\rangle = \left\langle \Psi \middle| \hat{T} + \hat{V}^{\mathrm{e\text{-}e}} \middle| \Psi \right\rangle = T[n] + V^{\mathrm{e\text{-}e}}[n] \,. \tag{3.6}$$

Now an energy functional for an **arbitrary external potential** $V_{\mathrm{ext}}(\mathbf{r})$ unrelated to the $\tilde{V}_{\mathrm{ext}}(\mathbf{r})$ determined by $n(\mathbf{r})$ can be defined:

$$E_V[n] = F[n] + \int \mathrm{d}^3\mathbf{r}\, V_{\mathrm{ext}}(\mathbf{r}) n(\mathbf{r}) \,. \tag{3.7}$$

Theorem 3.2 (The second Hohenberg–Kohn theorem). For an arbitrary external potential V_{ext} and electron densities $n(\mathbf{r})$ of N electrons,

$$E_V[n] = F[n] + \int \mathrm{d}^3\mathbf{r}\, V_{\mathrm{ext}}(\mathbf{r}) n(\mathbf{r}) \geq E_0 \,, \tag{3.8}$$

where E_0 is the ground state energy for N electrons in the external potential $V_{\mathrm{ext}}(\mathbf{r})$, and this functional reaches its minimum if and only if n is the true ground state electron density in such external potential V_{ext}.

Proof. By the first theorem, a given $n(\mathbf{r})$ determines its own external potential $\tilde{V}_{\mathrm{ext}}(\mathbf{r})$ and ground state $\left|\tilde{\Psi}\right\rangle$. If this state is used as a trial state for the Hamiltonian with external potential $V_{\mathrm{ext}}(\mathbf{r})$, we have

$$\left\langle \tilde{\Psi} \middle| \hat{H} \middle| \tilde{\Psi} \right\rangle = \left\langle \tilde{\Psi} \middle| \hat{F} \middle| \tilde{\Psi} \right\rangle + \left\langle \tilde{\Psi} \middle| \hat{V}_{\mathrm{ext}} \middle| \tilde{\Psi} \right\rangle = F[n] + \int \mathrm{d}^3\mathbf{r}\, V_{\mathrm{ext}}(\mathbf{r}) n(\mathbf{r}) = E_V[n] \geq E_0 \tag{3.9}$$

by the variational principle. For non-degenerate ground states, equality holds if $\left|\tilde{\Psi}\right\rangle$ is the ground state for the potential $V_{\mathrm{ext}}(\mathbf{r})$. □

This means that to find the true ground state energy of a molecular system, we only need to find the electron density $n(\mathbf{r})$ that minimises the functional $E_V[n]$!

3.2 Kohn–Sham Theory

Using the Hohenberg–Kohn theorems, we have reformulated the Schrödinger equation, i.e. finding Ψ that minimises

$$\left\langle \Psi \middle| \hat{T}^{\mathrm{e}} + \hat{V}^{\mathrm{n\text{-}e}} + \hat{V}^{\mathrm{e\text{-}e}} \middle| \Psi \right\rangle \,, \tag{3.10}$$

to finding the electron density n that minimises the energy functional

$$E_V[n] = T^{\text{e}}[n] + V^{\text{e-e}}[n] + \int \mathrm{d}^3\mathbf{r}\, V_{\text{ext}}(\mathbf{r}) n(\mathbf{r})\,, \tag{3.11}$$

which should be much simpler as n is a lower dimensional object.

This now seems like an easy problem as long as we know what the functionals $T^{\text{e}}[n]$ and $V^{\text{e-e}}[n]$ are — but we don't. We haven't changed the nature that this is a quantum many-body problem. The complicated correlated motions of the electrons made it impossible to give their kinetic and interaction energy in a simple closed form. To make progress, we need Kohn–Sham theory.

Although we don't know the kinetic energy of correlated electrons, we know the kinetic energy of a single electron with wavefunction ψ, which is

$$-\frac{1}{2}\int \mathrm{d}^3\mathbf{r}\, \psi^*(\mathbf{r})\nabla^2\psi(\mathbf{r})\,. \tag{3.12}$$

Hence, if we assume that the electrons are moving in their orbitals $\{\phi_i\}$ without interacting with each other[3], we have the kinetic energy

$$T_s^{\text{e}}[n] = -\frac{1}{2}\sum_i \int \mathrm{d}^3\mathbf{r}\, \phi_i^*(\mathbf{r})\nabla^2\phi_i(\mathbf{r})\,. \tag{3.13}$$

These $\{\phi_i\}$ are known as the *Kohn–Sham orbitals*. $T_s^{\text{e}}[n]$ is the kinetic energy functional of non-interacting electrons of density n, where the subscript s indicates "single-particle" electrons.

We also don't know the potential energy from electron interaction due to exchange interactions and correlations[4]. However, we are very familiar with the Coulomb interaction energy of the electrons. This contribution is known as the *Hartree energy* and has the simple form

$$E^{\text{Hartree}}[n] = \frac{1}{2}\int \mathrm{d}^3\mathbf{r}\, \mathrm{d}^3\mathbf{r}'\, \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{3.14}$$

Hence, by pulling these two contributions out, we obtain the formula

$$\begin{aligned} E_V[n] &= T^{\text{e}}[n] + V^{\text{e-e}}[n] + V^{\text{n-e}}[n] \\ &= T_s^{\text{e}}[n] + E^{\text{Hartree}}[n] + V^{\text{n-e}}[n] + \big(T^{\text{e}}[n] + V^{\text{e-e}}[n] - T_s^{\text{e}}[n] - E^{\text{Hartree}}[n]\big) \\ &= T_s^{\text{e}}[n] + E^{\text{Hartree}}[n] + V^{\text{n-e}}[n] + E^{\text{xc}}[n]\,, \end{aligned} \tag{3.15}$$

where

$$E^{\text{xc}}[n] \coloneqq T^{\text{e}}[n] + V^{\text{e-e}}[n] - T_s^{\text{e}}[n] - E^{\text{Hartree}}[n] \tag{3.16}$$

is the exchange and correlation energy. This is now the only term that is unknown. All the complicated quantum-mechanical many-body effects are placed in this term only.

Recall when we introduced the Kohn–Sham orbitals, we defined them as the "effective non-interacting one-electron orbitals". In order to solve them, we need to solve their Schrödinger equation. The effective potential that the electron is moving is

$$\begin{aligned} V^{\text{eff}}(\mathbf{r}) &= \frac{\delta(E^{\text{Hartree}}[n] + V^{\text{n-e}}[n] + E^{\text{xc}}[n])}{\delta n(\mathbf{r})} \\ &= \int \mathrm{d}^3\mathbf{r}'\, \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + V_{\text{ext}}(\mathbf{r}) + \frac{\delta E^{\text{xc}}[n]}{\delta n(\mathbf{r})} \\ &=: V^{\text{Hartree}}(\mathbf{r}) + V_{\text{ext}}(\mathbf{r}) + V^{\text{xc}}(\mathbf{r})\,, \end{aligned} \tag{3.17}$$

[3] Recall that orbitals are effective one-electron wavefunctions. Kohn–Sham theory essentially assumes the system to have a non-interacting wavefunction as a Slater determinant made from the Kohn–Sham orbitals.

[4] Correlation basically means that an electronic system can lower its energy by letting its electrons to avoid each other, so the motion of the electrons are "correlated". See more in later chapters on Hartree–Fock theory.

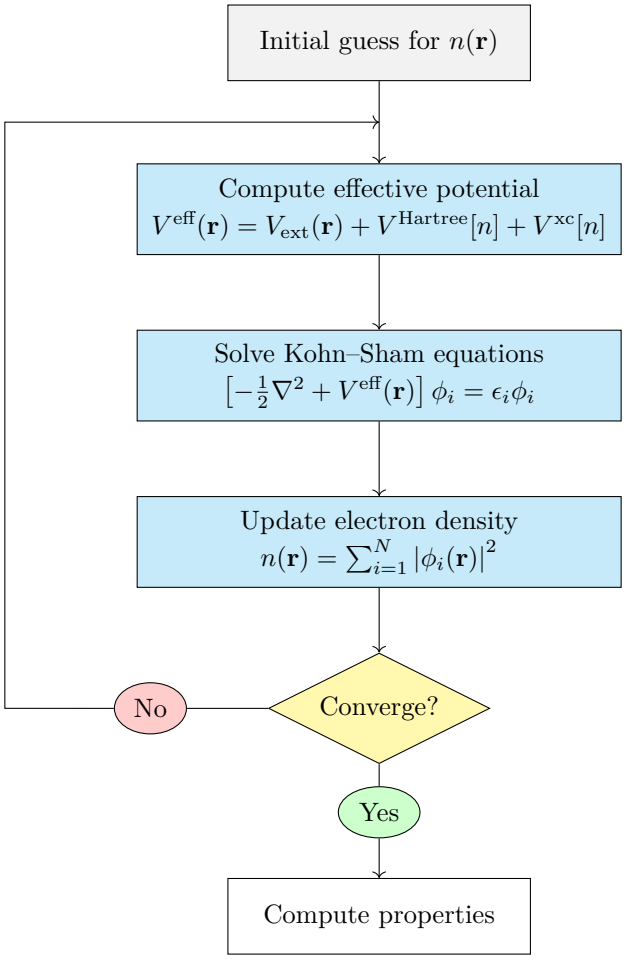


Figure 3.1: Illustration of the self-consistent field (SCF) cycle used to solve the Kohn–Sham equations and obtain the ground state density and total energy.

so the Kohn–Sham orbitals are solved from the Schrödinger equation

$$\left[-\frac{1}{2}\nabla^2 + V^{\text{eff}}\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r})\,. \tag{3.18}$$

This is a single electron Schrödinger equation so it is not painful to solve.

As you can see, the electron density for minimum energy depends on the energy functional, which depends on the Kohn–Sham orbitals, which depends on the effective potentials, which in turn depends on the electron density. Thus, we need to perform this in an iterative process. We must have an initial guess on the electron density $n(\mathbf{r})$, work everything out, and determine a new electron density $n(\mathbf{r})$, and repeat until the electron density $n(\mathbf{r})$ converges. This is known as a *self-consistent field* (SCF) method. A flow chart is shown in figure 3.1.

3.3 Exchange and Correlation

As we commented above, all the complexity of Kohn–Sham theory lies in the exchange correlation functional, where all the quantum many-body effects live in. It is impossible to give an exact and closed form expression of this functional, so approximations must be made.

As its name suggests, $E^{\text{xc}}[n]$ have two contributions

(i) Electron exchange: this arises because a many-body wavefunction must by antisymmetric under exchange of any two electrons since electrons are fermions. This leads to the Pauli exclusion principle, which reduces the probability of electrons with the same spin coming close to each other, and therefore reduces the Coulomb repulsion.

(ii) Electron correlation: it also reduces the Coulomb energy between electrons because the motion of each individual electron is correlated with the motion of all others, helping electrons to keep spatially separated.

Note that an important consequence of exchange interaction being approximated is that it leads to *self-interaction error*, i.e. an electron is repelling itself. In theories like Hartree–Fock where electron exchange are exact, the self repulsion in the Coulomb term is precisely cancelled by the self exchange in the exchange term, so there is no self-interaction. However in DFT, the Coulomb term (Hartree energy) is exact, while the exchange term is not, so the self-interaction is not completely cancelled. This significantly affects the accuracy of DFT for systems with strongly correlated electrons or localised electrons (e.g. transition metal complexes). More errors in DFT will be introduced in later sections.

An exchange correlation functional can be classified as being either "empirical" or "non-empirical". An empirical functional aims to reproduce experiment results and usually contain a lot of parameters fitted from experimental data. A non-empirical functional is designed to abide by a number of physical constraints which the exact E^{xc} functional should obey and do not include parameters other than the fundamental constants.

A major way of categorising exchange correlation functional is to group them in terms of their physical complexity, and it is known as the *Jacob's ladder*. As you climb up the ladder, the functional gets more complex and computationally more expensive, but will usually give higher accuracy. The five rungs of the Jacob's ladder, from low to high, are

(i) local density approximation (LDA);

(ii) generalised gradient approximation (GGA);

(iii) meta-GGA;

(iv) hybrid functionals;

(v) double-hybrid functionals.

3.3.1 Local Density Approximation

There is a system in which the exchange energy is known exactly, which is the *uniform electron gas* (a "jellium"). A uniform electron gas with constant electron density n has the exchange energy

$$E^{\text{x-UEG}}[n] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} \int \mathrm{d}^3\mathbf{r}\, n^{4/3}\,. \tag{3.19}$$

There are also some widely accepted expressions for the correlation energy in uniform electron gas, but they are significantly more complicated. It has the general form

$$E^{\text{c-UEG}}[n] = \int \mathrm{d}^3\mathbf{r}\, n\epsilon^{\text{c-UEG}}(n) \tag{3.20}$$

for some function $\epsilon^{\text{c-UEG}}$.

Knowing this, a natural thought of approximating the exchange and correlation energy in a general system is to treat each point locally as uniform electron gas, so that

$$E^{\text{xc-LDA}}[n] = \int \mathrm{d}^3\mathbf{r}\, n(\mathbf{r})\epsilon^{\text{xc-LDA}}(n(\mathbf{r}))\,, \tag{3.21}$$

where $\epsilon^{\text{xc-LDA}}$ is the exchange-correlation energy per particle of the homogeneous electron gas of density n. From our expressions for uniform electron gas, it should be

$$\begin{aligned}\epsilon^{\text{xc-LDA}}(n(\mathbf{r})) &= \epsilon^{\text{x-LDA}} + \epsilon^{\text{c-LDA}} \\ &= -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} n^{1/3}(\mathbf{r}) + \epsilon^{\text{c-UEG}}(n(\mathbf{r}))\,. \end{aligned} \tag{3.22}$$

This is called *local density approximation* (LDA).

Strictly, the LDA is valid only for slowly varying densities. However, experience with calculations of atoms, molecules, and solids shows that the LDA actually works surprisingly well, given its crudeness.

3.3.2 Generalised Gradient Approximation

To go beyond LDA, a natural thought is to include the local gradient of the electron density, in addition to the electron density itself, into the exchange correlation energy expression. Hence, this is known as *generalised gradient approximation* (GGA). It has the general form

$$E^{\text{xc-GGA}}[n] = \int \mathrm{d}^3\mathbf{r}\, n(\mathbf{r})\epsilon^{\text{xc-GGA}}(n(\mathbf{r}), \boldsymbol{\nabla} n(\mathbf{r}))\,. \tag{3.23}$$

Since it is extremely difficult to come up with a valid correlation functional, there is an extremely limited number of them. Hence, the main difference between the GGAs is usually found in the exchange energy density terms. Since the GGA should reduce to LDA when $|\boldsymbol{\nabla} n| = 0$, a lot of GGA exchange functionals can be written as a series expansion

$$\epsilon^{\text{x-GGA}}(n, s) = \epsilon^{\text{x-LDA}} F^{\text{x-GGA}}(s)\,, \tag{3.24}$$

where $F^{\text{x-GGA}}(s)$ is known as the *enhancement factor* and $s = |\boldsymbol{\nabla} n|\,/2(3\pi^2)^{1/3} n^{4/3}$ is a dimensionless reduced density gradient.

Example. Two related GGA functionals PBE and revPBE have the enhancement factor

$$F^{\text{x-PBE}} = 1 + \kappa - \frac{\kappa}{1 + \frac{\mu s^2}{\kappa}}\,, \tag{3.25}$$

where μ is some parameter that is the same for PBE and revPBE. The difference is only in the parameter κ: $\kappa^{\text{revPBE}} = 1.245$ while $\kappa^{\text{PBE}} = 0.804$. This makes F_x rise more rapidly with s for revPBE than for PBE. As a consequence, regions with large reduced density gradients are stabilised more with revPBE than PBE. This means that in a revPBE calculation, a rapid reduction of the electron density in the internuclear region is expected, which in turn leads to weaker interactions with revPBE.

DFT only became widely used in quantum chemistry after GGAs were developed. It generally has a significant improvement in accuracy compared with LDA, while it does not increase the computational costs significantly.

3.3.3 Meta-GGAs

Having included the local electron density gradients in GGAs, we can further include the second order terms such as the second derivative of electron density $\nabla^2 n(\mathbf{r})$, and/or kinetic energy densities $\tau_\sigma(n) = \frac{1}{2}\sum_i |\boldsymbol{\nabla}\phi_i|^2$. These are known as *meta-GGAs*.

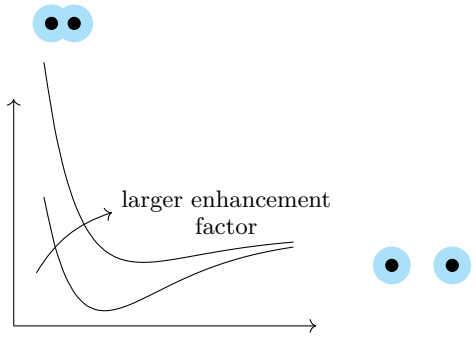


Figure 3.2: GGA with high enhancement factors encourage regions with high local gradient densities. Therefore they predict weaker binding.

3.3.4 Hybrid Functionals

The *hybrid functionals* (or *hybrid GGA*) add the exact exchange interaction calculated from Hartree–Fock (HF) theory to some conventional treatment of DFT exchange and correlation via a technique called adiabatic connection[5].

For example, the most widely used functional in quantum chemistry, B3LYP, takes the following form

$$E^{\text{xc}} = E^{\text{xc-GGA}} + a(E^{\text{x-HF}} - E^{\text{x-GGA}})\,. \tag{3.26}$$

Hybrid functionals have significantly better accuracy, along with a significantly longer computation time. Hybrid functionals are so successful because they benefit from error compensation between "underbinding" of (meta-)GGAs and the "overbinding" of Hartree–Fock. The underbinding of GGAs is due to the self-interaction error and the overbinding of HF is due to the lack of correlation.

3.3.5 Double-hybrid Functionals

Double-hybrid functionals incorporate virtual (unoccupied) orbitals into the functional. For example, the B2-PLYP functional has the form

$$E^{\text{xc}} = (1 - a_x)E^{\text{x-DFT}} + a_x E^{\text{x-HF}} + (1 - a_c)E^{\text{c-DFT}} + a_c E^{\text{c-PT2}}\,. \tag{3.27}$$

It can be seen that in addition to exact exchange (via HF), electron correlation is taken into account through perturbation theory (PT).

These functionals are very computationally expensive and are still not as widely used as other functionals.

3.4 Dispersion Forces in DFT

Another serious problem found for functionals on the first four rungs of Jacob's ladder is that they fail to account for van der Waals dispersion forces. This greatly limited the use of DFT in areas such as biology. Since this century, considerable progress has been made to correct this dispersion error.

[5]Detail in my course notes on B10 electronic structure.

3.4.1 Dispersion Force

The concept of dispersion forces should be familiar. It originates from the electrostatic attraction between the instantaneous dipole in one particle arising from electron density fluctuation and the induced dipole on another particle.

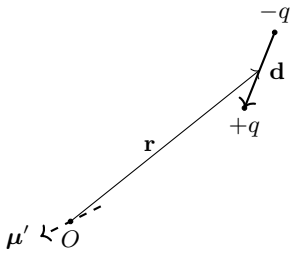
It is hard for DFT to account for dispersion forces because

(i) Under Kohn–Sham approximation, the DFT is a mean field theory, in which an electron interacts with the total electron density, rather than with instantaneous electron positions which is the basis of the dispersion interaction.

(ii) Most exchange correlation functionals used are local. They only use information about the local electron density at a single point, so there is no way electrons on one molecule know there are instantaneous dipole produced on another molecule.

This means that for two noble gas atoms, for example, these functionals can give binding or repulsion only when there is an overlap of the electron densities of the two atoms.

Theorem 3.3. For large r, the dispersion potential energy of two molecules scales as $-\frac{1}{r^6}$.



Proof. Consider a molecule at the origin and another molecule at $\mathbf{r}$. The molecule at $\mathbf{r}$ has an instantaneous dipole modeled by two charges of $+q$ and $-q$ separated by $\mathbf{d}$, so it has the dipole $\boldsymbol{\mu} = q\mathbf{d}$. By assumption, $|\mathbf{r}| \gg |\mathbf{d}|$. The potential at O generated by this dipole is

$$\begin{aligned}\phi &= \frac{1}{4\pi\epsilon_0}\left(\frac{q}{|\mathbf{r}|} - \frac{q}{|\mathbf{r}+\mathbf{d}|}\right)\\ &= \frac{q}{4\pi\epsilon_0}\left(\frac{1}{r} - \left(\frac{1}{r} + \mathbf{d}\cdot\boldsymbol{\nabla}\frac{1}{r} + \frac{1}{2}(\mathbf{d}\cdot\boldsymbol{\nabla})^2\frac{1}{r} + \dots\right)\right)\\ &= \frac{q}{4\pi\epsilon_0}\frac{\mathbf{d}\cdot\mathbf{r}}{r^3} = \frac{\boldsymbol{\mu}\cdot\mathbf{r}}{4\pi\epsilon_0 r^3}\,. \end{aligned} \tag{3.28}$$

Then the electric field at O is

$$\begin{aligned}\mathbf{E} &= -\boldsymbol{\nabla}\phi = -\frac{1}{4\pi\epsilon_0}\mathbf{e}_i\frac{\partial}{\partial x_i}\frac{\mu_j x_j}{(x_k x_k)^{3/2}}\\ &= -\frac{1}{4\pi\epsilon_0}\mathbf{e}_i\left[\frac{\mu_j\delta_{ij}(x_k x_k)^{3/2} - \mu_j x_j\cdot\frac{3}{2}(x_l x_l)^{1/2}\cdot 2\delta_{ik}x_k}{(x_m x_m)^3}\right]\\ &= \frac{1}{4\pi\epsilon_0}\mathbf{e}_i\frac{3(\mathbf{r}\cdot\boldsymbol{\mu})r\mathbf{r} - \boldsymbol{\mu}r^3}{r^6}\\ &= \frac{\mu}{4\pi\epsilon_0 r^3}[3(\hat{\mathbf{r}}\cdot\hat{\boldsymbol{\mu}})\hat{\mathbf{r}} - \hat{\boldsymbol{\mu}}]\,, \end{aligned} \tag{3.29}$$

which scales as r^{-3}. Let the molecule at O have a polarisability α, then the dipole induced by the electric field is

$$\boldsymbol{\mu}' = \alpha\mathbf{E}\,. \tag{3.30}$$

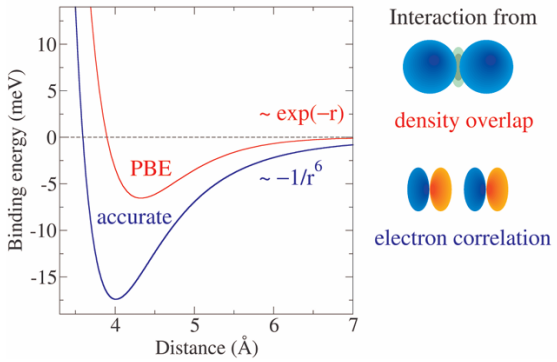


Figure 3.3: Binding curves for the Kr dimer obtained with the PBE exchange-correlation functional and an accurate model potential. PBE can not predict dispersion force, so all the interactions it predicts come from electron density overlap, which decays $\sim \exp(-r)$ at large r. It does predict some binding — but this does not arise from the correct physics. It is because the enhancement factor in this functional is not large enough i.e. it arises because the DFT's inaccuracy in electron density overlap interaction. Figure adapted from J. Chem. Phys. 137, 120901 (2012).

Note that the energy of a dipole-induced dipole pair isn't just $-\boldsymbol{\mu}\cdot\mathbf{E}$ — this would be the case for two permanent dipoles. For the induced dipole here, $\boldsymbol{\mu}$ itself depends on $\mathbf{E}$, so instead, we compute the work done in turning on the field from $\mathbf{E}=\mathbf{0}$ to its final value:

$$\begin{aligned} U &= -\int_{\mathbf{0}}^{\mathbf{E}} \mathrm{d}\mathbf{E}' \cdot \boldsymbol{\mu}'(\mathbf{E}') \\ &= -\frac{1}{2}\mathbf{E}\cdot\alpha\mathbf{E} \\ &= -\frac{\mu^2}{2(4\pi\epsilon_0)^2 r^6}[3(\hat{\mathbf{r}}\cdot\hat{\boldsymbol{\mu}})\hat{\mathbf{r}}-\hat{\boldsymbol{\mu}}]\cdot\alpha[3(\hat{\mathbf{r}}\cdot\hat{\boldsymbol{\mu}})\hat{\mathbf{r}}-\hat{\boldsymbol{\mu}}]\,. \end{aligned} \tag{3.31}$$

This scales as r^{-6} as claimed. □

Here we will introduce some of the most widely used approaches for dispersion force correction.

3.4.2 C_6 Correction Schemes

The basic requirement for any DFT-based dispersion scheme should be that it recovers the correct $-1/r^6$ asymptotic behaviour, so an obvious idea is to add a $-1/r^6$ term separately after the DFT calculation, so the total energy is

$$E = E_{\mathrm{DFT}} + E_{\mathrm{disp}}\,. \tag{3.32}$$

DFT-D Scheme

If we assume the dispersion force to be pairwise additive, isotropic, and constant for the same pair of elements, then we can assume the dispersion interaction between element A and element B is always C_6^{AB}/r^6, and so

$$E_{\mathrm{disp}} = -\sum_{A>B}\frac{C_6^{\mathrm{AB}}}{r_{\mathrm{AB}}^6}\,. \tag{3.33}$$

This very primitive scheme is known as *DFT-D*.

It has some obvious drawbacks:

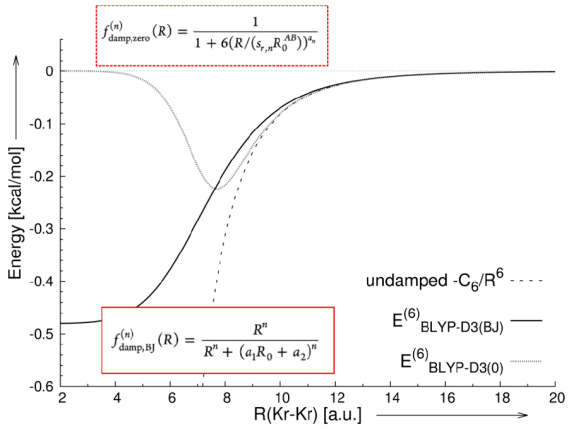


Figure 3.4: Examples of different damping functions. One function goes to zero at short distances and the other goes to a finite value. Figure adapted from Chem. Rev. 116, 5105 (2016); DOI: 10.1021/acs.chemrev.5b00533.

(i) It is not clear where one should obtain the C_6 coefficients: various formulae involving various quantities have been proposed.

(ii) Different chemical states of the atom in different environments have different dispersion interaction. This is neglected.

(iii) The divergence for $1/r^6$ at $r = 0$ must be removed.

DFT-D2 Scheme

In the second generation, DFT-D2, the dispersion coefficients are calculated from a formula which couples ionisation potentials and static polarisabilities of isolated atoms. Also, a *damping* term is introduced to avoid the divergence at $r = 0$. Furthermore, the dispersion energy, E_{disp}, is scaled according to the exchange functional used. This leads to the energy expression

$$E_{\text{disp}} = -\sum_{A>B} f(r_{\text{AB}}, \text{A}, \text{B}, \text{DFT}) \frac{C_6^{\text{AB}}}{r_{\text{AB}}^6} \,. \tag{3.34}$$

The shape of the underlying binding curve is sensitive to the XC functional used and so the damping functions must be adjusted so as to be compatible with each XC functional. This fitting is also sensitive to the definition of atomic size (van der Waals radii are usually used) and must be done carefully since the damping function can actually affect the binding energies even more than the C_6 coefficients.

Some problems still remain. The same coefficient will be assigned to an element no matter what its oxidation or hybridisation state. This error can be huge: the carbon C_6 coefficients can differ by 35% between the sp and sp^3 hybridized states.

DFT-D3 and DFT-D4

In DFT-D3 this environmental dependence of the C_6 coefficients is captured considering the number of neighbors (coordination number) each atom has. This effect is accounted for by having a range of precalculated C_6 coefficients for various pairs of elements in different reference (hybridization) states.

Despite being a simple scheme and having negligible computational cost, it obtains pretty accuracy results (mean error 8.4% for a reference set).

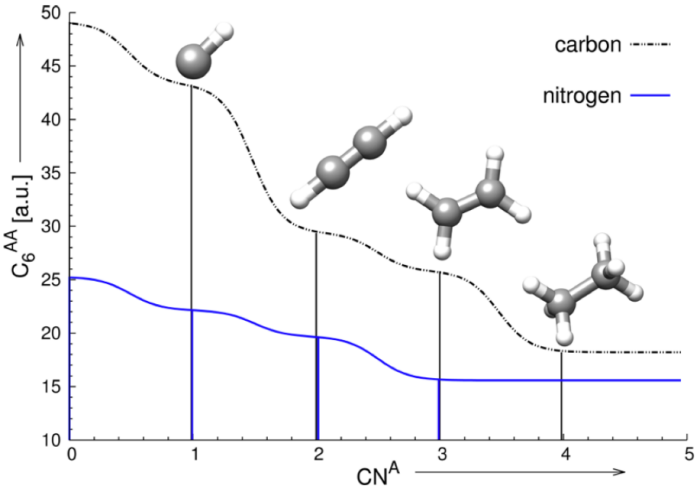


Figure 3.5: C_6^{AA} values for carbon and nitrogen with different coordination number in DFT-D3 scheme. Figure from Chem. Rev. 116, 5105 (2016).

While DFT-D3 takes only coordination number into account, DFT-D4 considers more information. The C_6 coefficients are calculated by numerical integration, considering atomic partial charge, polarisability and geometry.

Tkatchenko–Scheffler Scheme

The difference in C_6 coefficients for the same atom in different environments is mainly caused by their different volumes, leading to different polarisabilities. Therefore, in the *Tkatchenko–Scheffler* (TS) scheme, we have a volume-dependent C_6 coefficient for atoms, determined from

$$C_6^{\mathrm{A}} \propto C_6^{\mathrm{A,free}} \left(\frac{V_{\mathrm{A}}}{V_{\mathrm{A}}^{\mathrm{free}}} \right)^2 . \tag{3.35}$$

This gives a mean absolute error of 5.4% in the same data set.

There are some other C_6 dispersion correction schemes, like the Becke Johnson model. The method we discussed so far all require predetermined input parameters to calculate the dispersion interaction, so they are all empirical.

3.4.3 Long-range Density Functional

We want some method that does not require any predetermined data — everything can just be calculated using the electron density. This means that we need to write out a functional for such non-local electron correlations, known as *long-range density functional*. These have the form

$$E^{\text{c-nl}} = \int \mathrm{d}^3\mathbf{r}_1 \, \mathrm{d}^3\mathbf{r}_2 \, n(\mathbf{r}_1)\phi(\mathbf{r}_1, \mathbf{r}_2)n(\mathbf{r}_2) \, , \tag{3.36}$$

where ϕ is some integration kernel that gives the $\sim 1/r^6$ asymptotic behaviour. As you can see, the functional is non-local because it takes the electron densities at two locations and integrates. However, the formula has a pairwise form (it adds the contribution from $\mathbf{r}_1$ and $\mathbf{r}_2$ infinitesimally via integration) and thus ignores the medium between points $\mathbf{r}_1$ and $\mathbf{r_2}$. This rather severe limitation was removed by Dion et al. in 2004 who proposed a functional form which can be evaluated for overlapping molecules and for arbitrary geometries. Within this approach, the exchange correlation

functional is

$$E^{\text{xc}} = E^{\text{x-GGA}} + E^{\text{c-LDA}} + E^{\text{c-nl}} \,. \tag{3.37}$$

This method is called the *van der Waals density functional*, vdW-DF. It is a very important conceptual development since it adds a description of dispersion directly within a DFT functional and combines correlations of all ranges in a single formula.

3.4.4 Beyond pairwise approaches

So far we have only discussed dispersion as being pairwise additive. That is the interaction energy of two atoms or molecules remains constant no matter what material separates them and all atoms interact on their own. This is a simplification that does not account for higher order collective excitations.

Some methods can extend the pairwise addition to three-body interaction, e.g. Axilrod–Teller–Muto corrections. Such correction is often added to the DFT-D approach.

There are a lot of other methods: some approximate each atom as a quantum harmonic oscillator, some use orbitals to calculate the correlation energy. We shall not explain them further.

4 Molecular Orbitals

We've been trying to avoid solving the Schrödinger equation. It's time to face it.

4.1 The Orbital Approximation

Most of chemistry cares little about the idea of the total wavefunction because it is generally concerned with the making and breaking of chemical bonds whose nature seems inherently localised, and it is an appealing and highly useful picture to envision electrons allocated individually to *orbitals* when thinking about electronic structure.

This idea is supported by the fact that if a Hamiltonian can be separated as one-electron Hamiltonians

$$\hat{H}(\mathbf{r}_1, \mathbf{r}_2) = \hat{H}_1(\mathbf{r}_1) + \hat{H}_2(\mathbf{r}_2)\,, \tag{4.1}$$

and if we can find ψ_1 and ψ_2 solving these one-electron Schrödinger equation

$$\hat{H}_i\psi_i = \epsilon_i\psi_i\,, \tag{4.2}$$

then their product solves the total Hamiltonian

$$\hat{H}\Psi(\mathbf{r}_1, \mathbf{r}_2) = (\epsilon_1 + \epsilon_2)\Psi(\mathbf{r}_1, \mathbf{r}_2) \tag{4.3}$$

for $\Psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)$. Therefore, we want to separate the overall wavefunction as a *Hartree product*

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\ldots\psi_n(\mathbf{r}_n)\,. \tag{4.4}$$

Each one-electron wavefunction solves the Schrödinger equation with the *effective one-electron Hamiltonian*

$$\hat{H}_{\text{eff}}(\mathbf{r}) = -\frac{1}{2}\nabla^2 + V_{\text{eff}}\,, \tag{4.5}$$

where V_{eff} is the average potential that the electron feels from the attraction of the nuclei and average repulsion of the other electrons. These one-electron wavefunctions are known as the *molecular orbitals* and this approximation is known as *orbital approximation*.

The next question is: how to find these molecular orbitals. If we don't know anything about the electron, how do we find the effective potential and then solve the Schrödinger equation? It turns out that a useful way of constructing molecular orbitals is to think of them as being made up from the respective atomic orbitals of elements — this should be a familiar idea from IA and IB. This leads to the idea of *linear combination of atomic orbitals* (LCAO), where each molecular orbital is written as

$$\psi_i = \sum_j c_j^{(i)}\phi_j\,, \tag{4.6}$$

where ϕ_j are the atomic orbitals.

By making further approximations, it is actually possible to solve these coefficients c_{ij} and hence determine the molecular orbitals using pen and paper! This is *Hückel theory*.

4.2 Hückel Theory

This should be familiar from IB Chemistry A. We will do a quick derivation.

The first simplification in Hückel theory is that we are only interested in the delocalised π system in a molecule, so we only consider one orbital on each atom at one time. If the molecule has e.g. two orthogonal π systems, they can be computed separately.

Suppose we have n atomic orbitals from n atoms in total, and we want to form n orthonormal molecular orbitals. By the Rayleigh–Ritz principle of variation, finding the eigenvalues of a Hermitian operator $\hat{H}$ is equivalent to finding the stationary values of the Rayleigh quotient[6]

$$\epsilon_i = \frac{\left\langle \psi_i \middle| \hat{H} \middle| \psi_i \right\rangle}{\langle \psi_i | \psi_i \rangle} \,. \tag{4.7}$$

We expand it in our basis of atomic orbitals

$$\epsilon_i = \frac{\left\langle \sum_s c_s^{(i)} \phi_s \middle| \hat{H} \middle| \sum_t c_t^{(i)} \phi_t \right\rangle}{\left\langle \sum_s c_s^{(i)} \phi_s \middle| \sum_t c_t^{(i)} \phi_t \right\rangle} = \frac{\sum_{s,t} c_s^{(i)*} c_t^{(i)} \left\langle \phi_s \middle| \hat{H} \middle| \phi_t \right\rangle}{\sum_{s,t} c_s^{(i)*} c_t^{(i)} \langle \phi_s | \phi_t \rangle} \,. \tag{4.8}$$

We define the *Hamiltonian matrix* $H_{st} = \left\langle \phi_s \middle| \hat{H} \middle| \phi_t \right\rangle$ and the *overlap matrix* $S_{st} = \langle \phi_s | \phi_t \rangle$. Then we can rewrite this as

$$\sum_{s,t} c_s^{(i)*} c_t^{(i)} H_{st} = \epsilon_i \sum_{s,t} c_s^{(i)*} c_t^{(i)} S_{st} \,. \tag{4.9}$$

For the ϵ_i to be stationary with respect to all variations of coefficients c_{ir}, we differentiate both sides with respect to $c_r^{(i)*}$ and set $\partial \epsilon_i / \partial c_r^{(i)*} = 0$ for all $1 \le r \le n$ to get

$$\sum_{s,t} \delta_{rs} c_t^{(i)} H_{st} = \epsilon_i \sum_{s,t} \delta_{rs} c_t^{(i)} S_{st} \,. \tag{4.10}$$

Hence, we have the *secular equation*

$$\sum_t (c_t^{(i)} H_{rt} - \epsilon_i c_t^{(i)} S_{rt}) = 0 \,, \tag{4.11}$$

or in matrix form

$$(\mathsf{H} - \epsilon_i \mathsf{S}) \mathbf{c}_i = \mathbf{0} \,. \tag{4.12}$$

Hückel theory approximates the matrix elements to take the form

$$H_{st} = \begin{cases} \alpha_s & \text{if } s = t \\ \beta_{st} & \text{if } s, t \text{ are bonded} \\ 0 & \text{otherwise} \end{cases} \tag{4.13}$$

and

$$S_{st} = \delta_{st} = \begin{cases} 1 & \text{if } s = t \\ 0 & \text{otherwise} \,, \end{cases} \tag{4.14}$$

α_s can be thought of as the energy of the atomic orbital on atom s, and the *resonance integral* β_{st} is the interaction of orbital s and t. Note that they are both negative. What Hückel's approximation says is essentially that two atoms only interact if they are neighbouring, and the atomic orbitals on different atoms do not overlap, i.e. our basis of atomic orbitals is orthonormal. Our secular equation then simplifies to

$$(\mathsf{H} - \epsilon_i \mathsf{I}) \mathbf{c}_i = \mathbf{0} \,, \tag{4.15}$$

or

$$\mathsf{H} \mathbf{c}_i = \epsilon_i \mathbf{c}_i \,. \tag{4.16}$$

This is essentially diagonalising H, where the eigenvalues are the orbital energies and the eigenvectors are the molecular orbital coefficients. Since H is Hermitian, we are guaranteed to get n mutually orthonormal eigenvectors (or at least can be made orthonormal if degeneracy arises) with n real eigenvalues.

[6] A proof can be found in IB Mathematical Methods. It is obvious from the variational principles that the eigenvector with the lowest eigenvalue can be constructed in this way by minimising the expectation value of the energy. We are claiming that higher eigenvalues are also the stationary values (not a minimum though) of this expectation value of energy.

4.2.1 Symmetry

Diagonalising large matrices by hand is painful. Luckily, it is possible to reduce the problems using symmetry.

An important fact is that only orbitals with the same symmetry (transforming as the same irreducible representation) can interact. Therefore, we can create a set of normalised symmetry orbitals

$$\theta_\mu = \sum_s \phi_s D_{s\mu} \tag{4.17}$$

such that each θ_μ transforms as a particular irreducible representation. Then the Hamiltonian matrix element becomes

$$\tilde{H}_{\mu\nu} = \langle \theta_\mu | H | \theta_\nu \rangle = \sum_{s,t} D^*_{s\mu} H_{st} D_{t\nu} \,, \tag{4.18}$$

or in matrix form

$$\tilde{\mathsf{H}} = \mathsf{D}^\dagger \mathsf{H} \mathsf{D} \,. \tag{4.19}$$

This can be seen as a basis transformation, where D is the basis transformation matrix, transforming from the basis $\{\phi_s\}$ to the basis $\{\theta_\mu\}$. In this new basis, the Hamiltonian will necessarily be in the block diagonal form

$$\tilde{\mathsf{H}} = \begin{pmatrix} \tilde{\mathsf{H}}^{\Gamma_1} & 0 & \cdots & 0 \\ 0 & \tilde{\mathsf{H}}^{\Gamma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\mathsf{H}}^{\Gamma_m} \end{pmatrix} , \tag{4.20}$$

where each block $\tilde{\mathsf{H}}^{\Gamma_i}$ consists of the matrix elements spanned by the basis transforming as the same irreducible representation Γ_i. The off-block-diagonal elements are necessarily 0 because orbitals with different symmetries do not interact.

Then, instead of diagonalising the huge matrix H, we only need to diagonalise some small matrices $\tilde{\mathsf{H}}^{\Gamma_i}$. The eigenvectors obtained are now the molecular orbitals in the new, transformed basis $\{\theta_\mu\}$. We can easily convert it back to the original atomic orbital basis $\{\phi_s\}$ by

$$\psi_i = \sum_\mu \theta_\mu \tilde{c}^{(i)}_\mu \,. \tag{4.21}$$

4.2.2 Special Systems

Linear Polyenes

The π system of a linear polyene with N carbons is analogous to a particle-in-a-box problem, but the space is discretised. You also get a sine wave solution, and the boundary conditions become $c^{(n)}_0 = c^{(n)}_{N+1} = 0$ which correspond to orbitals "off-the-end" of the molecule. The orbital coefficients are

$$c^{(n)}_s = \sqrt{\frac{2}{N+1}} \sin \frac{n\pi s}{N+1} \,, \tag{4.22}$$

and the orbital energies are

$$\epsilon_n = \alpha + 2\beta \cos \frac{\pi n}{N+1} \tag{4.23}$$

for $1 \leq n \leq N$.

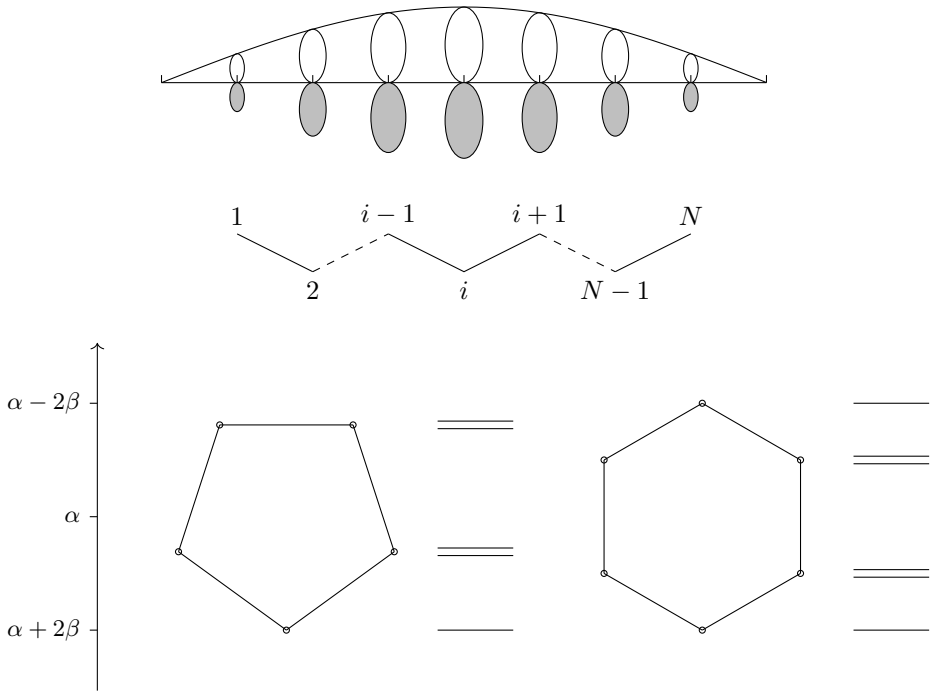


Figure 4.1: Frost circle construction for the energy levels of cyclic polyene.

Cyclic Polyenes

Consider a cyclic polyene of N atoms. By single-valuedness, we must have $c_s^{(n)} = c_{s+N}^{(n)}$. It is analogous to a particle on a ring, and has orbital coefficients

$$c_s^{(n)} = \sqrt{\frac{1}{N}} \mathrm{e}^{2\pi \mathrm{i} n s/N} \tag{4.24}$$

with energies

$$\epsilon_n = \alpha + 2\beta \cos \frac{2\pi n}{N} \,. \tag{4.25}$$

for $-N/2 < n \leq N/2$.

The energy levels can be conveniently viewed as the vertices of the N-gon inscribed in a circle with radius 2β and centered at α. This is known as the *Frost circle* construction.

This expression (4.24) of molecular orbitals is nice and compact, but the values are complex. We can make everything real by doing linear combinations of the degenerate orbitals and get

$$\tilde{c}_s^{(0)} = \sqrt{\frac{1}{N}} \tag{4.26}$$

$$\tilde{c}_s^{(ns)} = \sqrt{\frac{2}{N}} \sin \frac{2\pi n s}{N} \tag{4.27}$$

$$\tilde{c}_s^{(nc)} = \sqrt{\frac{2}{N}} \cos \frac{2\pi n s}{N} \tag{4.28}$$

$$\tilde{c}_s^{(\frac{N}{2})} = \sqrt{\frac{1}{N}} (-1)^s \,, \tag{4.29}$$

where the $N/2$ orbital exists only if N is even.

4.3 Is Hückel Theory Predictive?

4.3.1 Total Energies

By Hückel theory, we have worked out a set of molecular orbitals for a bunch of electrons. The next question is, how to put these electrons into the orbitals, and how to work out the total energy of the system?

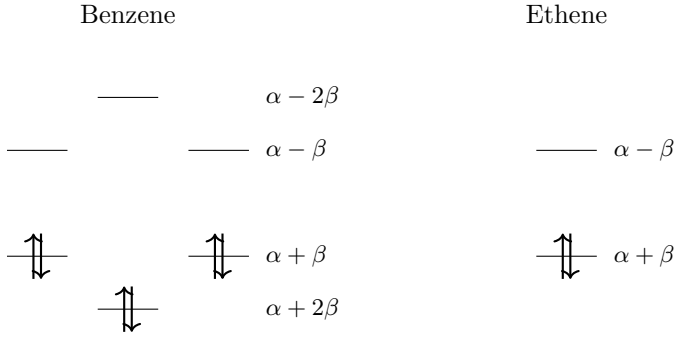
By Pauli's principle, we can only put two electrons with opposite spins in the same orbital. And a system always wants to keep its energy as low as possible. Therefore, we occupy the orbitals from the lowest to the highest until we have put in all of the electrons. To work out the total energy, we simply need to add up all the orbital energies that each electron sits in.[7] We define an *occupation number*, f_n, as the number of electrons in orbital n (it would be 0, 1 or 2), then since we assumed all electrons are independent under our effective potential approximation,

$$E_{\text{Hückel}} = \sum_n f_n \epsilon_n \,. \tag{4.30}$$

4.3.2 Delocalisation Energy

One of the triumphs of Hückel theory was to rationalise the stability of benzene and other aromatics, showing that they are made more stable by a significant delocalisation energy.

Let's take benzene as an example. The molecular orbital diagrams of benzene and a normal double bond (ethene) are shown in the figure.



The delocalisation energy is the energy difference between the total energy of the delocalised molecule and the energy if the π system were localised i.e. each π bond is treated separately. For benzene, this is

$$\begin{aligned} E_{\text{deloc}} &= [2(\alpha + 2\beta) + 4(\alpha + \beta)] - [6(\alpha + \beta)] \\ &= 2\beta \,. \end{aligned} \tag{4.31}$$

We see that Hückel theory predicts the extra stability of benzene through delocalisation i.e. *aromaticity*.

We can calculate the delocalisation energy for more aromatic compounds. The results are shown in the table below.

Two conclusions can be drawn from this. First, the close parallel between the Hückel theory prediction on the variation of the delocalisation energy for the series, as compared to experiment,

[7] The answer is actually not that simple... See later.

Molecule	Theory	Experiment / kJ mol^{-1}	Estimated $\lvert\beta\rvert$ / kJ mol^{-1}
benzene	2β	-155	77.7
naphthalene	3.68β	-315	85.7
anthracene	5.32β	-441	82.7
phenanthrene	5.45β	-462	84.8

which indicates that even such crude calculations are able to reproduce a significant trend. Taking β to be about -80 kJ mol^{-1} leads to a reasonable agreement between predicted and experimental delocalisation energies. Second, phenanthrene (three benzene rings fused in a "zig-zag" fashion) is slightly more stable than anthracene (three benzene rings fused in a straight line). This trend continues for larger systems: the annulation to give "zig-zag" forms are indeed experimentally more stable than the corresponding linear ones.

It is impressive that our very crude theory is able to give some interesting, semi-quantitative, results.

4.3.3 Orbital Energies

Matters look less rosy if we consider a different type of measurement of β, using ionisation energies.

Recall that the first ionisation energy is the minimum energy required to remove an electron from a molecule, and therefore it is reasonable to suppose that the electron removed comes from HOMO. Since, according to Hückel theory, the energy of an orbital can be written in the form

$$\epsilon = \alpha + x\beta\,, \tag{4.32}$$

where x is a suitable coefficient which depends on the molecule, one may suppose that if we take a series of molecules (e.g. benzene, naphthalene, anthracene, *etc.*), for which the x can be calculated, and plot the ionisation energy as a function of x, then the slope of such a (hopefully linear) curve would yield β, whereas the intercept would yield α. It turns out that a least-squares fit yields

$$\text{Experimental Orbital Energy} = -685 + (-239 \pm 16)x \text{ kJ mol}^{-1}\,. \tag{4.33}$$

We obtained a value of $\beta \approx -240$ kJ mol^{-1} — almost three times as large as what we've predicted using total energy!

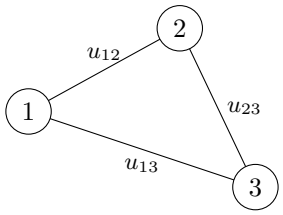
In fact this turns out to be a quite general feature of Hückel theory. Experiments which depend on the energy of single orbitals turn out to yield values of β which are always roughly a factor of two larger than those based on total energies, in which the energies of many orbitals are summed together.

The explanation for this behaviour can be found by considering the manner in which electron-electron interactions are dealt with in the Hamiltonian $\hat{H}$. Recall that the $\hat{H}$ is an effective Hamiltonian in which an electron sees the average field due to all other electrons. In other words, electron 1 sees a field due to the average of electrons 2, 3, *etc.*, and this field, in addition to the field due to the nuclear framework, determines the energy eigenvalue of electron 1. Similarly, electron 2 sees the average field of electron 1, electron 3, *etc.*, and its energy eigenvalue reflects these interactions as well.

Therefore, if we add the energy eigenvalue of electron 1 and electron 2, we have counted twice the average electrostatic interaction between electron 1 and electron 2. This is exactly what is done in our method of calculating the total π-energy: we simply add the energy of all occupied levels. On the other hand, if we are dealing with purely the energy of a single energy level, there is no double

counting. Therefore, it should not be surprising that methods used to estimate β based on the total energies yield values about 1/2 of that from ionisation potential experiments.[8]

Example. We will illustrate our idea with a toy model. Consider a 3-electron system with only repulsion.



The energy of each electron is

$$\begin{aligned} \epsilon_1 &= u_{12} + u_{13} \\ \epsilon_2 &= u_{12} + u_{23} \\ \epsilon_3 &= u_{13} + u_{23} \,. \end{aligned} \tag{4.34}$$

If we directly sum up the individual energy of each electron, we get

$$E = \epsilon_1 + \epsilon_2 + \epsilon_3 = 2u_{12} + 2u_{13} + 2u_{23} \,. \tag{4.35}$$

However, the actual total energy should be

$$E_{\text{total}} = u_{12} + u_{13} + u_{23} \,. \tag{4.36}$$

We have double-counted all the interactions!

4.4 Populations and Bond Orders

While the MOs produce a delocalised picture of bonding in a molecule, it is often helpful to be able to ask more specific questions, such as whether the electrons are concentrated on particular atoms, or to probe the relative strengths of bonding between two neighbouring atoms. With this information, we would, for example, be able to say whether in butadiene the canonical picture of two terminal double bonds is the most appropriate.

For some state $|\phi\rangle$, define the *projection operator*

$$\hat{q} = |\phi\rangle \langle\phi| \,. \tag{4.37}$$

What is the meaning of this? Let's act it on some general state $|\psi\rangle$. We get

$$\hat{q} |\psi\rangle = |\phi\rangle \langle\phi|\psi\rangle \,. \tag{4.38}$$

It returns the state $|\phi\rangle$ we are projecting onto, scaled by the inner product of two states $\langle\phi|\psi\rangle$. As its name tells, it "projects" the state $|\psi\rangle$ on to $|\phi\rangle$.

Let's see how it acts on the molecular orbitals. Let's evaluate the expectation value of a molecular orbital under the projection operator of an atomic orbital $\hat{q}_i = |\phi_i\rangle \langle\phi_i|$. Since under Hückel's

[8] To fix such a problem, one uses theories that treat the effective Hamiltonian more carefully. See Hartree–Fock theory.

approximation, the atomic orbitals $\{\phi_s\}$ are orthonormal, we have

$$\begin{aligned}\langle\psi_n|\hat{q}_i|\psi_n\rangle &= \left\langle \sum_s c_s^{(n)}\phi_s \middle| \phi_i \right\rangle \left\langle \phi_i \middle| \sum_t c_t^{(n)}\phi_t \right\rangle \\ &= \left(\sum_s c_s^{(n)*}\delta_{is}\right)\left(\sum_t c_t^{(n)}\delta_{it}\right) \\ &= c_i^{(n)*}c_i^{(n)} = \left|c_i^{(n)}\right|^2 . \end{aligned} \tag{4.39}$$

The effect is to extract the squared modulus of the coefficient of the given atomic orbital, which can in turn be thought of as the population on that atom from the orbital.

We may therefore work out the total population on atom s, by summing together such quantities

$$\begin{aligned} q_s &= \sum_n f_n \langle\psi_n|\hat{q}_s|\psi_n\rangle \\ &= \sum_n f_n \left|c_s^{(n)}\right|^2 , \end{aligned} \tag{4.40}$$

where recall f_n is the occupation number of orbital n.

Analogously, we may define the operator

$$\hat{p}_{st} = \frac{1}{2}\left(|\phi_s\rangle\langle\phi_t| + |\phi_t\rangle\langle\phi_s|\right) , \tag{4.41}$$

which combines the sign and amplitude information of the orbital coefficients from neighbouring sites, giving a measure of the bonding. The $|\phi_s\rangle\langle\phi_t|$ and its Hermitian conjugate $|\phi_t\rangle\langle\phi_s|$ are averaged to make this operator Hermitian. Summing over orbitals, the bond order between two neighbouring atoms s and t can be calculated by

$$\begin{aligned} p_{st} &= \sum_n f_n \langle\psi_n|\hat{p}_{st}|\psi_n\rangle \\ &= \sum_n f_n \,\mathrm{Re}[c_s^{(n)*}c_t^{(n)}] . \end{aligned} \tag{4.42}$$

The sign and magnitude of p_{st} indicate the bonding, with $+1$ indicating one bonding contribution, and -1 one anti-bonding contribution.

4.5 Alternant Hydrocarbons

A useful application of Hückel theory is to rationalise the differing properties of *alternant* and *non-alternant* hydrocarbons. An alternant hydrocarbon is one in which the carbon atoms can be divided into a "starred" class and a "non-starred" class, such that each starred atom is neighboured by only unstarred atoms, and vice versa. The starred class is generally taken to be the larger of the two classes of atoms. Two examples in figure 4.2 illustrate this.

In Hückel theory, alternant hydrocarbons have the following properties:

(i) The energy levels of alternant hydrocarbons are symmetrically paired around an α, such that if orbital n has energy $\epsilon_n = \alpha + k\beta$ then its pair, orbital $N - n + 1$ has energy $\epsilon_{N-n+1} = \alpha - k\beta$.

(ii) The AO coefficients in any pair of complementary orbitals are related by changing the sign of the coefficients on the unstarred atoms.

(iii) In a neutral alternant, the atomic populations in the ground state are precisely unity.

(iv) In odd-alternants, there is a non-bonding orbital, with nonzero coefficients only on the starred atoms, and the sum of coefficients on starred atoms which surround an unstarred site is zero.

Alternant Non-alternant

Figure 4.2: Examples of alternant and non-alternant hydrocarbons.

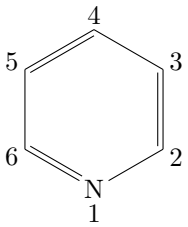


Figure 4.3: Numbering of pyridine.

4.6 Heteroconjugated Systems and Perturbation Theory

Suppose we want to use Hückel theory to solve a heteroconjugated system (a conjugated system involving more than one type of atoms) e.g. pyridine or pyrrole. Since nitrogen has a different energy than carbon, and a C-C π bond would have a different strength than a C-N π bond, we should set a different α and β value for them.

4.6.1 Non-degenerate Perturbation Theory

Suppose we want to solve for pyridine. We use the numbering shown in the figure. Then the π system has the Hamiltonian matrix

$$\mathsf{H} = \begin{pmatrix} \alpha_{\mathrm{N}} & \beta_{\mathrm{NC}} & 0 & 0 & 0 & \beta_{\mathrm{NC}} \\ \beta_{\mathrm{NC}} & \alpha & \beta & 0 & 0 & 0 \\ 0 & \beta & \alpha & \beta & 0 & 0 \\ 0 & 0 & \beta & \alpha & \beta & 0 \\ 0 & 0 & 0 & \beta & \alpha & \beta \\ \beta_{\mathrm{NC}} & 0 & 0 & 0 & \beta & \alpha \end{pmatrix}. \tag{4.43}$$

Let's set $\alpha_{\mathrm{N}} = \alpha + 0.5\beta$ and $\beta_{\mathrm{NC}} = 0.8\beta$.

Of course we can solve this Hamiltonian matrix directly, but isn't this system (and the Hamiltonian of course) too similar to benzene, which we already know the solution? It might be helpful to think of this Hamiltonian as the Hamiltonian for benzene, plus a small modification, which we call a *perturbation*

$$\begin{aligned} \mathsf{H} &= \mathsf{H}_0 + \Delta\mathsf{H} \\ &= \begin{pmatrix} \alpha & \beta & 0 & 0 & 0 & \beta \\ \beta & \alpha & \beta & 0 & 0 & 0 \\ 0 & \beta & \alpha & \beta & 0 & 0 \\ 0 & 0 & \beta & \alpha & \beta & 0 \\ 0 & 0 & 0 & \beta & \alpha & \beta \\ \beta & 0 & 0 & 0 & \beta & \alpha \end{pmatrix} + \begin{pmatrix} 0.5\beta & -0.2\beta & 0 & 0 & 0 & -0.2\beta \\ -0.2\beta & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -0.2\beta & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \end{aligned} \tag{4.44}$$

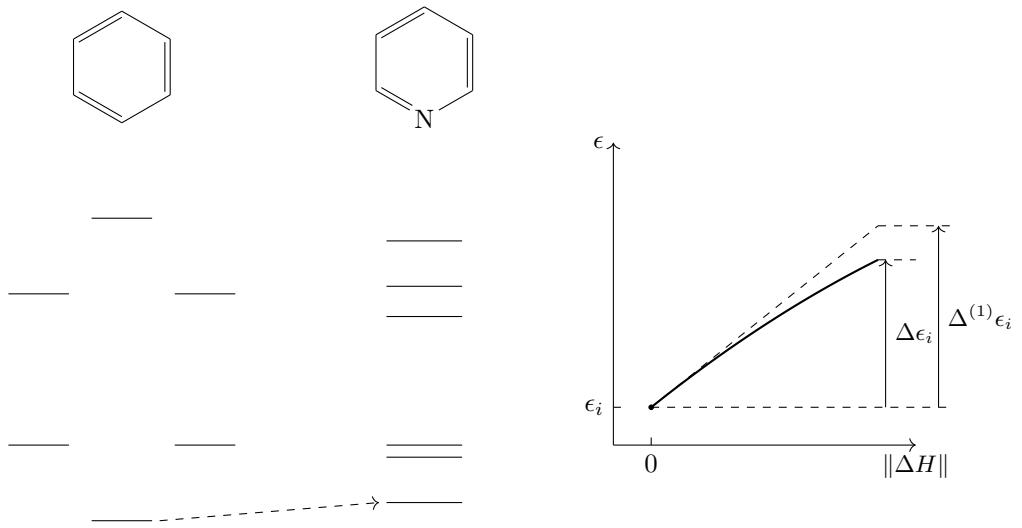


Figure 4.4: We expect that, by changing the Hamiltonian slightly, the molecular orbital energies will also shift only slightly. This shift should be smoothly dependent on the "size" of the perturbation. We can easily work out this change in first order.

where H_0 is the Hamiltonian of our reference system, benzene in this case, and ΔH is our perturbation. We hope that, by modifying the Hamiltonian only slightly, the molecular orbitals only change slightly as well. Then, it might be possible to express, even if only approximately, the energy levels and MOs of the perturbed Hamiltonian (e.g. pyridine) in terms of the unperturbed one (e.g. benzene).

This is the topic of *perturbation theory*. It will be covered in great depth in course *C7: Further Quantum Mechanics* (and in even greater depth in Mathematical Tripos Part II *Principles of Quantum Mechanics*). We will show some simple results here.

We will first work on the case where the molecular orbital energy levels are non-degenerate.

Proposition 4.1. Let $\hat{H}$ be a perturbed Hamiltonian that can be written as

$$\hat{H} = \hat{H}_0 + \Delta\hat{H}\,. \tag{4.45}$$

Suppose we know that the unperturbed Hamiltonian $\hat{H}_0$ has a normalised non-degenerate eigenstate $|\psi_i\rangle$ and corresponding eigenvalue ϵ_i. Then, after the perturbation, the corresponding eigenvalue is changed by

$$\Delta^{(1)}\epsilon_i = \left\langle \psi_i \middle| \Delta\hat{H} \middle| \psi_i \right\rangle \tag{4.46}$$

to first order.

Proof. We will show this result loosely. For a rigorous proof, see C7 or Principles of Quantum Mechanics.

Suppose when we change $\hat{H}_0$ to $\hat{H}_0 + \Delta\hat{H}$, ψ_i has changed to $\psi_i + \Delta\psi_i$, then the energy has changed to

$$\epsilon_i + \Delta\epsilon_i = \frac{\left\langle \psi_i + \Delta\psi_i \middle| \hat{H}_0 + \Delta\hat{H} \middle| \psi_i + \Delta\psi_i \right\rangle}{\langle \psi_i + \Delta\psi_i | \psi_i + \Delta\psi_i \rangle}\,. \tag{4.47}$$

Note that to change ψ_i meaningfully, $\Delta\psi_i$ must be orthogonal to ψ_i since scaling not does actually change the state, so $\langle \Delta\psi_i | \psi_i \rangle = 0$. Expand to the first order in Δ, we get

$$\begin{aligned}\epsilon_i + \Delta^{(1)}\epsilon_i &= \frac{\left\langle \psi_i \middle| \hat{H}_0 \middle| \psi_i \right\rangle + \left\langle \Delta\psi_i \middle| \hat{H}_0 \middle| \psi_i \right\rangle + \left\langle \psi_i \middle| \Delta\hat{H} \middle| \psi_i \right\rangle + \left\langle \psi_i \middle| \hat{H}_0 \middle| \Delta\psi_i \right\rangle}{\langle \psi_i | \psi_i \rangle + \langle \psi_i | \Delta\psi_i \rangle + \langle \Delta\psi_i | \psi_i \rangle} \\ &= \epsilon_i + \left\langle \Delta\psi_i \middle| \hat{H}_0 \middle| \psi_i \right\rangle + \left\langle \psi_i \middle| \Delta\hat{H} \middle| \psi_i \right\rangle + \left\langle \psi_i \middle| \hat{H}_0 \middle| \Delta\psi_i \right\rangle\,. \end{aligned} \tag{4.48}$$

Since $H_0\psi_i = \epsilon_i\psi_i$, we have

$$\left\langle \psi_i \middle| \hat{H}_0 \middle| \Delta\psi_i \right\rangle^* = \left\langle \Delta\psi_i \middle| \hat{H}_0 \middle| \psi_i \right\rangle = \epsilon_i \left\langle \Delta\psi_i | \psi_i \right\rangle = 0\,. \tag{4.49}$$

Hence,

$$\Delta^{(1)}\epsilon_i = \left\langle \psi_i \middle| \Delta\hat{H} \middle| \psi_i \right\rangle \tag{4.50}$$

as claimed. □

4.7 Degenerate Perturbation Theory

When applying this to degenerate orbitals you must take a little more care to recover a well-defined solution. Recall that in the reference system, any orthogonal linear combination of two degenerate orbitals would still have the same energy and can be regarded as equally valid. Once the perturbation has been applied however, the different linear combinations may produce different energy changes. In such cases, the perturbation is said to break the degeneracy. To recover a well-defined energy change, we resort to the variational principle again, asking: which combination of the degenerate states produces the lowest energy under the perturbation? This is none other than our old friend the secular problem, and we must therefore diagonalize a small matrix consisting of just the degenerate states to solve it.

If there are some degenerate eigenstates of $\hat{H}_0$, denoted ψ_i, then we may write out the perturbation $\Delta\hat{H}$ in the basis of these two orbitals

$$\Delta H_{ij} = \left\langle \psi_i \middle| \Delta\hat{H} \middle| \psi_j \right\rangle\,. \tag{4.51}$$

To determine the appropriate energies, we then solve

$$(\mathsf{\Delta H} - \Delta\epsilon \mathsf{I})\mathbf{c} = \mathbf{0}\,. \tag{4.52}$$

The eigenvalues, $\Delta\epsilon$ are the appropriate perturbed energies, and the eigenvectors give the appropriate linear combination of ψ_i which has that perturbed energy.

Example. Let's try this on the occupied orbitals of pyridine. We use the numbering as before, thinking it as a perturbation on the reference system benzene, whose orbital coefficients and energies are well-known

$$\mathbf{c}^{(0)} = \frac{1}{\sqrt{6}}\begin{pmatrix}1 & 1 & 1 & 1 & 1 & 1\end{pmatrix}^{\mathrm{T}} \qquad E^{(0)} = \alpha + 2\beta \tag{4.53}$$

$$\mathbf{c}^{(1c)} = \frac{1}{\sqrt{12}}\begin{pmatrix}2 & 1 & -1 & -2 & -1 & 1\end{pmatrix}^{\mathrm{T}} \qquad E^{(1c)} = \alpha + \beta \tag{4.54}$$

$$\mathbf{c}^{(1s)} = \frac{1}{2}\begin{pmatrix}0 & 1 & 1 & 0 & -1 & -1\end{pmatrix}^{\mathrm{T}} \qquad E^{(1s)} = \alpha + \beta\,. \tag{4.55}$$

The perturbation is

$$\mathsf{\Delta H} = \frac{\beta}{10}\begin{pmatrix} 5 & -2 & 0 & 0 & 0 & -2 \\ -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}\,. \tag{4.56}$$

The state $|\psi_0\rangle$ is non-degenerate, so

$$\begin{aligned}\Delta^{(1)}\epsilon_0 &= \left\langle \psi_0 \middle| \Delta\hat{H} \middle| \psi_0 \right\rangle \\ &= \mathbf{c}^{(0)\dagger}\Delta\mathsf{H}\mathbf{c}^{(0)} \\ &= -\frac{\beta}{20}\,. \end{aligned} \tag{4.57}$$

The next two orbitals are degenerate, so we need to use degenerate perturbation theory. We evaluate

$$\left\langle \psi_{1c} \middle| \Delta\hat{H} \middle| \psi_{1c} \right\rangle = \frac{\beta}{30} \tag{4.58}$$

$$\left\langle \psi_{1s} \middle| \Delta\hat{H} \middle| \psi_{1s} \right\rangle = 0 \tag{4.59}$$

$$\left\langle \psi_{1c} \middle| \Delta\hat{H} \middle| \psi_{1s} \right\rangle = 0\,, \tag{4.60}$$

so the matrix

$$\Delta\mathsf{H} = \begin{pmatrix} +\frac{\beta}{30} & 0 \\ 0 & 0 \end{pmatrix} \tag{4.61}$$

is already diagonalised. The variations in energies are

$$\Delta^{(1)}\epsilon = +\frac{\beta}{30}, 0\,. \tag{4.62}$$

It turns out that the matrix $\Delta\hat{\mathsf{H}}$ is automatically diagonalised for us because we have chosen a good basis. If we instead rotate the unperturbed orbitals to

$$\mathbf{c}^{(1c)\prime} = \frac{1}{\sqrt{12}}\begin{pmatrix} 1 & -1 & -2 & -1 & 1 & 2 \end{pmatrix}^{\mathrm{T}} \tag{4.63}$$

$$\mathbf{c}^{(1s)\prime} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & -1 & -1 & 0 \end{pmatrix}^{\mathrm{T}} \tag{4.64}$$

using the same perturbation matrix, then we would get

$$\left\langle \psi'_{1c} \middle| \Delta\hat{H} \middle| \psi'_{1c} \right\rangle = \frac{\beta}{120} \tag{4.65}$$

$$\left\langle \psi'_{1s} \middle| \Delta\hat{H} \middle| \psi'_{1s} \right\rangle = \frac{\beta}{40} \tag{4.66}$$

$$\left\langle \psi'_{1c} \middle| \Delta\hat{H} \middle| \psi'_{1s} \right\rangle = \frac{\beta}{20\sqrt{12}}\,, \tag{4.67}$$

giving

$$\Delta\mathsf{H}' = \begin{pmatrix} \frac{\beta}{120} & \frac{\beta}{20\sqrt{12}} \\ \frac{\beta}{20\sqrt{12}} & \frac{\beta}{40} \end{pmatrix}\,. \tag{4.68}$$

Diagonalising these will still give us the correct eigenvalues $\Delta^{(1)}\epsilon = +\frac{\beta}{30}, 0$.

A lesson is to put nodes on special atoms.

5 Many-electron Wavefunction

Throughout the last chapter, we assumed the wavefunction to take the form of a *Hartree product*

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \psi_1(\mathbf{r}_1) \ldots \psi_n(\mathbf{r}_n)\,. \tag{5.1}$$

This makes the problem simple and easy to solve, but it has a severe problem — a Hartree product does not obey Pauli principle. To take a step further, we must consider spins.

5.1 Spin

Recall that electron has spin $s = \frac{1}{2}$, so it has two possible spin states $s_z = +\frac{1}{2}, -\frac{1}{2}$. The two $s_z = +\frac{1}{2}, -\frac{1}{2}$ spin eigenfunctions are represented by $\alpha(\omega)$ and $\beta(\omega)$ respectively, where ω is a fictitious coordinate to keep track of which electron a spin belongs to (and is a simplification of some quantum field theory formulation). We have the spin magnitude squared operator $\hat{s}^2(\omega)$, and the operator for the z component of spin $\hat{s}_z(\omega)$, which acts on the spin eigenstate $\sigma(\omega)$ by

$$\hat{s}^2\sigma(\omega) = \hbar^2 s(s+1)\sigma(\omega) \tag{5.2}$$

$$\hat{s}_z\sigma(\omega) = \hbar s_z \sigma(\omega)\,. \tag{5.3}$$

A many-electron wavefunction also has spin, and corresponding spin operators. The total z component of spin operator takes the simple form

$$\hat{S}_z(\omega_1, \omega_2, \ldots, \omega_n) = \sum_{i=1}^{n} \hat{s}_z(\omega_i) \tag{5.4}$$

because the component of a vector is a scalar and we can add them simply. There is a corresponding total spin square magnitude operator, $\hat{S}^2$, but its form is not a simple sum and not trivial to apply, and will not be needed in this course.[9] For a total spin eigenfunction $\Sigma(\omega_1, \ldots, \omega_n)$, the eigenvalue equations are

$$\hat{S}_z\Sigma = \hbar M_S \Sigma \tag{5.6}$$

$$\hat{S}^2\Sigma = \hbar^2 S(S+1)\Sigma\,, \tag{5.7}$$

where for a given S, $M_S = -S, -S+1, \ldots, S-1, S$.

When writing a complete wavefunction, we must also consider the spin part of the wavefunction in addition to the spatial part. Therefore, a complete wavefunction should be

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_n; \omega_1, \ldots, \omega_n)\,. \tag{5.8}$$

The most important feature for electrons is that they are fermions. An important rule for fermions is that the wavefunction must be antisymmetric with respect to exchange of electrons, so that

$$\Psi(\ldots, \mathbf{r}_i, \ldots, \mathbf{r}_j, \ldots, \omega_i, \ldots, \omega_j, \ldots) = -\Psi(\ldots, \mathbf{r}_j, \ldots, \mathbf{r}_i, \ldots, \omega_j, \ldots, \omega_i, \ldots)\,. \tag{5.9}$$

Notice that by exchanging the electrons, we exchange both the spatial coordinates and the spin coordinates.

As one can easily see, a simple Hartree product does not obey this condition — we need to find some ways to construct antisymmetric total wavefunction from one-electron wavefunctions (orbitals).

[9] The most convenient form is written in terms of ladder operators, as

$$\hat{S}^2 = \hat{S}_+\hat{S}_- + (\hat{S}_z)^2 - \hbar\hat{S}_z\,. \tag{5.5}$$

As you have seen in IB, one strategy of constructing an antisymmetric wavefunction is to construct symmetric and antisymmetric spatial and spin wavefunctions separately, then combine the symmetric spatial wavefunction with the antisymmetric spin wavefunction and vice versa. For two electrons with different spatial wavefunctions, the possible wavefunctions are

$$\Psi_{\text{triplet}} = \frac{1}{\sqrt{2}}[\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) - \psi_2(\mathbf{r}_1)\psi_1(\mathbf{r}_2)] \times \begin{cases} \alpha(\omega_1)\alpha(\omega_2) \\ \frac{1}{\sqrt{2}}[\alpha(\omega_1)\beta(\omega_2) + \beta(\omega_1)\alpha(\omega_2)] \\ \beta(\omega_1)\beta(\omega_2) \end{cases} \tag{5.10}$$

$$\Psi_{\text{singlet}} = \frac{1}{\sqrt{2}}[\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) + \psi_2(\mathbf{r}_1)\psi_1(\mathbf{r}_2)] \times \frac{1}{\sqrt{2}}[\alpha(\omega_1)\beta(\omega_2) - \beta(\omega_1)\alpha(\omega_2)]\,. \tag{5.11}$$

The triplet state has $S = 1$ and $M_S = -1, 0, 1$. The singlet state has $S = 0$ and $M_S = 0$.

5.2 Slater Determinants

The above method worked well for systems with two electrons, but it becomes increasingly painful for more and more electrons. We hope to find a more general and convenient method of constructing antisymmetric wavefunctions. It turns out that in mathematics, there is an object that is always antisymmetric — that is a determinant.

Because the electron exchanges in Pauli principle affect both the spatial and spin coordinates of an electron, it is convenient to group them together as a single coordinate $\mathbf{x} = (\mathbf{r}, \omega)$. We can also combine the spatial and spin orbitals together to form a *spin-orbital*

$$\chi(\mathbf{x}) \equiv \psi^\sigma(\mathbf{x}) = \psi(\mathbf{r})\sigma(\omega)\,. \tag{5.12}$$

By the property that a determinant changes the sign when we exchange any two columns or rows, we are guaranteed that

$$\Psi = \frac{1}{\sqrt{2}}\begin{pmatrix} \chi_a(\mathbf{x}_1) & \chi_b(\mathbf{x}_1) \\ \chi_a(\mathbf{x}_2) & \chi_b(\mathbf{x}_2) \end{pmatrix} = \frac{1}{\sqrt{2}}[\chi_a(\mathbf{x}_1)\chi_b(\mathbf{x}_2) - \chi_b(\mathbf{x}_1)\chi_a(\mathbf{x}_2)] \tag{5.13}$$

is antisymmetric with respect to the exchange of coordinates $\mathbf{x}_1$ and $\mathbf{x}_2$, since exchanging the coordinates is equivalent to exchanging the two rows. Additionally, if the spin-orbitals are normalised, this determinant is also normalised. This can be easily generalised. For n spin-orbitals, a *Slater determinant* is defined as

$$\Psi = \frac{1}{\sqrt{n!}}\begin{vmatrix} \chi_a(\mathbf{x}_1) & \chi_b(\mathbf{x}_1) & \cdots & \chi_r(\mathbf{x}_1) \\ \chi_a(\mathbf{x}_2) & \chi_b(\mathbf{x}_2) & \cdots & \chi_r(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_a(\mathbf{x}_n) & \chi_b(\mathbf{x}_n) & \cdots & \chi_r(\mathbf{x}_n) \end{vmatrix}. \tag{5.14}$$

This is guaranteed to be antisymmetric, and since a $n \times n$ determinant can be expanded into $n!$ terms, it is normalised given that the spin orbitals are orthonormal. Sometimes, we denote it by the shorthand

$$\Psi = \frac{1}{\sqrt{n!}}\left|\chi_a\chi_b\cdots\chi_r\right|\,. \tag{5.15}$$

5.2.1 Two-Electron System

Finally we are going to solve the Schrödinger equation, starting from a simple two-electron system. Let's observe the form of the electronic Hamiltonian (in atomic units)[10]

$$\hat{H}(\mathbf{r}_1, \mathbf{r}_2) = -\frac{1}{2}\nabla^2_{\mathbf{r}_1} - \frac{1}{2}\nabla^2_{\mathbf{r}_2} + V(\mathbf{r}_1) + V(\mathbf{r}_2) + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} + V^{\text{n-n}}\,, \tag{5.16}$$

where V is the potential energy from electron-nucleus interaction. This splits very neatly into one-electron term for each electron

$$\hat{h}(\mathbf{r}) = -\frac{1}{2}\nabla^2_{\mathbf{r}} + V(\mathbf{r})\,, \tag{5.17}$$

a two-electron interaction

$$u(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = \frac{1}{r_{12}}\,, \tag{5.18}$$

and a constant term $V^{\text{n-n}}$. Therefore we can write

$$\hat{H}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i=1}^{2} \hat{h}(\mathbf{r}_i) + \hat{u}(\mathbf{r}_1, \mathbf{r}_2) + V^{\text{n-n}}\,. \tag{5.19}$$

More generally, for a system with n electrons, we include the one-electron Hamiltonian for all electrons, and repulsion between all pairs of electrons, so

$$\hat{H}(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \sum_{i=1}^{n} \hat{h}_i + \sum_{i<j}^{n} \frac{1}{r_{ij}} + V^{\text{n-n}}\,. \tag{5.20}$$

Now let's calculate the energies for some two-electron systems. First, consider a system where the electrons have the same spatial parts but different spins, with spin orbitals $\psi_a\alpha$ and $\psi_a\beta$. We assume the wavefunctions to be in the form of a Slater determinant

$$\begin{aligned}\Psi &= \begin{vmatrix} \psi_a(1)\alpha(1) & \psi_a(1)\beta(1) \\ \psi_a(2)\alpha(2) & \psi_a(2)\beta(2) \end{vmatrix} \\ &= \frac{1}{\sqrt{2}}\left[\psi_a(1)\alpha(1)\psi_a(2)\beta(2) - \psi_a(1)\beta(1)\psi_a(2)\alpha(2)\right]\,. \end{aligned} \tag{5.21}$$

Note we have abbreviated the coordinates $\mathbf{x}_i, \mathbf{r_i}, \omega_i$ as i — which one it refers to should be clear. An example of a molecule of this type is H_2 in ground state, or a Helium atom.

All we need to do then is to expand the energy expression

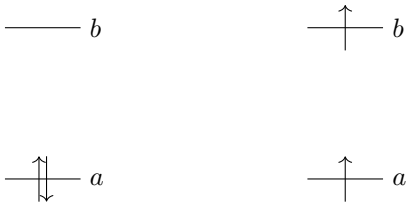
$$E = \left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle = \left\langle \Psi \middle| \hat{h}(1) + \hat{h}(2) + \frac{1}{r_{12}} + V^{\text{n-n}} \middle| \Psi \right\rangle\,. \tag{5.22}$$

We will do this by brute force — general expression of the energy exists, but the derivation requires some non-trivial mathematics. This will be done in *B10: Electronic Structure*. Before we proceed, let's introduce some shorthand notations. We will denote matrix elements of $\hat{h}$ as $\left\langle \psi_i \middle| \hat{h} \middle| \psi_j \right\rangle = h_{ij}$. We will also encounter two-electron integrals, which we denote as[11]

$$\left\langle \psi_a(1)\psi_b(2) \middle| \frac{1}{r_{12}} \middle| \psi_c(1)\psi_d(2) \right\rangle = \langle ab|cd\rangle\,. \tag{5.23}$$

[10] Again, you can either include $V^{\text{n-n}}$ to the electronic Hamiltonian or not — it doesn't matter. It's just a constant term.

[11] The notation we use here is known as the *physicists' notation*. There is also an alternative called *chemists' notation*. We will not show it here to not confuse you. If you are interested, see B10.



Case 1
$S = 0$, $M_S = 0$

Case 2
$S = 1$, $M_S = 1$

Finally, it's time to expand the energy expression.

$$\left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle = \frac{1}{2} \Big\langle \psi_a(1)\alpha(1)\psi_a(2)\beta(2) - \psi_a(1)\beta(1)\psi_a(2)\alpha(2) \Big| \hat{h}(1) + \hat{h}(2) + \frac{1}{r_{12}} + V^{\text{n-n}} \Big| \psi_a(1)\alpha(1)\psi_a(2)\beta(2) - \psi_a(1)\beta(1)\psi_a(2)\alpha(2) \Big\rangle \tag{5.24}$$

Since $\langle\alpha|\beta\rangle = 0$, while $\langle\alpha|\alpha\rangle = \langle\beta|\beta\rangle = 1$, the cross terms all vanish and the expression simplifies to

$$\begin{aligned} \left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle = \quad & \underbrace{\frac{1}{2}\left[\left\langle \psi_a \middle| \hat{h} \middle| \psi_a \right\rangle_1 \langle \psi_a | \psi_a \rangle_2 + \left\langle \psi_a \middle| \hat{h} \middle| \psi_a \right\rangle_1 \langle \psi_a | \psi_a \rangle_2 \right]}_{\text{from } \hat{h}(1)} \\ & + \underbrace{\frac{1}{2}\left[\langle \psi_a | \psi_a \rangle_1 \left\langle \psi_a \middle| \hat{h} \middle| \psi_a \right\rangle_2 + \langle \psi_a | \psi_a \rangle_1 \left\langle \psi_a \middle| \hat{h} \middle| \psi_a \right\rangle_2 \right]}_{\text{from } \hat{h}(2)} \\ & + \frac{1}{2}\left[\left\langle \psi_a(1)\psi_a(2) \middle| \frac{1}{r_{12}} \middle| \psi_a(1)\psi_a(2) \right\rangle_{1,2} \times 2 \right] \\ & + \frac{1}{2} \times 2 \times V^{\text{n-n}} \\ = \quad & 2h_{aa} + \langle aa|aa\rangle + V^{\text{n-n}}, \end{aligned} \tag{5.25}$$

where the subscripts of the bra-kets indicate the coordinate that is integrated over. The difference between this and the non-interacting electron picture is just the Coulomb interaction of the electrons $\langle aa|aa\rangle$. This is known as the *Coulomb integral*, often denoted J, which you have met in IB.

We now consider an alternative case. This time the two electrons are in different spatial orbitals, but their spins are aligned. An example would be the triplet excited state of an He atom. We will denote their spin-orbitals as $\psi_a\alpha$ and $\psi_b\alpha$, where the spatial orbitals ψ_a and ψ_b are orthogonal, so the Slater determinant is

$$\begin{aligned} \Psi &= \begin{vmatrix} \psi_a(1)\alpha(1) & \psi_b(1)\alpha(1) \\ \psi_a(2)\alpha(2) & \psi_b(2)\alpha(2) \end{vmatrix} \\ &= \frac{1}{\sqrt{2}} \left[\psi_a(1)\alpha(1)\psi_b(2)\alpha(2) - \psi_b(1)\alpha(1)\psi_a(2)\alpha(2) \right] . \end{aligned} \tag{5.26}$$

Again, we expand the energy expression. This time some terms will be eliminated due to the

orthogonality of spatial wavefunctions $\langle\psi_a|\psi_b\rangle = 0$. We get

$$\begin{aligned}
\left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle &= \frac{1}{2}\Big\langle \psi_a(1)\alpha(1)\psi_b(2)\alpha(2) - \psi_b(1)\alpha(1)\psi_a(2)\alpha(2) \Big| \\
&\qquad \hat{h}(1) + \hat{h}(2) + \frac{1}{r_{12}} + V^{\text{n-n}} \Big| \psi_a(1)\alpha(1)\psi_b(2)\alpha(2) - \psi_b(1)\alpha(1)\psi_a(2)\alpha(2) \Big\rangle \\
&= \underbrace{\frac{1}{2}\Big[\Big\langle \psi_a \Big| \hat{h} \Big| \psi_a \Big\rangle_1 \langle\psi_b|\psi_b\rangle_2 - \Big\langle \psi_a \Big| \hat{h} \Big| \psi_b \Big\rangle_1 \cancelto{0}{\langle\psi_b|\psi_a\rangle_2}}_{} \\
&\qquad \underbrace{- \Big\langle \psi_b \Big| \hat{h} \Big| \psi_a \Big\rangle_1 \cancelto{0}{\langle\psi_a|\psi_b\rangle_2} + \Big\langle \psi_b \Big| \hat{h} \Big| \psi_b \Big\rangle_1 \langle\psi_a|\psi_a\rangle_2 \Big]}_{\text{from } \hat{h}(1)} \\
&\qquad + \text{same for } \hat{h}(2) \\
&\qquad + \frac{1}{2}\left(\langle ab|ab\rangle - \langle ab|ba\rangle - \langle ba|ab\rangle + \langle ba|ba\rangle\right) + V^{\text{n-n}} \,. \qquad (5.27)
\end{aligned}$$

Note that by exchanging the label of integration,

$$\begin{aligned}
\langle ba|ab\rangle &= \left\langle b(1)a(2) \middle| \frac{1}{r_{12}} \middle| a(1)b(2) \right\rangle_{1,2} \\
&= \left\langle a(1)b(2) \middle| \frac{1}{r_{12}} \middle| b(1)a(2) \right\rangle_{12} = \langle ab|ba\rangle \,, \qquad (5.28)
\end{aligned}$$

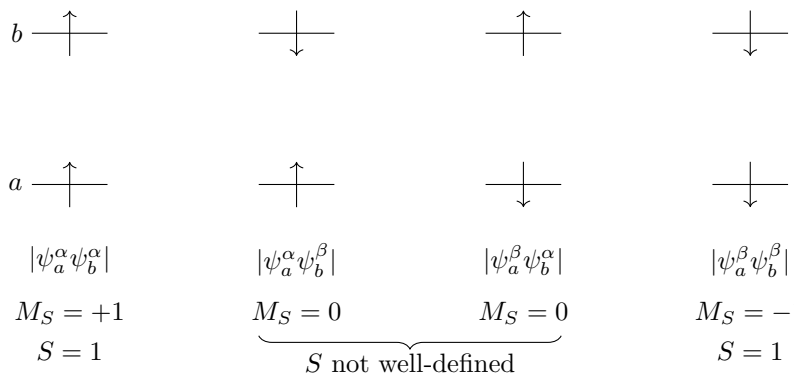
and similarly $\langle ba|ba\rangle = \langle ab|ab\rangle$. Therefore,

$$E = h_{aa} + h_{bb} + \langle ab|ab\rangle - \langle ab|ba\rangle + V^{\text{n-n}} \,. \qquad (5.29)$$

The electron-electron interaction is now split into two parts, the first is a classical Coulomb integral, denoted J, and the second the quantum mechanical *exchange integral* $-K$ you have seen last year.

Also note that this expression has no terms like $\langle aa|aa\rangle$ or $\langle bb|bb\rangle$, indicating that there is no *self-interaction error*. This is because we have dealt the exchange interactions exactly.

It is important to note that while Slater determinants always obey the antisymmetry of exchange of electrons, they do not necessarily obey the other symmetries of the Hamiltonian (and thus do not always have "good quantum numbers"). An example of this is again to consider the excited state of Helium, where the two spatial orbitals are ψ_a and ψ_b. There are four possible Slater determinants, with either electron being spin up or spin down, as shown in the figure below.



The two states with $M_S = +1$ and $M_S = -1$ clearly must correspond to $S = 1$, but the S values of the two $M_S = 0$ states are not well defined. This is because these two Slater determinants cannot be factorised as a product of a spatial part and a spin part, so we can't assign the S value based on

the symmetry of the spin part wavefunction (recall that the triplet $S = 1$ state has a symmetric spin wavefunction and the singlet $S = 0$ state has an antisymmetric wavefunction). On the other hand, e.g. the $M_S = +1$ wavefunction can

$$\begin{aligned} |\psi_a^\alpha \psi_b^\alpha| &= \frac{1}{\sqrt{2}}[\psi_a(1)\alpha(1)\psi_b(2)\alpha(2) - \psi_b(1)\alpha(1)\psi_a(2)\alpha(2)] \\ &= \underbrace{\frac{1}{\sqrt{2}}[\psi_a(1)\psi_b(2) - \psi_b(1)\psi_a(2)]}_{\text{antisymmetric spatial wavefunction}} \times \underbrace{\alpha(1)\alpha(2)}_{\text{symmetric spin wavefunction}} . \end{aligned} \tag{5.30}$$

In fact, you will show in question sheet that the two $M_S = 0$ determinants above are linear combinations of the triplet $M_S = 0$ state and the singlet $M_S = 0$ states, so they are clearly not eigenfunctions of $\hat{S}^2$.

5.2.2 More than Two Electrons

Our above calculations for systems with two electrons are already messy enough, and it will get even messier for more electrons. See *B10: Electronic Structure* for details. However, the energy expression is in a very simple form

$$E = \sum_i h_{ii} + \sum_{i<j} \langle ij|ij\rangle - \sum_{\substack{i<j \\ \text{same spin}}} \langle ij|ji\rangle + V^{\text{n-n}} . \tag{5.31}$$

This should be reasonably physically intuitive — we first sum over the one-electron energies, then each pair of electrons gives one Coulomb interaction, and each pair of electrons with the same spin gives one exchange integral.

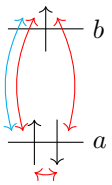


Figure 5.1: A three-electron system. The red arrows are the Coulomb interactions and the blue arrows are the exchange interactions.

For example, the three electron system in the figure would have energy

$$E = 2h_{aa} + h_{bb} + \langle aa|aa\rangle + 2\langle ab|ab\rangle - \langle ab|ba\rangle . \tag{5.32}$$

5.3 Hartree–Fock Theory

In the last section, we found an easy way of generating antisymmetric wavefunctions from orbitals using Slater determinants, and we have calculated the energies of these states. However, it seems that we have skipped the question of how to find these orbitals.

Although Hückel theory is a nice and easy way to do so, it actually misses a lot of things, such as the self-consistent interaction with the other electrons to get relative orbital energies correct. Instead, we will develop a method called the *Hartree–Fock theory* (HF).

Unlike Hückel theory, which finds the stationary values of the orbital energies, the Hartree–Fock theory assumes the total wavefunction to be a single Slater determinant and it aims to minimise the total energy.

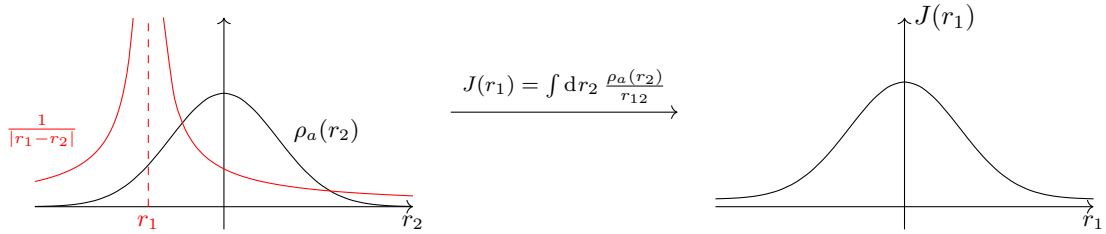


Figure 5.2: An illustration of forming the Coulomb operator. At first glance it may seem that this integral diverges — it does diverge in 1 dimension, but it nicely converges in 3 dimensions.

Let's use ground state helium atom as an example, where the wavefunction is

$$\Psi = \frac{1}{\sqrt{2}} \left| \psi_a^\alpha \psi_a^\beta \right| \tag{5.33}$$

and the energy is

$$E = 2h_{aa} + \langle aa|aa\rangle \, . \tag{5.34}$$

We want to find the ψ_a that minimises the total energy E.

To express the wavefunction, we need some basis function. A sensible basis would be the hydrogenic orbitals of He^+, so we write

$$\psi_a = \sum_{i=1}^{n} c_i^{(a)} \phi_i \, , \tag{5.35}$$

where $\{\phi_i\}$ are the basis functions. Our problem now reduces to finding the coefficients $c_i^{(a)}$ making the energy stationary. Let's take a closer look at the energy expression and write it more explicitly:

$$E = 2 \left\langle \psi_a \middle| \hat{h} \middle| \psi_a \right\rangle + \left\langle \psi_a(1)\psi_a(2) \middle| \frac{1}{r_{12}} \middle| \psi_a(1)\psi_a(2) \right\rangle \, . \tag{5.36}$$

The one-electron term looks fine, but the two-electron (Coulomb) terms seem a bit complicated — because it involves two electrons. We hope to simplify it to a one-electron integral, as electron 1 interacting with the average repulsion of electron 2. We have

$$\begin{aligned} \left\langle \psi_a(1)\psi_a(2) \middle| \frac{1}{r_{12}} \middle| \psi_a(1)\psi_a(2) \right\rangle &= \int \mathrm{d}^3\mathbf{r}_1 \, \mathrm{d}^3\mathbf{r}_2 \, \frac{\psi_a^*(\mathbf{r}_1)\psi_a^*(\mathbf{r}_2)\psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2)}{r_{12}} \\ &= \int \mathrm{d}^3\mathbf{r}_1 \, \psi_a^*(\mathbf{r}_1) \left[\int \mathrm{d}^3\mathbf{r}_2 \, \frac{\psi_a^*(\mathbf{r}_2)\psi_a(\mathbf{r}_2)}{r_{12}} \right] \psi_a(\mathbf{r}_1) \\ &= \int \mathrm{d}^3\mathbf{r}_1 \, \psi_a^*(\mathbf{r}_1) \left[\hat{J}(\mathbf{r}_1) \right] \psi_a(\mathbf{r}_1) \, . \end{aligned} \tag{5.37}$$

We defined a new operator, $\hat{J}(\mathbf{r})$, the Coulomb operator. It is the potential energy an electron would receive on average due to Coulomb repulsion from the other electron.

Therefore, we can write our total energy as

$$E = 2 \left\langle \psi_a \middle| \hat{h} \middle| \psi_a \right\rangle + \left\langle \psi_a \middle| \hat{J} \middle| \psi_a \right\rangle \, . \tag{5.38}$$

We now need to find the ψ_a that minimises this. Again, the rigorous derivation is in course B10. We will illustrate everything briefly. Consider a variation $\psi_a \to \psi_a + \delta\psi_a$. To meaningfully change ψ_a, we require $\langle \delta\psi_a | \psi_a \rangle = 0$. Then the variation in energy, to the first order, is

$$\delta E = 4 \left\langle \delta\psi_a \middle| \hat{h} \middle| \psi_a \right\rangle + 4 \left\langle \delta\psi_a \middle| \hat{J} \middle| \psi_a \right\rangle \, . \tag{5.39}$$

This first order variation in energy vanishes at the minimum, which is satisfied when

$$[\hat{h}(\mathbf{r}) + \hat{J}(\mathbf{r})]\psi_a(\mathbf{r}) = \epsilon_a \psi_a(\mathbf{r})\,. \tag{5.40}$$

We have therefore found a simple one-electron operator which takes into account electron-electron interactions. If we consider the more general case, with more electrons with different or same spins, we create what is known as the *Fock operator*,

$$\hat{F}(\mathbf{r}) = \hat{h}(\mathbf{r}) + \hat{J}(\mathbf{r}) - \hat{K}(\mathbf{r})\,, \tag{5.41}$$

where the *Coulomb operator* $\hat{J}$ accounts for the classical Coulomb repulsion between electrons, and $\hat{K}$ is the *exchange operator* which deals with exchange interactions, defined similarly as $\hat{J}$.[12]

The key result of these derivations is that we have created a new one-electron operator, the Fock operator, whose eigenfunctions are the molecular orbitals which minimise the energy of a Slater determinant. To find the orbitals we solve the *Fock equations*,

$$\hat{F}\psi_a = \epsilon_a \psi_a\,. \tag{5.42}$$

Since we constructed the molecular orbitals by linear combination of atomic orbitals

$$\psi_a = \sum_i c_i^{(a)} \phi_i\,, \tag{5.43}$$

we may solve the Fock equation in the matrix form. Define *Fock matrix* $F_{ij} = \left\langle \phi_i \middle| \hat{F} \middle| \phi_j \right\rangle$, the coefficients of the molecular orbitals are determined by the eigenvalue equation

$$\mathsf{F}\mathbf{c}^{(a)} = \epsilon_a \mathbf{c}^{(a)}\,. \tag{5.44}$$

We may interpret ϵ_a as the "orbital energy" of the orbital ψ_a, but rigorously there is no such thing as "orbital energy" in Hartree–Fock theory — there is only total energy. Moreover, it is absolutely wrong to obtain the total energy by adding up the orbital energies. It involves double-counting of interactions, just as what we discussed for Hückel theory.

5.3.1 The Self-Consistent Field Procedure

Because the Fock operator, $\hat{F}$ is both dependent on a set of orbitals (which are used to make $\hat{J}$ and $\hat{K}$), and is used to define the orbitals, it is not entirely straightforward to actually solve the Fock equations to achieve orbitals which are eigenfunctions of their own Fock operator. The simplest way of doing this is to start with a guess set of orbitals and iterate until self-consistency, just as what we did in Kohn–Sham theory. This kind of method is known as *self-consistent field* (SCF) method.[13] The general procedure goes as follows:

(i) We first need to:
   (a) Specify the molecular geometry (coordinates of nuclei)
   (b) Specify the charge (the number of electrons)
   (c) Specify spin
   (d) Pick atomic orbital basis set

[12] More properly, $\hat{J}$ includes the interaction with all electrons in the system, so $\left\langle \psi_a \middle| \hat{J} \middle| \psi_a \right\rangle$ actually includes the interaction of electron in orbital ψ_a with itself. This isn't a problem because this is exactly cancelled by an equal contribution in $\hat{K}$ where the electron exchanges with itself. This is the cancellation of self-interaction, which only occurs if we treat exchange interaction exactly (as opposed to DFT).

[13] Since the Hartree–Fock theory is the first theory in quantum chemistry that uses a SCF procedure, Hartree–Fock theory itself is sometimes also referred to as the SCF method.

so that we know the form of the Hamiltonian and the Slater determinant $\Psi = \frac{1}{\sqrt{n!}} |\chi_1 \chi_2 \dots \chi_n|$.

(ii) Make an initial guess of the orbitals for our system

$$\psi_n^{\text{guess}} = \sum_i c_i^{(n)} \phi_i \tag{5.45}$$

using e.g. Hückel theory.

(iii) Work out the Fock operator and hence determine the Fock matrix in the atomic orbital basis

$$F_{ij} = \left\langle \phi_i \middle| \hat{F} \middle| \phi_j \right\rangle . \tag{5.46}$$

(iv) Diagonalise F to get the new molecular orbitals

$$\mathsf{F}\mathbf{c}^{(n)} = \epsilon_n \mathbf{c}^{(n)} . \tag{5.47}$$

(v) Compare the new $\mathbf{c}^{(n)}$ to the previous $\mathbf{c}^{(n)}$. If they are close enough (i.e. converged), then stop. Otherwise, go back to (iii).

(vi) Calculate the total energy $\left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle$.

5.4 Electron Correlation

The key approximation in Hartree–Fock theory is the factorisation of the n electron wavefunction into one-electron orbitals. This is displayed in its most elementary form in equation (5.1), and the Slater determinant form keeps this basic assumption while accounting for antisymmetrisation. In this product form it can be seen that the distribution of one electron is effectively independent from all the others — their motion is not correlated with each other. In the true wavefunction, the singularity of the Coulomb interaction between electrons should give rise to a lowering of probability when they are close, and so this effect is missing in the Hartree product.

Within a single Slater determinant, the antisymmetrisation does produce a Fermi hole around electrons of the same spin as the wavefunction vanishes when they are at the same point, so this approximation does take some correlation into account, but no such effect is present for electrons of opposite spin. This is observed by the fact that in both Hartree product form $\Psi = \psi_1 \psi_2$ and the singlet Slater determinant form $\Psi = \frac{1}{\sqrt{2}} \left| \psi_1^\alpha \psi_2^\beta \right|$, we have the probability of finding electrons at $\mathbf{r}_1$ and $\mathbf{r}_2$ to be

$$P(\mathbf{r}_1, \mathbf{r}_2) = |\psi_1(\mathbf{r}_1)|^2 \, |\psi_2(\mathbf{r}_2)|^2 . \tag{5.48}$$

This lack of correlation is the origin of the error in the Hartree–Fock limit of equilibrium H_2.

The electron correlation error becomes more serious for stretched bonds, and grows to rather pathological proportions in the limit of full dissociation. The error in the dissociation limit can be quantified for H_2 by looking at the Hartree–Fock energy in the dissociated limit as $R \to \infty$, which you will show in one exercise.

$$\lim_{R \to \infty} E_{H_2}(R) = 2E_H + \frac{J_H}{2} . \tag{5.49}$$

We define the dissociation energy

$$\Delta_{\text{dissoc.}} = - \left(E_{H_2}(R_{\text{eq}}) - \lim_{R \to \infty} E_{H_2}(R) \right) \tag{5.50}$$

and the atomisation energy

$$\Delta_{\text{at.}} = - \left(E_{H_2}(R_{\text{eq}}) - 2E_H \right) . \tag{5.51}$$

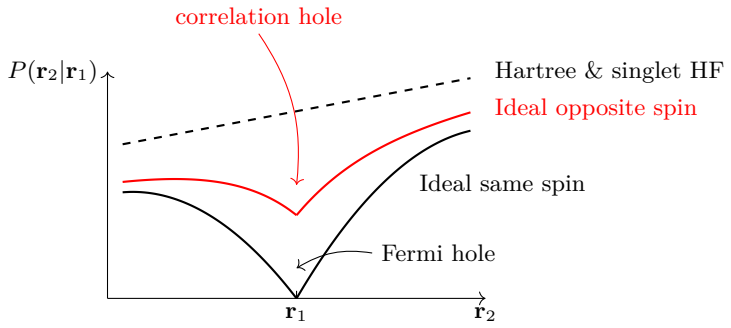


Figure 5.3: Dynamic correlation reduces the chances of electrons, even with the same spin, to approach.

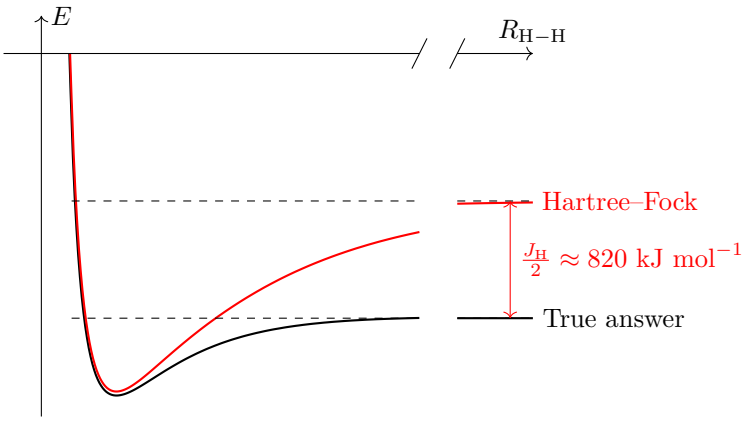


Figure 5.4: The Hartree–Fock theory over-predicts the dissociation energy significantly due to the lack of correlation.

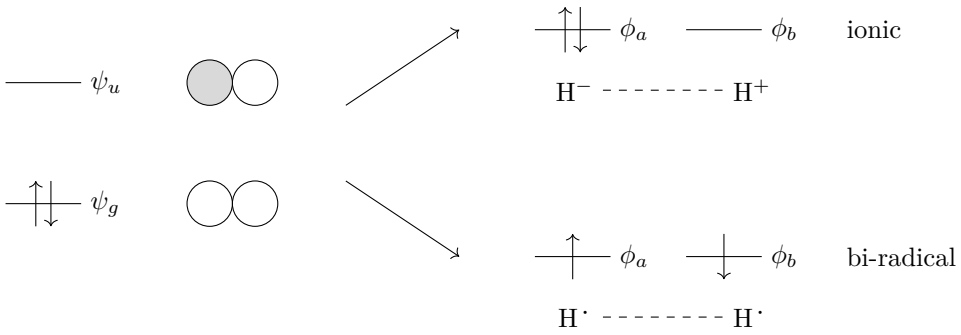


Figure 5.5: Due to the lack of electron correlation, Hartree–Fock predicts a dissociation into a superposition of bi-radical and ionic states. We can view this as two $\frac{1}{2}$-electrons are attached to each atom at dissociation.

These energies should be the same, but they are not in HF theory — Hartree–Fock theory over-predicts the dissociation energy by $\frac{1}{2}J_{\mathrm{H}}$, which turns out to be about 820 kJ mol^{-1}, a huge error.

What's wrong with the Hartree–Fock solution? Let's have a look again at the ground-state wavefunction written as a product of molecular orbitals

$$\begin{aligned}\Psi_{\mathrm{HF}}(1,2) &= \frac{1}{\sqrt{2}}\begin{vmatrix}\psi_g^\alpha(1) & \psi_g^\beta(1)\\ \psi_g^\alpha(2) & \psi_g^\beta(2)\end{vmatrix}\\ &= \frac{1}{2\sqrt{2}(1+S)}[\phi_a(1)+\phi_b(1)][\phi_a(2)+\phi_b(2)][\alpha(1)\beta(2)-\beta(1)\alpha(2)]\,, \end{aligned} \tag{5.52}$$

where $S = \langle\phi_a|\phi_b\rangle$ is the overlap integral. Expanding the product over the spatial orbitals, after omitting the spin orbitals which will not affect the final energy, we get

$$\Psi_{\mathrm{HF}}(1,2) \sim \underbrace{\phi_a(1)\phi_a(2)}_{\text{ionic}} + \underbrace{\phi_a(1)\phi_b(2)}_{\text{bi-radical}} + \underbrace{\phi_b(1)\phi_a(2)}_{\text{bi-radical}} + \underbrace{\phi_b(1)\phi_b(2)}_{\text{ionic}}\,. \tag{5.53}$$

We may interpret the terms as either ionic, with both electrons on one of the hydrogens, and a bare nucleus as the other, or bi-radical, with each hydrogen getting a single electron. Because of these ionic terms, the Hartree–Fock wavefunction cannot properly describe molecular dissociation (which should not contain the ionic terms in this limit).

Recall that Hartree–Fock arises from an approximation, and although Ψ_{HF} is the best possible single determinant which can describe our molecule, it is not an eigenfunction of the many-electron Hamiltonian. (This is because the orbitals in HF are eigenfunctions of the Fock operators, while the total Hamiltonian is not a sum of the Fock operators.) With this single-determinant approach, the electrons are effectively independent of each other and the probability distributions of electrons are independent. At dissociation this is very problematic, as we know the correct wavefunction has one electron on each atom.

In the exact wavefunction at dissociation, the probability of finding both electrons around the same nucleus is therefore very small indeed, but in this Hartree–Fock wavefunction it is 50%. In order to improve the description of molecular dissociation, one must introduce electron correlation, such that if one electron finds itself around one nucleus, the other electron has an enhanced probability of finding itself around the other nucleus.

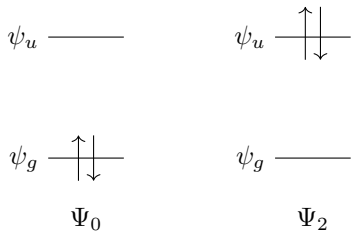
5.4.1 Configuration Interaction

How can one introduce such electron correlation? One approach is to go beyond the use of a single Slater determinant to represent the wavefunction, and use a linear combination of multiple Slater

determinants. For example, for the two-electron case, we may write the wavefunction as

$$\Psi(1,2) = \Psi_0 + \sum_I c_I \Psi_I \,, \tag{5.54}$$

where Ψ_0 is the ground state Slater determinant (Hartree–Fock solution) and Ψ_I are the Slater determinants with further excitations. This wavefunction Ψ is known as a *multi-determinantal wavefunction*. We can find the best possible coefficients c_I by variational principles. This method of introducing electron correlation is called *configuration interaction*.



Why is the introduction of excitation states helpful? Let's consider the H_2 molecule again. It turns out that it will be particularly useful to introduce a second determinant with doubly occupied anti-bonding orbital $\psi_u = \frac{1}{\sqrt{2(1-S)}}(\phi_a - \phi_b)$

$$\begin{aligned}
\Psi_2(1,2) &= \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_u^\alpha(1) & \psi_u^\beta(1) \\ \psi_u^\alpha(2) & \psi_u^\beta(2) \end{vmatrix} \\
&= \frac{1}{2\sqrt{2}(1-S)} [\phi_a(1) - \phi_b(1)][\phi_a(2) - \phi_b(2)][\alpha(1)\beta(2) - \beta(1)\alpha(2)] \\
&\sim \underbrace{\phi_a(1)\phi_a(2)}_{\text{ionic}} - \underbrace{\phi_a(1)\phi_b(2)}_{\text{bi-radical}} - \underbrace{\phi_b(1)\phi_a(2)}_{\text{bi-radical}} + \underbrace{\phi_b(1)\phi_b(2)}_{\text{ionic}} \,.
\end{aligned} \tag{5.55}$$

We call this Ψ_2 because we denote the Hartree–Fock ground state Ψ_0 and we preserve Ψ_1 for the singly excited state. Notice that in Ψ_2 the ionic components of the wavefunction come with the opposite sign to the bi-radical parts. Therefore if we subtract Ψ_2 from Ψ_0 (the Hartree–Fock solution (5.53)) we remove the troublesome ionic factors, leaving

$$c_0\Psi_0 - c_2\Psi_2 \sim [\phi_a(1)\phi_b(2) + \phi_b(1)\phi_a(2)][\alpha(1)\beta(2) - \beta(1)\alpha(2)] \,, \tag{5.56}$$

which dissociates correctly.

How do we solve for the coefficients of each Slater determinant in the configuration interaction? The answer is to use the variational principle. Define the Hamiltonian matrix in the basis of the Slater determinants $H_{IJ} = \left\langle \Psi_I \middle| \hat{H} \middle| \Psi_J \right\rangle$. Since the orbitals are orthonormal (as they are the eigenstates of the Fock operator), so too are the Slater determinants. The secular equation we need to solve is

$$(\mathsf{H} - E_A \mathsf{I})\mathbf{c}^{(A)} = \mathbf{0} \,. \tag{5.57}$$

The coefficients and energies are the eigenvalues and eigenvectors of H.

Back to our H_2 case, the secular equation is

$$\begin{pmatrix} E_0 - E & U \\ U & E_2 - E \end{pmatrix} = 0 \,, \tag{5.58}$$

where $E_0 = H_{00}, E_2 = H_{22}, U = H_{02} = H_{20}$. The solutions are

$$E = \frac{E_0 + E_2}{2} \pm \frac{1}{2}\sqrt{(E_0 - E_2)^2 + 4U^2} \,. \tag{5.59}$$

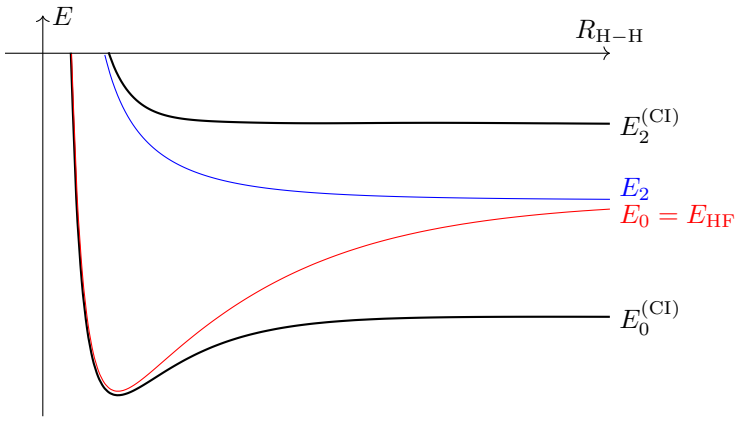


Figure 5.6: We can retain some of the electron correlation by configuration interaction.

In the limit of small U at large R, we have

$$E_0^{(\mathrm{CI})} \approx E_0 + \frac{U^2}{E_0 - E_2} \quad \text{and} \quad E_2^{(\mathrm{CI})} \approx E_2 - \frac{U^2}{E_0 - E_2}\,, \tag{5.60}$$

and similarly one can show that

$$\Psi_0^{(\mathrm{CI})} \approx \Psi_0 + \frac{U}{E_0 - E_2}\Psi_2\,. \tag{5.61}$$

One can further improve the wavefunction by enlarging the basis and constructing other (indeed all) possible excited determinants which have the same symmetry and spin ($^1\Sigma_g^+$) as Ψ_0. The limit of including all determinants it is possible to make with a given set of orbitals is known as *Full CI*, and is the best possible answer given our choice of basis set (the atomic orbitals). This provides, in principle, a systematic technique to introduce correlation. However, it is not in general a practical method, because the number of possible excited-state determinants grows in a factorial (i.e. exponential) manner (both with number of basis functions and the number of electrons). Methods to treat accurately electron correlation remain one of the most active and challenging areas of Theoretical Chemistry.

5.4.2 Wavefunction-based Methods of Correlation

We have seen that for even a simple molecule like H_2, there is a rich complexity to its wavefunction, and a technique like Hartree–Fock theory is not able to exactly reproduce the binding-curve — a treatment of electron correlation is essential. If the Full CI limit is computationally feasible we should use it, but in practice it can be used only for the very smallest systems because of the exponentially large size of Slater determinant space.

An appealing option would be to simply truncate the determinant space, including Slater determinants which differ by at most x electrons from the Hartree–Fock determinant; x is known as the excitation level. Unfortunately, performing CI in this naively truncated space leads to wavefunctions whose quality becomes significantly worse as the system size studied increases (because the x-excitation configuration takes a smaller and smaller proportion of the full CI as the number of electrons increases) so this is not a generally useful method.

All is not lost, however, and there exist two systematically improvable methods which are routinely used in electronic structure. *Møller–Plesset* (MP) Theory and *coupled cluster* (CC) Theory. They will be covered in the *B10: Electronic Structure* course, but key to their use is that they each form a family of methods where you can go to a more accurate level to be sure that you are sufficiently converged. Møller–Plesset methods are based on perturbation theory and labelled MP2,...,MPn where n labels the

Method	Scaling	Bond length Error / pm	Atomisation energy Error / kJ mol^{-1}
HF	N^3	2.7	423
MP2	N^5	0.5	36
MP3	N^6	1.2	
MP4	N^7	0.4	
MPn	N^{n+3}		
CCSD	N^6	0.8	37
CCSDT	N^8	0.1	3
CC…x	N^{2x+2}		
CCSD(T)	N^7	0.2	5
Experiment		0.2	4
Approx. value		150	1500

Table 1: The typical errors of different *ab initio* methods. CCSD(T) is seen as the golden standard of quantum chemistry.

level of perturbation theory (see also the *C7: Further Quantum Mechanics* course). Coupled Cluster methods are labelled with the level of excitation which is explicitly treated in the wavefunction: CCSD, CCSDT, CCSDTQ,… These differ from truncated CI methods in that quality of the energies for a given truncation level does not worsen as the system size increases. Hybrids of these methods exist where the final level of excitation is treated perturbatively, denoted: CCSD(T), CCSDT(Q),… These are routinely used as they have a lower computational scaling than the full method and produce extremely good approximations to the equivalent full CC method.

6 Reaction Theory

Most chemical reactions involve a change in the distribution of electrons, and the journey from reactants to products usually encounters some form of barrier on the potential energy surface. To understand the origin of these barriers we will consider how the multiple electronic states of a molecule change along a reaction pathway. Before this, however, we will consider again the initial assumptions that went into decoupling the nuclear and electronic wavefunctions in the Born–Oppenheimer approximation,

$$\Psi_{\mathrm{BO}}(\mathsf{r},\mathsf{R}) = \Psi_e(\mathsf{r};\mathsf{R})\Psi_n(\mathsf{R})\,. \tag{6.1}$$

6.1 Beyond the Born–Oppenheimer Approximation (Not Lectured)

The question arises, how good is the Born–Oppenheimer approximation, and when does it break down? To analyse this, we must consider a slightly more complicated form of the total wavefunction. Recall that there are in fact many solutions of the electronic Schrödinger equation, each corresponding to different electronic states, which we label k; these have energies $E_e^{(k)}(\mathsf{R})$ and wavefunctions $\Psi_e^{(k)}(\mathsf{r};\mathsf{R})$. The lowest electronic state, which has $k = 0$, corresponds to the very lowest energy configuration of electrons in the system for each nuclear state, as if the electrons were allowed to relax completely at each geometry. The electronic eigenfunctions are known as the *adiabatic electronic states*.[14]

To each electronic state, we can associate a nuclear wavefunction $\Psi_n^{(k)}(\mathsf{R})$, which could be obtained by solving the nuclear Schrödinger equation. To a first approximation, one could consider each state to be independent, but quantum mechanics has a nasty habit of coupling things together, and so to solve for the correct wavefunctions we must write the total wavefunction as a superposition of these states. Our first guess might take the form

$$\Psi_{\mathrm{total}}(\mathsf{r},\mathsf{R}) = \sum_k C_k \Psi_e^{(k)}(\mathsf{r};\mathsf{R})\Psi_n^{(k)}(\mathsf{R})\,, \tag{6.2}$$

which linearly combines the electronic states each with its own nuclear wavefunction. Though there are many solutions to the Schrödinger equation of this form, we are concerned with the total wavefunction with the lowest total energy. We see that the Born–Oppenheimer approximation amounts to taking a single term in this sum, and is sometimes called the *adiabatic approximation*. For most chemical purposes, the Born–Oppenheimer approximation proves sufficient, but to achieve spectroscopically accurate energies, additional corrections must be taken into account, though these are generally less than 1 kJ mol^{-1} in magnitude.

6.2 Avoided Crossing and Conical Intersections

Whilst the electronic states which are eigenfunctions of the electronic Schrödinger equation are convenient, they are not always so easy to interpret, as to minimise the energy the nature of the electronic state may change rapidly. An example of this is the dissociation of the diatomic LiF. In its ground equilibrium electronic state, it occurs mostly as $\mathrm{Li}^+\mathrm{F}^-$, but by comparing ionization energies and electron affinities we conclude that at infinite separation this molecule should separate

[14] In this context, a system being *adiabatic* means that the conditions are only changing gradually such that the system is allowed to adapt its configuration. Hence the Born–Oppenheimer approximation is also known as the *adiabatic approximation*. We assume that the electronic motion is faster than the nuclear motion, and for each nuclear geometry, the electrons are in their ground state. This is an example of the adiabatic theorem, which states that a physical system remains in its instantaneous eigenstate if a given perturbation is acting on it slowly enough and if there is a gap between the eigenvalue and the rest of the Hamiltonian's spectrum.

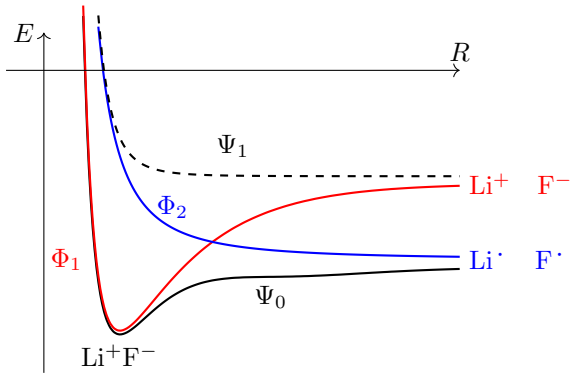


Figure 6.1: Φ_1, Φ_2 are *diabatic states*, where electrons do not respond to the nuclei. Ψ_0, Ψ_1 are *adiabatic states* where we allow electrons to relax at each nuclear position. When the ground state Ψ_0 LiF dissociates as R increases, it changes from predominantly Φ_1 to $\Phi_1 \pm \Phi_2$ (sign depends on $\left\langle \Phi_1 \middle| \hat{H} \middle| \Phi_2 \right\rangle$) then to Φ_2.

into neutral fragments, $\mathrm{Li}^{\cdot}$ and $\mathrm{F}^{\cdot}$.[15] A chemically intuitive view is to consider these two different pictures as different electronic states. These are known as *diabatic states*, loosely defined as the states where the nature of the electronic state changes less quickly as the geometry changes. The diabatic states are not electronic eigenfunctions, and provide an alternative representation to the adiabatic electronic states. For most purposes they can be regarded as an interpretative tool.

In LiF, we see that the ionic state will have a considerably higher energy than the bi-radical state at dissociation, and so the two diabatic states will cross somewhere along the dissociation curve. Recalling that the diabatic states are not electronic eigenfunctions, we may ask the following question: What combination of the diabatic states will give us the lowest energy? Calling these states Φ_1 and Φ_2, we see that we can simply solve the secular equations in this basis, and diagonalize $H_{ij} = \left\langle \Phi_i \middle| \hat{H} \middle| \Phi_j \right\rangle$. The ground state of this will be a linear combination of the two diabatic states with the lowest energy — the adiabatic state.

If the states are widely separated in energy ($|H_{11} - H_{22}| \gg 0$) the lowest of these will be a good approximation to the adiabatic state. When $|H_{11} - H_{22}| \approx 0$, however, we must consider role of the off-diagonal element H_{12}. This controls the mixing of the two diabatic states, and if it is non-zero, then when these states are close in energy, the adiabatic state will become an approximately equal mix of the two, smoothly changing from one to the other, splitting the two states apart. This effect is known as an *avoided crossing*. In general, unless there are symmetry reasons, H_{12} will not be zero, and so all such crossings will be avoided and the adiabatic ground state will not intersect with the excited state. This is rigorously true for diatomics, but for higher dimensional systems, there exist points of *accidental degeneracy* where $H_{11} = H_{22}$ and $H_{12} = 0$. These crossing points are known as *conical intersections* and have a dramatic effect on the excited electronic states of the molecule (especially on their relaxation). They are beyond the scope of this course.

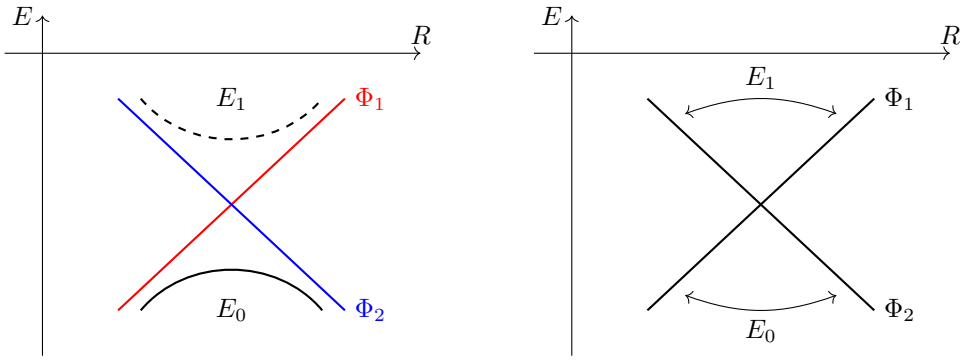
6.3 Chemical Reactions

In this section, we will apply what we previously discussed to predict whether a reaction would be feasible or not. We will use pericyclic reactions as examples. Look at the two reactions in the

[15] The ionisation energy of Li is 520 kJ mol^{-1} and the electron affinity of F is 320 kJ mol^{-1}, i.e.

$$\mathrm{Li} \longrightarrow \mathrm{Li}^+ + \mathrm{e}^- \qquad 520 \text{ kJ mol}^{-1}$$
$$\mathrm{F}^- \longrightarrow \mathrm{F} + \mathrm{e}^- \qquad 320 \text{ kJ mol}^{-1}.$$

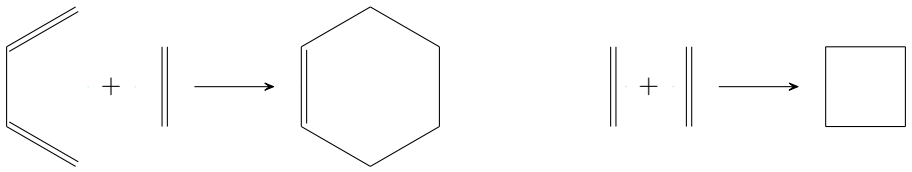
In fact, any diatomics would dissociate to atoms because all ionisation energies are larger than electron affinities.



$H_{12} \neq 0$
An avoided crossing

$H_{12} = 0$
A conical intersection

Figure 6.2: Normally when two diabatic states cross, there will be an avoided crossing. A conical intersection will occur when $H_{12} = 0$ for symmetry reasons or occasionally accidentally.



(a) (b)

Figure 6.3: (a) The [4+2] Diels–Alder cycloaddition; (b) a [2+2] forbidden cycloaddition

figure. Reaction (a) is the [4+2] (Diels–Alder) cycloaddition, and it occurs readily, whereas the corresponding [2+2] reaction is forbidden. Why should this be? Presumably, the latter reaction encounters a substantial barrier which is not present in the former — can the presence or absence of large barriers be understood solely in terms of the underlying MOs?

6.3.1 Orbital Correlation Diagrams

During a chemical reaction, we can parameterise the progress of the reaction by a *reaction coordinate*, along which the geometry of the system evolves smoothly from the reactants to the products. We want to track how the electronic structure of the system evolves as we move along the reaction coordinate. From now on, we will assume the wavefunctions of each molecule in the reactants and the products can be well-represented by a single Slater determinant. At the start and end, the electronic structure is determined by the MOs of the reactants and products respectively. We may trace the diabatic states along the reaction coordinate by considering Slater determinants made from specific occupations of these orbitals; along the reaction coordinate the orbitals will adapt to the new geometry, but the occupancy of the orbitals remains constant for a given diabatic state.

Along the reaction coordinate the molecular orbitals can, in principle, be calculated at every point, but in practice we only care about the energy ordering of the orbitals so can we may approximate how the orbitals change along the reaction by lines which join the orbitals of the reactants and the orbitals of the product. Crucially, orbitals of the same symmetry cannot cross in energy.

In general, to construct an orbital correlation diagram, we follow the following steps:

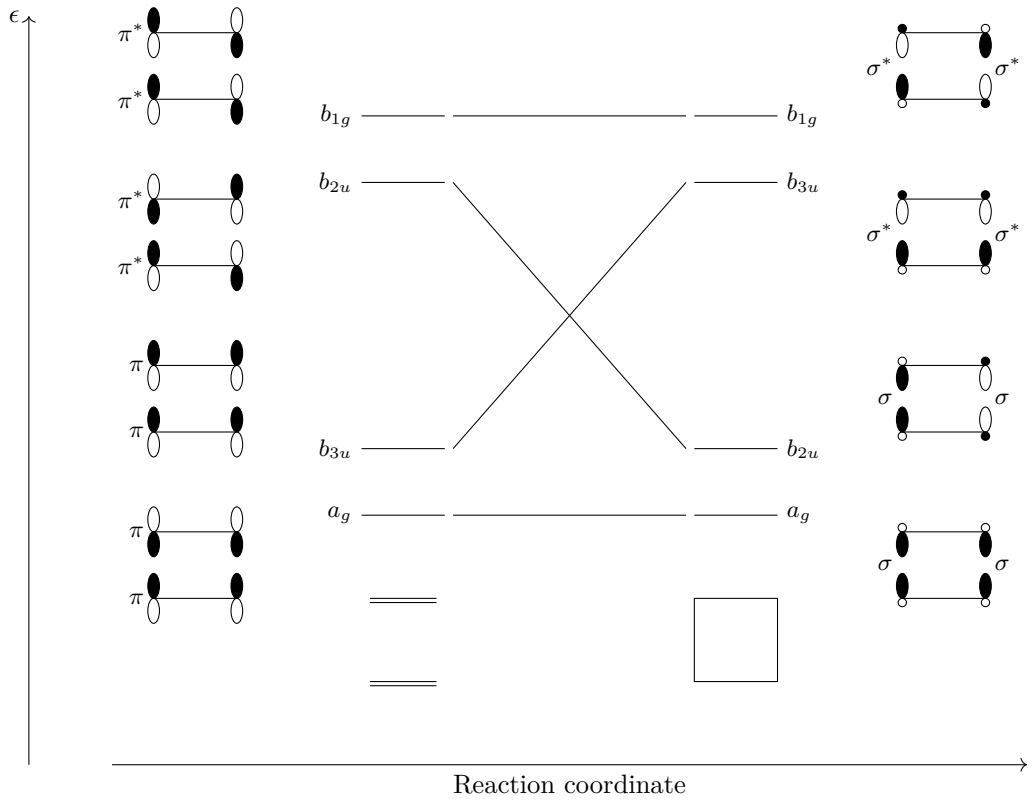


Figure 6.4: The orbital correlation diagram for ethene dimerisation.

(i) Specify the reactants and the products.

(ii) Find a symmetry (the point group) during the whole reaction process.

(iii) Draw energy-ordered MOs of the reactants and products. Occupied orbitals corresponding to bonds which do not change can be ignored.

(iv) Draw the energy levels for each MO (in products and reactants) and label them with the conserved symmetry.

(v) Working up the diagram from the lowest energy, for each orbital on the left, connect it to the lowest not-yet-connected orbital of the same symmetry on the right. Lines corresponding to orbitals of the same symmetry should not cross.

As an example, let us first consider the [2+2] system in figure 6.4. The unreacted ethene molecules have two sets of π-orbitals, whereas the cyclobutane product has two newly-formed σ-orbitals. In the unreacted state, the ground-state configuration is $(a_g)^2(b_{3u})^2$. If the electrons maintain in their original orbitals, then this will eventually lead to a doubly excited state in the product. Even if this double occupancy is not maintained all the way, but just to the transition state (approximately midway on the diagram), the energy of the system is substantially higher than the initial state. As a result, the reaction is forbidden if the starting molecule is in the ground state. On the other hand, if the molecule is photo-excited initially, promoting an electron from b_{3u} to b_{2u}, the reaction proceeds without an unduly large barrier. This reaction is photo-active.

6.3.2 State Correlation Diagram

The situation will be clearer by introducing the *state correlation diagram*. This is similar to the orbital correlation diagram, but we are correlating the electronic states between the reactants and the

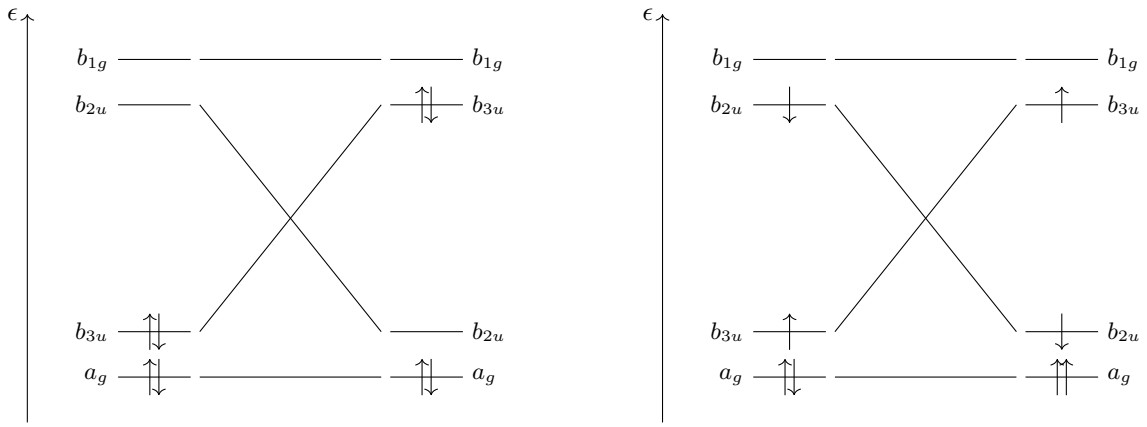


Figure 6.5: The reaction is forbidden if the reactant is in the electronic ground state. However this reaction is photo-active.

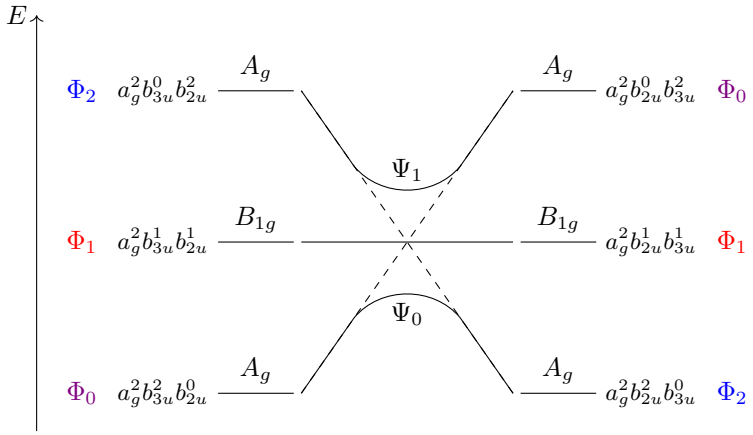


Figure 6.6: The state correlation diagram for ethene dimerisation. There is a high energy barrier so the reaction is thermally forbidden.

products. Again, there will be avoided crossing when two states with the same symmetry intersects.

In practice, we will consider only the ground electronic state and the first few excited states, and the states they corresponds to in the product. For our cycloaddition, we will consider

$$\text{Ground state } \Phi_0 \quad (A_g): \quad a_g^2 b_{3u}^2 \tag{6.3}$$
$$\text{Singly excited } \Phi_1 \quad (B_{1g}): \quad a_g^2 b_{3u}^1 b_{2u}^1 \tag{6.4}$$
$$\text{Doubly excited } \Phi_2 \quad (A_g): \quad a_g^2 b_{2u}^2\,. \tag{6.5}$$

Since the ground state and the doubly-excited state are of the same symmetry, they can mix, and do so quite substantially as they become energetically close. The energy and the coefficients of Ψ_i in terms of Φ_i can be obtained by solving secular equations. If we let

$$\begin{aligned} E_0 &= \left\langle \Phi_0 \middle| \hat{H} \middle| \Phi_0 \right\rangle \\ E_2 &= \left\langle \Phi_2 \middle| \hat{H} \middle| \Phi_2 \right\rangle \\ U &= \left\langle \Phi_0 \middle| \hat{H} \middle| \Phi_2 \right\rangle , \end{aligned} \tag{6.6}$$

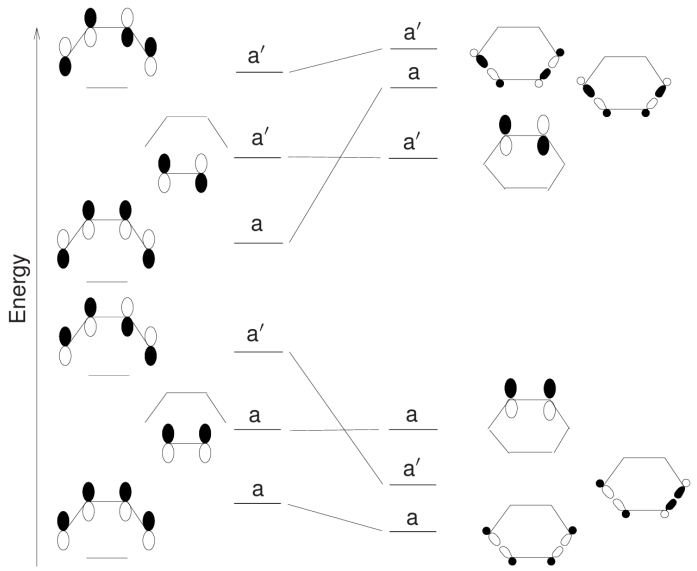


Figure 6.7: The orbital correlation diagrams for the Diels–Alder reaction. Figure adapted from the official course notes by Dr. Alex Thom.

then solving

$$\begin{vmatrix} E_0 - E & U \\ U & E_2 - E \end{vmatrix} = 0 \tag{6.7}$$

gives

$$E = \frac{E_0 + E_2}{2} \pm \frac{1}{2}\sqrt{(E_0 - E_2)^2 + 4U^2} \approx E_0 \pm \frac{U^2}{E_0 - E_2} \tag{6.8}$$

in the $|U| \ll |E_0 - E_2|$ limit.

It appears as if the two states repel each other, irrespective of the sign of U; the ground-state is lowered in energy and the excited state pushed up. This is the basis for the non-crossing rule for states with the same symmetry. The first excited state, being of a different symmetry, is unaffected. The magnitude of the repulsion depends both on U and on the energy difference of the two diabatic states. At the transition state, the energy difference is small and so the 'repulsion' is greatest.

Nevertheless, even with the above lowering of energy, there remains a large barrier (compared to the thermal energy available to the molecules) to reaction for the ground-state, and ethene dimerisation remains therefore thermally forbidden.

Turning now to the [4+2] cycloaddition. We notice that only a σ symmetry is preserved during the process, so we label everything under the C_s point group symmetry. Using the same process as before, we can see from the state correlation diagram that there is no energy barrier for the transitions of the electronic ground states between the reactants and the product, so the reaction is thermally allowed.

Similar considerations work for ring-opening and ring-closing, where different mechanisms, called *con-rotatory* and *dis-rotatory* can be allowed or forbidden, which leads to methods for selective synthesis of cis- or trans- products. You will investigate this in one of the exercises. These symmetry considerations are codified more generally in the *Woodward–Hoffmann rules* which you may encounter in the further study of pericyclic reactions.

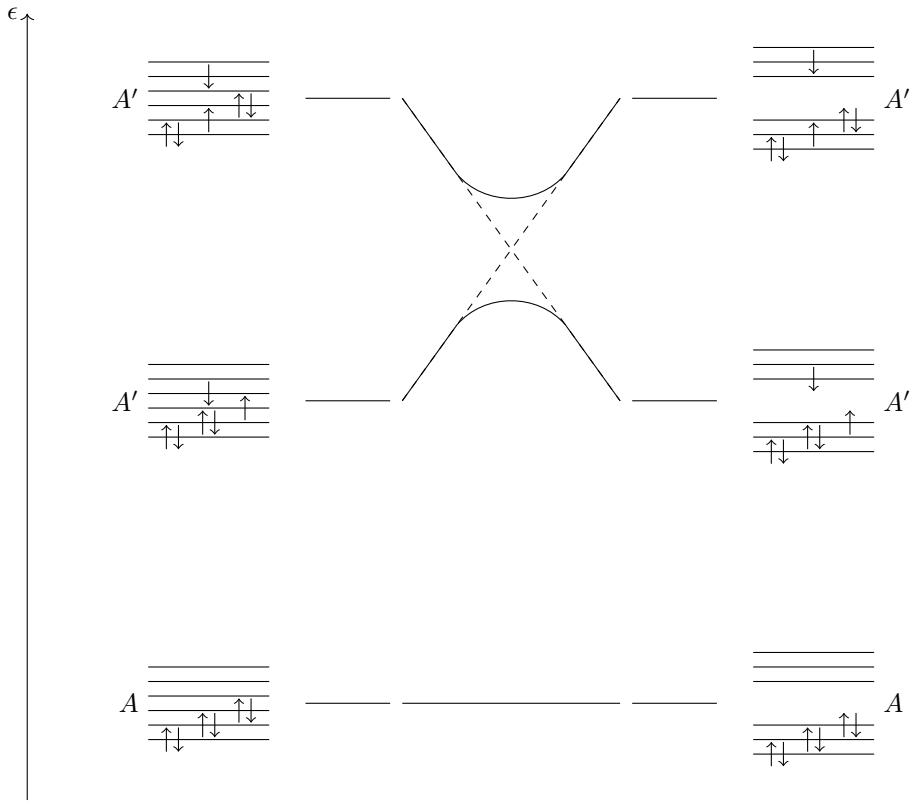


Figure 6.8: The state correlation diagram of the Diels–Alder reaction. The reaction of the ground state is thermally allowed.