Electronic Structure

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 original. They are largely based on the notes by Dr. Alex Thom, who lectured this course in 2023 and 2024. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

Contents

Part I

Ab Initio Quantum Chemistry

1 The Hamiltonian and the Wavefunction

In these lectures, we shall use atomic units. These are defined as

(i) unit of length is *Bohr* a_0, with dimension L;

(ii) unit of charge is the magnitude of the charge on the electron e, with dimension Q;

(iii) unit of action is *reduced Planck's constant* $\hbar$, with dimension $\mathrm{ML^2T^{-1}}$;

(iv) unit of mass is the mass of the electron m_e, with dimension M.

These definitions imply that $4\pi\epsilon_0 = 1$ because of the relation

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2}\,. \tag{1.1}$$

In these units the unit of energy is *Hartree* E_h. The speed of light in atomic units is 137.036.

In these units the molecular electronic Schrödinger Hamiltonian, ignoring all relativistic terms, is

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{n}\nabla_i^2 - \sum_{i=1}^{n}\sum_{A=1}^{N}\frac{Z_A}{\|\mathbf{r}_i - \mathbf{R}_A\|} + \sum_{i>j}^{n}\frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|} + \sum_{A>B}^{N}\frac{Z_A Z_B}{\|\mathbf{R}_A - \mathbf{R}_B\|}\,, \tag{1.2}$$

where i, j label electrons at $\mathbf{r}_i, \mathbf{r}_j$, and A, B label nuclei with charges Z_A, Z_B at $\mathbf{R}_A, \mathbf{R}_B$. This part of the course is concerned with the determination of solutions (energies and wavefunctions) of the Schrödinger equation

$$\hat{H}\Psi = E\Psi \tag{1.3}$$

for fixed positions of the nuclei. This is formally known as the *clamped-nucleus approximation*. Eigenvalue $E \equiv E(\{\mathbf{R_i}\})$ as a function of nuclear positions $\{\mathbf{R}_i\}$ is the *potential energy surface*. This part of the course is on *ab initio* quantum chemistry, meaning that we use no (semi-)empirical or experimental information.

The fundamental expansion functions which we use to find approximate solutions of Schrödinger's equation are *Slater determinants*, which have the form

$$\Psi = \frac{1}{\sqrt{n!}}\begin{vmatrix} \phi_1(1) & \phi_1(2) & \cdots & \phi_1(n) \\ \phi_2(1) & \phi_2(2) & \cdots & \phi_2(n) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_n(1) & \phi_n(2) & \cdots & \phi_n(n) \end{vmatrix}\,, \tag{1.4}$$

where i is the shorthand notation for spatial and spin coordinates $\mathbf{x}_i = (\mathbf{r}_i, \omega_i)$.

We can rewrite the Slater determinant into a form that is much easier to handle. But first, we need some knowledge of permutation groups.

Lemma 1.1. Let X be a set.

(i) The *permutations* on X, $\hat{P} : X \to X$ form a group under function composition, which is known as the *permutation group of* X.

(ii) If X has n elements, then the permutation group on X is denoted S_n. It has $n!$ elements.

(iii) A *k-cycle* is a permutation

$$(a_1\ a_2\ \ldots\ a_k)(i) = \begin{cases} a_{j+1} & \text{if } i = a_j \text{ for } j < k \\ a_1 & \text{if } i = a_k \\ i & \text{if } i \neq a_j \text{ for any } j\,. \end{cases} \tag{1.5}$$

A *transposition* is a 2-cycle.

(iv) A permutation has a unique decomposition into disjoint cycles, and a k-cycle can be decomposed into $k-1$ transpositions. Therefore, all permutations can be decomposed into transpositions.

(v) If a permutation $\hat{P}_u$ can be decomposed into an even number of transpositions, then the permutation is said to have an even *parity* $\sigma_u = +1$. If it is decomposed into an odd number of transpositions, then it has an odd parity $\sigma_u = -1$.

We may rewrite the Slater determinant as

$$\Psi = \hat{\mathcal{A}}(\phi_1(1)\phi_2(2)\ldots\phi_n(n))\,, \tag{1.6}$$

where $\hat{\mathcal{A}}$ is the *antisymmetriser* defined as

$$\hat{\mathcal{A}} \coloneqq \frac{1}{\sqrt{n!}} \sum_{u \in S_n}^{n!} \sigma_u \hat{P}_u\,, \tag{1.7}$$

in which $\hat{P}_u$ is a *permutation operator* that permutes the coordinates in the *Hartree product* $\Phi \coloneqq \phi_1(1)\phi_2(2)\ldots\phi_n(n)$. For example, the permutation $\hat{P}_{(i\ j)}$ permutes the coordinates of electron i and electron j, so that

$$\begin{aligned} \hat{P}_{(i\ j)}\Phi &= \hat{P}_{(i\ j)}\phi_1(1)\ldots\phi_i(i)\ldots\phi_j(j)\ldots\phi_n(n) \\ &= \phi_1(1)\ldots\phi_i(j)\ldots\phi_j(i)\ldots\phi_n(n)\,. \end{aligned} \tag{1.8}$$

The antisymmetriser therefore sums over all possible permutations of the electron coordinates, multiplied by the parity of the permutations so that the result obeys the *Pauli's exclusion principle*: a wavefunction changes sign on the interchange of the coordinates of two electrons. Note that these determinants are zero if

(i) two spin-orbitals are the same ($\phi_i = \phi_j$), or

(ii) the coordinates of two electrons are the same ($(\mathbf{r}_i, \omega_i) = (\mathbf{r}_j, \omega_j)$)

as desired from Pauli's principle.

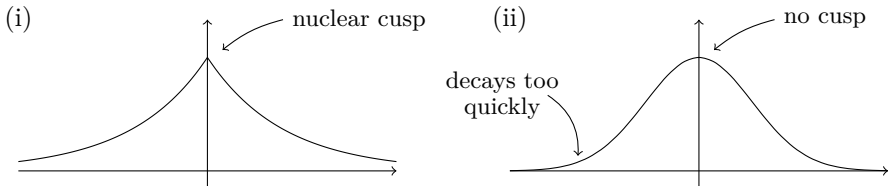


Figure 2.1: A comparison of (i) Slater type orbital and (ii) Gaussian function.

2 Basis Functions

Ideally it would be best to use hydrogenic type functions

$$R_{nl}(r)Y_{lm}(\theta,\phi) \tag{2.1}$$

as expansion functions for molecular orbitals. Equivalently, these functions can be written as

$$r^n x^p y^q z^s \exp(-Z_n r)\,, \tag{2.2}$$

from which we can see that such a set includes $\mathrm{s}, \mathrm{p}, \mathrm{d}, \mathrm{f}, \ldots$ hydrogenic atomic orbital wavefunctions. These functions are usually referred to as *Slater functions*. However it is impossible to analytically evaluate the one and two electron integrals which arise in the evaluation of matrix elements if these functions are used. This means that to use them in computation, we need time-consuming and often inaccurate numerical integrations.

2.1 Gaussian Type Orbitals

Instead, computational quantum chemistry usually uses *Gaussian basis functions*. They have the form

$$x^p y^q z^s \exp(-ar^2)\,, \tag{2.3}$$

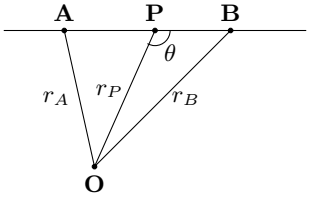
where p, q, s are integers. The angular parts of these functions are the same as the Slater functions, but the radial part is different. The derivative of an s Gaussian is zero at the origin, unlike the Slater function. Also the Gaussian dies off with $\sim \exp(-ar^2)$ dependence compared to the Slater function's $\sim \exp(-ar)$ dependence at large r. Thus Gaussian functions have a totally different behaviour from Slater functions at both small r and large r.

However the key advantage of Gaussians is that all the required integrals are very easy to evaluate. This follows from the fact that the product of two Gaussians is another Gaussian. This is known as the Gaussian product theorem.

Theorem 2.1 (Gaussian product theorem). Let E be a Euclidean space with Euclidean norm $\|\cdot\|$, $\mathbf{r}, \mathbf{r}_A, \mathbf{r}_B \in E$ and $\alpha, \beta \in \mathbb{R}$, $\alpha, \beta > 0$. Then

$$\mathrm{e}^{-\alpha\|\mathbf{r}-\mathbf{r}_A\|^2}\mathrm{e}^{-\beta\|\mathbf{r}-\mathbf{r}_B\|^2} = c\mathrm{e}^{-\gamma\|\mathbf{r}-\mathbf{r}_P\|^2} \tag{2.4}$$

for some $\gamma \in \mathbb{R}$, $\mathbf{r}_P \in E$ and normalising factor $c \in \mathbb{R}$.



Proof. Let $\mathbf{P}$ be a point along $\mathbf{AB}$. The cosine formula gives

$$r_A^2 = r_P^2 + \|\mathbf{PA}\|^2 + 2\,\|\mathbf{PA}\|\, r_P \cos\theta\,, \tag{2.5}$$

$$r_B^2 = r_P^2 + \|\mathbf{PB}\|^2 - 2\,\|\mathbf{PB}\|\, r_P \cos\theta\,. \tag{2.6}$$

If we choose the position $\mathbf{P}$ such that $\|\mathbf{PA}\| \,/\, \|\mathbf{PB}\| = \beta/\alpha$, that is

$$\mathbf{r}_P = \frac{\alpha\mathbf{r}_A + \beta\mathbf{r}_B}{\alpha+\beta}\,, \tag{2.7}$$

then

$$\alpha r_A^2 + \beta r_B^2 = (\alpha+\beta)r_P^2 + \alpha\,\|\mathbf{PA}\|^2 + \beta\,\|\mathbf{PB}\|^2\,. \tag{2.8}$$

Completing the algebra then gives the result

$$\mathrm{e}^{-\alpha r_A^2}\mathrm{e}^{-\beta r_B^2} = \mathrm{e}^{-\frac{\alpha\beta}{\alpha+\beta}\|\mathbf{AB}\|^2}\mathrm{e}^{-(\alpha+\beta)r_P^2} \tag{2.9}$$

as required. □

The integrals of Gaussian functions are easily evaluated analytically.

Lemma 2.2 (Gaussian integration lemma). For any constant $a, b \in \mathbb{R}$ with $a > 0$,

$$\int_{-\infty}^{\infty} \mathrm{d}u\, \mathrm{e}^{-a(u+b)^2} = \sqrt{\frac{\pi}{a}}\,. \tag{2.10}$$

It is then possible to evaluate the integrals that are needed for electronic structure calculation analytically. For example, the overlap integrals of two s Gaussians are

$$\begin{aligned}\langle\eta_{A\alpha}|\eta_{B\beta}\rangle &\coloneqq \int \mathrm{d}^3\mathbf{r}\, \mathrm{e}^{-\alpha r_A^2}\mathrm{e}^{-\beta r_B^2} \\ &= \left(\frac{\pi}{\alpha+\beta}\right)^{\frac{3}{2}} \mathrm{e}^{-\frac{\alpha\beta}{\alpha+\beta}\|\mathbf{AB}\|^2}\,.\end{aligned} \tag{2.11}$$

The four-centre two-electron integral can also be evaluated:

$$\begin{aligned}&\langle\eta_{A\alpha}\eta_{C\gamma}|\eta_{B\beta}\eta_{D\delta}\rangle \equiv (\eta_{A\alpha}\eta_{B\beta}|\eta_{C\gamma}\eta_{D\delta}) \\ &\coloneqq \iint \mathrm{d}^3\mathbf{r}_1\, \mathrm{d}^3\mathbf{r}_2\, \frac{\mathrm{e}^{-\alpha r_{1A}^2}\mathrm{e}^{-\beta r_{1B}^2}\mathrm{e}^{-\gamma r_{2C}^2}\mathrm{e}^{-\delta r_{2D}^2}}{r_{12}} \\ &= \langle\eta_{A\alpha}|\eta_{B\beta}\rangle\,\langle\eta_{C\gamma}|\eta_{D\delta}\rangle\,\mathrm{erf}\left(\left[\frac{(\alpha+\beta)(\gamma+\delta)}{\alpha+\beta+\gamma+\delta}\right]^{\frac{1}{2}}\|\mathbf{PQ}\|\right)\,,\end{aligned} \tag{2.12}$$

where

$$\mathbf{r}_P = \frac{\alpha\mathbf{r}_A + \beta\mathbf{r}_B}{\alpha+\beta}\,,\ \mathbf{r}_Q = \frac{\gamma\mathbf{r}_C + \delta\mathbf{r}_D}{\gamma+\delta} \tag{2.13}$$

and

$$\mathrm{erf}(x) \coloneqq \frac{2}{\sqrt{\pi}}\int_0^x \mathrm{d}t\, \exp(-t^2) \tag{2.14}$$

is the *error function*. This is more often quoted in an equivalent form in the literature

$$\langle\eta_{A\alpha}\eta_{C\gamma}|\eta_{B\beta}\eta_{D\delta}\rangle = \frac{2\pi^{5/2}}{pq\sqrt{p+q}}F_0(\rho\,\|\mathbf{PQ}\|^2)\,, \tag{2.15}$$

where $p = \alpha+\beta$, $q = \gamma+\delta$, $\rho = pq/(p+q)$ and

$$F_n(x) = \int_0^1 \mathrm{d}t\, \exp(-xt^2)t^{2n} \tag{2.16}$$

is the n^{th}-order Boys function. You don't need to know any of the formulae above. You only need to appreciate that they exist, and computers can use them to quickly get the results of the integrals.

2.2 Standard Basis Sets

To overcome the difficulty of the less desirable short and long range behaviour of Gaussians, it is common today to use fixed combinations of one to six Gaussians as basis functions, with the combinations chosen such that they look more like Slater functions.

These combinations are established, and there are a variety of different notations, usually referred to by the primary author who devised and published them. Here we will give a few examples:

- STO-3G (Pople) means the use of a contracted combination of three Gaussians to represent each Slater function. H requires 1 s basis function made from 3 Gaussian functions. Li – Ne require 2 s basis functions and one set of three p basis functions, so there are in total 5 of them, each made up from 3 Gaussian functions.

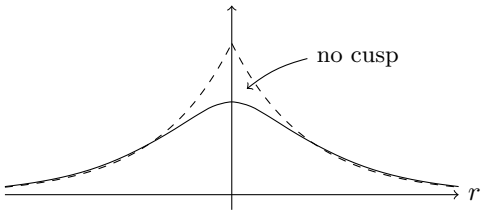


Figure 2.2: STO-3G uses 3 Gaussian functions to mimic the Slater type hydrogen 1s orbital.

You can find the parameters of these basis functions on Basis Set Exchange.[1] For example, if you search for the parameters of Li, you will get

```
#BASIS SET: (6s,3p) -> [2s,1p]
Li    S
    0.1611957475E+02       0.1543289673E+00
    0.2936200663E+01       0.5353281423E+00
    0.7946504870E+00       0.4446345422E+00
Li    SP
    0.6362897469E+00      -0.9996722919E-01       0.1559162750E+00
    0.1478600533E+00       0.3995128261E+00       0.6076837186E+00
    0.4808867840E-01       0.7001154689E+00       0.3919573931E+00
END
```

Let's have a close look at what everything means.

The first section gives the parameters of the 3 Gaussians composing Li 1s orbitals:

$$\eta_{\mathrm{Li,1s}} = \sum_{i=1}^{3} c_i \mathrm{e}^{-a_i r^2} \,, \tag{2.17}$$

where

i	$a_i \,/\, a_0^{-2}$	c_i
1	16.11957475	0.1543289673
2	2.936200663	0.5353281423
3	0.7946504870	0.4446345422

[1] https://www.basissetexchange.org/

The second section gives the 3 Gaussians shared by the 2s and 2p orbitals:

$$\eta_{\mathrm{Li,2s}} = \sum_{i=1}^{3} c_i^{(\mathrm{s})} \mathrm{e}^{-a_i r^2} \,, \tag{2.18}$$

and

$$\eta_{\mathrm{Li,2p}_q} = \sum_{i=1}^{3} q c_i^{(\mathrm{p})} \mathrm{e}^{-a_i r^2} \,, \tag{2.19}$$

where $q \in \{x, y, z\}$ are Cartesian functions and parameters are

i	$a_i \,/\, a_0^{-2}$	c_i^{s}	c_i^{p}
1	0.6362897469	−0.09996722919	0.1559162750
2	0.1478600533	0.3995128261	0.6076837186
3	0.04808867840	0.7001154689	0.3919573931

Note that the Gaussian functions are the same for 2s and 2p orbitals.

Such a basis set is described as a *minimal basis set* as it uses one basis function per atomic orbital.

- 6-31G* (Pople) is a *split valence basis set*, in which a valence atomic orbital is split into two basis functions to allow more flexibility. It has 6 Gaussians in the core orbital and 3 and 1 Gaussians in the two valence orbitals. Moreover, due to the chemical environments, the electron density in an atom may be distorted. This requires us to bring in some orbitals with higher angular momentum for polarisation. To distort an s orbital, we need to introduce a p orbital, and to distort a p orbital, we need to use d orbitals *etc.* The star means polarisation functions on all atoms except hydrogen, and ** would also add polarisation to hydrogen.

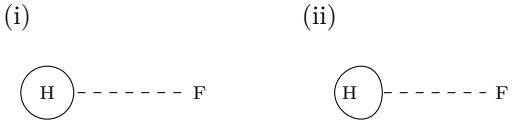


Figure 2.3: A hydrogen atom (i) without polarisation and (ii) with polarisation. The polarisation can be achieved by using an extra set of p orbitals.

A hydrogen atom has a 1s valence orbital, so 6-31G* would split it into two basis functions, composed of 3 and 1 Gaussians respectively. For Li-Ne, the 1s core orbital is made up from 6 Gaussians, while the 2s and 2p valence orbitals are each split into 2 basis functions composed of 3 and 1 Gaussians respectively. Moreover, polarisation introduces a set of d orbitals. Hence they are described by 15 basis functions (3s2p1d, and there are 6 d orbitals in 6-31G* — see later).

- cc-pVTZ (Dunning) is a large basis set for Hartree–Fock and is primarily designed for correlated methods. It stands for correlation consistent polarised valence triple zeta (e.g. H 3s2p1d (14 functions) Li-Ne 4s3p2d1f (30 functions)). Larger basis sets are available e.g. cc-pVQZ or cc-pV5Z.

- Augmented functions for diffuse basis sets for the description of anionic systems are e.g. aug-cc-pVTZ (Dunning) or 6-311++G(3df,3pd) (Pople).

- cc-pCVTZ (Dunning) includes functions for correlating core electrons.

2.2.1 Cartesian vs Spherical

For d functions and beyond, the Cartesian function $\{x_i x_j\}$ spans a larger space than the spherical harmonics $\{2z^2 - x^2 - y^2, x^2 - y^2, xy, xz, yz\}$. Care should be taken as to which definition is used within a basis set. The Pople 6-31G family use Cartesian d-functions, but spherical functions for f orbitals and beyond. Dunning style cc-pVXZ basis sets use purely spherical harmonics.

2.3 Slater Type Orbitals (Non-examinable)

A normalised Slater basis function is defined by

$$S(\zeta, n, \ell, m; r, \theta, \phi) = \frac{(2\zeta)^{n+\frac{1}{2}}}{[(2n)!]^{\frac{1}{2}}} r^{n-1} \mathrm{e}^{-\zeta r} Z_{\ell m}(\theta, \phi)\,. \tag{2.20}$$

Thus they are defined through their exponent ζ, principal quantum number n, angular momentum quantum number ℓ and the magnetic quantum number m. We have $n > \ell$ and $m = -\ell, -\ell+1, \ldots, \ell$. $Z_{\ell m}(\theta, \phi)$ are the real spherical harmonics given by

$$Z_{\ell,m}(\theta, \phi) = \left(\frac{(2\ell+1)(\ell - |m|)!}{2\pi(\ell + |m|)!} \right)^{\frac{1}{2}} \begin{cases} \cos(|m|\,\phi) P_\ell^{|m|}(\cos\theta) & \ell \geq m > 0 \\ \frac{1}{\sqrt{2}} P_\ell^0(\cos\theta) & m = 0 \\ \sin(|m|\,\phi) P_\ell^{|m|}(\cos\theta) & -\ell \leq m < 0\,, \end{cases} \tag{2.21}$$

where P_ℓ^m are the *associated Legendre functions*.

The formulae for the analytic potential of a STO are given by

$$V(\zeta, n, \ell, m; r, \theta, \phi) = \int_0^\infty \mathrm{d}r' \int_0^\pi \mathrm{d}\theta' \int_0^{2\pi} \mathrm{d}\phi' \, \frac{S(\zeta, n, \ell, m; r', \theta', \phi')}{\|\mathbf{r} - \mathbf{r}'\|} r'^2 \sin\theta'\,. \tag{2.22}$$

Using the well-known expansion of $\|\mathbf{r} - \mathbf{r}'\|$ in spherical harmonics, the following result is obtained

$$V(\zeta, n, \ell, m; r, \theta, \phi) = \frac{4\pi(2\zeta)^{n+\frac{1}{2}}}{[(2n)!]^{\frac{1}{2}}(2\ell+1)} Z_{\ell m}(\theta, \phi) I_{n\ell}(r)\,, \tag{2.23}$$

with

$$I_{n\ell}(r) = r^{-\ell-1} \int_0^r \mathrm{d}r' \, (r')^{n+\ell+1} \mathrm{e}^{-\zeta r'} + r^\ell \int_r^\infty \mathrm{d}r' \, (r')^{n-\ell} \mathrm{e}^{-\zeta r'}\,. \tag{2.24}$$

Both of these integrals may be evaluated recursively.

The only other formula that is often used is the Laplacian of an STO

$$\nabla^2 S(\zeta, n, \ell, m; r, \theta, \phi) = [[n(n-1) - \ell(\ell+1)] r^{-2} - 2n\zeta r^{-1} + \zeta^2] S(\zeta, n, \ell, m; r, \theta, \phi)\,. \tag{2.25}$$

3 Matrix Element of Determinants and Energy Expressions

We write the Slater determinant of a set of orthonormal spin-orbitals in the form

$$\begin{aligned}\Psi &= \frac{1}{\sqrt{n!}}\begin{vmatrix}\phi_1(1) & \phi_1(2) & \cdots & \phi_1(n)\\ \phi_2(1) & \phi_2(2) & \cdots & \phi_2(n)\\ \vdots & \vdots & \ddots & \vdots\\ \phi_n(1) & \phi_n(2) & \cdots & \phi_n(n)\end{vmatrix}\\ &= \hat{\mathcal{A}}(\phi_1\phi_2\ldots\phi_n)\\ &= \hat{\mathcal{A}}\,\Phi\,. \end{aligned} \tag{3.1}$$

We now work out some properties of $\hat{\mathcal{A}}$ that may help us evaluate the matrix elements.

Lemma 3.1.

(i) $\hat{\mathcal{A}}^2 = \sqrt{n!}\,\hat{\mathcal{A}}$.

(ii) $\hat{P}_u$ is unitary, and so $\hat{\mathcal{A}}$ is Hermitian, i.e. $\left\langle \hat{\mathcal{A}}\,\Phi\middle|\Psi\right\rangle = \left\langle \Phi\middle|\hat{\mathcal{A}}\,\Psi\right\rangle$ for any Φ, Ψ.

(iii) $\left\langle \hat{\mathcal{A}}\,\Phi\middle|\hat{H}\middle|\hat{\mathcal{A}}\,\Phi\right\rangle = \sqrt{n!}\left\langle \Phi\middle|\hat{H}\middle|\hat{\mathcal{A}}\,\Phi\right\rangle$.

Proof.

(i) We have

$$\begin{aligned}\hat{\mathcal{A}}^2 &= \frac{1}{n!}\sum_u^{n!}\sigma_u\hat{P}_u\sum_v^{n!}\sigma_v\hat{P}_v\\ &= \frac{1}{n!}\sum_u^{n!}\sum_v^{n!}(\sigma_u\sigma_v)(\hat{P}_u\hat{P}_v)\,. \end{aligned} \tag{3.2}$$

We need a simple result from group theory known as the *rearrangement theorem*. It states that for any element g in a group G, $gG = \{gh \mid h \in G\} = G$. So, for each $\hat{P}_u$

$$\sum_{v\in S_n}\sigma_u\sigma_v\hat{P}_u\hat{P}_v = \sum_{w\in S_n}\sigma_w\hat{P}_w\,. \tag{3.3}$$

This basically says that doing all possible permutations then do a specific permutation is equivalent to doing all possible permutations. Therefore,

$$\begin{aligned}\hat{\mathcal{A}}^2 &= \frac{1}{n!}\sum_{u,w}^{n!}\sigma_w\hat{P}_w\\ &= \frac{1}{n!}n!\sqrt{n!}\,\hat{\mathcal{A}}\\ &= \sqrt{n!}\,\hat{\mathcal{A}}\,. \end{aligned} \tag{3.4}$$

Note that this is different from a projection operator, for which $\hat{Q}^2 = \hat{Q}$.

(ii) $\hat{P}_u$ is unitary because

$$\begin{aligned}\left\langle \hat{P}_u\Phi\middle|\Psi\right\rangle &= \left\langle \hat{P}_u^{-1}\hat{P}_u\Phi\middle|\hat{P}_u^{-1}\Psi\right\rangle\\ &= \left\langle \Phi\middle|\hat{P}_u^{-1}\Psi\right\rangle \end{aligned} \tag{3.5}$$

as acting the same permutation on both sides only introduces a relabelling of the integration variable. Note that $\hat{P}_u$ and $\hat{P}_u^{-1}$ have the same parity since parity $\sigma : S_n \to C_2$ is a homomorphism and $\sigma(\hat{P}_u^{-1})\sigma(\hat{P}_u) = \sigma(\hat{E}) = 1$. Therefore we have

$$\begin{aligned}\left\langle \sum_u \sigma_u \hat{P}_u \Phi \middle| \Psi \right\rangle &= \left\langle \Phi \middle| \sum_u \sigma_u \hat{P}_u^{-1} \Psi \right\rangle \\ &= \left\langle \Phi \middle| \sum_u \sigma_u \hat{P}_u \Psi \right\rangle . \end{aligned} \tag{3.6}$$

Hence

$$\left\langle \hat{\mathcal{A}}\,\Phi \middle| \Psi \right\rangle = \left\langle \Phi \middle| \hat{\mathcal{A}}\,\Psi \right\rangle . \tag{3.7}$$

(iii) From (ii), we have

$$\left\langle \hat{\mathcal{A}}\,\Phi \middle| \hat{H} \middle| \hat{\mathcal{A}}\,\Phi \right\rangle = \left\langle \Phi \middle| \hat{\mathcal{A}}\,\hat{H}\,\hat{\mathcal{A}}\,\Phi \right\rangle . \tag{3.8}$$

$\hat{H}$ is a symmetric operator that is invariant under the swapping of any pair of labels, so $\hat{H}$ commutes with $\hat{\mathcal{A}}$. Thus we have

$$\left\langle \hat{\mathcal{A}}\,\Phi \middle| \hat{H} \middle| \hat{\mathcal{A}}\,\Phi \right\rangle = \left\langle \Phi \middle| \hat{H} \middle| \hat{\mathcal{A}}^2\,\Phi \right\rangle = \left\langle \Phi \middle| \hat{H} \middle| \sum_u \sigma_u \hat{P}_u \Phi \right\rangle \tag{3.9}$$

as claimed. This is a significant simplification as we have reduced a sum of $(n!)^2$ terms to a sum of $n!$ terms. □

Theorem 3.2. For a general Hamiltonian

$$\hat{H} = C + \sum_i \hat{h}(i) + \sum_{i>j} \frac{1}{r_{ij}} , \tag{3.10}$$

the expected value of $\hat{H}$ for the Slater determinant $\Psi = \hat{\mathcal{A}}\,\Phi$ is given by

$$\left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle = C + \sum_i h_{ii} + \sum_{i>j} [(ii|jj) - (ij|ij)] , \tag{3.11}$$

where h_{ij} is the matrix element $\left\langle \phi_i \middle| \hat{h} \middle| \phi_j \right\rangle$ and

$$\left\langle \phi_i(1)\phi_j(2) \middle| \frac{1}{r_{12}} \middle| \phi_k(1)\phi_l(2) \right\rangle \coloneqq \langle ij|kl\rangle \equiv (ik|jl) . \tag{3.12}$$

The notation $\langle \cdot \,|\, \cdot \rangle$ is the *physicists' notation* and $(\cdot \,|\, \cdot)$ is the *chemists' notation*.

Proof. We consider each term of the operator in turn.

(i) First the constant term C:

$$\langle \Psi | C | \Psi \rangle = C \left\langle \phi_1(1)\phi_2(2)\dots\phi_n(n) \middle| \sum_u \sigma_u \hat{P}_u \phi_1(1)\phi_2(2)\dots\phi_n(n) \right\rangle . \tag{3.13}$$

If $\hat{P}_u = \hat{E}$, the identity operation, then the contribution is

$$C \langle \phi_1|\phi_1\rangle_1 \langle \phi_2|\phi_2\rangle_2 \dots \langle \phi_n|\phi_n\rangle_n = C . \tag{3.14}$$

If $\hat{P}_u = \hat{P}_{12}$, we obtain the contribution

$$-C \cancelto{0}{\langle \phi_1|\phi_2\rangle_1} \cancelto{0}{\langle \phi_2|\phi_1\rangle_2} \dots \langle \phi_n|\phi_n\rangle_n = 0 . \tag{3.15}$$

Similarly, if $\hat{P}_u$ is any non-identity permutation, it will permute some i to some $j \neq i$, then the contribution is

$$\sigma_u C \ldots \cancelto{0}{\langle \phi_i | \phi_j \rangle_i} \ldots = 0 \,. \tag{3.16}$$

Therefore, only the identity permutation in the sum contributes and so

$$\langle \Psi | C | \Psi \rangle = C \,. \tag{3.17}$$

Moreover, let $C = 1$ then $\langle \Psi | \Psi \rangle = 1$ so we can check that the Slater determinant Ψ is normalised.

(ii) Next, consider the one-electron Hamiltonian term $\hat{h}(i)$:

$$\left\langle \Psi \middle| \hat{h}(i) \middle| \Psi \right\rangle = \left\langle \phi_1(1)\phi_2(2)\ldots\phi_n(n) \middle| \hat{h}(i) \middle| \sum_u \sigma_u \hat{P}_u \phi_1(1)\phi_2(2)\ldots\phi_n(n) \right\rangle . \tag{3.18}$$

If $\hat{P}_u = \hat{E}$, we obtain the contribution

$$\langle \phi_1 | \phi_1 \rangle_1 \langle \phi_2 | \phi_2 \rangle_2 \ldots \left\langle \phi_i \middle| \hat{h} \middle| \phi_i \right\rangle_i \ldots \langle \phi_n | \phi_n \rangle_n = \left\langle \phi_i \middle| \hat{h} \middle| \phi_i \right\rangle =: h_{ii} \,. \tag{3.19}$$

If $\hat{P}_u = \hat{P}_{ij}$, we obtain the contribution

$$- \langle \phi_1 | \phi_1 \rangle_1 \ldots \left\langle \phi_i \middle| \hat{h} \middle| \phi_j \right\rangle_i \ldots \cancelto{0}{\langle \phi_j | \phi_i \rangle_i} \ldots \langle \phi_n | \phi_n \rangle_n = 0 \,, \tag{3.20}$$

and the contribution from other terms vanish similarly. Thus summing over i, we obtain

$$\left\langle \Psi \middle| \sum_i \hat{h}(i) \middle| \Psi \right\rangle = \sum_i h_{ii} \,. \tag{3.21}$$

(iii) Finally, there is the $e^- - e^-$ repulsion term $\frac{1}{r_{ij}}$:

$$\left\langle \Psi \middle| \frac{1}{r_{ij}} \middle| \Psi \right\rangle = \left\langle \phi_1(1)\phi_2(2)\ldots\phi_n(n) \middle| \frac{1}{r_{ij}} \middle| \sum_u \sigma_u \hat{P}_u \phi_1(1)\phi_2(2)\ldots\phi_n(n) \right\rangle . \tag{3.22}$$

If $\hat{P}_u = \hat{E}$, we obtain the contribution

$$\langle \phi_1 | \phi_1 \rangle_1 \langle \phi_2 | \phi_2 \rangle_2 \ldots \left\langle \phi_i(i)\phi_j(j) \middle| \frac{1}{r_{ij}} \middle| \phi_i(i)\phi_j(j) \right\rangle_{ij} \ldots \langle \phi_n | \phi_n \rangle_n = (ii|jj) \,. \tag{3.23}$$

If $\hat{P}_u = \hat{P}_{ij}$, we obtain the contribution

$$- \langle \phi_1 | \phi_1 \rangle_1 \ldots \left\langle \phi_i(i)\phi_j(j) \middle| \frac{1}{r_{ij}} \middle| \phi_i(j)\phi_j(i) \right\rangle_{ij} \ldots \langle \phi_n | \phi_n \rangle_n = -(ij|ij) \,. \tag{3.24}$$

Any other permutation will involve the overlap of different orbitals and hence give zero contribution. Thus summing over $i > j$, we obtain

$$\left\langle \Psi \middle| \frac{1}{r_{ij}} \middle| \Psi \right\rangle = \sum_{i>j} [(ii|jj) - (ij|ij)] \,. \tag{3.25}$$

Hence, the Slater determinant constructed from a Hartree product of normalised spin-orbitals $\Psi = \hat{\mathcal{A}}\,\Phi$ is also normalised and

$$\left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle = C + \sum_i h_{ii} + \sum_{i>j} [(ii|jj) - (ij|ij)] \,. \tag{3.26}$$

$(ii|jj)$ is known as the *Coulomb integral* and $(ij|ij)$ is the *exchange integral*. □

Proposition 3.3. The expectation value of a Slater determinant for an arbitrary operator $\hat{Q}$ is invariant to a unitary rotation of its orbitals.

Proof. Let $\Psi = \hat{\mathcal{A}}\,\phi_1 \dots \phi_n$, where $\{\phi_i\}_{i=1}^n$ are orthonormal, and let $\boldsymbol{\psi} = \mathsf{U}\boldsymbol{\phi}$, where U is unitary so that

$$\mathsf{U}^\dagger = \mathsf{U}^{-1}\,,\ |\det \mathsf{U}| = 1\,. \tag{3.27}$$

In component form,

$$\psi_i = \sum_{j=1}^{n} U_{ij}\phi_j\,. \tag{3.28}$$

Let the Slater determinant of the transformed orbital be

$$\begin{aligned}
\Psi' &= \hat{\mathcal{A}}\,\psi_1 \dots \psi_n \\
&= \hat{\mathcal{A}} \sum_{j_1=1}^{n} U_{1j_1}\phi_{j_1}(1) \sum_{j_2=1}^{n} U_{2j_2}\phi_{j_2}(2) \cdots \sum_{j_n=1}^{n} U_{nj_n}\phi_{j_n}(n) \\
&= \underbrace{\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_n=1}^{n}}_{n^n \text{ terms}} U_{1j_1}U_{2j_2}\dots U_{nj_n}\,\hat{\mathcal{A}}\,\phi_{j_1}(1)\phi_{j_2}(2)\dots\phi_{j_n}(n)\,.
\end{aligned} \tag{3.29}$$

By the property of the antisymmetriser,

$$\hat{\mathcal{A}}\,\phi_{j_1}(1)\phi_{j_2}(2)\dots\phi_{j_n}(n) \neq 0 \tag{3.30}$$

if and only if $\{1, 2, \dots, n\} = \{j_1, j_2, \dots, j_n\}$. Thus, for a term in the sum to be non-zero, we must have

$$\phi_{j_1}(1)\phi_{j_2}(2)\dots\phi_{j_n}(n) = \hat{P}_{j_1\dots j_n}\phi_1(1)\phi_2(2)\dots\phi_n(n)\,. \tag{3.31}$$

Hence,

$$\begin{aligned}
\Psi' &= \sum_{\hat{P}_{j_1\dots j_n}}^{n!} U_{1j_1}\dots U_{nj_n} \underbrace{\hat{\mathcal{A}}\,\hat{P}_{j_1\dots j_n}}_{\hat{\mathcal{A}}\,\hat{P}_{j_1\dots j_n} = \sigma_{j_1\dots j_n}\,\hat{\mathcal{A}}} \phi_1(1)\dots\phi_n(n) \\
&= \underbrace{\sum_{\hat{P}_{j_1\dots j_n}}^{n!} U_{1j_1}\dots U_{nj_n}\sigma_{j_1\dots j_n}}_{\det \mathsf{U}}\,\Psi \\
&= \det \mathsf{U}\Psi\,.
\end{aligned} \tag{3.32}$$

Finally, we have

$$\begin{aligned}
\left\langle \Psi' \middle| \hat{Q} \middle| \Psi' \right\rangle &= \left\langle \det \mathsf{U}\Psi \middle| \hat{Q} \middle| \det \mathsf{U}\Psi \right\rangle \\
&= |\det \mathsf{U}|^2 \left\langle \Psi \middle| \hat{Q} \middle| \Psi \right\rangle \\
&= \left\langle \Psi \middle| \hat{Q} \middle| \Psi \right\rangle\,.
\end{aligned} \tag{3.33}$$

□

Proposition 3.4 (Slater–Condon rule). Let Ψ_i^a denote the *single replacement* Slater determinant defined as

$$\Psi_i^a \coloneqq \hat{\mathcal{A}}\,\phi_1(1)\phi_2(2)\dots\phi_{i-1}(i-1)\phi_a(i)\phi_{i+1}(i+1)\dots\phi_n(n)\,, \tag{3.34}$$

and let Ψ_{ij}^{ab} denote the *double replacement* Slater determinant

$$\Psi_{ij}^{ab} \coloneqq \hat{\mathcal{A}}\,\phi_1\phi_2\dots\phi_{i-1}\phi_a\phi_{i+1}\dots\phi_{j-1}\phi_b\phi_{j+1}\dots\phi_n\,, \tag{3.35}$$

where $\{\phi_1, \ldots, \phi_n, \phi_a, \phi_b\}$ are orthonormal. We have

$$\left\langle \Psi_i^a \middle| \hat{H} \middle| \Psi \right\rangle = h_{ai} + \sum_{j=1}^{n} [(ai|jj) - (aj|ij)] =: F_{ai} \,, \tag{3.36}$$

$$\left\langle \Psi_{ij}^{ab} \middle| \hat{H} \middle| \Psi \right\rangle = (ai|bj) - (aj|bi) \,. \tag{3.37}$$

Proof. (i) The matrix element is

$$\left\langle \Psi_i^a \middle| \hat{H} \middle| \Psi \right\rangle = \left\langle \phi_1 \phi_2 \ldots \phi_{i-1} \phi_a \phi_{i+1} \ldots \phi_n \middle| \hat{H} \middle| \sum_u \sigma_u \hat{P}_u \phi_1 \ldots \phi_{i-1} \phi_i \phi_{i+1} \ldots \phi_n \right\rangle . \tag{3.38}$$

The constant term in the Hamiltonian contributes nothing to the matrix element as

$$\langle \phi_a | \phi_j \rangle = 0 \quad \forall j \in \{1, \ldots, n\} \,. \tag{3.39}$$

Contribution from single electron hamiltonian must eliminate the zero overlap $\langle \phi_a | \phi_i \rangle$, so the only non-zero term comes from $\hat{h}(i)$ and $\hat{P}_u = \hat{E}$, where the contribution is

$$\left\langle \phi_a(i) \middle| \hat{h}(i) \middle| \phi_i(i) \right\rangle_i = h_{ai} \,. \tag{3.40}$$

For the two-electron term, we still need to sandwich the operator between orbitals whose overlap is zero. Thus the non-zero terms are from $\frac{1}{r_{ij}}$ for some $j \in \{1, \ldots, n\} \setminus \{i\}$ and when $\hat{P}_u = \hat{E}$ or $\hat{P}_{ij}$. We obtain

$$\left\langle \phi_a(i)\phi_j(j) \middle| \frac{1}{r_{ij}} \middle| \phi_i(i)\phi_j(j) \right\rangle_{ij} = \langle aj|ij \rangle = (ai|jj) \,, \tag{3.41}$$

$$-\left\langle \phi_a(i)\phi_j(j) \middle| \frac{1}{r_{ij}} \middle| \phi_j(i)\phi_i(j) \right\rangle_{ij} = -\langle aj|ji \rangle = -(aj|ij) \,. \tag{3.42}$$

Hence the full result is

$$\left\langle \Psi_i^a \middle| \hat{H} \middle| \Psi \right\rangle = h_{ai} + \sum_{j=1, j\neq i}^{n} [(ai|jj) - (aj|ij)] =: F_{ai} \,. \tag{3.43}$$

The $j \neq i$ condition may be omitted as this term vanishes.

(ii) For the double replacement matrix element, notice that there must be no contribution from the constant term and single-electron terms in the Hamiltonian, since there is no way to eliminate zero overlaps $\langle \phi_a | \phi_i \rangle$ and $\langle \phi_b | \phi_j \rangle$ at the same time. Thus we must use the term $\frac{1}{r_{ij}}$ from $\hat{H}$, and $\hat{P}_u = \hat{E}$ or $\hat{P}_{ij}$. These are

$$\left\langle \phi_a(i)\phi_b(j) \middle| \frac{1}{r_{ij}} \middle| \phi_i(i)\phi_j(j) \right\rangle_{ij} \langle \phi_1|\phi_1 \rangle_1 \ldots \langle \phi_n|\phi_n \rangle_n = \langle ab|ij \rangle = (ai|bj) \tag{3.44}$$

$$-\left\langle \phi_a(i)\phi_b(j) \middle| \frac{1}{r_{ij}} \middle| \phi_j(i)\phi_i(j) \right\rangle_{ij} \langle \phi_1|\phi_1 \rangle_1 \ldots \langle \phi_n|\phi_n \rangle_n = -\langle ab|ji \rangle = -(aj|bi) \tag{3.45}$$

Hence the full result is

$$\left\langle \Psi_{ij}^{ab} \middle| \hat{H} \middle| \Psi \right\rangle = (ai|bj) - (aj|bi) \,. \tag{3.46}$$

An application of this arises in Møller–Plesset theory. □

Remark. Since $\hat{H}$ contains at most two electron operators, any triple (or higher) replacement must have

$$\left\langle \Psi_{ijk}^{abc} \middle| \hat{H} \middle| \Psi \right\rangle = 0 \,. \tag{3.47}$$

Theorem 3.5.

(i) In general spin-orbitals $\Psi = \hat{\mathcal{A}}(\phi_1\phi_2\dots\phi_n)$ for an n electron system,

$$E = \left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle = \sum_i^n h_{ii} + \frac{1}{2}\sum_{i,j}^n [(ii|jj) - (ij|ij)]\,. \tag{3.48}$$

(ii) For a closed-shell $2n$ electron system $\Psi = \hat{\mathcal{A}}(\phi_1^\alpha\phi_1^\beta\phi_2^\alpha\phi_2^\beta\dots\phi_n^\alpha\phi_n^\beta)$,

$$E = 2\sum_i^n h_{ii} + \sum_{i,j}^n [2\,(ii|jj) - (ij|ij)]\,. \tag{3.49}$$

(iii) For the high spin open shell molecule $\Psi = \hat{\mathcal{A}}(\phi_1^\alpha\phi_1^\beta\phi_2^\alpha\phi_2^\beta\dots\phi_n^\alpha\phi_n^\beta\phi_{n+1}^\alpha\dots\phi_{n+p}^\alpha)$, denoting $i,j,\dots$ as doubly occupied orbitals and $l,m,\dots$ as singly occupied orbitals,

$$\begin{aligned} E = {} & 2\sum_{i=1}^n h_{ii} + \sum_{l=n+1}^{n+p} h_{ll} + \sum_{i,j=1}^n [2\,(ii|jj) - (ij|ij)] \\ & + \sum_{i=1}^n \sum_{l=n+1}^{n+p} [2\,(ll|ii) - (li|li)] + \frac{1}{2}\sum_{l,m=n+1}^{n+p} [(ll|mm) - (lm|lm)]\,. \end{aligned} \tag{3.50}$$

Proof.

(i) We already have the result

$$E = \sum_i^n h_{ii} + \sum_{i>j}^n [(ii|jj) - (ij|ij)] \tag{3.51}$$

from Theorem 3.2. We only need to note that when $i = j$, the contribution

$$(ii|ii) - (ii|ii) = 0 \tag{3.52}$$

vanishes. This is known as *self-interaction cancellation*. Thus we can rewrite the sum as

$$E = \sum_i^n h_{ii} + \frac{1}{2}\sum_{i,j=1}^n [(ii|jj) - (ij|ij)]\,. \tag{3.53}$$

(ii) This time the label $i,j,\dots$ is for spatial orbitals instead of spin-orbitals, so each i and j can have α or β spin. The one-electron Hamiltonian contribution is trivial. Now consider the two-electron interaction part. For each particular i,j in the sum, we have:

- four contributions to $(ii|jj)$, with spin parts

$$\langle\alpha|\alpha\rangle_i \langle\alpha|\alpha\rangle_j \ , \ \langle\alpha|\alpha\rangle_i \langle\beta|\beta\rangle_j \ , \ \langle\beta|\beta\rangle_i \langle\alpha|\alpha\rangle_j \ \text{ and } \ \langle\beta|\beta\rangle_i \langle\beta|\beta\rangle_j \ . \tag{3.54}$$

- two contributions to $(ij|ij)$, with spin parts

$$\langle\alpha|\alpha\rangle_i \langle\alpha|\alpha\rangle_j \ \text{ and } \ \langle\beta|\beta\rangle_i \langle\beta|\beta\rangle_j \ . \tag{3.55}$$

  Note that the terms $\left(\phi_i^\alpha\phi_j^\beta\middle|\phi_i^\alpha\phi_j^\beta\right)$ and $\left(\phi_i^\beta\phi_j^\alpha\middle|\phi_i^\beta\phi_j^\alpha\right)$ vanishes because they have vanishing spin inner products

$$\langle\alpha|\beta\rangle_i \langle\alpha|\beta\rangle_j \ \text{ and } \ \langle\beta|\alpha\rangle_i \langle\beta|\alpha\rangle_j \ . \tag{3.56}$$

- for the special case where $i = j$, we get two contributions in total, which are $\left(\phi_i^\alpha\phi_i^\alpha\middle|\phi_i^\beta\phi_i^\beta\right)$ and $\left(\phi_i^\beta\phi_i^\beta\middle|\phi_i^\alpha\phi_i^\alpha\right)$.

Hence we get the expression

$$E = 2\sum_{i}^{n} h_{ii} + \sum_{i,j}^{n} [2\,(ii|jj) - (ij|ij)]\,. \tag{3.57}$$

(iii) The one-electron Hamiltonian part is obviously

$$2\sum_{i=1}^{n} h_{ii} + \sum_{l=n+1}^{n+p} h_{ll}\,. \tag{3.58}$$

Now consider the two-electron integral. The contributions from the interactions between closed-shell electrons are the same as (ii), given by

$$\sum_{i,j=1}^{n} [2\,(ii|jj) - (ij|ij)]\,. \tag{3.59}$$

Next consider the interaction between a closed-shell electron in spatial orbital ϕ_i with a electron in open shell orbit ϕ_l^α. The ϕ_i electron can have either α or β spin, contributing two $(ll|ii)$ term and only one $(li|li)$ term from ϕ_i^α, as the other $(li|li)$ term

$$\left\langle \phi_i^\beta(i)\phi_l^\alpha(l)\left|\frac{1}{r_{12}}\right|\phi_i^\beta(l)\phi_l^\alpha(i)\right\rangle = (li|li)\,\langle\beta|\alpha\rangle_i\,\langle\alpha|\beta\rangle_j \tag{3.60}$$

has a vanishing spin part inner product. Hence, the contribution is

$$\sum_{i=1}^{n}\sum_{l=n+1}^{n+p} [2\,(ll|ii) - (li|li)]\,. \tag{3.61}$$

Finally, consider the interactions between two open shell electrons that both have spin α. The contribution is easily determined as

$$\frac{1}{2}\sum_{l,m=n+1}^{n+p} [(ll|mm) - (lm|lm)]\,. \tag{3.62}$$

Combining all the contributions, we have

$$\begin{aligned} E = 2\sum_{i=1}^{n} h_{ii} + \sum_{l=n+1}^{n+p} h_{ll} + \sum_{i,j=1}^{n} [2\,(ii|jj) - (ij|ij)] \\ + \sum_{i=1}^{n}\sum_{l=n+1}^{n+p} [2\,(ll|ii) - (li|li)] + \frac{1}{2}\sum_{l,m=n+1}^{n+p} [(ll|mm) - (lm|lm)] \end{aligned} \tag{3.63}$$

as claimed. □

4 Hartree–Fock Theory

4.1 Closed Shell Hartree–Fock

The energy of a closed-shell molecule is

$$E = \left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle = 2\sum_i h_{ii} + \sum_{i,j} [2\,(ii|jj) - (ij|ij)] \tag{4.1}$$

with $\Psi = \hat{\mathcal{A}}(\phi_1^2 \dots \phi_n^2)$. Pick a basis set $\{\eta_\mu\}_{\mu=1}^m$ such that each orbital ϕ_i is expressed in terms of the basis functions as

$$\phi_i = \sum_{\mu=1}^m \eta_\mu c_{\mu i}\,. \tag{4.2}$$

We want to find the orbitals which make the energy stationary with respect to variations of the molecular orbital coefficients $c_{\mu i}$ while maintaining the orbital orthonormality.

We may work out explicitly $\partial E/\partial c_{\mu i}$ and set it to zero while keeping $\langle \phi_i | \phi_j \rangle = \delta_{ij}$. Here we will use another method.

Suppose we have found these n orbitals $\{\phi_i\}_{i=1}^n$, then since there are m basis functions, there will be $(m-n)$ other orbitals $\{\phi_a\}_{a=n+1}^m$ (called *unoccupied* or *virtual orbitals*) which obey $\langle \phi_a | \phi_i \rangle = 0$, $1 \le i \le n$, $n+1 \le a \le m$. We therefore derive the condition that E is stationary with respect to the variation

$$\phi_k \to \phi_k + \epsilon\phi_a \quad (k = 1, \dots, n\,;\ a = n+1, \dots, m)\,. \tag{4.3}$$

Such variation will automatically maintain the orbital orthonormality to first order in ϵ since

$$\langle \phi_k + \epsilon\phi_a | \phi_i \rangle = \delta_{ik} + \epsilon \cancelto{0}{\langle \phi_a | \phi_i \rangle}\,. \tag{4.4}$$

We therefore substitute (4.3) into (4.1), pick out the coefficient of ϵ and set it to zero. Doing this for the one-electron part gives

$$h'_{kk} \coloneqq \left\langle \phi_k + \epsilon\phi_a \middle| \hat{h} \middle| \phi_k + \epsilon\phi_a \right\rangle = h_{kk} + \epsilon(h_{ak} + h_{ka}) + \epsilon^2 h_{aa}\,. \tag{4.5}$$

The coefficient is therefore $2h_{ak}$ by Hermiticity. Similarly,

$$(kk|jj)' \coloneqq (k+\epsilon a\ k+\epsilon a|jj) = (kk|jj) + \epsilon[(ka|jj) + (ak|jj)] + \mathcal{O}(\epsilon^2)\,, \tag{4.6}$$

$$(ii|kk)' \coloneqq (ii|k+\epsilon a\ k+\epsilon a) = (ii|kk) + \epsilon[(ii|ka) + (ii|ak)] + \mathcal{O}(\epsilon^2)\,, \tag{4.7}$$

$$(kj|kj)' \coloneqq (k+\epsilon a\ j|k+\epsilon a\ j) = (kj|kj) + \epsilon[(aj|kj) + (kj|aj)] + \mathcal{O}(\epsilon^2)\,, \tag{4.8}$$

$$(ik|ik)' \coloneqq (i\ k+\epsilon a|i\ k+\epsilon a) = (ik|ik) + \epsilon[(ik|ia) + (ia|ik)] + \mathcal{O}(\epsilon^2)\,. \tag{4.9}$$

Using the appropriate symmetries of two-electron integrals, and by replacing $\sum_i$ with $\sum_j$ when appropriate, we get the first order variation in energy

$$\delta E = 4h_{ak} + \sum_{j=1}^n [8\,(ak|jj) - 4\,(aj|kj)]\,. \tag{4.10}$$

Hence, the stationary condition is

$$h_{ak} + \sum_{j=1}^n [2\,(ak|jj) - (aj|kj)] = 0 \tag{4.11}$$

for every $k = 1, \dots, n$, $a = n+1, \dots, m$. We then define the *Fock Hamiltonian* $\hat{F}$ to be the operator such that

$$\left\langle \phi_a \middle| \hat{F} \middle| \phi_k \right\rangle \coloneqq h_{ak} + \sum_{j=1}^n [2\,(ak|jj) - (aj|kj)]\,. \tag{4.12}$$

Then the Fock Hamiltonian can be written as

$$\hat{F}(\mathbf{x}) = \hat{h}(\mathbf{x}) + 2\sum_{j=1}^{n} \int \mathrm{d}^3\mathbf{y}\, \frac{|\phi_j(\mathbf{y})|^2}{\|\mathbf{x}-\mathbf{y}\|} - \sum_{j=1}^{n} \int \mathrm{d}^3\mathbf{y}\, \frac{\phi_j^*(\mathbf{y})\phi_j(\mathbf{x})}{\|\mathbf{x}-\mathbf{y}\|} \hat{P}_{xy}\,, \tag{4.13}$$

where $\hat{P}_{xy}\phi_k(\mathbf{x}) = \phi_k(\mathbf{y})$.

The Fock Hamiltonian is therefore an effective one-electron Hamiltonian, including kinetic, nuclear attraction and an average potential term made up of Coulomb and exchange parts.

Therefore, from equations (4.11) and (4.12), the condition that the energy E is stationary with respect to variations of the molecular orbital coefficients is

$$F_{ak} \coloneqq \left\langle \phi_a \middle| \hat{F} \middle| \phi_k \right\rangle = 0 \qquad (k = 1, \ldots, n\,,\ a = n+1, \ldots, m)\,. \tag{4.14}$$

Written in matrix form, it is a necessary and sufficient condition that the *Fock matrix* is in the block diagonal form

$$\mathsf{F} = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ n{+}1 \\ \vdots \\ m \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & 2 & \cdots & n{+}1 & \cdots & m \end{array} \\ \begin{pmatrix} F_{ij} & & & 0 & & \\ & \ddots & & & & \\ & & \ddots & & & \\ 0 & & & F_{ab} & & \end{pmatrix} \end{array}\,. \tag{4.15}$$

Then a clearly sufficient condition is that the Fock matrix F is fully diagonalised, i.e.

$$\mathsf{F} = \begin{pmatrix} \epsilon_1 & & & & & \\ & \epsilon_2 & & 0 & & \\ & & \ddots & & & \\ & & & \epsilon_{n+1} & & \\ & 0 & & & \ddots & \\ & & & & & \epsilon_m \end{pmatrix}\,. \tag{4.16}$$

The resulting orbitals $\{\phi_i\}$, satisfying

$$\hat{F}\phi_i = \epsilon_i \phi_i\,, \tag{4.17}$$

or equivalently

$$F_{ji} = \left\langle \phi_j \middle| \hat{F} \middle| \phi_i \right\rangle = \epsilon_i \langle \phi_j | \phi_i \rangle = \epsilon_i \delta_{ij} \tag{4.18}$$

is known as the *canonical Hartree–Fock orbitals*. If we expand it in the non-orthogonal atomic orbital basis as $\phi_i = \sum_\mu \eta_\mu c_{\mu i}$, then

$$\begin{aligned} \sum_{\mu\nu} c_{\mu j}^* c_{\nu i} \left\langle \eta_\mu \middle| \hat{F} \middle| \eta_\nu \right\rangle &= \epsilon_i \sum_{\mu\nu} c_{\mu j}^* c_{\nu i} \langle \eta_\mu | \eta_\nu \rangle \\ \mathbf{c}_j^\dagger \mathsf{F} \mathbf{c}_i &= \epsilon_i \mathbf{c}_j^\dagger \mathsf{S} \mathbf{c}_i \\ \mathsf{F}\mathbf{c}_i &= \epsilon_i \mathsf{S}\mathbf{c}_i\,, \end{aligned} \tag{4.19}$$

where F is in the AO basis and $S_{\mu\nu} = \langle \eta_\mu | \eta_\nu \rangle$ is the overlap integral. This can be rewritten into the *canonical secular equations*

$$\sum_{\nu=1}^{m} \left\langle \eta_\mu \middle| \hat{F} - \epsilon_i \middle| \eta_\nu \right\rangle c_{\nu i} = 0\,. \tag{4.20}$$

Since $\hat{F}$ is an effective Hamiltonian, and ϵ_i are the corresponding energies of the orbitals ϕ_i, we choose the lowest n eigensolutions to be the occupied orbitals and the remaining $(m-n)$ as unoccupied virtual orbitals.

Note that $\hat{F}$ itself involves the occupied orbitals ϕ_j, so an iterative process is therefore required to solve the *self-consistent field* equations:

(i) Select the geometry of the molecule and the basis set $\{\eta_\mu\}$;

(ii) Evaluate the basis function integrals $h_{\mu\nu}$, $(\mu\nu|\sigma\tau)$ and $S_{\mu\nu}$;

(iii) Guess some coefficients $c_{\mu i}$ for the occupied orbitals;

(iv) Form the *density matrix* $D_{\mu\nu} = \sum_{i=1}^{n} c_{\mu i} c_{\nu i}^*$, where the sum is taken over occupied spatial orbitals;

(v) Construct the Fock matrix

$$F_{\mu\nu} = \left\langle \eta_\mu \middle| \hat{F} \middle| \eta_\nu \right\rangle = h_{\mu\nu} + \sum_{\sigma,\tau=1}^{m} D_{\tau\sigma}[2\,(\mu\nu|\sigma\tau) - (\mu\sigma|\nu\tau)]\,; \tag{4.21}$$

(vi) Solve the secular equations. Determine the new density matrix D. If D has changed (greater than a sufficiently small tolerance), return to (v). If not, proceed to (vii);

(vii) Calculate the energy through

$$E = 2\sum_{\mu,\nu=1}^{m} D_{\nu\mu} h_{\mu\nu} + \sum_{\mu,\nu,\sigma,\tau=1}^{m} D_{\nu\mu} D_{\tau\sigma}[2\,(\mu\nu|\sigma\tau) - (\mu\sigma|\nu\tau)]\,. \tag{4.22}$$

This is the *Hartree–Fock* (HF) or *self-consistent field* (SCF) method.[2]

Proposition 4.1. The computational scaling of an RHF cycle with $2n$ electrons in m spatial atomic orbitals is $\mathcal{O}(m^4)$.

Proof. We consider the computational scaling of each step.

(i) Select geometry and basis set.
This creates a list of atomic orbitals $\{\eta_\mu\}_{i=1}^{m}$. $\mathcal{O}(m)$

(ii) Evaluate atomic basis function integrals.
There are m^2 integrals of types $S_{\mu\nu}$ and $h_{\mu\nu}$. $\mathcal{O}(m^2)$
There are m^4 integrals of type $(\mu\nu|\sigma\tau)$. $\mathcal{O}(m^4)$

(iii) Guess the initial MO coefficients.
This is a $m \times n$ matrix. $\mathcal{O}(mn)$.

(iv) Form the density matrix

$$D_{\mu\nu} = \sum_{i=1}^{n} c_{\mu i} c_{\nu i}^*\,. \tag{4.24}$$

This is a $m \times m$ matrix, each element is a sum over n orbitals. $\mathcal{O}(m^2 n)$

(v) Construct the Fock matrix

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{\sigma,\tau}^{m} D_{\tau\sigma}[2\,(\mu\nu|\sigma\tau) - (\mu\sigma|\nu\tau)]\,. \tag{4.25}$$

This is a $m \times m$ matrix, each element is a sum over $m \times m$ terms. $\mathcal{O}(m^4)$

[2] You will prove these expressions in the exercises. We often define the auxiliary *interaction matrix*

$$G_{\mu\nu} = \sum_{\sigma,\tau=1}^{m} D_{\sigma\tau}[2\,(\mu\nu|\sigma\tau) - (\mu\sigma|\nu\tau)]\,, \tag{4.23}$$

so that $F_{\mu\nu} = h_{\mu\nu} + G_{\mu\nu}$ and $E = 2\sum_{\mu,\nu=1}^{m} D_{\mu\nu} h_{\mu\nu} + \sum_{\mu,\nu=1}^{m} D_{\mu\nu} G_{\mu\nu}$.

(vi) Solve the secular equation

$$\mathsf{Fc} = \epsilon \mathsf{Sc}\,. \tag{4.26}$$

This is a matrix diagonalisation. The best algorithm does it in $\mathcal{O}(\dim^3)$. $\mathcal{O}(m^3)$

The overall scaling is therefore $\mathcal{O}(m^4)$. Step (v) is the most computationally expensive step. Although step (ii) and step (v) are both $\mathcal{O}(m^4)$, (ii) only needs to be done once while (v) needs to be done in every iteration. □

Theorem 4.2 (Koopmans' theorem). In closed-shell restrictive Hartree–Fock theory, the first ionisation energy of a molecular system is equal to the negative of the orbital energy of the highest occupied molecular orbital.

Proof. We first use the *frozen orbital approximation*, which assumes that the orbitals do not relax after ionisation. Assume that n electrons in a closed-shell system, then by the definition of the Fock matrix (4.12), the energy of the i^{th} orbital is

$$\epsilon_i = \left\langle \phi_i \middle| \hat{F} \middle| \phi_i \right\rangle = h_{ii} + \sum_{j=1}^{n/2} [2\,(ii|jj) - (ij|ij)]\,. \tag{4.27}$$

This is a summation over spatial orbitals, and the Coulomb and exchange terms do not include any spin integrations. We may rewrite this as a sum over n spin-orbitals as

$$\epsilon_i = h_{ii} + \sum_{j=1}^{n} [(ii|jj) - (ij|ij)]\,. \tag{4.28}$$

The RHF energy of the system with all n electrons is

$$E_n = \sum_{i=1}^{n} h_{ii} + \sum_{i<j}^{n} [(ii|jj) - (ij|ij)]\,. \tag{4.29}$$

By assumption, the energy of the ionised system is the same summation up to the $(n-1)^{\text{th}}$ electron, which is

$$E_{n-1} = \sum_{i=1}^{n-1} h_{ii} + \sum_{i<j}^{n-1} [(ii|jj) - (ij|ij)]\,. \tag{4.30}$$

Then the ionisation energy is

$$\begin{aligned} \mathrm{IE}_1 &= E_{n-1} - E_n \\ &= -h_{nn} - \sum_{i=1}^{n-1} [(ii|nn) - (in|in)] \\ &= -h_{nn} - \sum_{i=1}^{n} [(nn|ii) - (in|in)] \\ &= -\epsilon_n\,. \end{aligned} \tag{4.31}$$

Note that this result does not include electron correlation or orbital relaxation, so it is far from being exact. □

Proposition 4.3. The density matrix obeys the *idempotency condition*

$$\mathsf{DSD} = \mathsf{D}\,, \tag{4.32}$$

and at the convergence of the SCF method,

$$\mathsf{SDF} = \mathsf{FDS}\,. \tag{4.33}$$

Proof.

$$\begin{aligned}(DSD)_{\mu\nu} &= c_{\mu i}c^*_{ji}\langle\eta_j|\eta_k\rangle c_{kl}c^*_{\nu l}\\ &= c_{\mu i}\langle c_{ji}\eta_j|c_{kl}\eta_k\rangle c^*_{\nu l}\\ &= c_{\mu i}\langle\phi_i|\phi_l\rangle c^*_{\nu l}\\ &= c_{\mu i}\delta_{il}c^*_{\nu l}\\ &= c_{\mu i}c^*_{\nu i}\\ &= D_{\mu\nu}\,. \end{aligned} \tag{4.34}$$

At convergence,

$$\mathsf{FC} = \epsilon\mathsf{SC}\,, \tag{4.35}$$

where ϵ is the diagonal matrix of orbital energies. Hence

$$\mathsf{FCC}^\dagger = \epsilon\mathsf{SCC}^\dagger \tag{4.36}$$

$$\mathsf{FD} = \epsilon\mathsf{SD}\,. \tag{4.37}$$

$$\begin{aligned}\mathsf{FDS} &= \epsilon\mathsf{SDS}\\ &= \mathsf{SD}\epsilon\mathsf{S}\,. \end{aligned} \tag{4.38}$$

$$\begin{aligned}\mathsf{FDSD} &= \mathsf{SD}\epsilon\mathsf{SD}\\ &= \mathsf{SDFD}\,. \end{aligned} \tag{4.39}$$

Therefore,

$$\mathsf{FDS} = \mathsf{SDF}\,. \tag{4.40}$$

□

4.2 Open Shell High Spin Hartree–Fock Theory (Non-examinable)

The energy of the high spin open shell Slater determinant of form $\hat{\mathcal{A}}(\phi_1^2\ldots\phi_N^2\phi_{N+1}^\alpha\ldots\phi_{N+P}^\alpha)$ is

$$\begin{aligned}E = &\sum_{i=1}^{N} h_{ii} + \sum_{i,j=1}^{N}\left[2\,(ii|jj) - (ij|ij)\right]\\ &+ \sum_{m=N+1}^{N+P} h_{mm} + \frac{1}{2}\sum_{m,n=N+1}^{N+P}\left[(mm|nn) - (mn|mn)\right]\\ &+ \sum_{i=1}^{N}\sum_{m=1}^{N+P}\left[2\,(mm|ii) - (im|im)\right]. \end{aligned} \tag{4.41}$$

We will use $i, j, \ldots$ to denote doubly occupied orbitals, $m, n, \ldots$ to denote singly occupied orbitals and $a, b, \ldots$ to denote the unoccupied orbitals.

To obtain the conditions for optimised orbitals, we consider the variations

(i) $$\phi_i \to \phi_i + \epsilon\phi_a\,, \tag{4.42}$$

(ii) $$\phi_m \to \phi_m + \epsilon\phi_a\,, \tag{4.43}$$

(iii) $$\phi_i \to \phi_i + \epsilon\phi_m\,,$$

$$\phi_m \to \phi_m - \epsilon\phi_i\,. \tag{4.44}$$

On substitution of these into the energy expression and the conditions $\frac{\partial E}{\partial \epsilon}\big|_{\epsilon=0}$ it follows that the self consistent conditions are

(i) $$F_{ai} = 0\,, \tag{4.45}$$

(ii) $$F_{ma} - \tfrac{1}{2}K^O_{ma} = 0\,, \tag{4.46}$$

(iii) $$F_{mi} + \tfrac{1}{2}K^O_{mi} = 0\,, \tag{4.47}$$

where

$$\mathsf{F} \coloneqq \mathsf{h} + 2\mathsf{J}^C - \mathsf{K}^C + \mathsf{J}^O - \frac{1}{2}\mathsf{K}^O\,, \tag{4.48}$$

and we have introduced the notations

$$J^C_{pq} \coloneqq \sum_{i=1}^{N} (pq|ii)\ , \tag{4.49}$$

$$K^C_{pq} \coloneqq \sum_{i=1}^{N} (pi|qi)\ , \tag{4.50}$$

$$J^O_{pq} \coloneqq \sum_{m=N+1}^{N+P} (pq|mm)\ , \tag{4.51}$$

$$K^O_{pq} \coloneqq \sum_{m=N+1}^{N+P} (pm|qm)\ . \tag{4.52}$$

We can define the *alpha* and *beta Fock matrices* as follows

$$\mathsf{F}^\alpha \coloneqq \mathsf{F} - \frac{1}{2}\mathsf{K}^O\,, \tag{4.53}$$

$$\mathsf{F}^\beta \coloneqq \mathsf{F} + \frac{1}{2}\mathsf{K}^O\,. \tag{4.54}$$

The iterative solutions for this ROHF model are the construction and diagonalisation of the following matrix

$$\begin{array}{c} \\ i,j,\ldots \\ m,n,\ldots \\ a,b,\ldots \end{array}
\begin{array}{c}
\overbrace{}^{i,j,\ldots}\ \overbrace{}^{m,n,\ldots}\ \overbrace{}^{a,b,\ldots} \\
\left(\begin{array}{ccc}
 & \mathsf{F}^\beta & \tfrac{1}{2}(\mathsf{F}^\alpha + \mathsf{F}^\beta) \\
\mathsf{F}^\beta & & \mathsf{F}^\alpha \\
\tfrac{1}{2}(\mathsf{F}^\alpha + \mathsf{F}^\beta) & \mathsf{F}^\alpha &
\end{array}\right)
\end{array}\ , \tag{4.55}$$

where the three rows and columns refer to doubly occupied, singly occupied and virtual orbitals. The diagonal blocks may be defined in any convenient ways — the standard choices are $(\mathsf{F} - \mathsf{K}^O)$, F and $(\mathsf{F} + \mathsf{K}^O)$ on the respective diagonal.

4.3 Unrestricted Hartree–Fock Theory

For a system with α spin-alpha electrons and β spin-beta electrons, *unrestricted Hartree–Fock* (UHF) finds the determinant with the lowest energy of the form

$$E^{\mathrm{UHF}} = \sum_{i}^{\alpha+\beta} h_{ii} + \frac{1}{2}\sum_{i,j=1}^{\alpha+\beta} (ii|jj) - \frac{1}{2}\sum_{i,j=1}^{\alpha} (ij|ij) - \frac{1}{2}\sum_{i,j=1}^{\beta} (ij|ij) \tag{4.56}$$

while allowing the alpha and beta orbitals to have different spatial parts. This expression accounts for the fact that both α and β electrons can interact with each other, while spin-alpha electrons only exchange with spin-alpha electrons, and spin-beta electrons only exchange with spin-beta electrons.

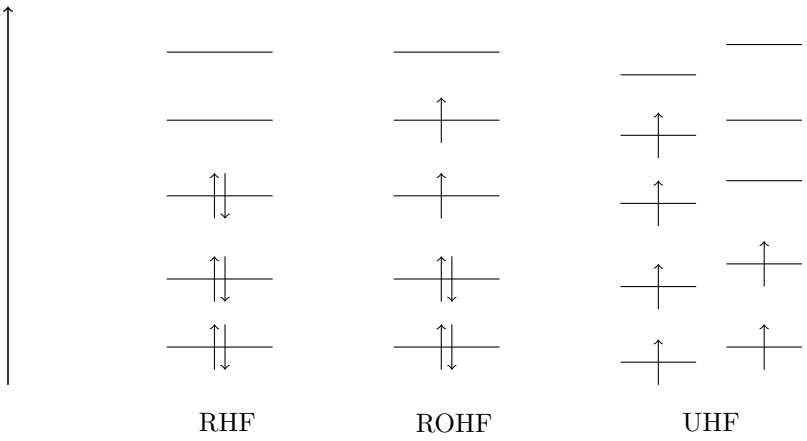


Figure 4.1: A schematic example of an RHF, an ROHF and a UHF system.

In UHF, two sets of SCF equations are simultaneously solved, using the alpha and beta Fock operators respectively

$$\hat{F}^{\alpha}(\mathbf{x}) = \hat{h}(\mathbf{x}) + \sum_{j=1}^{\alpha+\beta} \int \mathrm{d}^3\mathbf{y}\, \frac{\phi_j^2(\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|} - \sum_{j=1}^{\alpha} \int \mathrm{d}^3\mathbf{y}\, \frac{\phi_j(\mathbf{y})\phi_j(\mathbf{x})}{\|\mathbf{x}-\mathbf{y}\|} \hat{P}_{xy}\,, \tag{4.57}$$

$$\hat{F}^{\beta}(\mathbf{x}) = \hat{h}(\mathbf{x}) + \sum_{j=1}^{\alpha+\beta} \int \mathrm{d}^3\mathbf{y}\, \frac{\phi_j^2(\mathbf{y})}{\|\mathbf{x}-\mathbf{y}\|} - \sum_{j=1}^{\beta} \int \mathrm{d}^3\mathbf{y}\, \frac{\phi_j(\mathbf{y})\phi_j(\mathbf{x})}{\|\mathbf{x}-\mathbf{y}\|} \hat{P}_{xy}\,. \tag{4.58}$$

4.3.1 Symmetry Breaking

Note that since the α and β electrons are no longer required to share the same set of spatial orbitals in UHF, it has more degrees of freedom than RHF and ROHF. Therefore, E^{UHF} is generally lower than E^{RHF}. We take the dissociation of H_2 as an example. H_2 will dissociate into two H atoms. However, in RHF, we restrict the two electrons in the ground electronic state with opposite spins to have the same spatial wavefunction. It is sometimes described that there are two half electrons associated with each nucleus, so at dissociation ($R_{\mathrm{H-H}} \to \infty$), the two electrons still have Coulomb interaction. Therefore, RHF predicts a higher energy than reality at dissociation. This can be fixed by UHF, in which the two electrons may take different spatial wavefunctions. This is shown in the figure below.

We can see that the main advantage of the UHF wavefunction is its simplicity and also that since the alpha and beta orbitals may be different, alpha and beta electrons avoid one another and therefore said to be correlated. No such correlation is present in RHF wavefunctions.

An objection to UHF theory is that the wavefunction is not an eigenfunction of $\hat{S}^2$. The spin contamination may be measured from the expectation value of $\hat{S}^2$

$$\left\langle \Phi \middle| \hat{S}^2 \middle| \Phi \right\rangle = M_S(M_S+1) + n^{\beta} - \sum_{i}^{\alpha}\sum_{j}^{\beta} (S_{ij}^{\alpha\beta})^2\,, \tag{4.59}$$

where $M_S \coloneqq \frac{1}{2}(n^{\alpha} - n^{\beta})$ and $S_{ij}^{\alpha\beta} = \left\langle \phi_i^{\alpha} \middle| \phi_j^{\beta} \right\rangle$.

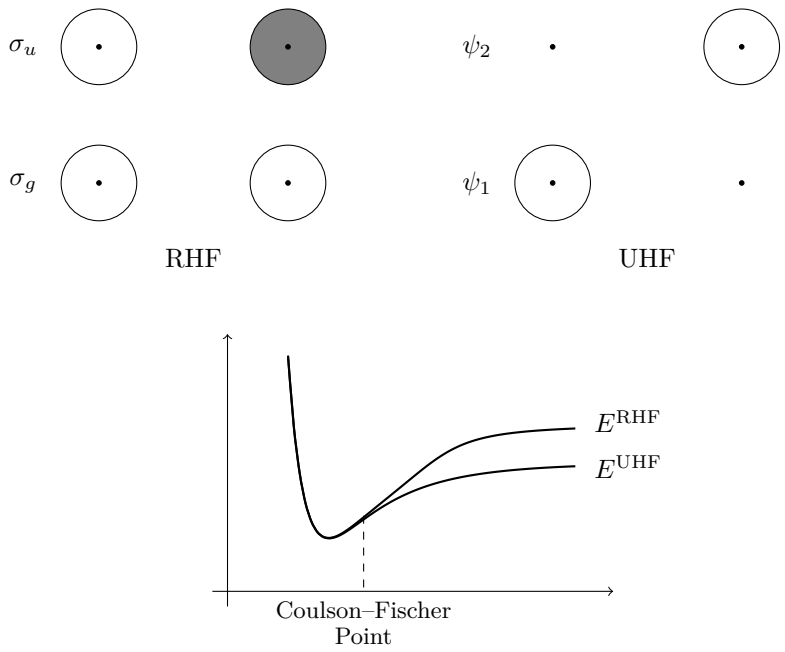


Figure 4.2: UHF allows α and β spin electrons to take different spatial orbitals. This makes UHF significantly better than RHF when calculating dissociation.

5 Electron Correlation

The self-consistent field method is an independent electron approximation. This is understood by recognising that each orbital is an eigenfunction of the same Fock Hamiltonian. Further, the wavefunction only depends upon the absolute coordinate of each electron, and not on the distance between the electrons. The motion of the electrons is therefore not correlated. The difference between the SCF energy using a large basis set (the so-called Hartree–Fock energy) and the exact energy of the Schrödinger Hamiltonian is called the correlation energy

$$E_{\mathrm{corr}} = E_{\mathrm{exact}} - E_{\mathrm{RHF}} < 0\,. \tag{5.1}$$

Correlation energy is usually pretty significant, being approximately 0.04 Hartrees (100 kJ mol^{-1}) for two electrons in a doubly occupied orbital. Therefore if a single bond is broken, this quantity is released. It is therefore important to be able to calculate the value of the correlation energy for a molecular system.

A simple way forward is to introduce interelectronic distances $r \equiv r_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|$ into the wavefunction. We can see why this is a good idea. Consider two electrons of opposite spin moving in space. Noting that their reduced mass is $m/2$ and denoting r as the distance between them, the Schrödinger equation for their motion is

$$\left(-\nabla^2 + \frac{1}{r}\right)\psi = E\psi\,. \tag{5.2}$$

Expanding the Laplacian in spherical polar coordinates yields

$$\left(-\frac{\partial^2}{\partial r^2} - \frac{2}{r}\frac{\partial}{\partial r} + \frac{\hat{\mathbf{L}}^2}{r^2} + \frac{1}{r}\right)\psi = E\psi\,, \tag{5.3}$$

where $\hat{\mathbf{L}}$ is the angular momentum operator. We now want to find a series expansion solution for an s orbital near $r = 0$, so that we can ignore the $\hat{L}$ part. We write $\psi(r) = a + br + \mathcal{O}(r^2)$, substitute in

the equation and let the coefficient of r^{-1} be zero. This yields

$$\psi(r_{ij}) = a\left(1 + \frac{1}{2}r_{ij} + \mathcal{O}(r_{ij}^2)\right) . \tag{5.4}$$

This expression demonstrates the reduction in probability as two electrons approach one another, although of course because the electrons have different spins, the wavefunction is not zero at coincidence. Thus we have shown that the exact wavefunction must contain terms which are linear in interelectronic distances.

However, we cannot proceed this way because the matrix elements of wavefunctions would then involve $3n$-dimensional integrals, where n is the number of electrons. The main current way forward is to continue with the use of determinants. We can easily show that by taking a linear combination of determinants we can introduce the square of the interelectronic distances. Consider a four-determinant wavefunction of the He atom

$$\Psi = \sum_{I=1}^{4} c_I \, \hat{\mathcal{A}}(\phi_I^\alpha \phi_I^\beta) \, , \tag{5.5}$$

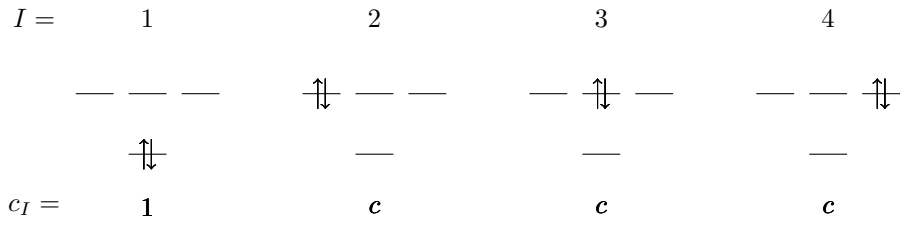
where

$$c_1 = 1 \, , \; c_2 = c_3 = c_4 = c \, . \tag{5.6}$$

$$\phi_1 = \exp(-2r) \, , \tag{5.7}$$

$$\phi_2 = x\phi_1 \, , \; \phi_3 = y\phi_1 \, , \; \phi_4 = z\phi_1 \, . \tag{5.8}$$

ϕ_1 corresponds to the 1s orbital and ϕ_2, ϕ_3, ϕ_4 corresponds to the $2p_x$, $2p_y$ and $2p_z$ orbitals respectively. Hence these four determinants correspond to the four states shown in the figure below.



Substitution yields

$$\Psi = \frac{1}{\sqrt{2}}(\alpha\beta - \beta\alpha)\mathrm{e}^{-2(r_1+r_2)}(1 + c(x_1x_2 + y_1y_2 + z_1z_2)) \, . \tag{5.9}$$

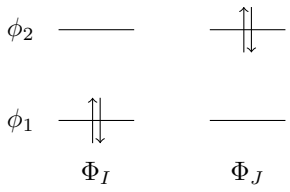
The result follows by noting

$$x_1x_2 + y_1y_2 + z_1z_2 = \mathbf{r}_1 \cdot \mathbf{r}_2 = \frac{1}{2}(r_1^2 + r_2^2 - r_{12}^2) \, . \tag{5.10}$$

The second, third and fourth determinants in the above example are each double replacements of the first determinant, and thus we have seen that double replacement determinants introduce the square of the interelectronic distances, and therefore introduce electron correlation effects into the wavefunction in the form

$$\Psi \sim (1 + r_{12}^2 + \dots) \, . \tag{5.11}$$

Note that there exist methods called *explicitly correlated*, often denoted f_{12} or $f(r_{12})$, which include r_{12} in linear order into the wavefunction, but they are beyond the scope of our course.



5.1 Full Configuration Interaction

The way forward is therefore to represent the wavefunction as a linear combination of many determinants. We now denote the Hartree–Fock determinant as Φ_0, and in general, we can make the excited determinant using the same set of orbitals.

Note that by the orthogonality of orbitals, we have

$$(\Phi_I|\Phi_J) = \delta_{IJ}\,. \tag{5.12}$$

In *configuration interaction* (CI), we write the wavefunction as

$$\Psi = \sum_I c_I \Phi_I\,. \tag{5.13}$$

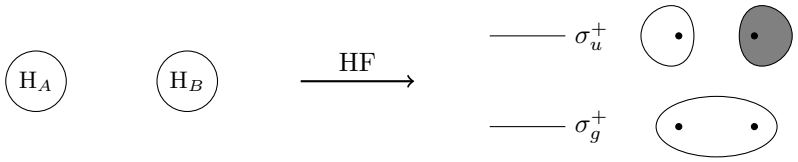
The CI coefficients c_I and the energy E are found by solving the secular equations

$$\sum_J \left\langle \Phi_I \middle| \hat{H} - E \middle| \Phi_J \right\rangle c_J = 0\,. \tag{5.14}$$

If one includes all possible determinants then this is called the *full configuration interaction* (FCI) method and it gives the best possible solution to the Schrödinger equation in the finite basis set we are using. However it is impractical for all but the tiniest system due to the exponentially large size of the matrix we have to diagonalise. Consider a system of m spatial orbitals and $2n$ electrons. The number of determinants we need to include is

$$\binom{2m}{2n} = \frac{(2m)!}{(2n)!(2m-2n)!} = \mathcal{O}(\mathrm{e}^{am})\,. \tag{5.15}$$

Example. Determine the symmetries of the ground and excited states in an FCI calculation for a minimal basis H_2 molecule.

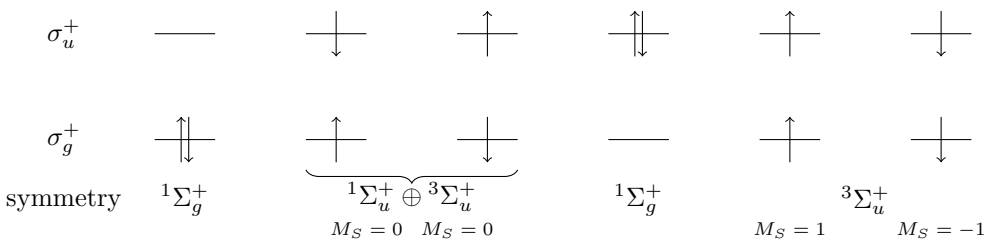


The number of determinants is

$$\binom{\#\text{ spin-orbitals}}{\#\, e^-} = \binom{4}{2} = 6\,. \tag{5.16}$$

5.2 Configuration Interaction with Singles and Doubles

The way forward is therefore to represent the wavefunction as a linear combination of many determinants, especially including all single and double replacements of the dominant SCF determinants



to give *configuration interaction with singles and doubles* (CISD). In practice it is found that the CI method is very slowly convergent, needing very large basis sets to generate a large number of virtual orbitals to use in the double replacement determinants. We can understand this: we have shown that the wavefunction requires terms linear in the interelectronic distance, but that CI only introduces the square of the interelectronic distance. It will then clearly take the CI wavefunction to involve many terms to approximate the true wavefunction. Even so, this is the only practical way forward. We have seen that the matrix element of the above secular equation can be evaluated in terms of one and two electron integrals. Almost all correlated wavefunction calculations are based on linear combination of determinants.

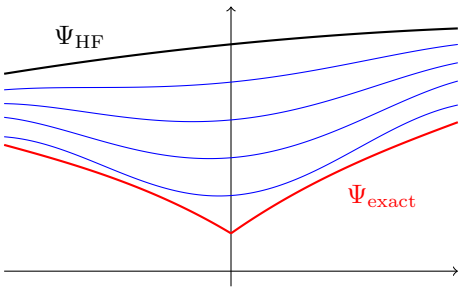


Figure 5.1: Increasing the size of basis set generates more virtual orbitals that allow us to include more terms in configuration interaction.

The introduction of higher angular momentum in basis results in better cusps. It can be shown from partial wave expansion that the truncation error is of the order

$$\Delta E \sim c(l_{\mathrm{trunc}} + 1)^{-3} \,. \tag{5.17}$$

We can extrapolate energies with increasing basis set.

5.3 Size Consistency

There is a further difficulty with straightforward truncated CI called the problem of *size consistency*. A method is said to be *size consistent* if the energy of a composite system A−B at infinite separation is the same as the energy of A plus the energy of B evaluated separately.

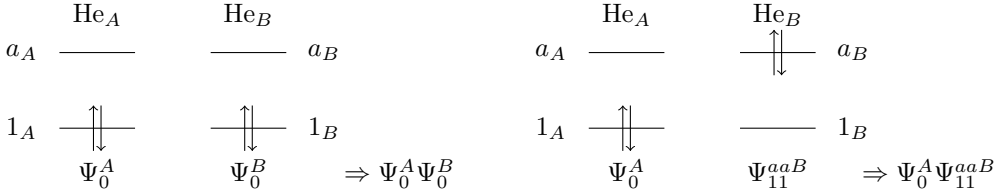


Figure 5.2: The ground state and an excited configuration of the He_2 system.

Consider a wavefunction for He involving one double replacement

$$\Psi = \Psi_0 + c\Psi_{11}^{aa} \,. \tag{5.18}$$

For two such He atoms represented by such a wavefunction a long way apart, the correct combined wavefunction is

$$(\Psi_0^A + c\Psi_{11}^{aaA})(\Psi_0^B + c\Psi_{11}^{aaB}) = \Psi_0^A\Psi_0^B + c(\Psi_{11}^{aaA}\Psi_0^B + \Psi_{11}^{aaB}\Psi_0^A) + c^2\Psi_{11}^{aaA}\Psi_{11}^{aaB} \,. \tag{5.19}$$

This wavefunction involves fourfold replacement of the reference wavefunction. Thus if the CI of the composite He_2 system is restricted to double replacements, the last term is removed, and then the wavefunction and its energy will not correspond to that of two separated He atoms. In practice it is very difficult to include all possible determinants in a full-CI calculation and therefore limited CI calculations suffer from this deficiency which is found to be significant numerically. On the other hand methods based on perturbation theory are size consistent as we shall see later.

There are two types of electron correlation effects, dynamic and non-dynamic. *Dynamic correlation* is like that met in the He atom, essentially arising from the reduction in probability of two electrons coming together. It is said to be short-range correlation. Electrons want to be further apart, and thus the electron repulsion contribution to the total energy is lowered. *Non-dynamic correlation* is best explained through CI. We have seen that to describe the dissociation of H_2 correctly, it is necessary to include the determinant $\hat{\mathcal{A}}(\sigma_u^2)$ in addition to the determinant $\hat{\mathcal{A}}(\sigma_g^2)$, and at infinite separation these two determinants are of equal weight because they are degenerate.[3] Non-dynamic electron correlation is therefore long range, and is introduced in CI by including those determinants which are essential to ensure correct dissociation. Such determinants are approximations to low-lying electronic states.

5.4 Møller–Plesset Theory

This is the simplest way to include the effects of electron correlation. It is an example of second order Rayleigh–Schrödinger perturbation theory.[4]

Theorem 5.1 (Rayleigh–Schrödinger perturbation theory). Suppose we have an exactly solvable reference eigenvalue problem

$$\hat{H}^{(0)}\Psi_i^{(0)} = E_i^{(0)}\Psi_i^{(0)} \,, \tag{5.20}$$

and we are interested in solving a related eigenvalue problem

$$\hat{H}\Psi_i = E_i\Psi_i \,, \tag{5.21}$$

where

$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)} \tag{5.22}$$

for some small *perturbation* $\hat{H}^{(1)}$. Then if the eigenstate $\Psi_i^{(0)}$ of $\hat{H}^{(0)}$ is non-degenerate,

$$E_i^{(1)} = \left\langle \Psi_i^{(0)} \middle| \hat{H}^{(1)} \middle| \Psi_i^{(0)} \right\rangle \,, \tag{5.23}$$

$$\Psi_i^{(1)} = \sum_{j\neq i} \frac{\left\langle \Psi_j^{(0)} \middle| \hat{H}^{(1)} \middle| \Psi_i^{(0)} \right\rangle}{E_i^{(0)} - E_j^{(0)}} \Psi_j^{(0)} \,, \tag{5.24}$$

$$E_i^{(2)} = \left\langle \Psi_i^{(0)} \middle| \hat{H}^{(1)} \middle| \Psi_i^{(1)} \right\rangle \,. \tag{5.25}$$

[3] See course *A4: Theoretical Techniques*.

[4] We will only introduce the non-degenerate time independent perturbation theory. See more details, including a discussion on degenerate perturbation theory and time dependent perturbation theory in course *C7: Further Quantum Mechanics*. If you are interested in some mathematical details, see my notes on Mathematical Tripos Part II: *Principles of Quantum Mechanics*.

Proof. Let the Hamiltonian depend smoothly on a parameter λ

$$\hat{H}(\lambda) = H^{(0)} + \lambda \hat{H}^{(1)} \,. \tag{5.26}$$

Then $\hat{H}(1)$ is the question we are interested in. The key is to assume that the eigenfunctions and eigenvectors of $\hat{H}(\lambda)$ are also analytic on the parameter λ, so that we can express them as a series expansion

$$\Psi_i = \Psi_i^{(0)} + \lambda \Psi_i^{(1)} + \lambda^2 \Psi_i^{(2)} + \mathcal{O}(\lambda^3) \,, \tag{5.27}$$

$$E_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \mathcal{O}(\lambda^3) \,, \tag{5.28}$$

such that

$$\hat{H}\Psi_i = E\Psi_i \tag{5.29}$$

with the conventional *intermediate normalisation*. Then

$$\left\langle \Psi_i^{(0)} \middle| \Psi_i \right\rangle = 1 \quad \Rightarrow \quad \left\langle \Psi_i^{(0)} \middle| \Psi_i^{(n)} \right\rangle = 0 \tag{5.30}$$

for all $n \neq 0$. Then we can expand $\hat{H}\Psi_i = E\Psi_i$ to get

$$\begin{aligned} &(\hat{H}^{(0)} + \lambda \hat{H}^{(1)})(\Psi_i^{(0)} + \lambda \Psi_i^{(1)} + \lambda^2 \Psi_i^{(2)} + \dots) \\ &\qquad = (E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \dots)(\Psi_i^{(0)} + \lambda \Psi_i^{(1)} + \lambda^2 \Psi_i^{(2)} + \dots) \,, \end{aligned} \tag{5.31}$$

which holds true for all λ. Hence, we gather terms in the same order of λ and get

$$\lambda^0 : \qquad \hat{H}^{(0)}\Psi_i^{(0)} = E_i^{(0)}\Psi_i^{(0)} \tag{5.32}$$

$$\lambda^1 : \qquad \hat{H}^{(1)}\Psi_i^{(0)} + \hat{H}^{(0)}\Psi_i^{(1)} = E_i^{(1)}\Psi_i^{(0)} + E_i^{(0)}\Psi_i^{(1)} \tag{5.33}$$

$$\lambda^2 : \qquad \hat{H}^{(1)}\Psi_i^{(1)} + \hat{H}^{(0)}\Psi_i^{(2)} = E_i^{(0)}\Psi_i^{(2)} + E_i^{(1)}\Psi_i^{(1)} + E_i^{(2)}\Psi_i^{(0)} \tag{5.34}$$

$$\dots$$

To extract $E^{(1)}$, we act $\left\langle \Psi_i^{(0)} \right|$ to the terms in the order of λ^1 to get

$$\left\langle \Psi_i^{(0)} \middle| \hat{H}^{(1)} \middle| \Psi_i^{(0)} \right\rangle + \underbrace{\left\langle \Psi_i^{(0)} \middle| \hat{H}^{(0)} \middle| \Psi_i^{(1)} \right\rangle}_{= \left\langle E_i^{(0)} \Psi_i^{(0)} \middle| \Psi_i^{(1)} \right\rangle = 0} = E_i^{(1)} \underbrace{\left\langle \Psi_i^{(0)} \middle| \Psi_i^{(0)} \right\rangle}_{=1} + E_i^{(0)} \underbrace{\left\langle \Psi_i^{(0)} \middle| \Psi_i^{(1)} \right\rangle}_{=0} , \tag{5.35}$$

which rearranges to give

$$E_i^{(1)} = \left\langle \Psi_i^{(0)} \middle| \hat{H}^{(1)} \middle| \Psi_i^{(0)} \right\rangle \,. \tag{5.36}$$

Next, to work out $\Psi_i^{(1)}$, we act $\left\langle \Psi_j^{(0)} \right|$ to the terms of order λ^1 to get

$$\left\langle \Psi_j^{(0)} \middle| \hat{H}^{(1)} \middle| \Psi_i^{(0)} \right\rangle + \underbrace{\left\langle \Psi_j^{(0)} \middle| \hat{H}^{(0)} \middle| \Psi_i^{(1)} \right\rangle}_{= \left\langle E_j^{(0)} \Psi_j^{(0)} \middle| \Psi_i^{(1)} \right\rangle} = E_i^{(1)} \underbrace{\left\langle \Psi_j^{(0)} \middle| \Psi_i^{(0)} \right\rangle}_{= \delta_{ij}} + E_i^{(0)} \left\langle \Psi_j^{(0)} \middle| \Psi_i^{(1)} \right\rangle \,, \tag{5.37}$$

$$\left\langle \Psi_j^{(0)} \middle| \hat{H}^{(1)} \middle| \Psi_i^{(0)} \right\rangle = (-E_j^{(0)} + E_i^{(0)}) \left\langle \Psi_j^{(0)} \middle| \Psi_i^{(1)} \right\rangle \,. \tag{5.38}$$

Since $\{\Psi_j^{(0)}\}$ is a complete orthonormal basis set and $\left\langle \Psi_i^{(0)} \middle| \Psi_i^{(1)} \right\rangle = 0$, we have

$$\Psi_i^{(1)} = \sum_j \left| \Psi_j^{(0)} \right\rangle \left\langle \Psi_j^{(0)} \middle| \Psi_i^{(1)} \right\rangle = \sum_{j \neq i} \frac{\left\langle \Psi_j^{(0)} \middle| \hat{H}^{(1)} \middle| \Psi_i^{(0)} \right\rangle}{E_i^{(0)} - E_j^{(0)}} \Psi_j^{(0)} \,. \tag{5.39}$$

Similarly, by acting $\left| \Psi_i^{(0)} \right\rangle$ on the λ^2 terms, we get

$$E_i^{(2)} = \left\langle \Psi_i^{(0)} \middle| \hat{H}^{(1)} \middle| \Psi_i^{(1)} \right\rangle \tag{5.40}$$

etc. □

Our system of interest is clearly the Schrödinger Hamiltonian $\hat{H}$, and we need to find a reference Hamiltonian that has known eigenstates and eigenvalues. Having just derived the Hartree–Fock theory, the sum of the Fock operators is obviously a good choice. Let

$$\hat{H}^{(0)} = \sum_{i=1}^{n} \hat{F}(i)\,, \tag{5.41}$$

$$\hat{H}^{(1)} = \hat{H} - \hat{H}^{(0)}\,. \tag{5.42}$$

In the restricted space defined by the basis set, the orbitals are eigenfunctions of the Fock operators

$$\hat{F}\phi_p = \epsilon_p \phi_p\,. \tag{5.43}$$

It follows that in zeroth order, the eigenstates of $\hat{H}^{(0)}$ are the Slater determinants made out of n of the m spin-orbitals, denoted $\Phi_I \coloneqq \hat{\mathcal{A}}(\phi_{I_1} \dots \phi_{I_n})$, such that

$$\hat{H}^{(0)}\Phi_I = E_I^{(0)}\Phi_I\,, \tag{5.44}$$

$$E_I^{(0)} = \sum_{i=1}^{n} \epsilon_{I_i}\,. \tag{5.45}$$

Particularly, the ground state

$$\hat{H}^{(0)}\Phi_0^{(0)} = E_0^{(0)}\Phi_0^{(0)}\,, \tag{5.46}$$

$$E_0^{(0)} = \sum_{i=1}^{n} \epsilon_i\,. \tag{5.47}$$

This method is the *Møller–Plesset theory*, MPn, where n labels the order of perturbation we consider.

The zeroth order energy is simply the sum of the orbital energies, which is a very poor estimate of the total energy because it involves double counting. Now from perturbation theory, the first order correction is

$$E_I^{(1)} = \left\langle \Phi_I^{(0)} \middle| \hat{H}^{(1)} \middle| \Phi_I^{(0)} \right\rangle \tag{5.48}$$

and hence to the first order, the total ground state energy is

$$E_0^{(0)} + E_0^{(1)} = \left\langle \Phi_0^{(0)} \middle| \hat{H} \middle| \Phi_0^{(0)} \right\rangle = E_{\mathrm{HF}}\,, \tag{5.49}$$

which is exactly the Hartree–Fock SCF energy.

Next, we want to calculate the second order energy correction, but first we need the ground state wavefunction corrected to the first order. By Rayleigh–Schrödinger perturbation theory, it can be expanded in the orthonormal set $\{\Phi_I\}$, which are the eigenstates of $\hat{H}^{(0)}$, as

$$\Phi^{(1)} = \sum_{I \neq 0} C_I \Phi_I\,, \tag{5.50}$$

where

$$C_I = -\frac{\left\langle \Phi_I \middle| \hat{H}^{(1)} \middle| \Phi_0^{(0)} \right\rangle}{H_{II}^{(0)} - E_0^{(0)}}\,. \tag{5.51}$$

Replacement of $\hat{H}^{(1)}$ by $\hat{H} - \hat{H}^{(0)}$ yields

$$C_I = -\frac{\left\langle \Phi_I \middle| \hat{H} \middle| \Phi_0^{(0)} \right\rangle}{H_{II}^{(0)} - E_0^{(0)}}\,. \tag{5.52}$$

Now we consider the various possibilities for Φ_I.

(i) Single replacement determinants Φ_i^a.
By Proposition 3.4 and (4.14), we get

$$\left\langle \Phi_i^a \middle| \hat{H} \middle| \Phi_0^{(0)} \right\rangle = F_{ai} = 0\,. \tag{5.53}$$

This is known as Brillouin's theorem.

(ii) Double replacement determinants Φ_{ij}^{ab}.
By Proposition 3.4,

$$\left\langle \Phi_{ij}^{ab} \middle| \hat{H} \middle| \Phi_0^{(0)} \right\rangle = (ai|bj) - (aj|bi)\ . \tag{5.54}$$

(iii) Triple and higher replacement determinants $\Phi_{ijk\ldots}^{abc\cdots}$.

$$\left\langle \Phi_{ijk\ldots}^{abc\ldots} \middle| \hat{H} \middle| \Phi_0^{(0)} \right\rangle = 0 \tag{5.55}$$

because $\hat{H}$ has no three electron operators.

Thus the only coefficients in (5.52) that are non-vanishing are those for double replacements. This shouldn't be surprising following our CI discussion. For a particular Φ_{ij}^{ab} the denominator is

$$H_{II}^{(0)} - E_0^{(0)} = \epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j\,, \tag{5.56}$$

and note that Φ_{ij}^{ab}, $-\Phi_{ji}^{ab}$, $-\Phi_{ij}^{ba}$ and Φ_{ji}^{ba} are the same, the second order ground state energy correction hence takes the form

$$\begin{aligned} E_0^{(2)} &= -\sum_I \frac{\left|\left\langle \Phi_I \middle| \hat{H} \middle| \Phi_0^{(0)} \right\rangle\right|^2}{H_{II}^{(0)} - E_0^{(0)}} \\ &= -\frac{1}{4} \sum_{i,j=1}^{n} \sum_{a,b=n+1}^{m} \frac{[(ia|jb) - (ib|ja)]^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \\ &=: \sum_{ij} \epsilon_{ij}\,, \end{aligned} \tag{5.57}$$

where ϵ_{ij} may be interpreted as the correlation energy of a pair of electrons.

For a closed-shell system $\Phi_0 = \hat{\mathcal{A}}(\phi_1^\alpha \phi_1^\beta \phi_2^\alpha \phi_2^\beta \ldots \phi_n^\alpha \phi_n^\beta)$ with $2n$ electrons, we can work out the terms vanishing due to spins in (5.57). Consider the possible spin combinations for a particular set of spatial orbital i, j, a, b:

Spatial	i	a	j	b	$(ia\|jb)$	$(ib\|ja)$
Spin	α	α	α	α	✓	✓
	β	β	β	β	✓	✓
	α	α	β	β	✓	✗
	β	β	α	α	✓	✗
	α	β	β	α	✗	✓
	β	α	α	β	✗	✓

and all other combinations of spins vanish. Summing these up gives

$$E_0^{(2)} = -\frac{1}{4} \sum_{i,j=1}^{n} \sum_{a,b=n+1}^{m} \frac{4\,(ia|jb)^2 + 4\,(ib|ja)^2 - 4\,(ia|jb)\,(ib|ja)}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j}\,. \tag{5.58}$$

Since $\sum_{ijab} (ib|ja)^2 = \sum_{ijab} (ia|jb)^2$, we can further simplify this to get

$$E_0^{(2)} = -\sum_{i,j,a,b} \frac{(ia|jb)\,[2\,(ia|jb) - (ib|ja)]}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j}\,. \tag{5.59}$$

$$E_{\mathrm{MP2}} = E_0^{(0)} + E_0^{(1)} + E_0^{(2)} = E_{\mathrm{HF}} + E_0^{(2)}\,. \tag{5.60}$$

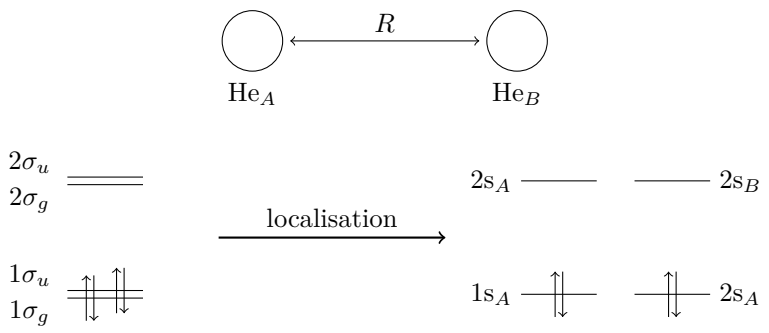


Figure 5.3: A localisation of a He_2 system that shows that MP2 is size consistent.

5.4.1 Size Consistency

Finally we need to explain why MP2 is a size consistent energy expression. We will assert, without proof, the following lemma.

Lemma 5.2. The Møller–Plesset energy is invariant to occupied-occupied and virtual-virtual orbitals unitary rotations of degenerate orbitals.

Remark. This is analogous to Proposition 3.3.

Consider two He atoms, each with 1s and 2s basis. At infinity, they give a pair of degenerate $1\sigma_g$, $1\sigma_u$ orbitals

$$|1\sigma_g\rangle = \frac{|1s_A\rangle + |1s_B\rangle}{\sqrt{2}}\,, \tag{5.61}$$

$$|1\sigma_u\rangle = \frac{|1s_A\rangle - |1s_B\rangle}{\sqrt{2}}\,. \tag{5.62}$$

By the above lemma, they can be localised on each atom. Thus a term in the MP2 energy expression $[(ai|bj) - (aj|bi)](\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j)^{-1}$ is only non-zero if the orbitals ϕ_a, ϕ_b, ϕ_i and ϕ_j are all on the same atom, in which case the term will be a part of the energy of the separate system A or B. If the orbitals are not all on A, or all on B, then the two electron integrals will be zero by zero overlap. Thus the sum of the separate energies of A and B is obtained. If we consider a large molecule, then the MP2 method will obtain the pair energies for all pairs of electrons, and thus there will be no degradation with system size.

On the other hand for CI, the percentage of the wavefunction accounted for by the reference configuration, single excitation and double excitations rapidly decreases as the number of electrons increases because of the increasingly large proportion of all the other excitation, i.e.

$$\langle \Psi_{\mathrm{CISD}} | \Psi_{\mathrm{FCI}} \rangle \downarrow \text{ as } n_{e^-} \uparrow\,. \tag{5.63}$$

Thus CI cannot be used for the study of large molecules.

5.4.2 Higher Orders of Perturbation Theory

The Møller–Plesset perturbation theory can be extended to include not only the second, but also the third and fourth order perturbation, named MP2, MP3, MP4 *etc.* MP2 is always useful, because for the ground state the energy correction is always negative (clear by the form of (5.57)) as it picks up electron corrections. However, it is not guaranteed that the MP2 energy is higher than the true ground state energy since the correction term

$$E_0^{(2)} = \left\langle \Phi_0^{(0)} \middle| \hat{H}^{(1)} \middle| \Phi_0^{(1)} \right\rangle \tag{5.64}$$

Figure 5.4: Dissociation of H_2 using different orders of Møller–Plesset perturbation theory with cc-pVTZ basis. The result of full-CI is also included, which can be seen as the correct answer. We see that a higher order of perturbation does not necessarily give a better answer, especially when the perturbation is large. In the $r \to \infty$ limit, the correlation is so large that the $1\sigma_g^+$ and $1\sigma_u^+$ orbitals of H_2 become degenerate, so the MP2 energy diverges to infinity. This is also the case for higher-order perturbations. (In fact in this case CCSD(T) will also break down for molecules apart from hydrogen. Multi-reference methods (e.g. CASSCF, CASPT2) should be used in these cases.)

is not the expectation value of $\hat{H}$ for a single state, so the variational principles do not apply.

It has been argued that the higher orders, especially MP4, will give greater accuracy, because they introduce higher order excitations (e.g. MP4 introduces three- and four- replacements). However, it has been demonstrated, by using larger basis sets in full-CI calculations, in many cases that the perturbation series for $\hat{H}(\lambda) = \hat{H}^{(0)} + \lambda\hat{H}^{(1)}$ has singularities inside $|\lambda| = 1$, and thus the series is not convergent for $\lambda = 1$. For example, it is explicitly shown that for Ne, the ground state and first excited state of $\hat{H}(\lambda)$ become degenerate when $\lambda = -0.82$. This problem is more severe if the electron correlation is strong so that the perturbation is large. For example, if we add a quartic potential perturbation to a harmonic oscillator's quadratic potential, then the radius of convergence of the perturbation energy will have a radius of convergence of zero.

5.5 Coupled Cluster Theory

Consider the full-CI wavefunction

$$\Psi_{\text{FCI}} = \sum_I c_I \Phi_I \,. \tag{5.65}$$

We now introduce the *excitation operator* $\hat{a}_I$ such that

$$\hat{a}_I \Phi_0 \coloneqq \Phi_I \tag{5.66}$$

excites the ground electron configuration to an excited configuration. An example would be

$$\hat{a}_i^a \Phi_0 = \Phi_i^a \,. \tag{5.67}$$

We can excite an already excited state

$$\hat{a}_i^a \Phi_j^b = \Phi_{ij}^{ab} \,, \tag{5.68}$$

and we need to introduce the rule

$$\hat{a}_i^a \Phi_i^b = 0 \,. \tag{5.69}$$

Then we may write

$$\begin{aligned} \Psi_{\text{FCI}} &= \sum_I c_I \hat{a}_I \Phi_0 \\ &= \hat{c}\Phi_0 \,, \end{aligned} \tag{5.70}$$

where we defined

$$\hat{c} \coloneqq \sum_I c_I \hat{a}_I \,. \tag{5.71}$$

We further define

$$\hat{T}_1 \coloneqq \sum_i \sum_a t_i^a \hat{a}_i^a \,, \tag{5.72}$$

$$\hat{T}_2 \coloneqq \sum_{i<j} \sum_{a<b} t_{ij}^{ab} \hat{a}_{ij}^{ab} \,, \tag{5.73}$$

$$\dots$$

for some $\{t_i^a\}$ and $\{t_{ij}^{ab}\}$, and

$$\hat{T} \coloneqq \hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_n \,, \tag{5.74}$$

then

$$\Psi_{\mathrm{FCI}} = \underbrace{(1 + \hat{T})}_{\hat{c}} \Phi_0 \tag{5.75}$$

if we choose $\{t\}$ appropriately. If we truncate $\hat{T}$ to

$$\hat{T}_{\mathrm{SD}} = \hat{T}_1 + \hat{T}_2 \,, \tag{5.76}$$

then this is exactly CISD. Hence, we want to rewrite it a little bit.

Now, consider the *coupled cluster* (CC) wavefunction

$$\Psi_{\mathrm{CC}} = \exp(\hat{T})\Phi_0 \,, \tag{5.77}$$

where the exponential of an operator is defined as

$$\begin{aligned} \exp(\hat{T}) &\coloneqq 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \hat{T}^n \\ &= 1 + \hat{T} + \frac{1}{2}\hat{T}\hat{T} + \dots \end{aligned} \tag{5.78}$$

This also includes all possible excitations of Φ_0, and so

$$\exp(\hat{T})\Phi_0 = \Psi_{\mathrm{FCI}} \tag{5.79}$$

for well-chosen $\{t\}$. But now if we truncate $\hat{T}$ to the second excitation, we get

$$\exp(\hat{T}_{SD}) = 1 + \hat{T}_1 + \hat{T}_2 + \frac{1}{2}(\hat{T}_1 + \hat{T}_2)(\hat{T}_1 + \hat{T}_2) + \dots \,, \tag{5.80}$$

which includes the terms

$$\hat{T}_2^2 \Phi_0 = \sum_{i<j} \sum_{k<l} \sum_{a<b} \sum_{c<d} t_{ij}^{ab} t_{kl}^{cd} \Phi_{ijkl}^{abcd} \tag{5.81}$$

$$\dots$$

We note immediately that it introduces quadruple excitations (and above) in a natural way. This method with $\hat{T}_{SD}$ truncated up to $\hat{T}_2$ is known as *coupled cluster with single and double excitation* (CCSD).

The CC wavefunction is size consistent. Consider a supermolecule AB, where A and B are infinitely separated, then

$$\Psi_{AB} = \exp(\hat{T}_A + \hat{T}_B)[\Phi_A \Phi_B] = \exp(\hat{T}_A)\Phi_A \exp(\hat{T}_B)\Phi_B = \Psi_A \Psi_B \,, \tag{5.82}$$

and hence

$$\hat{H}_{AB}\Psi_{AB} \quad = (\hat{H}_A + \hat{H}_B)\Psi_A\Psi_B = (E_A + E_B)\Psi_{AB} \,. \tag{5.83}$$

To determine the amplitudes $\{t_i^a\}, \{t_{ij}^{ab}\}$ and the energy E, we need to solve the Schrödinger equation

$$\hat{H}\Psi = E\Psi \quad \Rightarrow \quad (\hat{H} - E)\,|\Psi\rangle = |0\rangle \ . \tag{5.84}$$

The inner product of $|0\rangle$ with any state is 0, as so obtained the *projected Schrödinger equation*

$$\left\langle \chi \middle| \hat{H} - E \middle| \Psi \right\rangle = 0 \,, \tag{5.85}$$

where $|\chi\rangle$ can be any vector. We choose to project the Schrödinger equation with Φ_0, Φ_i^a and Φ_{ij}^{ab}, and use the normalisation

$$\left\langle \Phi_0 \middle| \mathrm{e}^{\hat{T}_{SD}} \middle| \Phi_0 \right\rangle = 1 \,. \tag{5.86}$$

(i) Projecting onto Φ_0, the Hartree–Fock solution:

$$\left\langle \Phi_0 \middle| \hat{H} - E \middle| \mathrm{e}^{\hat{T}_{SD}} \Phi_0 \right\rangle = 0 \tag{5.87}$$

$$\Rightarrow \quad \left\langle \Phi_0 \middle| \hat{H} \left(1 + \hat{T}_2 + \frac{1}{2}\hat{T}_1^2 \right) \Phi_0 \right\rangle = E \,, \tag{5.88}$$

where no terms higher than double replacements enter by the usual theorem for matrix elements for Slater determinants. The term $\hat{T}_i$ vanishes by Brillouin's theorem as SCF orbitals are used.

(ii) Projecting onto Φ_i^a:

$$\left\langle \Phi_i^a \middle| \hat{H} \left(1 + \hat{T}_1 + \hat{T}_2 + \frac{1}{2}\hat{T}_1^2 + \hat{T}_1\hat{T}_2 + \frac{1}{6}\hat{T}_1^3 \right) \Phi_0 \right\rangle = E t_i^a \,. \tag{5.89}$$

Note that at most triple replacements enter the ket.

(iii) Projecting onto Φ_{ij}^{ab}:

$$\begin{aligned} \left\langle \Phi_{ij}^{ab} \middle| \hat{H} \Big(1 + \hat{T}_1 + \frac{1}{2}\hat{T}_1^2 + \frac{1}{6}\hat{T}_1^3 + \frac{1}{24}\hat{T}_1^4 \\ + \hat{T}_2 + \frac{1}{2}\hat{T}_2^2 + \hat{T}_2\hat{T}_1 + \frac{1}{2}\hat{T}_1^2\hat{T}_2 \Big) \Phi_0 \right\rangle = E(t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a) \,. \end{aligned} \tag{5.90}$$

The most costly term to evaluate is $\left\langle \Phi_{ij}^{ab} \middle| \hat{H} \middle| \hat{T}_2^2 \Phi_0 \right\rangle$.

5.5.1 Brueckner Coupled Cluster Theory (Non-Examinable)

These equations are greatly simplified if the concept of the *Brueckner orbital* is introduced. This is non-examinable. The Brueckner orbital is usually defined in terms of a full-CI expansion of the wavefunction.

$$\Psi = \Phi_0 + \sum c_i^a \Phi_i^a + \sum c_{ij}^{ab} \Phi_{ij}^{ab} + \ldots \tag{5.91}$$

The occupied Brueckner orbitals ϕ_i are defined such that c_i^a are zero. Thus in the Brueckner coupled cluster theory (called BD theory), we can remove $\hat{T}_1$ from the CCSD equations to obtain

(i) $$\left\langle \Phi_0 \middle| \hat{H}(1 + \hat{T}_2)\Phi_0 \right\rangle = E \,, \tag{5.92}$$

(ii) $$\left\langle \Phi_i^a \middle| \hat{H}(1 + \hat{T}_2)\Phi_0 \right\rangle = 0 \,, \tag{5.93}$$

(iii) $$\left\langle \Phi_{ij}^{ab} \middle| \hat{H} \left(1 + \hat{T}_2 + \tfrac{1}{2}\hat{T}_2^2 \right) \Phi_0 \right\rangle = E t_{ij}^{ab} \,. \tag{5.94}$$

In practice the equations have to be iterated until the orbitals satisfy (ii). The precise form of these equations are

(i) $$E = F_{ii} + \tfrac{1}{2}\,(ij\|ij) + \tfrac{1}{4}\,(ij\|ab)\, t_{ij}^{ab} \,, \tag{5.95}$$

(ii) $$0 = F_{ia} + t_{ik}^{ac} F_{kc} + u_i^a \,, \tag{5.96}$$

(iii) $$0 = (ab\|ij) + F_{bc} t_{ij}^{ac} + F_{ac} t_{ij}^{cb} - F_{jk} t_{ik}^{ab} - F_{ik} t_{kj}^{ab} + u_{ij}^{ab} + v_{ij}^{ab} \,, \tag{5.97}$$

	R_e	$1.5R_e$	$2.0R_e$
SCF	216	271	370
MP2	11	30	70
MP4	2	12	30
CCSD	4.122	10.158	21.404
CCSD(T)	0.717	1.998	-4.634
Chemical	1 kcal mol^{-1} = 1.6 mE_h		
Spectroscopic	1 cm^{-1} = 4.6×10^{-3} mE_h		

Table 1: The calculated $E - E_{\text{Full-CI}}$ in mE_h using DZP basis set for water.

where the following notations have been introduced:

$$(pq\|rs) = (pr|qs) - (ps|qr) \tag{5.98}$$

$$F_{pq} = h_{pq} + (pk\|qk) \tag{5.99}$$

$$u_i^a = -\frac{1}{2}(ja\|bc)\, t_{ij}^{bc} - \frac{1}{2}(jk\|ib)\, t_{jk}^{ab} \tag{5.100}$$

$$\begin{aligned} u_{ij}^{ab} = {} & \frac{1}{2}(ab\|cd)\, t_{ij}^{cd} + \frac{1}{2}(kl\|ij)\, t_{kl}^{ab} \\ & - (kb\|jc)\, t_{ik}^{cd} - (ka\|jc)\, t_{ik}^{cb} - (kb\|ic)\, t_{kj}^{ac} - (ka\|ic)\, t_{jk}^{ab} \end{aligned} \tag{5.101}$$

$$\begin{aligned} v_{ij}^{ab} = {} & \frac{1}{4}(kl\|cd)\,[t_{ij}^{cd} t_{kl}^{ab} - 2(t_{ij}^{ac} t_{kl}^{bd} + t_{ij}^{bd} t_{kl}^{ac}) \\ & - 2(t_{ik}^{ab} t_{jl}^{cd} + t_{ik}^{cd} t_{jl}^{ab}) + 4(t_{ik}^{ac} t_{jl}^{bd} + t_{ik}^{bd} t_{jl}^{ac})]\,. \end{aligned} \tag{5.102}$$

Summation convention applies. The costs of these equations are n^3N^3 and n^2N^4, where n and N are the numbers of occupied and virtual orbitals. Thus, the computational cost of CCSD is $\mathcal{O}(N^6)$. They are best evaluated in the A.O. basis.

The principal omission is the inclusion of the effect of triple replacements. It is best to include these perturbatively through terms like the following

$$t_{ij}^{ab} \left\langle \Phi_{ij}^{ab} \middle| \hat{H} \middle| \Phi_{ijk}^{acd} \right\rangle \left\langle \Phi_{ijk}^{acd} \middle| \hat{H} \middle| \Phi_{il}^{ac} \right\rangle t_{il}^{ac} (\epsilon_a + \epsilon_c + \epsilon_d - \epsilon_i - \epsilon_j - \epsilon_k)^{-1}\,. \tag{5.103}$$

The inclusion of these terms is denoted CCSD(T) or BD(T). CCSD(T) is the gold standard of quantum chemistry, with computational cost $\mathcal{O}(N^7)$.

An example is shown in table 1, where $E - E_{\text{Full-CI}}$ calculated using DZP basis is tabulated in millihartrees at various bond lengths for H_2O, C_{2v} point group. The importance of triple replacements is observed. Note that the energy obtained using CCSD(T) falls below the actual (full-CI) energy since this method is not variational.

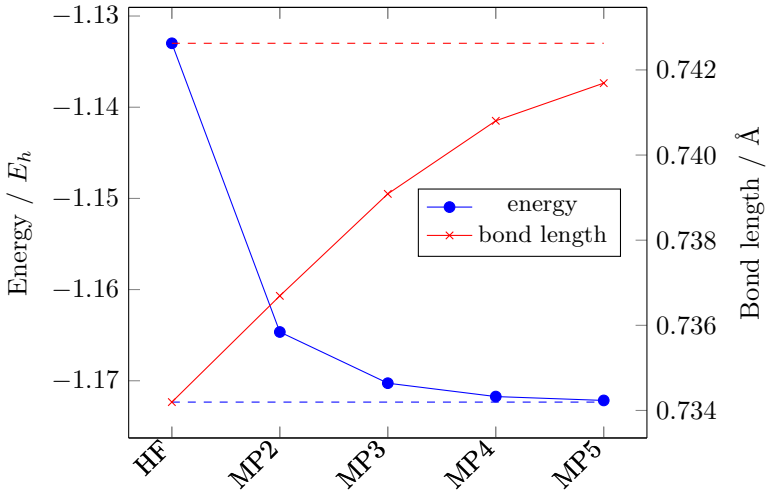


Figure 5.5: The equilibrium geometry and bond length calculated using various methods with cc-pVTZ basis set. The full-CI result is included as the dashed line. Note that CCSD is equivalent to full-CI in two-electron systems.

6 Self Consistent Field Gradient Theory (Non-examinable)

We are going to derive a formula for the gradient of the SCF energy, at a point $\mathbf{X}$ on the SCF potential energy surface. Therefore we wish to calculate

$$\lim_{\delta\mathbf{X}\to 0} \frac{E(\mathbf{X}+\delta\mathbf{X}) - E(\mathbf{X})}{\delta\mathbf{X}} = \frac{\mathrm{d}E}{\mathrm{d}\mathbf{X}}\,. \tag{6.1}$$

At $\mathbf{X}+\delta\mathbf{X}$, the orbitals ϕ_i will have changed through their basis functions η_α and through their molecular orbital coefficients $c_{\alpha i}$.

The change in the basis functions is easily evaluated. Consider a basis function η_α which is an s-Gaussian $\exp(-\alpha(\mathbf{x}-\mathbf{X})^2)$. Differentiating it with respect to centre $\mathbf{X}$, we obtain

$$\frac{\mathrm{d}\eta_\alpha}{\mathrm{d}\mathbf{X}} = -2\alpha(\mathbf{X}-\mathbf{x})\exp(-\alpha(\mathbf{x}-\mathbf{X})^2)\,. \tag{6.2}$$

This is a p-Gaussian. We denote it as $\eta_\alpha^{\mathbf{X}}$.

It is simplest to proceed using Lagrange multipliers. We wish to differentiate the energy maintaining orbital orthogonality. We therefore consider the function $G(\mathbf{c},\mathbf{X})$, where

$$G(\mathbf{c},\mathbf{X}) = E(\mathbf{c},\mathbf{X}) - \sum_{i,j}\epsilon_{ij}(S_{ij}-\delta_{ij})\,. \tag{6.3}$$

Here $E(\mathbf{c},\mathbf{X})$ is the SCF energy expression in terms of molecular orbital integrals, S_{ij} is the overlap integral $\langle\phi_i|\phi_j\rangle$, and ϵ_{ij} are the Lagrange multipliers. We know that

$$\frac{\partial G}{\partial \mathbf{c}} = 0 \tag{6.4}$$

because this is the SCF condition. Indeed it can be easily shown that (6.4) is equivalent to (4.20) if we use orbitals which diagonalise the Lagrangian matrix, and in this case $\epsilon_{ii} = \epsilon_i$, the orbital energies. In this case

$$G(\mathbf{c},\mathbf{X}) = E(\mathbf{c},\mathbf{X}) - \sum_{i}\epsilon_{i}(S_{ii}-1)\,. \tag{6.5}$$

The expression for the gradient of the energy is

$$\frac{\mathrm{d}G}{\mathrm{d}\mathbf{X}} = \frac{\partial G}{\partial \mathbf{c}}\frac{\mathrm{d}\mathbf{c}}{\mathrm{d}\mathbf{X}} + \frac{\partial G}{\partial \mathbf{X}}\,. \tag{6.6}$$

The first term is zero by (6.4), and thus the expression for the energy gradient is

$$\frac{\mathrm{d}G}{\mathrm{d}\mathbf{X}} = \frac{\partial G}{\partial \mathbf{X}} = \frac{\partial E}{\partial \mathbf{X}} - \sum_i \epsilon_i \frac{\partial S_{ii}}{\partial \mathbf{X}}\,. \tag{6.7}$$

Using the superscript notation introduced above to denote differentiation with respect to basis functions only, we have the expression for the energy gradient

$$\frac{\mathrm{d}E}{\mathrm{d}\mathbf{X}} = E^{\mathbf{X}} - \sum_i \epsilon_i S_{ii}^{\mathbf{X}}\,, \tag{6.8}$$

where, for example,

$$\begin{aligned} S_{ii}^{\mathbf{X}} &= 2\left\langle \phi_i^{\mathbf{X}} \middle| \phi_i \right\rangle = \sum_{\alpha,\beta} c_{\alpha i} c_{\beta i} S_{\alpha\beta}^{\mathbf{X}} \\ &= \sum_{\alpha,\beta} c_{\alpha i} c_{\beta i} [\left\langle \eta_\alpha^{\mathbf{X}} \middle| \eta_\beta \right\rangle + \left\langle \eta_\alpha \middle| \eta_\beta^{\mathbf{X}} \right\rangle]\,. \end{aligned} \tag{6.9}$$

We can now write the expression for the gradient of the closed-shell SCF energy in terms of differentiated basis function integrals and some density matrices

$$\frac{\mathrm{d}E}{\mathrm{d}\mathbf{X}} = 2\sum_{\alpha,\beta} D_{\alpha\beta} h_{\alpha\beta}^{\mathbf{X}} + \sum_{\alpha,\beta,\gamma,\delta} D_{\alpha\beta} D_{\gamma\delta} [2\,(\alpha\beta|\gamma\delta)^{\mathbf{X}} - (\alpha\gamma|\beta\delta)^{\mathbf{X}}] - 2\sum_{\alpha\beta} E_{\alpha\beta} S_{\alpha\beta}^{\mathbf{X}}\,, \tag{6.10}$$

where

$$(\alpha\beta|\gamma\delta)^{\mathbf{X}} = \left(\alpha^{\mathbf{X}}\beta\middle|\gamma\delta\right) + \left(\alpha\beta^{\mathbf{X}}\middle|\gamma\delta\right) + \left(\alpha\beta\middle|\gamma^{\mathbf{X}}\delta\right) + \left(\alpha\beta\middle|\gamma\delta^{\mathbf{X}}\right)\,, \tag{6.11}$$

$$h_{\alpha\beta}^{\mathbf{X}} = \left\langle \alpha^{\mathbf{X}} \middle| h \middle| \beta \right\rangle + \left\langle \alpha \middle| h \middle| \boldsymbol{\beta}^{\mathbf{X}} \right\rangle + \left\langle \alpha \middle| h^{\mathbf{X}} \middle| \beta \right\rangle\,, \tag{6.12}$$

$$D_{\alpha\beta} = \sum_i c_{\alpha i}^* c_{\beta i}\,, \tag{6.13}$$

$$E_{\alpha\beta} = \sum_i \epsilon_i c_{\alpha i}^* c_{\beta i}\,. \tag{6.14}$$

Equation (6.10) shows that a knowledge of the density matrices D, E (available following the SCF calculation) and differentiated basis function integrals lead directly to the energy gradient; it is therefore a very easy computational procedure and an extremely important one for quantum chemistry, the Hartree–Fock geometry of a molecule can now be determined efficiently.

7 Multi-Configurational Methods

Not all systems are even moderately well described by a single Slater determinant, and these are known as *multi-reference systems*, as there is more than one sensible reference Slater determinant. Such situations commonly arise when bond breakings or degeneracies are involved (e.g. for transition metals). In such cases, the single-determinant Hartree–Fock equations usually have multiple solutions close in energy to each other.

The way to proceed in such systems is to do *multi-configurational SCF* (MCSCF) theory where the wavefunction is expressed as a limited Slater determinant expansion, $\Psi = \sum_I C_I \Phi_I$, and we find

$$\min_{\{c_{\mu_i}\},\{C_I\}} \left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle . \tag{7.1}$$

This method can be considered a combination between configuration interaction (where the molecular orbitals are not varied but the expansion of the wavefunction is) and Hartree–Fock (where there is only one determinant, but the molecular orbitals are varied).

The determinants Φ_I are made from the set of orbitals, and can be chosen by hand or as a *complete active space* CAS(e,o), where e is the number of active electrons and o is the number of active orbitals. Such calculation is known as *complete active space self-consistent field* (CASSCF).

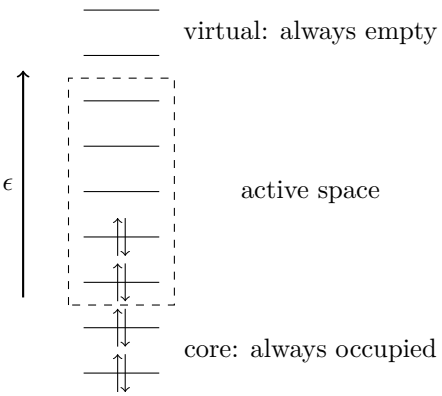


Figure 7.1: In this example, we would like to consider all possible Slater determinants in CAS(4e,5o). It picks up the non-dynamic correlation.

7.1 Correlation

We need to extend MCSCF to include dynamic correlations. Most single-reference methods have an extension to a multi-reference version, where MCSCF wavefunctions are often used as reference states, but usually with some drawbacks.

Complete active space configuration interaction (CASCI) or *multi-reference configuration interaction* (MRCI) makes a space of single and double (and sometimes triple) excitations of all the determinants in the reference wavefunction, and diagonalises the Hamiltonian in this space. As for single-reference CI, it is not size-consistent and is quite computationally expensive.

Complete active space perturbation theory (CASPT2) extends Møller–Plesset perturbation theory to a CAS reference function. Among such calculations it is relatively widely used, but requires a lot of skill to set up calculations and suffers from *intruder state problems* (if there exists E_I for an excited state Φ_I that has similar energy to one in the active space, the denominator in the perturbation energy will be close to zero, and the energy will blow up).

Multi-reference coupled cluster (MRCC) method suffers from a difficulty of defining what excitations to include in $\hat{T}$ and are generally quite computationally expensive as well as losing the size-consistency inherent in the CC ansatz.

7.2 Valence Bond Methods

Valence bond (VB) methods are an attempt to form the wavefunction explicitly in terms of bonds. A simple form of the hydrogen molecule wavefunction is

$$\Psi_{\mathrm{VB}} = \frac{1}{\sqrt{2(1+S^2)}}(\phi_1(1)\phi_2(2) + \phi_2(1)\phi_1(2))\frac{1}{\sqrt{2}}(\alpha(1)\beta(2) - \beta(1)\alpha(2))\,, \tag{7.2}$$

where ϕ_1 and ϕ_2 are non-orthogonal atomic orbitals centred on nucleus 1 and 2 respectively, so overlap $S = \langle\phi_1|\phi_2\rangle$ is nonzero.

A more general form of this allows variational flexibility, forming non-orthogonal orbitals u and v expanded in a basis set

$$\Psi_{\mathrm{GVB}} = \frac{1}{\sqrt{2(1+S^2)}}(u(1)v(2) + v(1)u(2))\frac{1}{\sqrt{2}}(\alpha(1)\beta(2) - \beta(1)\alpha(2))\,, \tag{7.3}$$

with $S = \langle u|v\rangle$.

The GVB has a simple form that dissociates H_2 correctly, but generalisations to many more electrons quickly become complicated. These equations are also hard to solve, with multiple solutions.

7.3 Non-Orthogonal CI

The Hartree–Fock SCF equations also have multiple solutions, corresponding to stationary points and local minima on the potential energy surface. It is possible to retain the attractive features of HF theory (size-consistency and low computational cost), and include multi-reference character by enumerating the low-energy SCF solutions, $\{\Phi_x\}$, and performing configuration interaction with them. Define

$$H_{xy} = \left\langle \Phi_x \middle| \hat{H} \middle| \Phi_y \right\rangle\,, \tag{7.4}$$

$$S_{xy} = \langle \Phi_x | \Phi_y \rangle\,, \tag{7.5}$$

then we need to solve the secular equation

$$\sum_y H_{xy}C_y = E\sum_y S_{xy}C_y\,. \tag{7.6}$$

The orbitals in state x are different and not orthogonal to those in state y, so the Slater–Condon rules cannot be used. Löwdin's method of corresponding orbitals allows for a generalisation of the Slater–Condon rules, and the generalised eigenvalue problem (including S_{xy}) must be solved. The solutions are however variational and size-consistent where the requisite SCF states can be located.

Part II

Density Functional Theory

8 Schrödinger Equation

The basic equation of Quantum Mechanics is given by the Schrödinger equation

$$\hat{H}\Psi = E\Psi\,, \tag{8.1}$$

with the Hamiltonian

$$\hat{H} = \sum_i -\frac{1}{2}\nabla_i^2 + \sum_i v(\mathbf{r}_i) + \sum_{i>j} \frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|}\,. \tag{8.2}$$

The ground state energy E_0 of the system is given by finding the ground state many electron wavefunction $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$. This can also be written as

$$E_0 = \min_{\|\Psi\|=1} \left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle\,. \tag{8.3}$$

For example the Hartree–Fock solution is by minimising the energy with the constraint that the wavefunction is given by a single Slater determinant

$$E_0 = \min_{\Psi \in \{\Psi_{\text{SD}}\}} \left\langle \Psi \middle| \hat{H} \middle| \Psi \right\rangle\,. \tag{8.4}$$

To go beyond the Hartree–Fock approximation, the wavefunction has to be improved; generally this is done by including other determinants. Methods such as Møller–Plesset theory and coupled cluster theory have become very widely used for standard chemical problems. For more complicated electronic structure problems multi-reference methods have also been widely used.

However the expansion of the wavefunction is costly. Hartree–Fock scales like N^4, MP2 like N^5, CCSD like N^6 and full-CI scales like e^N. For small molecules the computational requirements are modest, though the scaling of the methods means that they quickly become prohibitively expensive and even the great advances made in computing and development of algorithms still make the need for other computational methods.

Converging full-CI to the basis set limit is the ultimate benchmark method, though in general unnecessarily expensive compared to basis-set-extrapolated coupled cluster calculations. Nevertheless CI calculations are still performed on multi-configurational systems, generally using a Complete Active Space (CAS). The largest conventional CI calculation to date has only been carried out on the (22e,22o) active space of pentacene, with some 1.2×10^{11} Slater determinants of the appropriate symmetry, though much larger systems have been studied by approximating the solutions (see section 17.3.1). This still makes the prospect of solving the Schrödinger equation for condensed phase chemical reactions look impossibly far in the future.

> *"The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved."*
>
> Paul A. M. Dirac (1929)

Computers have helped greatly but still the direct solution of the Schrödinger equation is a big challenge and a current area of research as outlined in the first half of the course. The complexity is

all in the wavefunction, (ignoring spin) and the probability of finding electron 1 at $\mathbf{r}_1$, electron 2 at $\mathbf{r}_2$ and so on is given by $|\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)|^2$. The electrons are indistinguishable so the probability to find any electron at a given point $\mathbf{r}$ is given by

$$\rho(\mathbf{r}) = \int \prod_{i=1}^{N} \mathrm{d}^3\mathbf{r}_i \; |\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)|^2 \sum_{j=1}^{N} \delta(\mathbf{r}_j - \mathbf{r}) \,. \tag{8.5}$$

This probability is the *electron density*, $\rho(\mathbf{r})$. It is a physical observable and no matter how many electrons are in the system it is only a real function of three dimensions $\rho : \mathbb{R}^3 \to \mathbb{R}$.

9 Density Functional Theory

Density functional theory (DFT) is an alternative to wavefunction-based theories and is grounded on the fact that the electron density, which is a physically measurable quantity, contains all the information required to determine the ground-state electronic energy of a system. More formally, this means that given a ground state electron density $\rho(\mathbf{r})$, there is a unique value of the electronic energy, E. This means that the energy is a *functional* of the density, and is denoted

$$E[\rho(\mathbf{r})]\,. \tag{9.1}$$

We use the properties of the density and of functionals to give a proof of this below.

9.1 Functionals and Functional Derivatives

Definition 9.1. A *functional* is a special type of function that maps functions to values. Let X be a space of functions, then

$$\begin{aligned} F : X &\to \mathbb{R} \\ y &\mapsto F[y] \end{aligned} \tag{9.2}$$

is a functional.

The analogue to differentiation of a function is the *functional derivative*. It can be thought of as the change in the value of the functional when its function argument is changed by an infinitesimal amount.

Definition 9.2. Let $F : V \to \mathbb{R}$ be a functional and $\rho, \phi \in V$. The *functional differential* of F at ρ in the direction of ϕ is defined as the directional derivative

$$\begin{aligned} \delta F[\rho, \phi] &= \lim_{\epsilon \to 0} \frac{F[\rho + \epsilon\phi] - F[\rho]}{\epsilon} \\ &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} F[\rho + \epsilon\phi] \end{aligned} \tag{9.3}$$

In most cases, the domain of the functional F is a space of differentiable functions ρ defined on some space $\Omega \subset \mathbb{R}^n$ and F is of the form

$$F[\rho] = \int_\Omega \mathrm{d}^n\mathbf{x}\, L(\mathbf{x}, \rho(\mathbf{x}), D\rho(\mathbf{x})) \tag{9.4}$$

for some function L, where $D\rho$ denotes the derivative of ρ. If this is the case, then we may define the *functional derivative* of F.

Definition 9.3. A functional of the form

$$F[\rho] = \int_\Omega \mathrm{d}^n\mathbf{x}\, L(\mathbf{x}, \rho(\mathbf{x}), D\rho(\mathbf{x})) \tag{9.5}$$

is *differentiable* if there exists a function $\delta F/\delta\rho$ such that

$$\delta F[\rho, \phi] = \int_\Omega \mathrm{d}^n\mathbf{x}\, \frac{\delta F}{\delta\rho}(\mathbf{x})\phi(\mathbf{x})\,. \tag{9.6}$$

Then $\delta F/\delta\rho$ is the *functional derivative* of F at ρ.

If F is restricted to only certain functions ρ (for example, if there are some boundary conditions imposed) then ϕ is restricted to functions such that $\rho + \epsilon\phi$ continues to satisfy these conditions. Heuristically, ϕ is the change in ρ, so we formally have $\phi = \delta\rho$.

Lemma 9.4. Let

$$G[y] = \int_\alpha^\beta \mathrm{d}x\, f(x, y, y') \tag{9.7}$$

be a functional of $y : \mathbb{R} \to \mathbb{R}$ smooth and restricted to some constant values at its boundary. Then the functional derivative of G is

$$\frac{\delta G}{\delta y} = \frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right). \tag{9.8}$$

Proof. For $\|\delta y\| \to 0$,

$$\begin{aligned}
\delta G &= \int_\alpha^\beta \mathrm{d}x\, f(y + \delta y, y' + \delta y'; x) - \int_\alpha^\beta \mathrm{d}x\, f(y, y'; x) \\
&= \int_\alpha^\beta \mathrm{d}x\, \left[f(y, y'; x) + \delta y \frac{\partial f}{\partial y} + (\delta y)' \frac{\partial f}{\partial y'} + \dots \right] - \int_\alpha^\beta \mathrm{d}x\, f(y, y'; x) \\
&= \int_\alpha^\beta \mathrm{d}x\, \delta y \frac{\partial f}{\partial y} + \left[\delta y \frac{\partial f}{\partial y'} \right]_\alpha^\beta - \int_\alpha^\beta \mathrm{d}x\, \delta y \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right) + \dots
\end{aligned} \tag{9.9}$$

Therefore, when $\delta y(\alpha) = \delta y(\beta) = 0$,

$$\delta G = \int_\alpha^\beta \mathrm{d}x\, \delta y \left[\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right) \right]. \tag{9.10}$$

□

Remark. We may alternatively define the functional derivative of a function G as

$$\frac{\delta G}{\delta y(x)} \coloneqq \lim_{\epsilon \to 0} \frac{G[y(u) + \epsilon\delta(u - x)] - G[y(u)]}{\epsilon}, \tag{9.11}$$

if this limit exists. Although this definition is a bit awkward and involves the delta function, it depends only on the functional G and its argument y. It can be shown that this definition is consistent with the theorem above.

$$\begin{aligned}
\int_\alpha^\beta \mathrm{d}x\, \delta y(x) \frac{\delta G}{\delta y(x)} &= \int_\alpha^\beta \mathrm{d}x\, \delta y(x) \lim_{\epsilon \to 0} \frac{G[y + \epsilon\delta(x - y)] - G[y]}{\epsilon} \\
&= \int_\alpha^\beta \mathrm{d}x\, \delta y(x) \lim_{\epsilon \to 0} \left[\int_\alpha^\beta \mathrm{d}u\, f(u, y(u, \epsilon\delta(u - x)), y'(u) + \epsilon\delta'(u - x)) \right. \\
&\quad \left. - f(u, y(u), y'(u)) \right] \frac{1}{\epsilon} \\
&= \int_\alpha^\beta \mathrm{d}x\, \delta y(x) \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left[\int_\alpha^\beta \mathrm{d}u\, \frac{\partial f}{\partial y} \epsilon\delta(u - x) + \frac{\partial f}{\partial y'} \delta'(u - x) \right] \\
&= \int_\alpha^\beta \mathrm{d}x\, \delta y(x) \left[\frac{\partial f}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right) \right].
\end{aligned} \tag{9.12}$$

Theorem 9.5 (Euler–Lagrange equation). Let $y(x)$ be a real, smooth function with fixed values at $x = \alpha$ and $x = \beta$. The functional

$$G[y] = \int_\alpha^\beta \mathrm{d}x\, f(x, y, y') \tag{9.13}$$

is stationary if and only if

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial f}{\partial y'}\right) = \frac{\partial f}{\partial y}. \tag{9.14}$$

Proof. Trivial from above lemma. □

Proposition 9.6. The functional derivative of

$$G[y] = \int \mathrm{d}^n x \, v(x) y(x) \,, \tag{9.15}$$

is given by

$$\frac{\delta G}{\delta y(x)} = v(x) \,. \tag{9.16}$$

Proof.

$$\frac{\delta G}{\delta y} = \frac{\partial v y}{\partial y} = v \,. \tag{9.17}$$

□

Proposition 9.7. The functional

$$F[\rho] = \int \mathrm{d}^3\mathbf{r} \, f(\mathbf{r}, \rho(\mathbf{r}), \boldsymbol{\nabla}\rho(\mathbf{r})) \tag{9.18}$$

has a functional derivative

$$\frac{\delta F}{\delta \rho(\mathbf{r})} = \frac{\partial f}{\partial \rho(\mathbf{r})} - \boldsymbol{\nabla} \cdot \frac{\partial f}{\partial \boldsymbol{\nabla}\rho(\mathbf{r})} \,. \tag{9.19}$$

Proof. For $\phi(\mathbf{r})$ that vanishes on the boundary of the region of integration,

$$\begin{aligned}
\int \mathrm{d}^3\mathbf{r} \, \phi(\mathbf{r}) \frac{\partial F}{\partial \rho(\mathbf{r})} &= \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \int \mathrm{d}^3\mathbf{r} \, f(\mathbf{r}, \rho + \epsilon\phi, \boldsymbol{\nabla}\rho + \epsilon\boldsymbol{\nabla}\phi) \\
&= \int \mathrm{d}^3\mathbf{r} \left(\frac{\partial f}{\partial \rho}\phi + \frac{\partial f}{\partial \boldsymbol{\nabla}\rho} \cdot \boldsymbol{\nabla}\phi \right) \\
&= \int \mathrm{d}^3\mathbf{r} \left[\frac{\partial f}{\partial \rho}\phi + \boldsymbol{\nabla} \cdot \left(\frac{\partial f}{\partial \boldsymbol{\nabla}\rho}\phi \right) - \left(\boldsymbol{\nabla} \cdot \frac{\partial f}{\partial \boldsymbol{\nabla}\rho} \right) \phi \right] \\
&= \int \mathrm{d}^3\mathbf{r} \left[\frac{\partial f}{\partial \rho}\phi - \left(\boldsymbol{\nabla} \cdot \frac{\partial f}{\partial \boldsymbol{\nabla}\rho} \right) \phi \right] \\
&= \int \mathrm{d}^3\mathbf{r} \left(\frac{\partial f}{\partial \rho} - \boldsymbol{\nabla} \cdot \frac{\partial f}{\partial \boldsymbol{\nabla}\rho} \right) \phi \,,
\end{aligned} \tag{9.20}$$

where the second line is obtained using the total derivative, the third line is obtained by use of a product rule for divergence, and the fourth line is obtained using the divergence theorem and the condition that $\phi = 0$ on the boundary of the region of integration. □

9.2 The Density

We define $\mathbf{r}_i$ to be the spatial coordinate, s_i to be the abstract spin coordinate and $\mathbf{x}_i \coloneqq (\mathbf{r}_i, s_i)$.

As a result of the Born–Oppenheimer approximation, the Coulomb potential arising from the nuclei is treated as a static external potential

$$V_{\text{ext}}(\mathbf{r}) = v(\mathbf{r}) \equiv -\sum_\alpha \frac{Z_\alpha}{\|\mathbf{r} - \mathbf{r}_\alpha\|} \,. \tag{9.21}$$

We use V_{ext} for a general potential and v for the nuclear-electron potential specifically. We define the remainder of the electronic Hamiltonian as

$$\hat{F} \coloneqq \hat{T} + \hat{V}_{\text{e-e}} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{i>j} \frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|} \tag{9.22}$$

such that

$$\hat{H} = \hat{F} + \hat{V}_{\text{ext}}\,, \tag{9.23}$$

where

$$\hat{V}_{\text{ext}} = \sum_i V_{\text{ext}}(\mathbf{r}_i)\,. \tag{9.24}$$

$\hat{F}$ is the same for all N-electron systems, so the Hamiltonian, and hence the ground state $|\Psi_0\rangle$ are completely determined by N and $V_{\text{ext}}(\mathbf{r})$. The ground state $|\Psi_0\rangle$ of this Hamiltonian gives rise to a ground state electronic density

$$\rho(\mathbf{r}_1) = N \int \mathrm{d}s_1 \prod_{i=2}^{N} \mathrm{d}^4\mathbf{x}_i \; |\Psi|^2 \;. \tag{9.25}$$

Note that $\mathrm{d}^4\mathbf{x}_i \equiv \mathrm{d}s_i\,\mathrm{d}^3\mathbf{r}_i$ integrates over both the spin and spatial coordinate of electron i. We clearly have

$$\int \mathrm{d}^3\mathbf{r}\,\rho(\mathbf{r}) = N\,. \tag{9.26}$$

Thus the ground-state $|\Psi_0\rangle$ and density $\rho_0(\mathbf{r})$ are both functionals of the number of electrons N and the external potential $V_{\text{ext}}(\mathbf{r})$.

9.3 Hohenberg–Kohn Theorems

Density functional theory, introduced in 1964 by Hohenberg and Kohn, makes two remarkable statements.

Theorem 9.8 (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 $\rho(\mathbf{r})$ 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 $\rho(\mathbf{r})$. 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,

$$\begin{aligned} E_1^0 < \left\langle \Psi_2 \middle| \hat{H}_1 \middle| \Psi_2 \right\rangle &= \left\langle \Psi_2 \middle| \hat{H}_2 \middle| \Psi_2 \right\rangle + \left\langle \Psi_2 \middle| \hat{H}_2 - \hat{H}_1 \middle| \Psi_2 \right\rangle \\ &= E_2^0 + \int \mathrm{d}^3\mathbf{r}\,\rho(\mathbf{r})[V_1(\mathbf{r}) - V_2(\mathbf{r})]\,, \end{aligned} \tag{9.27}$$

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

$$E_1^0 + E_2^0 < E_2^0 + E_1^0\,, \tag{9.28}$$

which is a contradiction. □

Thus, at least in principle, the ground-state density determines (within a constant) the external potential of the Schrödinger equation of which it is a solution. The external potential and number of electrons

$$N = \int \mathrm{d}^3\mathbf{r}\,\rho(\mathbf{r}) \tag{9.29}$$

determine all the ground-state properties of the system since the Hamiltonian and ground state wavefunction are determined by them.

So for all densities $\rho(\mathbf{r})$ which are ground state densities for some external potential (such a density is known as *v-representable*), the functional

$$F[\rho] = \left\langle \Psi \middle| \hat{F} \middle| \Psi \right\rangle = \left\langle \Psi \middle| \hat{T} + \hat{V}_{\text{e-e}} \middle| \Psi \right\rangle = T[\rho] + V_{\text{e-e}}[\rho] \tag{9.30}$$

is unique and well-defined, since $\rho(\mathbf{r})$ determines the external potential and N (and therefore $\hat{F}$) and thence $|\Psi\rangle$. Now a functional for an arbitrary external potential $V_{\text{ext}}(\mathbf{r})$ unrelated to the $\tilde{V}_{\text{ext}}(\mathbf{r})$ determined by $\rho(\mathbf{r})$ can be defined:

$$E_V[\rho] = F[\rho] + \int \mathrm{d}^3\mathbf{r}\, V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})\,. \tag{9.31}$$

Theorem 9.9 (The second Hohenberg–Kohn theorem). For an arbitrary external potential V_{ext} and v-representable densities $\rho(\mathbf{r})$ of N electrons,

$$E_V[\rho] = F[\rho] + \int \mathrm{d}^3\mathbf{r}\, V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) \geq E_0\,, \tag{9.32}$$

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

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

$$\left\langle\tilde{\Psi}\middle|\hat{H}\middle|\tilde{\Psi}\right\rangle = \left\langle\tilde{\Psi}\middle|\hat{F}\middle|\tilde{\Psi}\right\rangle + \left\langle\tilde{\Psi}\middle|\hat{V}_{\text{ext}}\middle|\tilde{\Psi}\right\rangle = F[\rho] + \int \mathrm{d}^3\mathbf{r}\, V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) = E_V[\rho] \geq E_0 \tag{9.33}$$

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

Thus the problem of solving the Schrödinger equation for non-degenerate ground-states can be recast into a variational problem of minimising the functional $E_V[\rho]$ with respect to v-representable densities.

- An aside: Bright–Wilson proof of the first Hohenberg–Kohn theorem.

  If one knew the exact electron density, $\rho(\mathbf{r})$, then the cusps of $\rho(\mathbf{r})$ would occur at the positions of the nuclei. Furthermore, a knowledge of $\|\boldsymbol{\nabla}\rho(\mathbf{r})\|$ at the nuclei would give their nuclear charges, as it may be shown that the nuclear cusp condition gives

  $$\left.\frac{\partial}{\partial r}\right|_{r=0} \bar{\rho}(r) = -2Z\bar{\rho}(0)\,, \tag{9.34}$$

  where $\bar{\rho}(r)$ is the average of $\rho(\mathbf{r})$ on a sphere of radius r around the nucleus. Thus the full Schrödinger Hamiltonian is known because it is completely defined once the position and charge of the nuclei are given. Hence, in principle (through full-CI), the wavefunction and energy are known, and thus everything is known. In conclusion, a knowledge of the density was all that was necessary for a complete determination of all molecular properties.

9.3.1 Levy Constrained Search

The original Hohenberg–Kohn proof is restricted to non-degenerate ground states and ρ being v-representable. However, this is not a problem if we consider another alternative proof given by Levy. We define a functional of the density $\rho(\mathbf{r})$ for the operator $\hat{F}$ defined above as

$$F[\rho] = \min_{|\Psi\rangle\mapsto\rho} \left\langle\Psi\middle|\hat{F}\middle|\Psi\right\rangle\,, \tag{9.35}$$

i.e. the functional taking the minimum value of the expectation value with respect to all states $|\Psi\rangle$ which give the density ρ. For a system with external potential $V_{\text{ext}}(\mathbf{r})$ and ground state $|\Psi_0\rangle$ with

energy E_0, consider a state $\left|\Psi_{[\rho]}\right\rangle$, an N-electron state which yields density $\rho(\mathbf{r})$ and minimises $F[\rho]$. Define $E_V[\rho]$ as

$$E_V[\rho] = F[\rho] + \int \mathrm{d}^3\mathbf{r}\, \rho(\mathbf{r}) V_{\mathrm{ext}}(\mathbf{r}) = \left\langle \Psi_{[\rho]} \middle| \hat{F} + \hat{V}_{\mathrm{ext}} \middle| \Psi_{[\rho]} \right\rangle . \tag{9.36}$$

Since $\hat{H} = \hat{F} + \hat{V}_{\mathrm{ext}}$, by the variational principle we obtain

$$E_V[\rho] \geq E_0 \,, \tag{9.37}$$

with equality if and only if $\left|\Psi_{[\rho]}\right\rangle = |\Psi_0\rangle$. This holds for all densities which can be obtained from an N-electron wavefunction (such a density is said to be N-representable — this is a weaker condition than v-representable). From the definition of $F[\rho]$ we must also have

$$F[\rho_0] \leq \left\langle \Psi_0 \middle| \hat{F} \middle| \Psi_0 \right\rangle , \tag{9.38}$$

since $|\Psi_0\rangle$ must be one of the states which yield $\rho_0(\mathbf{r})$. Adding $\int \mathrm{d}^3\mathbf{r}\, \rho_0(\mathbf{r}) V_{\mathrm{ext}}(\mathbf{r})$ gives

$$E_V[\rho_0] \leq E_0 \,. \tag{9.39}$$

Hence, combining with (9.37) we have

$$E_V[\rho] \geq E_V[\rho_0] = E_0 \tag{9.40}$$

as desired.

Thus the ground-state density $\rho_0(\mathbf{r})$ minimises the functional $E_V[\rho]$ and the minimum value is the ground-state electronic energy. Note that the requirement for non-degeneracy of the ground-state has disappeared, and further that instead of considering only v-representable densities, we can now consider N-representable densities. The requirements of N-representability are much weaker and satisfied by any well-behaved density, subjected to only a few conditions like non-negativity and proper differentiability.

This variational principle allows us to write down the condition that the energy (9.32) is stationary with respect to changes in the density, subjected to the constraint that (9.26) holds:

$$\delta E_V[\rho] - \mu\delta \left[\int \mathrm{d}^3\mathbf{r}\, \rho(\mathbf{r}) - N \right] = 0 \,, \tag{9.41}$$

for which the stationary condition is, in terms of the functional derivative,

$$\mu = V_{\mathrm{ext}}(\mathbf{r}) + \frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} \,. \tag{9.42}$$

The importance of this theorem is its practicality, namely that if we assume that we have a good functional representation, then we can minimise it to get best density and structure. It is also the place where the ground state restriction enters, because the second Hohenberg–Kohn theorem can only be proved for the ground state.

10 Pure Density Functional Theory

10.1 Thomas–Fermi Theory

The simplest example of a density functional theory came long before the Hohenberg–Kohn theorem and came not long after the introduction of the Schrödinger equation. Thomas and Fermi used the density as the basic variable, using a model system of the uniform electron gas to define functionals with explicit approximations for $T[\rho]$ and $V_{\text{e-e}}[\rho]$. For the T they took the kinetic energy of the uniform electron gas (see later), and the classical (or Hartree) electron-electron repulsion term for $V_{\text{e-e}}$

$$J[\rho] = \frac{1}{2}\int \mathrm{d}^3\mathbf{r}\,\mathrm{d}^3\mathbf{r}'\,\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{\|\mathbf{r}-\mathbf{r}'\|}\,. \tag{10.1}$$

This gives an energy expression

$$E^{\text{TF}}[\rho] = \underbrace{\frac{3}{10}(3\pi^2)^{2/3}\int \mathrm{d}^3\mathbf{r}\,\rho^{5/3}}_{T[\rho]} + \underbrace{\frac{1}{2}\int \mathrm{d}^3\mathbf{r}\,\mathrm{d}^3\mathbf{r}'\,\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{\|\mathbf{r}-\mathbf{r}'\|}}_{V_{\text{e-e}}[\rho]=J[\rho]} + \int \mathrm{d}^3\mathbf{r}\,V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})\,. \tag{10.2}$$

Minimisation of the energy expression leads to the *Thomas–Fermi equation*

$$\frac{5}{3}c_{\text{F}}\rho^{2/3}(\mathbf{r}) + \int \mathrm{d}^3\mathbf{r}'\,\frac{\rho(\mathbf{r}')}{\|\mathbf{r}-\mathbf{r}'\|} + V_{\text{ext}}(\mathbf{r}) = \mu\,, \tag{10.3}$$

where $c_F = 3(3\pi^2)^{2/3}/10$.

This is a beautiful and simple theory along the lines of the exact density functional theory however with approximate functionals. There was much interest in understanding this theory, especially in the solid state community, where it was thought to be a good basis for the description of metals. However for chemistry Thomas–Fermi theory suffers from a devastating issue — *Teller's non-binding theorem*. The energy of a molecular system as a function of the nuclear coordinates

$$E(\mathbf{R}) = E^{\text{TF}}[\rho;\mathbf{R}] + \sum_{A>B}\frac{Z_A Z_B}{\|\mathbf{R}_A-\mathbf{R}_B\|} \tag{10.4}$$

decreases as $\|\mathbf{R}_A-\mathbf{R}_B\| \to \infty$. No molecule is bound with respect to dissociation to its constituent atoms — meaning that Thomas–Fermi model is not applicable to any molecular system. This illustrates the difficulty of using approximate functionals and here the especial difficulty in approximating T. For atoms and molecules near equilibrium the *virial theorem* can show that $\langle T\rangle = -E$. The kinetic energy is a massive term and any approximation to it has a very large error.

10.2 Kinetic Energy Functionals

Aside from the uniform electron gas kinetic energy

$$T^{\text{UEG}}[\rho] = \frac{3}{10}(3\pi^2)^{2/3}\int \mathrm{d}^3\mathbf{r}\,\rho^{5/3}(\mathbf{r})\,, \tag{10.5}$$

which is exact when ρ is constant, it has proven challenging to derive functionals for kinetic energy. For a one-electron system, the wavefunction is simply the square root of its density, and from this von Weizsäcker gave the functional form

$$T^{\text{vW}}[\rho] = \frac{1}{8}\int \mathrm{d}^3\mathbf{r}\,\frac{\|\boldsymbol{\nabla}\rho\|^2}{\rho}\,. \tag{10.6}$$

The problem of finding an accurate kinetic energy functional for more than one electron is a subject of ongoing research in *orbital-free density functional theory*, but little progress has been made over the years and effort has generally been focused on the Kohn–Sham formulation instead.

11 Kohn–Sham Theory

The major step forward for the implementation of density functional theory came with the introduction of the Kohn–Sham equations. The key idea here is to consider a fictitious system of non-interacting electrons which has the same density as the fully interacting system. The non-interacting wavefunction is a single determinant of N non-interacting electrons in N orbitals ϕ_i

$$\Phi_{\mathrm{KS}} = \hat{\mathcal{A}}\,\phi_1\phi_2\ldots\phi_N\,. \tag{11.1}$$

The key part here is that for this non-interacting system the kinetic energy is known exactly. The non-interacting kinetic energy (or Kohn–Sham kinetic energy) T_{s} can be expressed in terms of the Kohn–Sham orbitals

$$T_{\mathrm{s}}[\{\phi_i\}] = \sum_i^N \left\langle \phi_i \middle| \hat{T} \middle| \phi_i \right\rangle\,, \tag{11.2}$$

and the density is given by

$$\rho(\mathbf{r}) = \sum_i^N |\phi_i(\mathbf{r})|^2\,. \tag{11.3}$$

In this form T_{s} is not strictly a functional of ρ, but instead of $\{\phi_i\}$ and hence we denote it as $T_{\mathrm{s}}[\{\phi_i\}]$. For any ρ, we can partition it into orbitals $\{\phi_i\}$ in many ways. Since we wish to find the minimal $E[\rho]$, we define

$$T_{\mathrm{s}}[\rho] = \min_{\{\phi_i\}\mapsto\rho} T_{\mathrm{s}}[\{\phi_i\}]\,. \tag{11.4}$$

The Kohn–Sham orbitals $\phi_i(\mathbf{r})$ are eigenfunctions of an effective one-electron Hamiltonian

$$\hat{h}_{\mathrm{s}}\phi_i = \epsilon_i\phi_i\,, \tag{11.5}$$

where

$$\hat{h}_{\mathrm{s}} = -\frac{1}{2}\nabla^2 + V_{\mathrm{s}}(\mathbf{r})\,, \tag{11.6}$$

and V_{s} is an as-yet determined potential. Hence, the Kohn–Sham Hamiltonian is

$$\hat{H}_{\mathrm{KS}} = \sum_{i=1}^N \hat{h}_{\mathrm{s}}(\mathbf{r}_i)\,. \tag{11.7}$$

Φ_{KS} is an eigenfunction of $\hat{H}_{\mathrm{KS}}$ not $\hat{H}$, but Φ_{KS} yields the same electron density as the true ground state wavefunction of $\hat{H}$.

We can therefore rewrite the total energy as

$$\begin{aligned} E[\rho] &= T[\rho] + V_{\text{e-e}}[\rho] + V_{\mathrm{ext}}[\rho] \\ &= T_{\mathrm{s}}[\rho] + J[\rho] + V_{\mathrm{ext}}[\rho] + [(T[\rho] - T_{\mathrm{s}}[\rho]) + (V_{\text{e-e}}[\rho] - J[\rho])]\,. \end{aligned} \tag{11.8}$$

The last term is the important exchange-correlation functional which contains all the complicated many-body quantum effects

$$E_{\mathrm{xc}}[\rho] = (T[\rho] - T_{\mathrm{s}}[\rho]) + (V_{\text{e-e}}[\rho] - J[\rho])\,. \tag{11.9}$$

If we take the total energy expression

$$E[\rho] = T_{\mathrm{s}}[\rho] + \int \mathrm{d}^3\mathbf{r}\,\rho(\mathbf{r})v(\mathbf{r}) + J[\rho] + E_{\mathrm{xc}}[\rho] \tag{11.10}$$

and minimise with respect to orbitals we get the famous Kohn–Sham equations

$$\left(-\frac{1}{2}\nabla^2 + v(r) + v_J(\mathbf{r}) + v_{\mathrm{xc}}(\mathbf{r})\right)\phi_i = \epsilon_i\phi_i\,, \tag{11.11}$$

where the Coulomb potential is

$$v_J(\mathbf{r}) = \int \mathrm{d}^3\mathbf{r}' \, \frac{\rho(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|} \tag{11.12}$$

and the exchange-correlation potential is given by the functional derivative

$$v_{\mathrm{xc}} = \frac{\delta E_{\mathrm{xc}}}{\delta \rho(\mathbf{r})} \, . \tag{11.13}$$

If, for example, the exchange-correlation functional takes the form

$$E_{\mathrm{xc}}[\rho] = \int \mathrm{d}^3\mathbf{r} \, f_{\mathrm{xc}}(\rho, \boldsymbol{\nabla}\rho) \, , \tag{11.14}$$

then the calculus of variations gives

$$v_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}}{\delta \rho(\mathbf{r})} = \frac{\partial f_{\mathrm{xc}}}{\partial \rho} - \boldsymbol{\nabla} \cdot \frac{\partial f_{\mathrm{xc}}}{\partial \boldsymbol{\nabla}\rho} \, . \tag{11.15}$$

All the complexity is now in one term: the exchange-correlation potential.

As you can see, the electron density is solved by Kohn–Sham equations, which in turn depends on the electron density. Thus, we again use the self-consistent field method to iteratively solve ρ.

12 The Density and Exact Relations

12.1 One and Two Particle (Reduced) Density Matrices

Since the true Hamiltonian is made up from at most two-particle operators, it is possible to express the energy in terms of at most two-particle quantities, rather than the full many-particle wavefunction Ψ. We form these by integrating over the many coordinates of all but a few electrons, producing *reduced density matrices*. The one particle (real space) reduced density matrix is formed from the wavefunction

$$\rho_1(\mathbf{r}_1', \mathbf{r}_1) = N \int \mathrm{d}s_1 \prod_{i=2}^{N} \mathrm{d}^4\mathbf{x}_i \, \Psi^*(\mathbf{r}_1', s_1, \mathbf{x}_2, \dots)\Psi(\mathbf{r}_1, s_1, \mathbf{x}_2, \dots) \,. \tag{12.1}$$

The 'diagonal' of this is its value for $\mathbf{r}_1' = \mathbf{r}_1$,

$$\rho_1(\mathbf{r}_1, \mathbf{r}_1) = N \int \mathrm{d}s_1 \prod_{i=2}^{N} \mathrm{d}^4\mathbf{x}_i \, |\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots)|^2 \,, \tag{12.2}$$

which is the usual electron density

$$\rho(\mathbf{r}) = \rho_1(\mathbf{r}, \mathbf{r}) \,. \tag{12.3}$$

This is the real-space analogue of density matrix $D_{\mu\nu}$, related by

$$\rho_1(\mathbf{r}', \mathbf{r}) = \sum_{\mu,\nu} \eta_\mu^*(\mathbf{r}') D_{\mu\nu} \eta_\nu(\mathbf{r}) \,. \tag{12.4}$$

To calculate general expectation values we need the two-particle density matrix $\Gamma_2(\mathbf{r}_1', \mathbf{r}_2', \mathbf{r}_1, \mathbf{r}_2)$. For our purposes, we only need its diagonal

$$\begin{aligned} \rho_2(\mathbf{r}_1, \mathbf{r}_2) &= \Gamma_2(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_1, \mathbf{r}_2) \\ &= \frac{N(N-1)}{2} \int \mathrm{d}s_1 \, \mathrm{d}s_2 \prod_{i=3}^{N} \mathrm{d}^4\mathbf{x}_i \, |\Psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots)|^2 \,. \end{aligned} \tag{12.5}$$

The factor $N(N-1)/2$ in front is the number of equivalent pairs.

The energy, E, given as the expectation value of the Hamiltonian can be expressed in terms of these matrices

$$E = -\frac{1}{2} \int \mathrm{d}^3\mathbf{r} \; [\nabla_{\mathbf{r}}^2 \rho_1(\mathbf{r}', \mathbf{r})]\big|_{\mathbf{r}'=\mathbf{r}} + \int \mathrm{d}^3\mathbf{r} \, v(\mathbf{r})\rho(\mathbf{r}) + \int \mathrm{d}^3\mathbf{r}_1 \, \mathrm{d}^3\mathbf{r}_2 \, \frac{\rho_2(\mathbf{r}_1, \mathbf{r}_2)}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \,. \tag{12.6}$$

In closed-shell restricted Hartree–Fock or Kohn–Sham theory the wavefunction is a single Slater determinant, and the one particle density matrix and the density are expressed in terms of spatial orbitals

$$\rho_1(\mathbf{r}', \mathbf{r}) = 2 \sum_i \phi_i^*(\mathbf{r}')\phi_i(\mathbf{r}) \,, \tag{12.7}$$

$$\rho(\mathbf{r}) = 2 \sum_i |\phi_i(\mathbf{r})|^2 \,. \tag{12.8}$$

In Hartree–Fock theory, the exchange energy is

$$\begin{aligned} K &= -\int \mathrm{d}^3\mathbf{r}_1 \, \mathrm{d}^3\mathbf{r}_2 \sum_{i,j} \frac{\phi_i^*(\mathbf{r}_1)\phi_i(\mathbf{r}_2)\phi_j^*(\mathbf{r}_2)\phi_j(\mathbf{r}_1)}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \\ &= -\frac{1}{4} \int \mathrm{d}^3\mathbf{r}_1 \, \mathrm{d}^3\mathbf{r}_2 \, \frac{|\rho_1(\mathbf{r}_1, \mathbf{r}_2)|^2}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \,. \end{aligned} \tag{12.9}$$

12.2 Scaling Relations

Scaling relations have proven important in the development of some exact conditions for the exchange and correlation functionals. If we consider the normalisation of the density

$$\int \mathrm{d}^3\mathbf{r}\,\rho(\mathbf{r}) = N \tag{12.10}$$

and we introduce a change of variables $\mathbf{r} = \lambda\mathbf{r'}$, then we have

$$\begin{aligned} &\int \mathrm{d}^3\mathbf{r}'\,\lambda^3\rho(\lambda\mathbf{r}') = N \\ \Rightarrow \quad &\int \mathrm{d}^3\mathbf{r}\,\lambda^3\rho(\lambda\mathbf{r}) = N\,. \end{aligned} \tag{12.11}$$

We can define the *scaled density*

$$\rho_\lambda(\mathbf{r}) = \lambda^3\rho(\lambda\mathbf{r}) \tag{12.12}$$

such that

$$\int \mathrm{d}^3\mathbf{r}\,\rho_\lambda(\mathbf{r}) = N\,. \tag{12.13}$$

We now consider the coordinate scaling of the exact exchange energy used in the scaled orbitals $\phi_{i\lambda}(\mathbf{r}) = \lambda^{3/2}\phi_i(\lambda\mathbf{r})$. We begin by writing the total exchange in terms of a contribution from pairs of orbitals

$$E_\mathrm{x}[\{\phi_{i\lambda}\}] = \sum_{i,j} E_\mathrm{x}[\phi_{i\lambda},\phi_{j\lambda}]\,. \tag{12.14}$$

The scaling of each orbital-pair component is determined from

$$\begin{aligned} E_\mathrm{x}[\phi_{i\lambda},\phi_{j\lambda}] &= \frac{1}{2}\int \mathrm{d}^3\mathbf{r}_1\,\mathrm{d}^3\mathbf{r}_2\,\frac{\phi^*_{i\lambda}(\mathbf{r}_1)\phi_{j\lambda}(\mathbf{r}_1)\phi_{i\lambda}(\mathbf{r}_2)\phi^*_{j\lambda}(\mathbf{r}_2)}{\|\mathbf{r}_1-\mathbf{r}_2\|} \\ &= \frac{\lambda^6}{2}\int \mathrm{d}^3\mathbf{r}_1\,\mathrm{d}^3\mathbf{r}_2\,\frac{\phi^*_i(\lambda\mathbf{r}_1)\phi_j(\lambda\mathbf{r}_1)\phi_i(\lambda\mathbf{r}_2)\phi^*_j(\lambda\mathbf{r}_2)}{\|\mathbf{r}_1-\mathbf{r}_2\|} \\ &= \frac{\lambda}{2}\int \lambda^3\mathrm{d}^3\mathbf{r}_1\,\lambda^3\mathrm{d}^3\mathbf{r}_2\,\frac{\phi^*_i(\lambda\mathbf{r}_1)\phi_j(\lambda\mathbf{r}_1)\phi_i(\lambda\mathbf{r}_2)\phi^*_j(\lambda\mathbf{r}_2)}{\|\lambda\mathbf{r}_1-\lambda\mathbf{r}_2\|} \\ &= \frac{\lambda}{2}\int \mathrm{d}^3\lambda\mathbf{r}_1\,\mathrm{d}^3\lambda\mathbf{r}_2\,\frac{\phi^*_i(\lambda\mathbf{r}_1)\phi_j(\lambda\mathbf{r}_1)\phi_i(\lambda\mathbf{r}_2)\phi^*_j(\lambda\mathbf{r}_2)}{\|\lambda\mathbf{r}_1-\lambda\mathbf{r}_2\|} \\ &= \lambda E_\mathrm{x}[\phi_i,\phi_j]\,. \end{aligned} \tag{12.15}$$

Now we can use the above scaling relation to find out more about possible forms of the exchange density functionals. If we examine a functional of the form

$$E_\mathrm{x}^\mathrm{local}[\rho] = c\int \mathrm{d}^3\mathbf{r}\,\rho^k(\mathbf{r})\,, \tag{12.16}$$

where 'local' means that the integrand does not depend on the density at two different positions simultaneously. Then

$$\begin{aligned} E_\mathrm{x}^\mathrm{local}[\rho_\lambda] &= c\int \mathrm{d}^3\mathbf{r}\,\lambda^{3k}\rho^k(\lambda\mathbf{r}) \\ &= c\lambda^{3k-3}\int \mathrm{d}^3\lambda\mathbf{r}\,\rho^k(\lambda\mathbf{r})\,. \end{aligned} \tag{12.17}$$

To satisfy the scaling requirement, we need $3k-3=1$, and therefore $k=4/3$.

12.3 Adiabatic Connection

Langreth and Perdew defined a family of Hamiltonians interpolating between the fully interacting Hamiltonian and the Kohn–Sham Hamiltonian which retain the same ground-state density. They parameterised this with constant λ which scales the electron-electron repulsion term

$$\hat{H}_\lambda = \hat{T} + \lambda \hat{V}_{\text{e-e}} + \hat{V}_\lambda \,, \tag{12.18}$$

where $\hat{V}_\lambda = \sum_i v_\lambda(\mathbf{r}_i)$ is the result of a one-electron local potential $v_\lambda(\mathbf{r})$. For each $\hat{H}_\lambda$ there is an energy-minimising wavefunction Ψ_λ satisfying

$$\hat{H}_\lambda \Psi_\lambda = E_\lambda \Psi_\lambda \,. \tag{12.19}$$

At each λ, the one-electron potential $v_\lambda(\mathbf{r})$ is chosen such that the density is kept the same as the full interacting density:

$$\int \prod_{i=2}^{N} \mathrm{d}^3\mathbf{r}_i \; |\Psi_\lambda(\mathbf{r}, \mathbf{r}_2, \dots)|^2 = \rho_\lambda(\mathbf{r}) = \rho(\mathbf{r}) \tag{12.20}$$

for all λ; in particular $\rho_0 = \rho_1 = \rho$. At $\lambda = 1$ we have the full interacting Hamiltonian

$$\hat{H}_1 = \hat{T} + \hat{V}_{\text{e-e}} + \sum_i v(\mathbf{r}_i) \,, \tag{12.21}$$

and at $\lambda = 0$ we have the Kohn–Sham system

$$\hat{H}_0 = \hat{T} + \sum_i v_{\text{s}}(\mathbf{r}_i) \,, \tag{12.22}$$

where $v_{\text{s}} = v + v_{\text{J}} + v_{\text{xc}}$. At each value of λ, the wavefunction Ψ_λ will be different, even though it corresponds to the same density $\rho(\mathbf{r})$. Since Ψ_λ is the minimum-energy eigenfunction of $\hat{H}_\lambda$, we can think of this as a search over the wavefunctions which give density ρ

$$\begin{aligned} E_\lambda &= \min_{\Psi_\lambda \mapsto \rho} \left\langle \Psi_\lambda \middle| \hat{H}_\lambda \middle| \Psi_\lambda \right\rangle \\ &= \min_{\Psi_\lambda \mapsto \rho} \left(\left\langle \Psi_\lambda \middle| \hat{T} + \lambda \hat{V}_{\text{e-e}} \middle| \Psi_\lambda \right\rangle + \left\langle \Psi_\lambda \middle| \hat{V}_\lambda \middle| \Psi_\lambda \right\rangle \right) . \end{aligned} \tag{12.23}$$

For a given λ, since $\hat{V}_\lambda = \sum_i v_\lambda(\mathbf{r}_i)$ contains only local one-electron operators, its expectation value is dependent only on the density, $\rho(\mathbf{r})$, and does not vary with Ψ_λ. Ψ_λ therefore minimises $\left\langle \Psi_\lambda \middle| \hat{T} + \lambda \hat{V}_{\text{e-e}} \middle| \Psi_\lambda \right\rangle$.

Consider the derivative of this expectation value

$$\begin{aligned} \frac{\partial}{\partial \lambda} \left\langle \Psi_\lambda \middle| \hat{T} + \lambda \hat{V}_{\text{e-e}} \middle| \Psi_\lambda \right\rangle &= \frac{\partial}{\partial \lambda'} \left\langle \Psi_{\lambda'} \middle| \hat{T} + \lambda \hat{V}_{\text{e-e}} \middle| \Psi_{\lambda'} \right\rangle + \frac{\partial}{\partial \lambda'} \left\langle \Psi_\lambda \middle| \hat{T} + \lambda \hat{V}_{\text{e-e}} \middle| \Psi_\lambda \right\rangle \\ &= 0 + \left\langle \Psi_\lambda \middle| \hat{V}_{\text{e-e}} \middle| \Psi_\lambda \right\rangle , \end{aligned} \tag{12.24}$$

where the first term is zero by the Hellmann–Feynman theorem as Ψ_λ is the minimising wavefunction. Now integrate between 0 and 1 gives

$$\begin{aligned} \int_0^1 \mathrm{d}\lambda \left\langle \Psi_\lambda \middle| \hat{V}_{\text{e-e}} \middle| \Psi_\lambda \right\rangle &= \int_0^1 \mathrm{d}\lambda \, \frac{\partial}{\partial \lambda} \left\langle \Psi_\lambda \middle| \hat{T} + \lambda \hat{V}_{\text{e-e}} \middle| \Psi_\lambda \right\rangle \\ &= \left\langle \Psi_1 \middle| \hat{T} + \hat{V}_{\text{e-e}} \middle| \Psi_1 \right\rangle - \left\langle \Psi_0 \middle| \hat{T} \middle| \Psi_0 \right\rangle \\ &= T[\rho] + V_{\text{e-e}}[\rho] - T_{\text{s}}[\rho] \,. \end{aligned} \tag{12.25}$$

Now subtract the Coulomb energy from both sides

$$T[\rho] + V_{\text{e-e}}[\rho] - T_{\text{s}}[\rho] - J[\rho] = \int_0^1 \mathrm{d}\lambda \left\langle \Psi_\lambda \middle| \hat{V}_{\text{e-e}} \middle| \Psi_\lambda \right\rangle - J[\rho] \,. \tag{12.26}$$

Define

$$W_\lambda \coloneqq \left\langle \Psi_\lambda \middle| \hat{V}_{\text{e-e}} \middle| \Psi_\lambda \right\rangle - J[\rho]\,, \tag{12.27}$$

then we get

$$E_{\text{xc}}[\rho] = \int_0^1 \mathrm{d}\lambda\, W_\lambda\,. \tag{12.28}$$

Thus if we know $W_\lambda[\rho]$, we can use it to create an exchange-correlation functional. In the $\lambda = 0$ limit, we have

$$\begin{aligned} W_0 &= \left\langle \Psi_0 \middle| \hat{V}_{\text{e-e}} \middle| \Psi_0 \right\rangle - J[\rho] \\ &= J[\rho] - K[\rho] - J[\rho] \\ &= -K[\rho] =: E_{\text{x}}[\rho]\,, \end{aligned} \tag{12.29}$$

which is the Hartree–Fock exchange energy of density ρ. For $\lambda = 1$,

$$\begin{aligned} W_1 &= \left\langle \Psi_1 \middle| \hat{V}_{\text{e-e}} \middle| \Psi_1 \right\rangle - J[\rho] \\ &= V_{\text{e-e}}[\rho] - J[\rho]\,, \end{aligned} \tag{12.30}$$

where $V_{\text{e-e}}(\rho)$ contains all two-electron Coulomb, exchange and correlation contributions. To make headway we consider that, for the exact exchange-correlation functional, the energy is identical to the FCI energy, $E_0[\rho] = E_1[\rho]$, so

$$E_0[\rho] = T_{\text{s}}[\rho] + J[\rho] + E_{\text{xc}}[\rho] = E_1[\rho] = T[\rho] + V_{\text{e-e}}[\rho]\,. \tag{12.31}$$

Rearranging this we have

$$W_1 = V_{\text{e-e}}[\rho] - J[\rho] = T_{\text{s}}[\rho] - T[\rho] + E_{\text{xc}}[\rho]\,, \tag{12.32}$$

and so

$$W_1 - E_{\text{xc}}[\rho] = T_{\text{s}}[\rho] - T[\rho] =: -T_{\text{c}}[\rho]\,, \tag{12.33}$$

where $T_{\text{c}}[\rho]$ is the correction to the non-interacting kinetic energy due to correlation.

This form of the exchange-correlation functional is instructive as W_λ can be evaluated for model systems and approximated.

12.4 Asymptotic Behaviour of the Density

An exact result which is of value for the electron density is that asymptotically

$$\rho \sim \exp(-2\sqrt{2I_{\text{min}}}r)\,, \tag{12.34}$$

where I_{min} is the exact first ionisation potential. We shall now prove this.

Consider a point $\mathbf{r}_N$ which is a long way from the molecule. The exact wavefunction for this molecule, which has N electrons may be written in this situation as

$$\Psi_N = \Psi_{N-1}\phi_N(\mathbf{r}_N)\,, \tag{12.35}$$

since there is no chance of having two electrons at such a large distance. Ψ_{N-1} is the wavefunction of the cation. We may write the N electron Hamiltonian as $\hat{H}_N = \hat{H}_{N-1} + \hat{h}_N$, with $\hat{h}_N$ being the one electron Hamiltonian for the last electron.

$$\hat{H}_N = \underbrace{-\sum_{i=1}^{N-1}\frac{1}{2}\nabla_i^2 + \sum_I\sum_{i=1}^{N-1}\frac{-Z_I}{r_{iI}} + \sum_{i<j}^{N-1}\frac{1}{r_{ij}}}_{\hat{H}_{N-1}} \underbrace{-\frac{1}{2}\nabla_N^2 + \sum_I\frac{-Z_I}{r_{NI}} + \sum_{i=1}^{N-1}\frac{1}{r_{iN}}}_{\hat{h}_N}\,. \tag{12.36}$$

An application of $\hat{H}_N$ to Ψ_N yields

$$\hat{H}_N \Psi_{N-1}\phi_N = (\hat{H}_{N-1}\Psi_{N-1})\phi_N + \Psi_{N-1}\hat{h}_N\phi_N \tag{12.37}$$

$$E_N \Psi_{N-1}\phi_N = (E_{N-1}\Psi_{N-1})\phi_N + \Psi_{N-1}\hat{h}_N\phi_N \tag{12.38}$$

$$E_N\phi_N = E_{N-1}\phi_N + \hat{h}_N\phi_N\,, \tag{12.39}$$

so

$$(E_N - E_{N-1})\phi_N = \hat{h}_N\phi_N\,, \tag{12.40}$$

or

$$-\frac{1}{2}\nabla_N^2\phi_N - \frac{1}{r_N}\phi_N = -I_{\min}\phi_N\,. \tag{12.41}$$

At a large distance the behaviour of ϕ_N is exponential, i.e. $\phi_N \sim \exp(-\alpha r)$. Substitution yields $\alpha^2 = 2I_{\min}$. The asymptotic form of the density will be given by $\phi_N(\mathbf{r}_N)^2$ and the result is obtained.

13 Approximate Exchange-Correlation Functionals

13.1 Local Density Approximation

There is no straightforward way in which the exchange-correlation functional $E_{\rm xc}[\rho]$ can be systematically improved, unlike *ab initio* quantum chemistry. In *ab initio* quantum chemistry SCF can be systematically improved in principle to unlimited accuracy through configuration interaction or perturbation theory. One way forward in DFT is to start from a model for which there is an exact solution — the uniform electron gas.

It is defined as a large number N of electrons in a cube of volume $V = l^3$, throughout which there is uniformly spread out a positive charge sufficient to make the system neutral. The uniform electron gas corresponds to the limit $N \to \infty$, $V \to \infty$, with density $\rho = N/V$ remaining finite.

We will now briefly indicate how the expressions for $T_{\rm s}[\rho]$ and $E_{\rm x}[\rho]$ are derived. The Kohn–Sham equations for the uniform electron gas are satisfied by plane wavefunctions

$$\phi_{\mathbf{k}} = \frac{1}{\sqrt{V}} \mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}} \tag{13.1}$$

where periodic boundary conditions require

$$\mathbf{k} = \frac{2\pi}{l}\mathbf{n}\,,\ \mathbf{n} \in \mathbb{Z}^3\,. \tag{13.2}$$

The one particle density matrix is given by

$$\rho_1(\mathbf{r}_1, \mathbf{r}_2) = \frac{2}{V} \sum_{\mathbf{k}\ \text{occupied}} \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2)}\,, \tag{13.3}$$

which, on replacing the sum by an integral with $\mathrm{d}^3\mathbf{n} = (V/8\pi^3)\mathrm{d}^3\mathbf{k}$ yields

$$\begin{aligned} \rho_1(\mathbf{r}_1, \mathbf{r}_2) &= \frac{1}{4\pi^3} \int \mathrm{d}^3\mathbf{k}\, \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} \\ &= \frac{1}{4\pi^3} \int_0^{k_{\rm F}} \mathrm{d}k\, k^2 \int \mathrm{d}\theta\, \mathrm{d}\phi\, \sin\theta \mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{r}_{12}}\,. \end{aligned} \tag{13.4}$$

The *Fermi level* k_F is defined by evaluating the density $\rho(\mathbf{r}) = \rho_1(\mathbf{r}, \mathbf{r})$

$$\rho(\mathbf{r}) = \frac{k_{\rm F}^3}{3\pi^2} \quad \Rightarrow \quad k_F = [3\pi^2\rho(\mathbf{r})]^{1/3}\,. \tag{13.5}$$

The kinetic energy

$$T_{\rm s} = \int \mathrm{d}^3\mathbf{r} \left[-\frac{1}{2}\nabla^2_{\mathbf{r}}\rho_1(\mathbf{r}', \mathbf{r})\right]_{\mathbf{r}'=\mathbf{r}} \tag{13.6}$$

is easily obtained. It is

$$\begin{aligned} T_{\rm s}[\rho] &= \frac{1}{8\pi^2} \int \mathrm{d}^3\mathbf{r} \int_0^{k_{\rm F}} \mathrm{d}k\, 4\pi k^4 \\ &= \frac{1}{10\pi^2} \int \mathrm{d}^3\mathbf{r}\, k_F^5 \\ &= \frac{3}{10}(3\pi^2)^{2/3} \int \mathrm{d}^3\mathbf{r}\, \rho^{5/3}(\mathbf{r})\,. \end{aligned} \tag{13.7}$$

To evaluate $\rho_1(\mathbf{r}_1, \mathbf{r}_2)$ generally we introduce the coordinates

$$\mathbf{r} = \frac{\mathbf{r}_1 + \mathbf{r}_2}{2}\,,\ \mathbf{s} = \mathbf{r}_1 - \mathbf{r}_2 \tag{13.8}$$

and choose $\mathbf{s}$ to lie on the k_z axis. The integral may then be evaluated to give the exact first order spinless density matrix

$$\rho_1(\mathbf{r}_1, \mathbf{r}_2) = 3\rho(\mathbf{r}) \frac{\sin t - t\cos t}{t^3}\,, \tag{13.9}$$

where $t \coloneqq k_{\mathrm{F}}(\mathbf{r})\,\|\mathbf{s}\|$. The complete result is

$$T_{\mathrm{s}}[\rho] = C_{\mathrm{F}} \int \mathrm{d}^3\mathbf{r}\, \rho^{5/3}(\mathbf{r}) \tag{13.10}$$

$$E_{\mathrm{x}}^{\mathrm{LDA}}[\rho] = -C_{\mathrm{x}} \int \mathrm{d}^3\mathbf{r}\, \rho^{4/3}(\mathbf{r})\,, \tag{13.11}$$

where we defined the constants

$$C_{\mathrm{F}} = \frac{3}{10}(3\pi^2)^{2/3} \approx 2.8712 \tag{13.12}$$

$$C_{\mathrm{x}} = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \approx 0.7386\,. \tag{13.13}$$

To derive the result for $E_{\mathrm{x}}^{\mathrm{LDA}}$, we have used the result

$$\int_0^\infty \mathrm{d}t\, (\sin t - t\cos t)^2 t^{-5} = \frac{1}{4}\,. \tag{13.14}$$

The exchange energy is called the *Dirac exchange energy* or sometimes *Slater exchange energy*.

It is usual to introduce the exchange energy per particle ϵ_{x} as a function of r_s, the radius of a sphere whose volume is the effective volume of an electron,

$$\frac{4}{3}\pi r_s^3 = \frac{1}{\rho}\,. \tag{13.15}$$

The results are

$$E_{\mathrm{x}}^{\mathrm{LDA}}[\rho] = \int \mathrm{d}^3\mathbf{r}\, \rho(\mathbf{r})\epsilon_{\mathrm{x}}(r_s)\,, \tag{13.16}$$

where

$$\epsilon_{\mathrm{x}}^{\mathrm{LDA}} = -\frac{0.4582}{r_s}\,. \tag{13.17}$$

In the case when the alpha spin density is not equal to the beta spin density, the kinetic and exchange energies are

$$T_{\mathrm{s}}[\rho_\alpha, \rho_\beta] = 2^{2/3} C_{\mathrm{F}} \int \mathrm{d}^3\mathbf{r}\, [\rho_\alpha^{5/3} + \rho_\beta^{5/3}]\,, \tag{13.18}$$

$$E_{\mathrm{x}}^{\mathrm{LSDA}}[\rho_\alpha, \rho_\beta] = -2^{1/3} C_{\mathrm{x}} \int \mathrm{d}^3\mathbf{r}\, [\rho_\alpha^{4/3} + \rho_\beta^{4/3}]\,, \tag{13.19}$$

as may be verified by putting $\rho_\alpha = \rho_\beta = \rho/2$ for the closed-shell case.

The ground state total energy of the uniform electron gas is

$$E[\rho] = T_{\mathrm{s}}[\rho] + \int \mathrm{d}^3\mathbf{r}\, \rho(\mathbf{r}) V_{\mathrm{ext}}(\mathbf{r}) + J[\rho] + E_{\mathrm{xc}}[\rho] + E_b\,, \tag{13.20}$$

where E_b being the electrostatic energy of the positive background, which is equal to the Coulomb energy because the magnitude of the positive charge density $n(\mathbf{r})$ is equal to $\rho(\mathbf{r})$. Because the external potential is defined by

$$V_{\mathrm{ext}}(\mathbf{r}) = -\int \mathrm{d}^3\mathbf{r}'\, \frac{n(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|}\,, \tag{13.21}$$

it follows that the second, third and fifth terms of (13.20) add to zero, and therefore

$$\begin{aligned} E[\rho] &= T_\mathrm{s}[\rho] + E_\mathrm{xc}[\rho] \\ &= T_\mathrm{s}[\rho] + E_\mathrm{x}[\rho] + E_\mathrm{c}[\rho]\,, \end{aligned} \tag{13.22}$$

where we have now split up the exchange-correlation term into an exchange term plus a correlation term.

For the correlation functional of the electron gas, reliance is placed on the numerical simulations of the uniform electron gas by Ceperley and Alder using the quantum Monte-Carlo method for several different values of r_s. (They studied $r_\mathrm{s} = 1, 2, 5, 10, 20, 50, 100$ for $\zeta = 0$ and 1). The correlation energy was obtained by subtracting the kinetic and exchange energies from (13.22). Also using analytic information for the high and the low density limit, Vosko, Wilk and Nusair gave the first accepted form for the correlation electron density per electron $\epsilon_\mathrm{c}(r_\mathrm{s})$, which is defined such that

$$E_\mathrm{c}[\rho] = \int \mathrm{d}^3\mathbf{r}\, \rho(\mathbf{r})\epsilon_\mathrm{c}(r_\mathrm{s}(\rho(\mathbf{r})))\,, \tag{13.23}$$

which covers both the spin polarised and spin compensated cases.

$$\begin{aligned} \epsilon_\mathrm{c}^\mathrm{LDA}(r_\mathrm{s}) = \frac{A}{2}\Bigg(&\ln\frac{x^2}{X(x)} + \frac{2b}{Q}\tan^{-1}\frac{Q}{2x+b} \\ &- \frac{bx_0}{X(x_0)}\left[\ln\frac{(x-x_0)^2}{X(x)} + \frac{2(b+2x_0)}{Q}\tan^{-1}\frac{Q}{2x+b}\right]\Bigg)\,, \end{aligned} \tag{13.24}$$

where $x = r_\mathrm{s}^{1/2}$, $X(x) = x^2 + bx + c$, $Q = (4c - b^2)^{1/2}$, and $A = 0.0621814$, $x_0 = -0.409286$, $b = 13.0720$, $c = 42.7198$. It was later updated by Perdew and Wang to give the PW91 local correlation functional.

This completes the description of the uniform electron gas and its functionals. The application of these functionals in chemistry is often referred to as the *local density approximation* (LDA) or the *local spin density approximation* (LSDA). For molecular chemistry the errors from LDA energies are often too large to be useful. The LDA also does not bind anions as $\epsilon_\mathrm{x}(\mathbf{r})$ decays too rapidly. Typical errors in E_x is usually 5% for being not negative enough, and E_c are usually too negative by 100%. The mean absolute error in atomisation energies is around 180 kJ mol^{-1}.

The main step forward for the introduction of DFT into chemistry was the introduction of the density gradient $\boldsymbol{\nabla}\rho$.

13.2 Generalised Gradient Approximation

Let's first introduce the dimensionless reduced gradient density

$$x = \frac{\|\boldsymbol{\nabla}\rho\|}{\rho^{4/3}}\,, \tag{13.25}$$

and a related additional factor

$$s = \frac{x}{2(3\pi^2)^{1/3}}\,. \tag{13.26}$$

It can be shown from the gradient expansion that for a slightly inhomogeneous electron gas with a slowly varying density, the exchange energy behaves like

$$\begin{aligned} E_\mathrm{x}^\mathrm{GEA2} &= -\int \mathrm{d}^3\mathbf{r}\, c_\mathrm{D}\rho^{4/3} + \frac{7}{432\pi(3\pi^2)^{1/3}}\frac{\|\boldsymbol{\nabla}\rho\|^2}{\rho^{4/3}} \\ &=: -\int \mathrm{d}^3\mathbf{r}\, c_\mathrm{D}\rho^{4/3} + c_{x^2}\rho^{4/3}x^2\,, \end{aligned} \tag{13.27}$$

where the Dirac exchange constant $c_{\mathrm{D}} = \frac{3}{4}(\frac{3}{\pi})^{1/3}$ and $c_{x^2} = 7/432(3\pi^2)^{1/3}$. This idea stems from the solid state however it did not take off in chemistry until the introduction of the *generalised gradient approximation* (GGA) as the density in molecules is a long way from slowly varying. In fact in molecular tails where the density has an exponential decay then the reduced gradient diverges. This causes big problems for the direct use of the gradient expansion as can be clearly seen for the use of the fourth order correction where the energy diverges for any atom.

Note that although GGA functionals make use of $\boldsymbol{\nabla}\rho$ (via x or s), they are still local. The functionals don't require ρ at two different $\mathbf{r}$ simultaneously. Using the gradient alleviates some of this locality.

- Becke86
$$E_{\mathrm{x}}^{\mathrm{B86}} = E_{\mathrm{x}}^{\mathrm{LDA}} - \sum_{\sigma} \int \mathrm{d}^3\mathbf{r}\, \rho_\sigma^{4/3} \left(\frac{\beta\gamma x_\sigma^2}{1+\gamma x_\sigma^2} \right) . \tag{13.28}$$
A simple form makes sure that the energy does not diverge as $x \to \infty$. This is fitted to atomic exchange energies with $\beta = 0.0036$ and $\gamma = 0.004$.

- Becke88
$$E_{\mathrm{x}}^{\mathrm{B88}} = E_{\mathrm{x}}^{\mathrm{LDA}} - \sum_{\sigma} \int \mathrm{d}^3\mathbf{r}\, \rho_\sigma^{4/3} \left(\frac{\beta x_\sigma^2}{1+6\beta x_\sigma \sinh^{-1}(x_\sigma)} \right) . \tag{13.29}$$
This is one of the most widely used exchange functionals in the literature. Its form was designed to give the correct asymptotic behaviour of the exchange energy density. However, its success is viewed more likely to be due to the procedure used to determine the constant $\beta = 0.0042$, a least squares fit to the Hartree–Fock exchange energies of the rare gas atoms (He-Kr).

- PBE
$$E_{\mathrm{x}}^{\mathrm{PBE}} = E_{\mathrm{x}}^{\mathrm{LDA}} - c_{\mathrm{D}} \sum_{\sigma} \int \mathrm{d}^3\mathbf{r}\, \rho_\sigma^{4/3} \left(\kappa - \frac{\kappa}{1+\frac{\mu s_\sigma^2}{\kappa}} \right) , \tag{13.30}$$
where $\mu = 0.21951$ and $\kappa = 0.804$. These parameters are chosen to satisfy the linear response of the uniform electron gas and also to make sure that the functional never breaks the Lieb–Oxford bound.

- OPTX
$$E_{\mathrm{x}}^{\mathrm{OPTX}} = 1.05151 E_{\mathrm{x}}^{\mathrm{LDA}} - 1.43169 \sum_{\sigma} \int \mathrm{d}^3\mathbf{r}\, \rho_\sigma^{4/3} \left(\frac{\gamma x_\sigma^2}{1+\gamma x_\sigma^2} \right)^2 . \tag{13.31}$$
The coefficients were obtained from fitting the HF energies of the atoms H-Ar. In the uniform gas limit ($x \to 0$) OPTX does not reproduce LDA.

13.3 Correlation Functionals

The development of correlation functionals has been more difficult with less firm theoretic basis on which to base the correlation functional. Even the scaling relation which has been too complicated to help simply in development of functionals.

There are two main correlation functionals that are used in the literature today.

13.3.1 PBE Correlation Functional

In the paper by Perdew, Burke and Ernzerhof they developed a correlation functional based on previous work of Perdew for the correlation energy of the uniform electron gas

$$E_{\mathrm{c}}^{\mathrm{PBE}} = \int \mathrm{d}^3\mathbf{r}\, \rho(\mathbf{r}) \left(\epsilon_{\mathrm{c}}^{\mathrm{LDA}}(r_{\mathrm{s}}, \zeta) + H(r_{\mathrm{s}}, \zeta, t) \right) , \tag{13.32}$$

where r_s is the local Seitz radius $\frac{1}{\rho} = \frac{4}{3}\pi r_\mathrm{s}^3$, $\zeta = (\rho_\alpha - \rho_\beta)/\rho \in [-1, 1]$ is the spin polarisation and $t = \|\boldsymbol{\nabla}\rho\| / (2\phi k_\mathrm{s}\rho)$ is a dimensionless density gradient, in which $\phi(\zeta) = [(1+\zeta)^{2/3} + (1-\zeta)^{2/3}]/2$ is a spin scaling factor and $k_\mathrm{s} = \sqrt{4k_\mathrm{F}/\pi}$ is a screening factor. The gradient correction part is given by

$$H = \gamma\phi^3 \ln\left[1 + \frac{\beta}{\gamma}t^2\left(\frac{1+At^2}{1+At^2+A^2t^4}\right)\right], \tag{13.33}$$

where

$$A = \frac{\beta}{\gamma}\left[\exp\left(\frac{-\epsilon_\mathrm{c}^\mathrm{LDA}}{\gamma\phi^3}\right) - 1\right]^{-1}. \tag{13.34}$$

13.3.2 LYP Correlation Functional

Colle and Salvetti have presented an approximated correlation energy formula for helium in terms of the electron density and a Laplacian of the second order Hartree–Fock density matrix. Their very clever analysis involved (a) a Jastrow representation for the correlation hole (b) a correlation energy expression in terms of this density matrix (c) assuming that the one-particle density matrix maintains its HF form (d) introducing a Taylor expansion for the correlation energy integrand (e) obtaining a good formula for the correlation energy, and then (f) fixing its parameters such that it holds for He. Lee, Yang and Parr restated this formula in terms of the density and the local kinetic energy density. Using gradient expansions for the latter term, they turned this formula into an explicit functional of ρ, involving the gradient and the Laplacian. Miehlich, Savin, Stoll and Preuss eliminated the Laplacian terms by integration by parts, and for a closed-shell system the result is

$$\begin{aligned} E_\mathrm{c} = &-a\int \mathrm{d}^3\mathbf{r}\,\frac{\rho}{1+d\rho^{-1/3}} \\ &- ab\int \mathrm{d}^3\mathbf{r}\,\omega\rho^2\left[C_\mathrm{F}\rho^{8/3} + \|\boldsymbol{\nabla}\rho\|^2\left(\frac{5}{12} - \delta\frac{7}{72}\right) - \frac{11}{24}\rho^2\|\boldsymbol{\nabla}\rho\|^2\right], \end{aligned} \tag{13.35}$$

where

$$\omega = \frac{\exp(-c\rho^{-1/3})}{1+d\rho^{-1/3}}\rho^{-11/3} \text{ and } \delta = c\rho^{-1/3} + \frac{d\rho^{-1/3}}{1+d\rho^{-1/3}} \tag{13.36}$$

and $a = 0.04918$, $b = 0.132$, $c = 0.2533$, $d = 0.349$, which are the Colle–Salvetti parameters from their fit to the helium atom. The great advantage of this functional is that it was derived from an actual correlated wavefunction for a two electron system, and has no relation to the uniform electron gas.

13.4 Hybrid Functionals

The next major advance in the development of functionals came from Becke with the introduction of hybrid functionals. Consider again the formula for the adiabatic connection

$$E_\mathrm{xc} = \int_0^1 \mathrm{d}\lambda\, W_\lambda\,. \tag{13.37}$$

Becke tried a simple model for the adiabatic connection integrand

$$W_\lambda^\mathrm{model} = a + b\lambda\,. \tag{13.38}$$

For this to be correct at $\lambda = 0$ and $\lambda = 1$ we need

$$a = W_0\,,\; b = W_1 - W_0\,. \tag{13.39}$$

The initial point of the adiabatic connection is exact exchange and Becke used LDA approximation W_1^LDA at $\lambda = 1$. The simple integration of the model λ and substitution gives

$$E_\mathrm{xc} = \frac{1}{2}E_\mathrm{x}^\mathrm{HF} + \frac{1}{2}W_1^\mathrm{LDA}\,. \tag{13.40}$$

This is the famous introduction of exact exchange into the functionals. The same year Becke tried a more empirical mixing of Hartree–Fock exchange into the functional and he found a good fit to a small set of atoms and molecules with the following form

$$E_{\mathrm{xc}} = aE_{\mathrm{x}}^{\mathrm{HF}} + (1-a)E_{\mathrm{x}}^{\mathrm{LDA}} + bE_{\mathrm{x}}^{\Delta\mathrm{B88X}} + cE_{\mathrm{c}}^{\mathrm{LDA}} + (1-c)E_{\mathrm{c}}^{\mathrm{GGA}}\,. \tag{13.41}$$

He found the parameters $a = 0.2$, $b = 0.72$ and $c = 0.81$. This leads to the most widely used functional B3LYP.

More recent work has tried to develop the adiabatic connection work with a more in depth study of possible models. Some success was had with a [1,1] Padé model

$$W_\lambda^{\mathrm{model}} = a + \frac{b\lambda}{1+c\lambda} \tag{13.42}$$

using some sort of information from W_0, W_0' and W_p. This gives rise to an expression

$$E_{\mathrm{xc}} = a + \frac{b}{c}\left(1 - \frac{\ln(1+c)}{c}\right)\,, \tag{13.43}$$

where

$$a = W_0 \tag{13.44}$$

$$b = W_0' \tag{13.45}$$

$$c = \frac{W_p - pW_0' - W_0}{p(W_p - W_0)}\,. \tag{13.46}$$

13.5 Fitted Exchange-Correlation Functionals

The exact properties of the exchange-correlation functionals can be very difficult to impose and also they may not necessarily help! For example the uniform electron gas condition that has been a driving force for the development of functionals — how is it helpful for chemistry when we never have a density that looks uniform? This is obviously an open question for debate. Many people have adopted a more pragmatic approach to the development of functionals and taken a flexible functional form that is fitted to a large amount of highly accurate experimental data. This is typified by the G2 set from Pople and co-workers where they try to find accurate experimental heats of formation for molecules to within 1 kcal/mol (the standard of chemical accuracy). For example the 10 parameter B97 functional and 15 parameter HCTH were some early examples fitted to data from the G2 set (with 93 systems). More recent work especially from the group of Truhlar in Minnesota have come up with a family of functionals, of which the best general performance functional is probably M06-2X with around 41 parameters fitted to more than 300 systems.

13.6 Other Ideas in Functionals

There have been several recent ideas in functional development

- Range separation

$$\frac{1}{r_{12}} \equiv \frac{\mathrm{erf}(\mu r_{12})}{r_{12}} + \frac{1-\mathrm{erf}(\mu r_{12})}{r_{12}}\,, \tag{13.47}$$

  so we can use DFT to short range and HF for long range. It is also the other way for screened hybrids. Examples of such functionals are the long-range corrected CAM-B3LYP and the screened HSE functional.

- Dispersion (Van der Waals) correction.

  DFT performs poorly for many dispersion related problems and is thought to be incorrect in the asymptotic limit. An r^{-6} empirical correction of the form

$$E_{\text{disp}} = -s_6 \sum_{i<j}^{N_{\text{at}}} \frac{C_6^{ij}}{R_{ij}^6} f_{\text{damp}}(R_{ij}) \tag{13.48}$$

  is added to give methods often denoted DFT-D.

- RPA

  RPA makes use of the unoccupied orbitals and eigenvalues from methods based on the many-body Green's function. It also helps to correct dispersion related problems.

- OEP

  Any energy expression can be minimised with respect to a local potential in their minimising equations, using a basis set expansion for the potential as well

$$\left(-\frac{1}{2}\nabla^2 + v_0(\mathbf{r}) + \sum_t b_t f_t(\mathbf{r}) \right) \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \,, \tag{13.49}$$

$$\frac{\partial E}{\partial v(\mathbf{r})} = 0 \,. \tag{13.50}$$

  This can place methods such as B3LYP firmly within Kohn–Sham theory. Note, the OEP potential can be determined exactly (up to a constant) if one knows any of the Kohn–Sham orbitals. For example, this is the case given the density in any one orbital system (one electron or closed-shell two electrons)

$$v_{\text{s}}(\mathbf{r}) = -\frac{\nabla^2 \phi_i(\mathbf{r})}{2\phi_i(\mathbf{r})} + C \,. \tag{13.51}$$

14 Energy versus Number of Electrons

The behaviour of energy versus number of electrons is of key importance in understanding some very important processes in electronic structure as well as more recently helping to understand the behaviour of approximate functionals.

14.1 Kohn–Sham Eigenvalues and Ionisation Energies

The interpretation of the Kohn–Sham eigenvalues is intimately related to the dependence of total energy on the number of electrons, i.e. ionisation potentials I and electron affinities A. These quantities are rigorously defined as the differences in total energy

$$I = E_{N-1} - E_N \tag{14.1}$$

$$A = E_N - E_{N+1}\,, \tag{14.2}$$

where E_N is the total energy of an N-electron system. In DFT, ionisation potentials and electron affinities are estimated from fully self-consistent total energies using approximate density functionals. I and A values computed this way are referred to as Δ-SCF estimates. In Hartree–Fock, Koopmans' theorem says I and A are related to the HOMO and LUMO energy eigenvalues

$$I = -\epsilon_{\mathrm{HOMO}} = -\epsilon_N(N)\,, \tag{14.3}$$

$$A = -\epsilon_{\mathrm{LUMO}} = -\epsilon_{N+1}(N)\,, \tag{14.4}$$

where $\epsilon_n(N)$ is the eigenvalue of spin-orbital number n of an N-electron system (which are generally negative). These numbers are derived using a frozen orbital approximation, where orbitals are not relaxed in response to removal or addition of electrons. In the Kohn–Sham approach we normally work with relaxed orbitals and (14.3, 14.4) does not hold. In general one finds I is underestimated and A is overestimated. For GGA functionals, provided molecules are small, Δ-SCF estimates are usually good (errors on the order of $0.1 \sim -0.5$ eV). Estimates from HOMO and LUMO are usually much worse (errors ~ 1 eV).

14.2 Non-Integral Numbers of Electrons

The work of Perdew, Parr, Levy and Balduz demonstrated using the grand canonical ensemble the simple result that the behaviour of the exact energy in density functional theory for non-integer numbers of electrons is a straight line interpolation between the nearest integers

$$E(N+\delta) = (1-\delta)E(N) + \delta E(N+1)\,. \tag{14.5}$$

14.3 Meaning of the Eigenvalues

It is easy to show from an extension to non-integer occupation numbers n_i with a density

$$\rho = \sum_i n_i \left|\phi_i\right|^2\,, \tag{14.6}$$

the energy is

$$E[\rho] = \sum_i n_i \left\langle \phi_i \middle| -\frac{1}{2}\nabla^2 \middle| \phi_i \right\rangle + \int \mathrm{d}^3\mathbf{r}\,\rho(\mathbf{r})v(\mathbf{r}) + J[\rho] + E_{\mathrm{xc}}[\rho]\,. \tag{14.7}$$

By taking derivatives, Janak showed that for simple functionals such as LDA and GGA,

$$\frac{\partial E}{\partial n_i} = \epsilon_i\,. \tag{14.8}$$

14.4 Band Gap

The band gap is defined as the energy difference between electron removal and electron addition

$$\begin{aligned} E_{\text{gap}} &= [E(N+1) - E(N)] - (E(N) - E(N-1)) \\ &= I - A\,. \end{aligned} \tag{14.9}$$

However due to the behaviour of the E versus N curve, we can also get the band gap from the derivatives at the integers

$$E_{\text{gap}} = \left.\frac{\partial E}{\partial N}\right|_{+} - \left.\frac{\partial E}{\partial N}\right|_{-}\,. \tag{14.10}$$

The Kohn–Sham expression for the bandgap is given by

$$\begin{aligned} E_{\text{gap}} &= \epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}} + \left.\frac{\partial E}{\partial N}\right|_{+} - \left.\frac{\partial E}{\partial N}\right|_{-}\,. \\ &= \epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}} + \Delta_{\text{xc}}\,. \end{aligned} \tag{14.11}$$

This expression has led to a fair bit of confusion in the literature. In part this is due to the very poor performance of approximate functionals for the E vs N curve. For all functionals of the LDA or GGA form, $\Delta_{\text{xc}} = 0$. This means that the difference of derivatives at N is given by the difference of the Kohn–Sham eigenvalues. However if we look at the typical values for the HOMO-LUMO gap in atoms or small molecules they underestimate the experimental gap by around 100 kcal mol^{-1} (5 eV) (In solids the difference is smaller though LDA/GGA typically underestimates the gap by $30-50\%$).

The problems of approximate density functionals for the band-gap indicate much bigger problems for calculation of the ground state energy of molecules. This is extremely important for the future improvement of DFT.

15 Practical Information

15.1 Numerical Quadrature

This is one of the difficulties of DFT. It is quite clear that the new integrals which arise in the Kohn–Sham equations may not be evaluated by analytic means because of the fractional powers of the density which arise. There are various ways forward, and they all involve a grid of points in the 3D real space. Some people favour a least squares fit of $v_{\rm xc}$ to an auxiliary Gaussian basis set, some people favour a completely numerical approach. Here we shall simply evaluate the required integrals using quadrature, and we now describe a quadrature scheme which we have found satisfactory.

- Space is divided into atomic Voronoi polyhedra such that the integral can be written as a sum over atomic contributions.
- Around each atom a radial and angular quadrature is chosen. For example the radial quadrature can be determined using Euler–Maclaurin, Treutler–Ahlrichs or several other schemes. For the 2D angular quadrature originally Gauss–Legendre was used but Lebedev has devised a very good angular scheme which exactly integrates the spherical harmonics $Y_{lm}(\theta,\phi)$ for all $-l \le m \le l$, $0 \le l \le L$ for some L. There are $(L+1)^2$ such functions. Lebedev described 194, 302 point quadrature schemes for which $L = 23, 29$ and are based on the symmetry of the octahedron.

In the end the integral is represented as

$$\begin{aligned} E_{\rm xc}[\rho] &= \int {\rm d}^3\mathbf{r}\, f_{\rm xc}(\rho(\mathbf{r}), \boldsymbol{\nabla}\rho(\mathbf{r})) \\ &= \sum_m {\rm wt}(m) f_{\rm xc}(\rho(m), \boldsymbol{\nabla}\rho(m))\,. \end{aligned} \tag{15.1}$$

The standard grids have approximately 30,000 points per atom. It is possible to substantially reduce the cost of the quadrature by (i) using symmetry (ii) reducing the number of angular points near nuclei (iii) carrying out Kohn–Sham iterations using low cost quadrature, and obtaining the energy by higher accuracy quadrature (iv) inserting several tests to eliminate the evaluation of basis functions which are far from grid points (v) evaluating the Becke weights at the start and sorting them.

15.2 Implementation of the Kohn–Sham Equations

For example the Kohn–Sham scheme has been implemented as part of the CADPAC program. It is not relevant to give the details here, but for typical molecular calculations the cost is approximately twice that of an SCF calculation. Essentially the structure for the extra part is

DO 2 Atomic r, θ, ϕ points. This determines grid points
DO 3 basis functions
evaluate η_α and $\boldsymbol{\nabla}\eta_\alpha$
END 3
DO 4 basis functions
DO 5 basis functions
form ρ etc
END 5,4
Form v_{xc}
DO 6 basis functions
DO 7 basis functions
construct Kohn–Sham matrix
END 7,6
END 2,1

From the looping structure in this outline it is easy to see that the basis cost is $\mathcal{O}(N^3)$, where N is a measure of the size of the molecule. It is a relatively trivial matter to further reduce the cost, by predetermining those basis functions which have a significant value at the particular grid point. The evaluation is then formally linear.

We observe that there are a large number of DFT codes in the literature under names such as GAUSSIAN/DFT, CADPAC, TURBOMOLE, DGAUSS, DeMon, ADF, Molpro, QCHEM, DALTON, each of which treats the evaluation of the Kohn–Sham matrix in a slightly different way.

Finally in this section we stress that for a given functional, it is possible to obtain the exact Kohn–Sham solution, provided a complete basis and exact quadrature are used. Our experience is that this is in practice achievable provided a TZ2P(+f) type basis is used, that is a good SCF basis. We can see every reason for using as reliable a quadrature as possible. The great advantage of DFT is that no configuration interaction calculation is performed, and therefore we are not trying to describe the electron-electron cusp in the wavefunction, which is well established to be the reason why correlated calculations are so slowly convergent with respect to basis set. This is the overriding advantage of modern density functional theory.

16 Challenges for DFT

16.1 Multi-Reference and Strongly Correlated Systems

As the Kohn–Sham equations rely on a single Slater determinant to produce the Kohn–Sham orbitals, systems where there are many different configurations close in energy (e.g. transition metals in low spin states) can be particularly challenging. A very revealing example is the lack of ability of most functionals to dissociate both H_2 and H_2^+, and similar problems are seen breaking other bonds. The nature of the problem is often described in terms of self-interaction, where the Coulomb contribution of an electron interacting with itself can only be removed by including 100% exact Hartree–Fock exchange, which is not usually included in most functionals.

DFT is particularly widely used in periodic (solid-state) systems where it has generally proven very successful. However in systems where there are atoms with multiple oxidation states (d- and f-block elements), there is a tendency for density functionals to delocalise electrons. This is often patched up (unsatisfactorily) by adding an empirical interaction penalising two electrons in the same spatial orbitals. This is known as DFT+U, where the U parameter stems from fitting a Hubbard model behaviour. This unfavourable U interaction encourages electrons to localise singly on atoms, though is a rather blunt instrument, and this approach needs careful benchmarking.

16.2 Charge Transfer

There are extensions to DFT to describe excited states (time-dependent DFT) which have been shown to be reasonably successful at describing localised excitations, but the same self-interaction errors above mean that charge-transfer excitations are very poorly described, and have energies which are too low.

16.3 Dispersion and Weak Interactions

The local form of the correlation functionals means that they cannot describe the spontaneous-dipole induced-dipole between separated atoms or molecules which give rise to London dispersion forces. In some weakly bonded systems this can be the majority of the binding energy, and so empirical corrections have been added (DFT-D) which include these explicitly. The difficulty of performing more accurate *ab initio* calculations on systems of sufficient size for this to be important means these corrections have not been widely tested beyond some benchmark systems.

16.4 Systematically Improvable Functionals

Though there has been a great deal of work parameterising functionals over the past thirty years, creating a pseudo-systematically improvable Jacobs Ladder of functionals, there has been very little successful development in the form of the functional, and once at the top of the ladder, for systems which still defy description at that level there is currently no direction to take towards more accurate descriptions (unlike the successive levels of accuracy in e.g. coupled cluster theory). As computer power has increased, there may be ways to include non-local effects into the functionals, which (aside from the inclusion of Hartree–Fock exchange) are not generally used in present paradigms.

17 Quantum Monte Carlo Methods

17.1 Variational Monte Carlo

To include e-e and e-n cusps, we write the wavefunction as

$$\Psi = \|\phi_1 \dots \phi_N\| \exp(J(\{\mathbf{r}_i\}, \{\mathbf{R}_i\})) \,, \tag{17.1}$$

where J is the Jastrow factor

$$J = \sum_{i<j} u(r_{ij}) + \sum_{i,I} \chi_I(r_{iI}) \,. \tag{17.2}$$

The explicit forms of u and χ_I vary but allow the e-e and e-n cusp conditions to be built in. An example is

$$u(r_{ij}) = \frac{a r_{ij}}{a + b r_{ij}} \,. \tag{17.3}$$

Unfortunately, the energy of this form of the wavefunction involves integrals which cannot be performed analytically. Instead we can use Monte Carlo sampling to calculate energy. Denote $\mathsf{r} = \{\mathbf{r}_i\}_{i=1}^N$.

$$\begin{aligned} E &= \frac{\int \mathrm{d}^{3n}\mathsf{r}\, \Psi^*(\mathsf{r}) \hat{H} \Psi(\mathsf{r})}{\int \mathrm{d}^{3n}\mathsf{r}\, \Psi^*(\mathsf{r}) \Psi(\mathsf{r})} \\ &= \int \mathrm{d}^{3n}\mathsf{r}\, \frac{\Psi^*(\mathsf{r})\Psi(\mathsf{r})}{N} \frac{(\hat{H}\Psi(\mathsf{r}))}{\Psi(\mathsf{r})} \\ &= \int \mathrm{d}^{3n}\mathsf{r}\, P(\mathsf{r}) E_{\mathrm{loc}}(\mathsf{r}) \\ &= \frac{1}{N_{\mathrm{T}}} \sum_{\mathsf{r}_\mu \sim P(\mathsf{r})}^{N_{\mathrm{T}}} E_{\mathrm{loc}}(\mathsf{r}_\mu) \,. \end{aligned} \tag{17.4}$$

17.1.1 Metropolis Sampling

To generate a distribution according to $P(\mathsf{r})$, we start from a point r consider the move $\mathsf{r} \to \mathsf{r}'$, which we accept with probability

$$P_{\mathrm{accept}} = \begin{cases} 1 & \text{if } P(\mathsf{r}') > P(\mathsf{r}) \\ \frac{P(\mathsf{r}')}{P(\mathsf{r})} & \text{otherwise} \,. \end{cases} \tag{17.5}$$

We perform this with a set of 'walkers', each with its own value of r, and sample E accordingly. We can also sample, e.g. $\frac{\partial E}{\partial b}$ and use this to optimize our wavefunction variationally.

17.2 Diffusion Monte Carlo

For Hamiltonian

$$\hat{H} = -\frac{1}{2}\nabla_{\mathsf{r}}^2 + V(\mathsf{r}, \mathsf{R}) \,, \tag{17.6}$$

we consider the time-dependent Schrödinger equation

$$\mathrm{i}\hbar \frac{\partial}{\partial t} |\Psi\rangle = \hat{H} |\Psi\rangle \tag{17.7}$$

$$\Rightarrow \qquad \frac{\partial \Psi}{\partial \tau} = -\hat{H} |\Psi\rangle \,, \tag{17.8}$$

where $\tau = \mathrm{i}t$ is the imaginary time. The state $|\psi\rangle$ is expanded in eigenstates of the Hamiltonian

$$|\psi\rangle = \sum_{i=1}^{\infty} c_i \, |\phi_i\rangle \ , \tag{17.9}$$

where

$$\hat{H} \, |\phi_i\rangle = \epsilon_i \, |\phi_i\rangle \ . \tag{17.10}$$

A formal solution of the imaginary time Schrödinger equation is

$$|\Psi(\tau_0 + \delta\tau)\rangle = \mathrm{e}^{-\hat{H}\delta\tau} \, |\psi(\tau_0)\rangle \ , \tag{17.11}$$

where the state $|\psi\rangle$ evolve from imaginary time τ_0 to a later time $\tau_0 + \delta\tau$. If the initial state $|\tau_0\rangle$ is expanded in energy ordered eigenstates, then

$$|\psi(\delta\tau)\rangle = \sum_{i=0}^{\infty} c_i \mathrm{e}^{-\epsilon_i \delta\tau} \, |\phi_i\rangle \ . \tag{17.12}$$

Hence any initial state $|\psi\rangle$ that is not orthogonal to the ground state $|\phi_0\rangle$ will evolve to the ground state in the long time limit

$$\lim_{\tau\to\infty} |\Psi(\tau)\rangle = c_0 \mathrm{e}^{-\epsilon_0 \tau} \, |\phi_0\rangle \ . \tag{17.13}$$

To find the ground state, we therefore need to solve the imaginary-time Schrödinger equation

$$\frac{\partial \Psi}{\partial \tau} = \frac{1}{2} \nabla_{\mathsf{r}}^2 \Psi - V(\mathsf{r}, \mathsf{R}) \Psi \, . \tag{17.14}$$

This is the diffusion equation with an extra source/sink term.

We may therefore solve the Schrödinger equation by creating a set of walkers, each containing a complete configuration of electrons, r_μ , which diffuse classically, so each walkers position is τ-dependent, $\mathsf{r}_\mu(\tau)$. The potential term requires these walkers to also contain a weight, $w_\mu(\mathsf{r})$, which is changed by the local value of $V(\mathsf{r})$. The energy is calculated from the (imaginary) time-average of the local energy of all the walkers (once they have reached sufficient imaginary time, τ_0, to be considered relaxed to the ground state):

$$E = \int_0^{\infty} \mathrm{d}\tau \sum_{\mu} E_{\mathrm{local}}(\mathsf{r}_\mu(\tau)) w_\mu(\tau) \, . \tag{17.15}$$

To perform this computationally we create a set of rules which simulate diffusion of walkers, as well as including processes to allow them to increase/decrease their population when the potential is very low/high.

This is known as *diffusion Monte Carlo* (DMC). In essence, the problem has been recast by two steps:

- representing the wavefunction as a time-averaged population of moving walkers rather than a $3n$-dimensional function;
- evolving this population with a stochastic process whose long-(imaginary) time limit is the ground state wavefunction.

17.2.1 Problems with Diffusion Monte Carlo

For electrons, Ψ has nodes as it is antisymmetric, so is not positive everywhere. Naively using DMC will produce the ground state bosonic wavefunction. To force this to represent fermionic electrons, we impose a boundary condition, forcing a nodal surface from a trial wavefunction Ψ_{T}. This is the

fixed node approximation. For a given nodal surface defined by $\Psi_\mathrm{T} = 0$, DMC will give the lowest energy under those conditions, so the problem reduces to finding an appropriate Ψ_T.

To get Ψ_T correct, we go back to quantum chemistry, and generally use a Slater determinant, or a Slater–Jastrow wavefunction (which may have been optimized by Variational Monte Carlo), or possibly a multi-configurational wavefunction from e.g. CASSCF. $E[\Psi_\mathrm{T}]$ will in general give an error, but this is systematically improvable. Overall, DMC is in general very accurate (better than CCSD(T)) and potentially low-scaling ($\mathcal{O}(N^3)$), albeit with a large prefactor. It is also embarrassingly parallel, and has seen a small, but growing usage over the past 20 years both in benchmarking and for real-world-scale calculations.

17.3 Fock Space Quantum Monte Carlo

One potential route to avoid approximations like the fixed-node approximation is to instead perform the Monte Carlo sampling in a space which is already antisymmetric, so the lowest state is fermionic. An appropriate such space is for example Slater determinant space, where the wavefunction expressed in terms of only Slater determinants can never collapse to the bosonic wavefunction.

17.3.1 Full Configuration Interaction Quantum Monte Carlo

Full configuration interaction quantum Monte Carlo (FCIQMC) is essentially the same problem as full-CI, but instead of finding the lowest eigenfunction by diagonalizing a matrix, a stochastic process is used to evolve the coefficients of the ground-state wavefunction. We may consider a conventional CI parameterization of the wavefunction

$$\Psi = \sum_I C_I \Phi_I \,. \tag{17.16}$$

When substituted into the imaginary-time Schrödinger equation as in DMC, we get

$$\frac{\partial \Psi}{\partial \tau} = -\hat{H}\Psi \,. \tag{17.17}$$

We would like to achieve a stable stationary ground state of this, when

$$\hat{H}\Psi = E\Psi \,, \tag{17.18}$$

i.e. $(\hat{H} - E)\Psi = 0$, and so we can in general shift the Hamiltonian by an as yet undetermined constant, E, to achieve stability. (This is also done in DMC).

$$\frac{\mathrm{d}\Psi}{\mathrm{d}\tau} = -(\hat{H} - E)\Psi \tag{17.19}$$

$$\sum_I \frac{\mathrm{d}C_I}{\mathrm{d}\tau}\Phi_I = -\sum_I C_I(\hat{H} - E)\Phi_I \tag{17.20}$$

$$\sum_I \frac{\mathrm{d}C_I}{\mathrm{d}\tau}\langle \Phi_J|\Phi_I\rangle = -\sum_I C_I \left\langle \Phi_J \middle| \hat{H} - E \middle| \Phi_I \right\rangle \tag{17.21}$$

$$\frac{\mathrm{d}C_J}{\mathrm{d}\tau} = -\sum_I (H_{JI} - E\delta_{JI})C_I \,. \tag{17.22}$$

This may be solved by finite-(imaginary)-timestep integration

$$\begin{aligned} C_J(\tau + \delta\tau) &= C_J(\tau) - \delta\tau \sum_I (H_{JI} - E\delta_{JI})C_I(\tau) \\ &= [1 - \delta\tau(H_{JJ} - E)]C_J(\tau) - \delta\tau \sum_{I\neq J} H_{JI}C_I(\tau) \,. \end{aligned} \tag{17.23}$$

The crucial second step in this process is to discretize the problem. Instead of C_J being represented by an arbitrary-precision real number, we choose to represent it as a population of particles (psi-particles or psips, also known as walkers akin to DMC) which are discrete. To obey (17.23) we must probabilistically kill particles (first term) or spawn new particles at J from other particles at I (second term).

The resulting dynamics are quite complicated, but it turns out that for many wavefunctions the wavefunction can be well-represented by orders of magnitude fewer particles than there are Slater determinants. Slater determinant spaces of 10^{38} have been simulated, and with further approximations up to 10^{108}. The wavefunction is still on-average exact however, and can be evaluated from the (imaginary-)time-average

$$\Psi = \sum_I \langle C_I(\tau) \rangle_\tau \Phi_I \,. \tag{17.24}$$

The occupied Slater determinants change over (imaginary-)time, so for very low-C_I determinants, there is seldom a particle there.

This method is not without challenges however, and a critical population of particles is needed for it to remain stable. This size of this critical population is usually a small fraction of the overall Slater determinant space, but has been shown to still scale exponentially with system size.

17.3.2 Coupled Cluster Monte Carlo

To avoid the inherent exponential scaling, it is possible to choose a polynomially scaling ansatz for the wavefunction and apply Monte Carlo to its solution. *Coupled cluster quantum Monte Carlo* (CCQMC) follows essentially the same dynamics as FCIQMC, but instead of representing the configuration interaction amplitudes (C_I) discretely, it represents excitation amplitudes (e.g. t_i^a in $\hat{T}_1 = \sum_{i,a} t_i^a \hat{a}_i^a$). The dynamics of this are even more complicated, but it can recover the polynomial scaling of coupled cluster methods along with the reduced storage requirements from the discrete representation.