Lecture Notes
in
Computational Chemistry

Electronic Structure Theory

Jürg Hutter

Physical Chemistry Institute

University of Zurich

Winterthurerstrasse 190

8057 Zurich, Switzerland

`hutter@pci.unizh.ch`

August 11, 2005

Contents

Part I

Lectures

Lecture 1

Basic Quantum Mechanics

1.1 Operators in quantum mechanics

An observable is a dynamical, measurable variable of a system. In classical physics observables are described using functions, in quantum mechanics operators are used. An operator is a symbol for a recipe how an operation on a function has to be performed.

1.1.1 Eigenfunctions and eigenvalues

In general an operator creates a new function when it is applied to a function. In special cases the new function is the same as the original one multiplied by a constant. A function f is an eigenfunction of an operator, if

$$\hat{\Omega} f = \omega f \ ,$$

where ω is a constant. Such an equation is called an eigenvalue equation. ω is the eigenvalue of the operator $\hat{\Omega}$.

The set of all eigenfunctions of an operator is a complete basis, this means an arbitrary function can be written as a linear combination of eigenfunctions.

$$\hat{\Omega} f_n = \omega_n f_n \ .$$

A function g is expanded in eigenfunctions

$$g = \sum_n c_n f_n \ ,$$

with c_n the expansion coefficients and the sum runs over all functions f_n. We can easily calculate the effect of the operator $\hat{\Omega}$ on the function g

$$\hat{\Omega} g = \hat{\Omega} \sum_n c_n f_n = \sum_n c_n \hat{\Omega} f_n = \sum_n c_n \omega_n f_n$$

3

Quantum mechanics can be developed without explicitly defining operators. However, it is instructive and intuitive to use special forms of operators. The most common uses the following form for position and momentum operator (position representation)

$$x \longrightarrow x \cdot \quad , \qquad p_x \longrightarrow \frac{h}{2\pi i} \frac{\partial}{\partial x} \quad .$$

The special form of the momentum operator follows from the postulates introduced in the next section.

1.1.2 Commutators

The order in which operators are applied to a function is important. Generally, we have

$$ABf \neq BAf \quad ,$$

or simply $AB \neq BA$. The quantity $AB - BA$ is called the commutator of A and B.

$$[A, B] = AB - BA \quad .$$

Operators with the same set of eigenfunctions commute

$$\begin{aligned}
ABg &= BAg \\
A \sum_n \omega_n^B c_n f_n &= B \sum_n \omega_n^A c_n f_n \\
\sum_n \omega_n^A \omega_n^B c_n f_n &= \sum_n \omega_n^B \omega_n^B c_n f_n
\end{aligned}$$

The opposite is also true; operators that commute can be diagonalized with the same set of functions.

The commutator for position and momentum operator in the special representation introduced above is

$$\begin{aligned}
[x, p_x]f = (xp_x - p_x x)f &= x \cdot \frac{h}{2\pi i} \frac{\partial f}{\partial x} - \frac{h}{2\pi i} \frac{\partial}{\partial x} \cdot xf \\
&= x \cdot \frac{h}{2\pi i} \frac{\partial f}{\partial x} - \frac{hf}{2\pi i} - x \cdot \frac{h}{2\pi i} \frac{\partial f}{\partial x} \\
&= \frac{ih}{2\pi} \cdot f
\end{aligned}$$

This is true for any function f, therefore we write

$$[x, p_x] = \frac{ih}{2\pi}$$

4

1.1.3 Construction of operators

Many operators we will need can be constructed from the operators for position and momentum. For example the operator of the kinetic energy $T = p^2/2m$ is in one dimension

$$\hat{T} = \frac{p_x^2}{2m} = \frac{1}{2m}\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)^2 = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}$$

and in three dimensions

$$
\begin{aligned}
\hat{T} &= \frac{p_x^2 + p_y^2 + p_z^2}{2m} = \frac{1}{2m}\left\{\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right)^2 + \left(\frac{\hbar}{i}\frac{\partial}{\partial y}\right)^2 + \left(\frac{\hbar}{i}\frac{\partial}{\partial z}\right)^2\right\} \\
&= -\frac{\hbar^2}{2m}\left\{\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right\} \\
&= -\frac{\hbar^2}{2m}\nabla^2
\end{aligned}
$$

where $\hbar = h/2\pi$ and the operator ∇^2 is the Laplacian.

The potential energy of a particle $V(x)$ is a multiplicative function in the position representation. This is also the case for the operator of the potential energy, for example the operator of the Coulomb energy of an electron in the field of a nuclei with charge Z

$$\hat{V} = -\frac{Ze^2}{4\pi\epsilon_0 r}.$$

The operator of the total energy of a system is the Hamilton operator

$$\hat{H} = \hat{T} + \hat{V}$$

The general recipe for the construction of quantum mechanical operators is:

1. Write down the classical expression for the observable using the position and momentum variables of the system.

2. Replace the different functions with their quantum mechanical operators.

1.2 Postulates of quantum mechanics

1.2.1 Postulates

The following postulates are not complete and in their most strict form, but they are sufficient for our purposes.

Table 1.1: Transition from classical expressions to quantum mechanical operators.

classical	quantum mechanical
Position	
e.g. x, y	$x\cdot$, $y\cdot$
Functions of position	multiplicative operators
e.g. x^2 y	$x^2 y\cdot$
or general f(x,y,z)	$f(x,y,z)\cdot$
Momentum	
e.g. p_x	$\frac{\hbar}{i}\frac{\partial}{\partial x}$
Functions of momentum	differential operators
e.g. $(p_x)^2$	$-\hbar^2\frac{\partial^2}{\partial x^2}$
or general g(p_x,p_y,p_z)	$g(\frac{\hbar}{i}\frac{\partial}{\partial x}, \frac{\hbar}{i}\frac{\partial}{\partial y}, \frac{\hbar}{i}\frac{\partial}{\partial z})$
Functions of position and momentum	
e.g. $\vec{l} = \vec{r} \times \vec{p}$	$(x,y,z) \times \frac{\hbar}{i}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$
Time t	$t\cdot$
Energy E	$i\hbar\frac{\partial}{\partial t}$

Postulate 1

The state of a system is completely characterized by its wave function $\Psi(r_1, r_2, \ldots, t)$.

The positions $r_1, r_2, \ldots$ are associated with particles 1, 2, ... In general additional internal coordinates (spin) are needed. From the wave function all information on the properties of a system can be derived, that are also accessible by experiments.

Postulate 2

Observables are represented by operators which have the following properties

$$[q, p_{q'}] = i\hbar\delta_{q,q'} \ , \qquad [q, q'] = 0 \ , \qquad [p_q, p_{q'}] = 0 \ ,$$

where q and q' are one of the coordinates x, y, z and p_q and $p_{q'}$ are the corresponding momenta.

With this postulate all operators that are functions of position and momentum coordinates are defined. Internal (spin) coordinates have to be defined separately.

Postulate 3

For a system characterized by its wave function Ψ the average value of a series of measurements of the observable Ω is given by the expectation value of the corresponding operator.

The expectation value of an operator $\hat{\Omega}$ for a general state Ψ is defined by

$$\langle\Omega\rangle = \frac{\int \Psi^*\hat{\Omega}\Psi d\tau}{\int \Psi^*\Psi d\tau} \quad .$$

If the wave function is normalized

$$\int \Psi^*\Psi d\tau = 1 \quad ,$$

we get for the expectation value

$$\langle\Omega\rangle = \int \Psi^*\hat{\Omega}\Psi d\tau \quad .$$

We will assume that state functions are normalized. If Ψ is an eigenstate of Ω with eigenvalue ω, we have

$$\langle\Omega\rangle = \int \Psi^*\hat{\Omega}\Psi d\tau = \int \Psi^*\omega\Psi d\tau = \omega\int \Psi^*\Psi d\tau = \omega \quad .$$

A series of measurements on identical systems will give the averaged value ω. Assuming that Ψ is an eigenstate of the system, but not of $\hat{\Omega}$, we can expand Ψ in the eigenfunctions of $\hat{\Omega}$

$$\Psi = \sum_n c_n\Psi_n \quad , \qquad \text{where } \Omega\Psi_n = \omega_n\Psi_n \text{ gilt.}$$

The expectation value is

$$\begin{aligned}
\langle\Omega\rangle &= \int \left(\sum_m c_m\Psi_m\right)^* \Omega \left(\sum_n c_n\Psi_n\right) d\tau = \sum_{nm} c_m^* c_n \int \Psi_m^*\hat{\Omega}\Psi_n d\tau \\
&= \sum_{nm} c_m^* c_n \omega_n \int \Psi_m^*\Psi_n d\tau \quad .
\end{aligned}$$

7

Because the eigenfunctions are an orthonormal basis, the last integral vanishes for $m \neq n$ and we get

$$\langle \Omega \rangle = \sum_{nm} c_m^* c_n \omega_n \int \Psi_m^* \Psi_n d\tau = \sum_n c_n^* c_n \omega_n = \sum_n |c_n|^2 \omega_n \ .$$

The expectation value is therefore a weighted sum of the eigenvalues of Ω. The contribution of a certain eigenvalue is given by the square of the absolute value of the corresponding expansion coefficient.

Postulate 3'
If Ψ is an eigenfunction of $\hat{\Omega}$ all measurements will give the value ω. Is Ψ not an eigenfunction of $\hat{\Omega}$ one will get with each measurement a value ω_n, an eigenvalue of $\hat{\Omega}$. The probability to get a certain eigenvalue is proportional to $|c_n|^2$, where c_n is the overlap of Ψ with the eigenfunction Ψ_n of $\hat{\Omega}$.

Postulate 4
The probability to find a particle in the volume element $d\tau$ at position r is proportional to $|\Psi(r)|^2 d\tau$.

This probability interpretation of the wave function is due to Max Born. From this it follows that $|\Psi(r)|^2$ is a probability density. The wave function is a probability amplitude. From this interpretation it also follows that the wave functions have to be square integrable

$$\int |\Psi|^2 d\tau < \infty \ ,$$

Because the probability to find a particle at any place in space has to be finite.

Postulate 5
The time evolution of a wave function $\Psi(r_1, r_2, \ldots, t)$ is governed by the Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H} \Psi \ .$$

This is a partial differential equation. $\hat{H}$ is the Hamilton operator of the system. For example the Schrödinger equation for a particle in one dimension is

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V(x) \Psi \ .$$

8

1.2.2 Separation of the Schrödinger equation

Often the Schrödinger equation can be separated into an equation for the positions and an equation for the time evolution. If the Hamilton operator does not depend on time, we get in one dimension

$$\hat{H}\Psi = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + V(x)\Psi = E\Psi = i\hbar\frac{\partial\Psi}{\partial t}$$

For stationary states the energy is a constant and we can solve the ordinary differential equation using

$$\Psi(x,t) = \psi(x)\exp\left(\frac{E}{i\hbar}t\right)$$

Inserting this into the Schrödinger equation we get

$$\hat{H}\psi = E\psi \ ,$$

the time-independent Schrödinger equation which determines the function $\psi(x)$.

1.3 Hydrogen Atom

The hydrogen atom consists of a proton of mass m_p and an electron of mass m_e. Reducing the two particle problem to a single particle equation can be achieved by separation of the center of mass movement:

- 3 center of mass coordinates (translation)

- 3 inner coordinates (relative movement)

This separation leads to the reduced mass for the relative movement:

$$\mu = \frac{m_e m_p}{m_e + m_p} = 0.999456 m_e \approx m_e$$

$$\hat{H} = -\frac{\hbar^2}{2\mu}\nabla^2 + V(r)$$

The relative movement can be cast into spherical coordinates (r, θ, ϕ). The Coulomb potential for this special case is a central field problem

$$V(r) = -\frac{Ze^2}{r} \ .$$

The nuclear charge is $Z = 1$ for the hydrogen atom.

1.3.1 Angular momentum operators

1.3.2 Definitions

Classical angular momentum:

$$
\begin{aligned}
\vec{l} &= \vec{r} \times \vec{p} = (l_x, l_y, l_z) \\
l_x &= y p_z - z p_y \\
l_y &= z p_x - x p_z \\
l_z &= x p_y - y p_x
\end{aligned}
$$

Operators in quantum mechanics:

$$
\hat{l}_x = \frac{\hbar}{i} \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right)
$$

$$
\hat{l}_y = \frac{\hbar}{i} \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right)
$$

$$
\hat{l}_z = \frac{\hbar}{i} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right)
$$

Total angular momentum :
classical:

$$
| \vec{l} |^2 = \vec{l} \cdot \vec{l} = l_x^2 + l_y^2 + l_z^2
$$

in quantum mechanics:

$$
\hat{l}^2 = \hat{l}_x^2 + \hat{l}_y^2 + \hat{l}_z^2 = \hat{l}_x \hat{l}_x + \hat{l}_y \hat{l}_y + \hat{l}_z \hat{l}_z
$$

Commutators

For two components:

$$
\begin{aligned}
\hat{l}_x \hat{l}_y f &= -\hbar^2 \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) f \\
&= -\hbar^2 \left(yz \frac{\partial^2 f}{\partial z \partial x} + y \frac{\partial f}{\partial x} - yx \frac{\partial^2 f}{\partial z^2} - z^2 \frac{\partial^2 f}{\partial y \partial x} + zx \frac{\partial^2 f}{\partial y \partial z} \right) \\
\hat{l}_y \hat{l}_x f &= -\hbar^2 \left(zy \frac{\partial^2 f}{\partial x \partial z} + x \frac{\partial f}{\partial y} - xy \frac{\partial^2 f}{\partial z^2} - z^2 \frac{\partial^2 f}{\partial x \partial y} + xz \frac{\partial^2 f}{\partial z \partial y} \right)
\end{aligned}
$$

Differences

$$
\begin{aligned}
(\hat{l}_x \hat{l}_y - \hat{l}_y \hat{l}_x) f &= -\hbar^2 \left(y \frac{\partial f}{\partial x} - x \frac{\partial f}{\partial y} \right) \\
&= i\hbar \frac{\hbar}{i} \left(x \frac{\partial f}{\partial y} - y \frac{\partial f}{\partial x} \right) \\
&= i\hbar \hat{l}_z f
\end{aligned}
$$

10

Commutator-equations:

$$[\hat{l}_z, \hat{l}_x] = i\hbar \hat{l}_y$$

For a single component and l^2:

$$
\begin{aligned}
[\hat{l}_x, \hat{l}^2] &= \hat{l}_x \hat{l}^2 - \hat{l}^2 \hat{l}_x \\
&= \hat{l}_x(\hat{l}_x\hat{l}_x + \hat{l}_y\hat{l}_y + \hat{l}_z\hat{l}_z) - (\hat{l}_x\hat{l}_x + \hat{l}_y\hat{l}_y + \hat{l}_z\hat{l}_z)\hat{l}_x \\
&= (\hat{l}_x\hat{l}_y)\hat{l}_y + (\hat{l}_x\hat{l}_z)\hat{l}_z - \hat{l}_y(\hat{l}_y\hat{l}_x) - \hat{l}_z(\hat{l}_z\hat{l}_x) \\
&= (i\hbar\hat{l}_z + \hat{l}_y\hat{l}_x)\hat{l}_y + (-i\hbar\hat{l}_y + \hat{l}_z\hat{l}_x)\hat{l}_z - \hat{l}_y(-i\hbar\hat{l}_z + \hat{l}_x\hat{l}_y) - \hat{l}_z(i\hbar\hat{l}_y + \hat{l}_x\hat{l}_z) \\
&= 0
\end{aligned}
$$

$$[\hat{l}_x, \hat{l}^2] = [\hat{l}_y, \hat{l}^2] = [\hat{l}_z, \hat{l}^2] = 0$$

From the commutators we learn that it is possible to measure the absolute value of the angular momentum and the component along one special direction at the same time. However, it is not possible to measure all three components at the same time.

Central field problems: The potential $V(r)$ is spherical symmetric and only depends on the distance from the origin. In this case we have

$$[\hat{l}_x, \hat{T}] = [\hat{l}_y, \hat{T}] = [\hat{l}_z, \hat{T}] = 0$$
$$[\hat{l}_x, \hat{V}(r)] = [\hat{l}_y, \hat{V}(r)] = [\hat{l}_z, \hat{V}(r)] = 0$$
$$[\hat{l}_x, \hat{H}] = [\hat{l}_y, \hat{H}] = [\hat{l}_z, \hat{H}] = 0$$
$$[\hat{l}^2, \hat{H}] = 0$$

For central field problems $\hat{H}, \hat{l}^2$ and $\hat{l}_z$ have a common set of eigenfunctions. Every stationary state is characterized by unique expectation values for $\hat{H}, \hat{l}^2$ and $\hat{l}_z$.

Angular momentum and Hamilton operators in spherical coordinates.

Central field problems are best treated in spherical coordinates (R, θ, ϕ). The operators in these coordinates are:

$$(x, y, z) \longrightarrow (R, \theta, \phi) \ .$$

We have to consider the rules of calculus, especially the chain rule.
Results (without deduction):

$$\hat{l}_z = \frac{\hbar}{i}\frac{\partial}{\partial\phi}$$

$$\hat{l}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]$$

$$\hat{H} = -\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right] + \frac{\hat{l}^2}{2mr^2} + V(r)$$

Therefore

$$\left.\begin{array}{c}\hat{l}_z \\ \hat{l}^2 \\ \hat{H}\end{array}\right\} \text{ is a function of } \left\{\begin{array}{c}\phi \\ \theta, \phi \\ r, \theta, \phi\end{array}\right.$$

Common set of eigenfunctions $\Psi(r, \theta, \phi)$.

$$\left.\begin{array}{c}\phi \\ \theta \\ r\end{array}\right\} \text{ independent part from eigenvalue problem for } \left\{\begin{array}{c}\hat{l}_z \\ \hat{l}^2 \\ \hat{H}\end{array}\right.$$

$$\Psi(r, \theta, \phi) = R(r)\,\Theta(\theta)\,\underbrace{\underbrace{\underbrace{\Phi(\phi)}_{\hat{l}_z}}_{\hat{l}^2}}_{\hat{H}}$$

Eigenvalue problem for the operator $\hat{l}_z$

$$\hat{l}_z\Phi(\phi) = \frac{\hbar}{i}\frac{\partial}{\partial\phi}\Phi(\phi) = l_z\Phi(\phi)$$

Eigenfunctions for $\Phi(\phi)$ can be found easily:

$$\Phi(\phi) = Ce^{im\phi} \qquad\qquad \text{constants } C, m$$

$$\hat{l}_z\Phi(\phi) = \frac{\hbar}{i}imCe^{im\phi} = m\hbar\Phi(\phi) \qquad\qquad \text{eigenvalue } m\hbar$$

Initial conditions:

$$\Phi(\phi) = \Phi(\phi + 2\pi) \qquad\qquad \text{periodicity}$$

$$Ce^{im\phi} = Ce^{im(\phi+2\pi)} = Ce^{im\phi}e^{im2\pi}$$

$$e^{im2\pi} = 1 \qquad\qquad m \text{ integer}$$

Normalization:

$$\int_0^{2\pi} \Phi^* \Phi d\phi = C^2 \int_0^{2\pi} e^{-im\phi} e^{im\phi} d\phi = C^2 \int_0^{2\pi} d\phi = 2\pi C^2 = 1$$
$$C = (2\pi)^{-1/2}$$
$$\Phi(\phi) = (2\pi)^{-1/2} e^{im\phi}$$

The z-component of the angular momentum can take the values $m\hbar$ (m integer).

Eigenvalue problem for the operator $\hat{l}^2$

Eigenfunctions $Y(\theta, \phi)$ and eigenvalues $\lambda \hbar^2$

$$\hat{l}^2 Y(\theta, \phi) = -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \phi^2} \right] Y(\theta, \phi)$$
$$= \lambda \hbar^2 Y(\theta, \phi)$$
$$Y(\theta, \phi) = \Theta(\theta) \Phi(\phi)$$

In the product function the eigenfunctions $\Phi(\phi)$ of $\hat{l}_z$ are known. We get:

$$\frac{\partial^2}{\partial \phi^2} \Phi(\phi) = -m^2 \Phi(\phi)$$

$$-\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \phi^2} \right] \Theta(\theta)\Phi(\phi) = \lambda \hbar^2 \Theta(\theta)\Phi(\phi)$$

The operator is no longer dependent on ϕ and therefore we can divide by $\Phi(\phi)$. We get a differential equation for Θ. This equation can be solved using similar methods as used for the harmonic oscillator.

- Substitution $x = \cos\theta$

- Power series ansatz

- Recursion formula

- Investigation of convergence, i.e. are there divergencies

- Divergencies can only be avoided at special values of λ where we get a finite power series.

We get $\lambda = l(l+1)$ with integer $l \geq 0$ and $l \geq |m|$. The square of the absolute value of the angular momentum can only take values $l(l+1)\hbar^2$ with ($l = 0, 1, 2, 3, \ldots$ and $l \geq |m|$).

Angular momentum eigenfunctions $Y_l^m(\theta, \phi)$

Spherical harmonics $Y_l^m(\theta, \phi)$ in product form:

$$Y_l^m(\theta, \phi) = \Theta(\theta)\Phi(\phi)$$

The θ dependent part of $\Theta(\theta)$ is proportional to the Legendre polynomials P_l for $m = 0$ and the associated Legendre functions $P_l^{|m|}$ for $m \neq 0$.
We get:

$$x = \cos\theta$$

$$P_l(x) = \frac{1}{2^l l!}\frac{d^l}{dx^l}(x^2 - 1)^l \qquad\qquad (l = 0, 1, \ldots)$$

$$P_l^{|m|}(x) = (1 - x^2)^{|m|/2}\frac{d^{|m|}}{dx^{|m|}}P_l(x) \qquad (|m| = 0, 1, 2, \ldots, l)$$

Eigenvalue equation:

$$\hat{l}^2 Y_l^m(\theta, \phi) = l(l+1)\hbar^2 Y_l^m(\theta, \phi) \qquad\qquad (l = 0, 1, \ldots)$$

$$\hat{l}_z Y_l^m(\theta, \phi) = m\hbar Y_l^m(\theta, \phi) \qquad\qquad (-l \leq m \leq l)$$

Spherical harmonics

$$Y_l^m(\theta, \phi) = \left[\frac{2l+1}{4\pi}\frac{(l - |m|)!}{(l + |m|)!}\right]^{1/2} P_l^{|m|}(\cos\theta)e^{im\phi}$$

Some special spherical harmonics:

$$Y_0^0 = \left(\frac{1}{4\pi}\right)^{1/2} \qquad\qquad\qquad s$$

$$Y_1^0 = \left(\frac{3}{4\pi}\right)^{1/2}\cos\theta \qquad\qquad\qquad p$$

$$Y_1^{\pm 1} = \left(\frac{3}{8\pi}\right)^{1/2}\sin\theta e^{\pm i\phi} \qquad\qquad\qquad p$$

$$Y_2^0 = \left(\frac{5}{16\pi}\right)^{1/2}(3\cos^2\theta - 1) \qquad\qquad d$$

$$Y_2^{\pm 1} = \left(\frac{15}{8\pi}\right)^{1/2}\sin\theta\cos\theta e^{\pm i\phi} \qquad\qquad d$$

$$Y_2^{\pm 2} = \left(\frac{15}{32\pi}\right)^{1/2}\sin^2\theta e^{\pm 2i\phi} \qquad\qquad d$$

14

Linear combinations of complex spherical harmonics Y_l^m give the real spherical harmonics $S_l^{(m)}$.

$$S_l^{(m)}(\theta, \phi) = \begin{cases} \frac{1}{\sqrt{2}}\left[Y_l^{-m}(\theta, \phi) + Y_l^m(\theta, \phi)\right] & m > 0 \\ \frac{1}{i\sqrt{2}}\left[Y_l^{-m}(\theta, \phi) - Y_l^m(\theta, \phi)\right] & m < 0 \end{cases}$$

The pair Y_l^m, Y_l^{-m} differs only in the factor $e^{im\phi}, e^{-im\phi}$. Using the Euler formulas:

$$\cos m\phi = \frac{1}{\sqrt{2}}\left[e^{-im\phi} + e^{im\phi}\right] \qquad m > 0$$

$$\sin m\phi = \frac{1}{i\sqrt{2}}\left[e^{-im\phi} - e^{im\phi}\right] \qquad m < 0$$

we get using the definition of Y_l^m the real functions

$$S_l^{(m)}(\theta, \phi) = N_{lm}P_l^{|m|}(\cos\theta)\begin{cases} \cos m\phi & m > 0 \\ \sin m\phi & m < 0 \end{cases}$$

General ansatz:

$$\hat{H}(r, \theta, \phi)\Psi(r, \theta, \phi) = E\Psi(r, \theta, \phi)$$
$$\Psi(r, \theta, \phi) = R(r)Y_l^m(\theta, \phi)$$

The angular part of the eigenfunctions is given by the spherical harmonics $Y_l^m(\theta, \phi)$. With this ansatz and using the eigenvalue equation for $\hat{l}^2$ we get the radial Schrödinger equation for the hydrogen atom.

$$-\frac{\hbar^2}{2\mu}\left[\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right]R(r) + \frac{l(l+1)\hbar^2}{2\mu r^2}R(r) - \frac{Ze^2}{r}R(r) = ER(r)$$

Solution of the radial Schrödinger equation

- Using a new radial function $u(r) = rR(r)$ results in the new equation

$$-\frac{\hbar^2}{2\mu}\frac{d^2u(r)}{dr^2} + \left[\frac{l(l+1)\hbar^2}{2\mu r^2} - \frac{Ze^2}{r}\right]u(r) = Eu(r)$$

- Investigation of the asymptotic behavior for $r \to \infty$ and $r \to 0$ leads to the following form of the solution

$$u(r) = (\kappa r)^{l+1}e^{-\kappa r}w(\kappa r)$$

$$\kappa = \sqrt{\frac{-2\mu E}{\hbar^2}}$$

κ is real for $E < 0$.

- Inserting this definition into the radial Schrödinger equation leads to the Laguerre differential equation for $w(\kappa r)$.

- Solutions of this differential equation can be found with a power series (recursion formula).

- The power series is not diverging only for special values of E: Quantization of energy.
$$\frac{Ze^2\kappa}{2|E|} = k + l + 1 = n$$

k, l, n integer; $k \geq 0, l \geq 0, n > l$

- Solution for E
$$E_n = -\frac{Z^2\mu e^4}{2\hbar^2 n^2}$$

$n \geq 1$ integer, $n > l$

Energies

$$E_n = -\frac{Z^2 e^2}{2a}\frac{1}{n^2} \qquad\qquad (n = 1, 2, 3, \ldots)$$
$$a = \frac{\hbar^2}{\mu e^2} \qquad\qquad \text{Bohr Radius} = 0.529 \text{ Å}$$

There is an infinite number of bound states with $E_n < 0$.
$E = 0$ corresponds to proton and electron infinitely separated.
$E > 0$ non-bound states with additional kinetic energy.
Ground state of hydrogen atom ($Z = 1, n = 1$):

$$E_1 = -\frac{e^2}{2a} \approx -13.6 \text{ eV}$$

Excited states of hydrogen atom ($Z = 1, E_n \propto -\frac{1}{n^2}$):

$$E_2 \approx -3.4\text{eV}, \qquad E_3 \approx -1.5\text{eV}, \qquad E_4 \approx -0.85\text{eV}$$

Each state Ψ_{nlm} is characterized by three quantum numbers, however, the energy depends only on n.
Energies are n^2 times degenerate.

n	l	m	Notation	States	Degeneracy
1	0	0	1s	1	1
2	0	0	2s	1	
	1	-1,0,1	2p	3	4
3	0	0	3s	1	
	1	-1,0,1	3p	3	
	2	-2,-1,0,1,2	3d	5	9

1.3.3   Radial function of hydrogen atom

$$R_{nl}(r) = - \left[\frac{(n-l-1)!}{2n[(n+l)!]^3} \left(\frac{2Z}{na} \right)^3 \right]^{1/2} \left(\frac{2Zr}{na} \right)^l e^{-\frac{Zr}{na}} L_{n+l}^{2l+1} \left(\frac{2Zr}{na} \right)$$

Laguerre polynomials

$$L_p(x) = e^x \frac{d^p}{dx^p} \left(x^p e^{-x} \right)$$

Generalized Laguerre polynomials

$$L_p^q(x) = \frac{d^q}{dx^q} L_p(x)$$

Explicite formulas

$$\rho = \frac{Zr}{na}$$

$$R_{10} = 2 \left(\frac{Z}{a} \right)^{3/2} e^{-\rho} \qquad\qquad 1s$$

$$R_{20} = 2^{-1/2} \left(\frac{Z}{a} \right)^{3/2} (1 - \rho) e^{-\rho} \qquad\qquad 2s$$

$$R_{21} = 6^{-1/2} \left(\frac{Z}{a} \right)^{3/2} \rho e^{-\rho} \qquad\qquad 2p$$

$$R_{30} = \frac{2}{9\sqrt{3}} \left(\frac{Z}{a} \right)^{3/2} (3 - 6\rho + 2\rho^2) e^{-\rho} \qquad\qquad 3s$$

$$R_{31} = \frac{4}{9\sqrt{6}} \left(\frac{Z}{a} \right)^{3/2} (2\rho - \rho^2) e^{-\rho} \qquad\qquad 3p$$

$$R_{32} = \frac{4}{9\sqrt{30}} \left(\frac{Z}{a} \right)^{3/2} \rho^2 e^{-\rho} \qquad\qquad 3d$$

Electron density distribution:

Probability to find an electron within the volume $d\tau$:

$$\Psi_{nlm}^* \Psi_{nlm} d\tau = \Psi_{nlm}^*(r,\theta,\phi) \Psi_{nlm}(r,\theta,\phi) r^2 \sin\theta dr d\theta d\phi$$

Probability to find an electron within r and $r+dr$, independent of the angles (θ,ϕ):

$$P_{nl}(r)dr = R_{nl}^* R_{nl} r^2 dr \int\int Y_l^{m*}(\theta,\phi) Y_l^m(\theta,\phi) \sin\theta d\theta d\phi$$

The integral is one (Normalization). For real R_{nl} one gets for the radial probability density

$$P_{nl}(r) = R_{nl}^2 r^2$$

Probability to find an electron within the space angle $d\Omega$ (within θ and $\theta+d\theta$, ϕ and $\phi+d\phi$), independent of the distance (r):

$$W_{nl}(\theta,\phi)d\Omega = Y_l^{m*}(\theta,\phi) Y_l^m(\theta,\phi) d\Omega \int R_{nl}(r) r^2 dr$$

The integral is one (Normalization). With the well known formulas for the spherical harmonics Y_l^m we get:

$$W_{nl}(\theta,\phi) = N_{lm}^2 [P_l^{|m|}(\cos\theta)]^2 = W_{lm}(\theta)$$

The angular distribution of the probability density is for the eigenfunctions of the hydrogen atom independent of ϕ and therefore rotational symmetric w.r.t. the z axis.

18
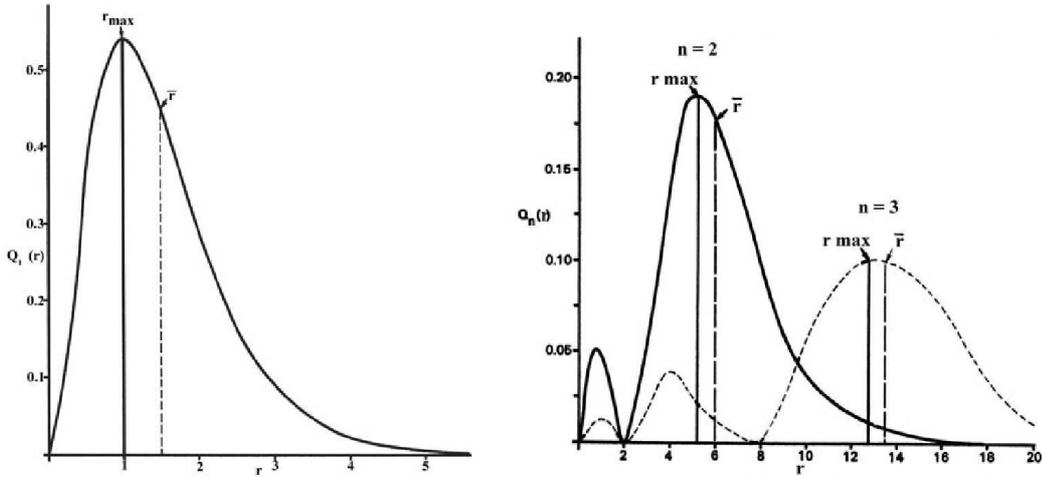
	Unit	values in a.u.	values in SI-units
Charge	elemental charge	$e = 1$	e= $1.602189 \cdot 10^{-19}$ C
Mass	electron mass	$m_e = 1$	m$_e$ =$9.109534 \cdot 10^{-31}$ kg
Angular momentum	Planck constant $\hbar$	$\hbar = 1$	$\hbar = 1.054589 \cdot 10^{-34}$ Js
Length	Bohr radius	$a_0 = \frac{\hbar^2}{m_e e^2} = 1$	$a_0 = 0.529177 \cdot 10^{-10}$ m
Energy	$2|E_1(H)|$ with $\mu = m_e$	Hartree= $\frac{m_e e^4}{\hbar^2} = 1$	Hartree $= 4.359814 \cdot 10^{-18}$ J

Lecture 2

Basic Mathematical Review

2.1 Linear algebra

2.1.1 Vector spaces

A n-dimensional vector can be represented by its components w.r.t. a set of orthogonal and normalised basis vectors $\{\vec{e}_i\}$

$$\vec{a} = \vec{e}_1 a_1 + \vec{e}_2 a_2 + \cdots + \vec{e}_n a_n = \sum_{i=1}^{n} \vec{e}_i a_i$$

For a given basis a vector is specified by its components. We represent a vector

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

as a column matrix. The scalar product of two vectors $\vec{a}$ and $\vec{b}$ is defined as

$$\vec{a} \cdot \vec{b} = \sum_{i=1}^{n} a_i^* b_i$$

This is due to $\vec{e}_i \cdot \vec{e}_j = \delta_{ij}$ where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

2.1.2   Operators

An operator acts on a vector and transforms it into another vector

$$\mathcal{O}\vec{a} = \vec{b}$$

An operator is linear if

$$\mathcal{O}\left(x\vec{a} + y\vec{b}\right) = x\mathcal{O}\vec{a} + y\mathcal{O}\vec{b}$$

A linear operator is completely determined if its effect on every possible vector is known. Because any vector can be expressed in terms of the basis, it is sufficient to know what $\mathcal{O}$ does to the basis vectors. $\mathcal{O}\vec{e}_i$ is a vector and can be expanded in the basis

$$\mathcal{O}\vec{e}_i = \sum_{j=1}^{n} \vec{e}_j O_{ji} \qquad i = 1, 2, \ldots n$$

The numbers O_{ji} form a $n \times n$ matrix

$$\mathbf{O} = \begin{pmatrix} O_{11} & O_{12} & \cdots & O_{1n} \\ O_{21} & O_{22} & \cdots & O_{2n} \\ \vdots & \vdots & & \vdots \\ O_{n1} & O_{n2} & \cdots & O_{nn} \end{pmatrix}$$

$\mathbf{O}$ is the matrix representation of the operator $\mathcal{O}$ in the basis $\{\vec{e}_i\}$.
If $\mathbf{A}$ and $\mathbf{B}$ are the matrix representations of the operators $\mathcal{A}$ and $\mathcal{B}$, the matrix representation of the operator $\mathcal{C} = \mathcal{AB}$ can be found as follows:

$$\mathcal{C}\vec{e}_j = \sum_{i=1}^{n} \vec{e}_i C_{ij}$$
$$= \mathcal{AB}\vec{e}_j$$
$$= \mathcal{A}\sum_{k=1}^{n} \vec{e}_k B_{kj}$$
$$= \sum_{i,k=1}^{n} \vec{e}_i A_{ik} B_{kj}$$

so that

$$C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$$

which is the definition of matrix multiplication and hence

$$\mathbf{C} = \mathbf{AB}$$

The order in which two operators or two matrices are multiplied is crucial. In general $\mathcal{AB} \neq \mathcal{BA}$ and $\mathbf{AB} \neq \mathbf{BA}$. This means two operators do not necessarily commute. The commutator of two operators is defined by

$$[\mathcal{A}, \mathcal{B}] = \mathcal{AB} - \mathcal{BA}$$

and can easily also be written in the matrix representation

$$[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$$

2.1.3 Matrices

A general rectangular $n \times m$ Matrix $\mathbf{A}$ has n rows and m columns

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{pmatrix}$$

If $n = m$ the matrix is square. Two matrices $\mathbf{A}$ and $\mathbf{B}$ can be multiplied $\mathbf{C} = \mathbf{AB}$ if the number of columns of $\mathbf{A}$ is equal to the number of rows of $\mathbf{B}$.

$$C_{ij} = \sum_{k=1}^{M} A_{ik} B_{kj}$$

Matrix-vector multiplication is a special case of matrix-matrix multiplication. Some important definitions and properties of matrices.

The adjoint of an $N \times M$ matrix $\mathbf{A}$, denoted by $\mathbf{A}^{\dagger}$, is a $M \times N$ matrix with elements

$$(\mathbf{A}^{\dagger})_{ij} = A_{ji}^{\star}$$

i.e. we take the complex conjugate of each of the matrix elements of $\mathbf{A}$ and interchange rows and columns. If the elements of $\mathbf{A}$ are real, then the adjoint of $\mathbf{A}$ is called the transpose of $\mathbf{A}$.

The adjoint of the product of two matrices, is equal to the product of the adjoint matrices in reversed order.

$$(\mathbf{AB})^{\dagger} = \mathbf{B}^{\dagger} \mathbf{A}^{\dagger}$$

Proof

$$(\mathbf{AB})^\dagger_{ij} = (\mathbf{AB})^\star_{ji}$$
$$= \sum_k (A_{jk} B_{ki})^\star = \sum_k A^\star_{jk} B^\star_{ki}$$
$$= \sum_k A^\dagger_{kj} B^\dagger_{ik} = (\mathbf{B}^\dagger \mathbf{A}^\dagger)_{ij}$$

The adjoint of a column matrix (vector) is a row matrix containing the complex conjugates of the elements.

$$\mathbf{a}^\dagger = (a^\star_1, a^\star_2, \cdots, a^\star_m)$$

If $\mathbf{a}$ and $\mathbf{b}$ are two column matrices with m elements, then

$$\mathbf{a}^\dagger \mathbf{b} = (a^\star_1, a^\star_2, \cdots, a^\star_m) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \sum_{i=1}^{m} a^\star_i b_i$$

This is the scalar product of two vectors.
Special properties of square matrices

- A matrix $\mathbf{A}$ is diagonal if all its off-diagonal elements are zero

$$A_{ij} = A_{ii} \delta_{ij} \ .$$

- The trace of $\mathbf{A}$ is the sum of its diagonal elements

$$\mathrm{tr}\mathbf{A} = \sum_i A_{ii} \ .$$

- The unit matrix is defined by $(\mathbf{1})_{ij} = \delta_{ij}$ and

$$\mathbf{1A} = \mathbf{A1} = \mathbf{A}$$

- The inverse of $\mathbf{A}$, denoted $\mathbf{A}^{-1}$ is a matrix such that

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{1}$$

- A unitary matrix $\mathbf{A}$ is one whose inverse is its adjoint, a real unitary matrix is called orthogonal

$$\mathbf{A}^{-1} = \mathbf{A}^\dagger$$

- A Hermitian matrix is self-adjoint, a real Hermitian matrix is called symmetric

$$\mathbf{A}^\dagger = \mathbf{A}$$

Further properties

- $\text{tr}\mathbf{AB} = \text{tr}\mathbf{BA}$

- $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

- If $\mathbf{U}$ is unitary and $\mathbf{B} = \mathbf{U}^\dagger\mathbf{AU}$, then $\mathbf{A} = \mathbf{UBU}^\dagger$

- If $\mathbf{A}, \mathbf{B}, \mathbf{C}$ Hermitian matrices, and $\mathbf{C} = \mathbf{AB}$, then $\mathbf{A}$ and $\mathbf{B}$ commute.

2.1.4 Determinants

The determinant of an $n \times n$ matrix $\mathbf{A}$ is a number obtained from

$$\det(\mathbf{A}) = |\mathbf{A}| = \begin{vmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{vmatrix} = \sum_{i=1}^{n!} (-1)^{p_i} \mathcal{P}_i A_{11} A_{22} \cdots A_{nn}$$

where $\mathcal{P}_i$ is a permutation operator that permutes the column indices $1, 2, 3, \ldots, n$ and the sum runs over all $n!$ permutations of the indices; p_i is the number of transpositions required to restore a given permutation to natural order. Important properties of determinants

- If each element in a row or in a column is zero the value of the determinant is zero.

- If $(\mathbf{A})_{ij} = A_{ii}\delta_{ij}$, then $|\mathbf{A}| = \prod_i A_{ii} = A_{11} A_{22} \ldots A_{nn}$.

- A single interchange of any two rows (or columns) of a determinant changes its sign.

- $|\mathbf{A}| = (|\mathbf{A}^\dagger|)^\star$

- $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$

- If any two rows (or columns) of a determinant are equal, the value of the determinant is zero.

- $|\mathbf{A}^{-1}| = (|\mathbf{A}|)^{-1}$

- If $\mathbf{AA}^\dagger = \mathbf{1}$, then $|\mathbf{A}|(|\mathbf{A}|)^\star = 1$.

- If $\mathbf{U}^\dagger\mathbf{OU} = \Omega$ and $\mathbf{U}^\dagger\mathbf{U} = \mathbf{UU}^\dagger = \mathbf{1}$, then $|\mathbf{O}| = |\mathbf{\Omega}|$.

2.1.5 Dirac notation

Dirac introduced a notation for vector spaces that allows to express results in a concise and simple manner. We consider n basis vectors denoted by the symbol $|i\rangle, i = 1, 2, \ldots, n$, which are called ket vectors or simply kets. We assume this basis is complete so that any ket vector $|a\rangle$ can be written as

$$|a\rangle = \sum_{i=1}^{n} |i\rangle \, a_i \ .$$

The column matrix $\mathbf{a}$ is the matrix representation of the abstract vector $|a\rangle$ in the basis $\{|i\rangle\}$. We introduce an abstract bra vector $\langle a|$ whose matrix representation is $\mathbf{a}^\dagger$. The scaler product between a bra $\langle a|$ and a ket $|b\rangle$ is defined as

$$\langle a||b\rangle \equiv \langle a \mid b \rangle = \mathbf{a}^\dagger \mathbf{b} = (a_1^\star, a_2^\star, \ldots, a_n^\star) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \sum_{i=1}^{n} a_i^\star b_i$$

The square of the length of a vector

$$\langle a \mid a \rangle = \sum_{i=1}^{n} a_i^\star a_i = \sum_{i=1}^{n} |a_i|^2$$

is always real and positive. In analogy to the ket basis we can introduce the bra basis $\{\langle i|\}$ that is complete in the sense that any bra vector can be written as a linear combination of the bra basis vectors.

$$\langle a| = \sum_i a_i^\star \langle i| \ .$$

The scalar product now becomes

$$\langle a \mid b \rangle = \sum_{ij} a_i^\star \langle i \mid j \rangle b_j \ .$$

For this to be identical with the previous definition we must have that

$$\langle i \mid j \rangle = \delta_{ij}$$

which is a statement of the orthonormality of the basis.

For a given ket $|a\rangle$ or bra $\langle a|$ the components of its matrix representation with respect to a basis $\{|i\rangle$ or $\{\langle i|\}$ can be determined by multiplication from the left or right with a general basis vector $|j\rangle$ or $\langle j|$

$$\langle j \mid a \rangle = \sum_i \langle j \mid i \rangle \, a_i = \sum_i \delta_{ji} \, a_i = a_j$$

$$\langle a \mid j \rangle = \sum_i a_i^\star \, \langle i \mid j \rangle = \sum_i a_i^\star \, \delta_{ij} = a_j^\star$$

Using these results we can write

$$|a\rangle = \sum_i |i\rangle \, a_i = \sum_i |i\rangle \langle i \mid a \rangle$$

and

$$\langle a| = \sum_i a_i^\star \, \langle i| = \sum_i \langle a \mid i \rangle \langle i|$$

which suggests we write

$$1 = \sum_i |i\rangle \langle i|$$

which is a statement of the completeness of the basis.

An operator $\mathcal{O}$ is defined as an entity which when acting on a ket $|a\rangle$ converts it into a ket $|b\rangle$

$$\mathcal{O} \mid a \rangle = |b\rangle \ .$$

The operator is completely determined by its action on all basis vectors

$$\mathcal{O} \mid i \rangle = \sum_j O_{ji} \mid j \rangle = \sum_j (\mathbf{O})_{ji} \mid j \rangle$$

so that $\mathbf{O}$ is the matrix representation of the operator $\mathcal{O}$ in the basis $\{|i\rangle\}$. The matrix representation is found by multiplying on the left by $\langle k|$

$$\langle k \mid \mathcal{O} \mid i \rangle = \sum_j \langle k \mid j \rangle (\mathbf{O})_{ji} = \sum_j \delta_{kj} (\mathbf{O})_{ji} = (\mathbf{O})_{ki}$$

which provides a useful expression for the matrix elements of $\mathbf{O}$.

2.1.6   Change of basis

We specify two different basis sets, one, $\{|i\rangle\}$ where we use latin letters $i, j, k, \ldots$ for the vectors and another $\{|\alpha\rangle\}$ where we use Greek letters $\alpha, \beta, \gamma, \ldots$. Thus we have

$$\langle i \mid j \rangle = \delta_{ij} \ , \qquad \sum_i |i\rangle \langle i| = 1$$

and

$$\langle \alpha \mid \beta \rangle = \delta_{\alpha\beta} \ , \qquad \sum_\alpha |\alpha\rangle\langle\alpha| = 1 \ .$$

Since the basis $\{|i\rangle\}$ is complete we can express any ket in the basis $\{|\alpha\rangle\}$ as a linear combination of kets in the basis $\{|i\rangle\}$ and vice versa. That is,

$$|\alpha\rangle = 1 \mid \alpha\rangle = \sum_i |i\rangle\langle i \mid \alpha\rangle = \sum_i |i\rangle U_{i\alpha} = \sum_i |i\rangle (\mathbf{U})_{i\alpha}$$

where we have defined the elements of a transformation matrix $\mathbf{U}$ as

$$\langle i \mid \alpha \rangle = U_{i\alpha} = (\mathbf{U})_{i\alpha} \ .$$

Transforming in the opposite direction, we have

$$|i\rangle = 1 \mid i \rangle = \sum_\alpha |\alpha\rangle\langle\alpha \mid i\rangle = \sum_\alpha |\alpha\rangle U_{i\alpha}^\star = \sum_\alpha |\alpha\rangle (\mathbf{U}^\dagger)_{\alpha i}$$

where we have used

$$\langle \alpha \mid i \rangle = \langle i \mid \alpha \rangle^\star = U_{i\alpha}^\star = (\mathbf{U}^\dagger)_{\alpha i} \ .$$

As a consequence of the orthonormality of the bases the transformation matrix $\mathbf{U}$ is unitary

$$\delta_{ij} = \langle i \mid j \rangle = \sum_\alpha \langle i \mid \alpha \rangle \langle \alpha \mid j \rangle$$
$$= \sum_\alpha U_{i\alpha}(\mathbf{U}^\dagger)_{\alpha j} = (\mathbf{U}\mathbf{U}^\dagger)_{ij}$$

which is in matrix notation

$$1 = \mathbf{U}\mathbf{U}^\dagger \ .$$

Starting from $\delta_{\alpha\beta} = \langle \alpha \mid \beta \rangle$ one can show that

$$1 = \mathbf{U}^\dagger\mathbf{U}$$

and hence $\mathbf{U}$ is unitary. Thus we arrive at the important result that two orthonormal bases are related by a unitary matrix. The elements of the transformation matrix $\mathbf{U}$ are scalar products between the two bases.

We now investigate the matrix representations of an operator for the two different bases. Suppose $\mathbf{O}$ is the matrix representation of $\mathcal{O}$ in the basis $\{|i\rangle\}$, while $\mathbf{\Omega}$ is its matrix representation in the basis $\{|\alpha\rangle\}$

$$O_{ij} = \langle i \mid \mathcal{O} \mid j \rangle \ ; \qquad \Omega_{\alpha\beta} = \langle \alpha \mid \mathcal{O} \mid \beta \rangle \ .$$

We find the relation between $\mathbf{O}$ and $\boldsymbol{\Omega}$ from

$$
\begin{aligned}
\Omega_{\alpha\beta} = \langle \alpha \mid \mathcal{O} \mid \beta \rangle &= \langle \alpha \mid 1 \, \mathcal{O} \, 1 \mid \beta \rangle \\
&= \sum_{ij} \langle \alpha \mid i \rangle \langle i \mid \mathcal{O} \mid j \rangle \langle j \mid \beta \rangle \\
&= \sum_{ij} (\mathbf{U}^\dagger)_{\alpha i} (\mathbf{O})_{ij} (\mathbf{U})_{j\beta} \ .
\end{aligned}
$$

Thus

$$
\boldsymbol{\Omega} = \mathbf{U}^\dagger \mathbf{O} \mathbf{U}
$$

or, multiplying on the left by $\mathbf{U}$ and on the right by $\mathbf{U}^\dagger$

$$
\mathbf{O} = \mathbf{U} \boldsymbol{\Omega} \mathbf{U}^\dagger \ .
$$

These equations show that the matrices $\mathbf{O}$ and $\boldsymbol{\Omega}$ are related by a unitary transformation. The importance lies in the fact that for any operator with a Hermitian matrix representation which is not diagonal in a given basis $\{|i\rangle\}$, it is always possible to find a basis $\{|\alpha\rangle\}$ in which the matrix representation of the operator is diagonal, i.e.

$$
\Omega_{\alpha\beta} = \omega_\alpha \delta_{\alpha\beta} \ .
$$

2.1.7 Eigenvalue problem

If the result of the operation $\mathcal{O} \mid \alpha \rangle$ is simply a constant times $|\alpha\rangle$, i.e.

$$
\mathcal{O} \mid \alpha \rangle = \omega_\alpha |\alpha\rangle
$$

then we say that $|\alpha\rangle$ is an eigenvector of the operator $\mathcal{O}$ with an eigenvalue ω_α. Without loss of generality we can choose the eigenvectors to be normalised

$$
\langle \alpha \mid \alpha \rangle = 1
$$

- The eigenvalues of a Hermitian operator are real.
  This follows from

  $$
  \langle \alpha \mid \mathcal{O} \mid \alpha \rangle = \langle \alpha \mid \mathcal{O}^\dagger \mid \alpha \rangle = \langle \alpha \mid \mathcal{O} \mid \alpha \rangle^\star
  $$

  Multiplying the eigenvalue relation by $\langle \alpha |$ and using the above relations we have

  $$
  \omega_\alpha = \omega_\alpha^\star
  $$

  which is the required result.

- The eigenvectors of a Hermitian operator are orthogonal.
  We consider

$$\mathcal{O}\mid\beta\rangle = \omega_\beta|\beta\rangle$$

and its adjoint

$$\langle\beta\mid\mathcal{O}^\dagger = \langle\beta|\omega_\beta^\star$$

Since $\mathcal{O}$ is Hermitian and ω_β is real, we have

$$\langle\beta\mid\mathcal{O} = \langle\beta|\omega_\beta$$

Multiplying $\mathcal{O}\mid\alpha\rangle = \omega_\alpha|\alpha\rangle$ by $\langle\beta|$ and $\langle\beta\mid\mathcal{O} = \langle\beta|\omega_\beta$ by $|\alpha\rangle$ and subtracting the results we obtain

$$(\omega_\beta - \omega_\alpha)\langle\beta\mid\alpha\rangle = 0$$

so that $\langle\beta\mid\alpha\rangle = 0$ if $\omega_\alpha \neq \omega_\beta$. Thus orthogonality follows immediately if the two eigenvalues are not the same (nondegenerate). It can be shown, that degenerate eigenvectors can always be chosen to be orthogonal (we omit this here).

The eigenvalue problem can be posed as follows. Given the matrix representation, $\mathbf{O}$, of a Hermitian operator $\mathcal{O}$ in the orthonormal basis $\{|i\rangle\}$, we wish to find the orthonormal basis $\{|\alpha\rangle\}$ in which the matrix representation, $\mathbf{\Omega}$ of $\mathcal{O}$ is diagonal. In other words we wish to diagonalise the matrix $\mathbf{O}$. The two representations of the operator $\mathcal{O}$ are related by an unitary transformation

$$\mathbf{\Omega} = \mathbf{U}^\dagger \mathbf{O} \mathbf{U}$$

Thus the problem of diagonalising the Hermitian matrix $\mathbf{O}$ is equivalent to the problem of finding the unitary matrix $\mathbf{u}$ that converts $\mathbf{O}$ into a diagonal matrix

$$\mathbf{U}^\dagger \mathbf{O} \mathbf{U} = \omega = \begin{pmatrix} \omega_1 & & & \\ & \omega_2 & \mathbf{0} & \\ & \mathbf{0} & \ddots & \\ & & & \omega_n \end{pmatrix} \ .$$

The problem can be reformulated as follows. Given the $n \times n$ Hermitian matrix $\mathbf{O}$, find all n distinct column vectors $\mathbf{c}$ (the eigenvectors) and corresponding n numbers ω (the eigenvalues) such that

$$\mathbf{O}\mathbf{c} = \omega\mathbf{c} \ .$$

This equation can be rewritten as

$$(\mathbf{O} - \omega\mathbf{1})\,\mathbf{c} = \mathbf{0} \ .$$

Nontrivial solutions ($\mathbf{c} \neq \mathbf{0}$) only exist when

$$|\mathbf{O} - \omega\mathbf{1}| = 0 \ .$$

This determinant is a polynomial of degree n in the unknown ω. It has n roots ω_α which in this case will be the eigenvalues of the matrix $\mathbf{O}$. Once we have found the eigenvalues, we can find the corresponding eigenvectors by substituting each ω_α into the above equation and solving the resulting equation for $\mathbf{c}^\alpha$. In this way $\mathbf{c}^\alpha$ can be found within a multiplicative constant, which is finally determined by requiring the eigenvectors to be normalised

$$\sum_i (c_i^\alpha)^\star c_i^\alpha = 1 \ .$$

In this way we find n solutions to

$$\mathbf{O}\mathbf{c}^\alpha = \omega_\alpha \mathbf{c}^\alpha \qquad \alpha = 1, 2, \ldots, n$$

Since $\mathbf{O}$ is Hermitian, the eigenvalues are real and the eigenvectors are orthogonal. We construct a matrix $\mathbf{U}$ defined as $U_{i\alpha} = c_i^\alpha$, i.e.

$$\mathbf{U} = \begin{pmatrix} c_1^1 & c_1^2 & \cdots & c_1^n \\ c_2^1 & c_2^2 & \cdots & c_2^n \\ \vdots & \vdots & & \vdots \\ c_n^1 & c_n^2 & \cdots & c_n^n \end{pmatrix} = (\mathbf{c}^1, \mathbf{c}^2, \cdots, \mathbf{c}^n)$$

Thus the αth colum of $\mathbf{U}$ is just the column matrix $\mathbf{c}^\alpha$. Then from the orthogonality relation we find

$$\sum_i U_{i\alpha}^\star U_{i\beta} = \sum_i (\mathbf{U}^\dagger)_{\alpha i}(\mathbf{U})_{i\beta} = \delta_{\alpha\beta}$$

which in matrix notation is
$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{1}$$

Multiplying $\mathbf{O}\mathbf{U} = \mathbf{U}\omega$ by $\mathbf{U}^\dagger$ we have

$$\mathbf{U}^\dagger \mathbf{O} \mathbf{U} = \omega$$

which shows that $\mathbf{U}$ is the unitary transformation between the original and new basis for the diagonal representation of $\mathcal{O}$.

2.1.8 Functions of matrices

Given a Hermitian matrix $\mathbf{A}$, we can define a function of $\mathbf{A}$, i.e. $f(\mathbf{A})$, in much the same way we define a function $f(x)$ of a simple variable x. For example, the square root of a matrix $\mathbf{A}$, which we denote by $\mathbf{A}^{1/2}$, is simply that matrix which when multiplied by itself gives $\mathbf{A}$

$$\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$$

The sine or the exponential of a matrix are defined by the Taylor series of the function

$$\exp(\mathbf{A}) = \mathbf{1} + \frac{1}{1}\mathbf{A} + \frac{1}{2}\mathbf{A}^2 + \frac{1}{3}\mathbf{A}^3 + \cdots$$

or in general

$$f(\mathbf{A}) = \sum_{k=0}^{\infty} c_k \mathbf{A}^k$$

We are still faced with the problem of calculating $\mathbf{A}^{1/2}$ or $\exp(\mathbf{A})$. If $\mathbf{A}$ is a diagonal matrix

$$(\mathbf{A})_{ij} = a_i \delta_{ij}$$

everything is simple, since

$$\mathbf{A}^k = \begin{pmatrix} a_1^k & & & \\ & a_1^k & \mathbf{0} & \\ & \mathbf{0} & \ddots & \\ & & & a_1^k \end{pmatrix}$$

so that

$$f(\mathbf{A}) = \sum_{k=0}^{\infty} c_k \mathbf{A}^k = \begin{pmatrix} \sum_k c_k a_1^k & & & \\ & \sum_k c_k a_2^k & \mathbf{0} & \\ & \mathbf{0} & \ddots & \\ & & & \sum_k c_k a_n^k \end{pmatrix}$$

$$= \begin{pmatrix} f(a_1) & & & \\ & f(a_2) & \mathbf{0} & \\ & \mathbf{0} & \ddots & \\ & & & f(a_n) \end{pmatrix}$$

Similarly, the square root of a diagonal matrix is

$$\mathbf{A}^{1/2} = \begin{pmatrix} a_1^{1/2} & & & \\ & a_2^{1/2} & \mathbf{0} & \\ & \mathbf{0} & \ddots & \\ & & & a_n^{1/2} \end{pmatrix}$$

If $\mathbf{A}$ is Hermitian we can always find a unitary transformation that diagonalises it

$$\mathbf{U}^\dagger \mathbf{A} \mathbf{U} = \mathbf{a}$$

The reverse transformation is

$$\mathbf{A} = \mathbf{U} \mathbf{a} \mathbf{U}^\dagger$$

Now notice that

$$\mathbf{A}^2 = \mathbf{U} \mathbf{a} \mathbf{U}^\dagger \mathbf{U} \mathbf{a} \mathbf{U}^\dagger = \mathbf{U} \mathbf{a}^2 \mathbf{U}^\dagger$$

or in general

$$\mathbf{A}^k = \mathbf{U} \mathbf{a}^k \mathbf{U}^\dagger$$

so that

$$f(\mathbf{A}) = \sum_{k=0}^{\infty} c_k \mathbf{A}^k = \mathbf{U} \left(\sum_k c_k \mathbf{a}^k \right) \mathbf{U}^\dagger = \mathbf{U} \, f(\mathbf{a}) \, \mathbf{U}^\dagger$$

$$= \mathbf{U} \begin{pmatrix} f(a_1) & & & \\ & f(a_2) & \mathbf{0} & \\ & \mathbf{0} & \ddots & \\ & & & f(a_n) \end{pmatrix} \mathbf{U}^\dagger$$

Thus to calculate any function of a Hermitian matrix $\mathbf{A}$, we first diagonalise $\mathbf{A}$ to obtain $\mathbf{a}$, the diagonal matrix containing all the eigenvalues of $\mathbf{A}$. We then calculate $f(\mathbf{a})$, which is easy because $\mathbf{a}$ is diagonal. Finally we transform $f(\mathbf{a})$ back using the matrix $\mathbf{U}$ formed by the eigenvectors of $\mathbf{A}$. For example, we can find the square root of a matrix $\mathbf{A}$ as

$$\mathbf{A}^{1/2} = \mathbf{U} \mathbf{a}^{1/2} \mathbf{U}^\dagger$$

since

$$\mathbf{A}^{1/2} \mathbf{A}^{1/2} = \mathbf{U} \mathbf{a}^{1/2} \mathbf{U}^\dagger \mathbf{U} \mathbf{a}^{1/2} \mathbf{U}^\dagger = \mathbf{U} \mathbf{a}^{1/2} \mathbf{a}^{1/2} \mathbf{U}^\dagger = \mathbf{U} \mathbf{a} \mathbf{U}^\dagger = \mathbf{A}$$

2.2 Perturbation theory

For an un-perturbed model system with the Hamiltonian $\hat{H}^{(0)}$ the exact wave functions $\Psi_n^{(0)}$ and exact energies $E_n^{(0)}$ are known.

$$\hat{H}^{(0)} \Psi_n^{(0)} = E_n^{(0)} \Psi_n^{(0)}$$

For the system of interest, a perturbed system with the Hamiltonian $\hat{H}$, the wave functions Ψ_n and energies E_n cannot be calculated exactly.

$$\hat{H}\Psi_n = E_n\Psi_n$$

Perturbation theory describes how the exact solutions of the model system change under the influence of a perturbation $\lambda\hat{H}^{(1)}$ (perturbation parameter λ, perturbation operator $\hat{H}^{(1)}$). In a systematic procedure corrections to the energy and the wave functions in first, second , third, ... order are defined:

$$E_n^{(1)}, E_n^{(2)}, E_n^{(3)}, \ldots \qquad \text{and} \qquad \Psi_n^{(1)}, \Psi_n^{(2)}, \Psi_n^{(3)}, \ldots$$

The perturbation $\lambda\hat{H}^{(1)}$ has to be small and we have

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

$$\left.\begin{array}{c} E_n \longrightarrow E_n^{(0)} \\ \Psi_n \longrightarrow \Psi_n^{(0)} \end{array}\right\} \qquad \text{for } \lambda\hat{H}^{(1)} \longrightarrow 0$$

2.2.1 Derivation

Ansatz:

$$E = E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \cdots$$
$$\Psi = \Psi^{(0)} + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \cdots$$
$$0 = \langle \Psi^{(0)} \mid \Psi^{(p)} \rangle \qquad \text{with } p = 1, 2, \ldots$$

Formal power series in the perturbation parameter λ. Orders of magnitude:

$$|E^{(0)}| \gg |\lambda E^{(1)}| \gg |\lambda^2 E^{(2)}| \gg \ldots$$

Insert in the Schrödinger equation:

$$\left(\hat{H}^{(0)} + \lambda\hat{H}^{(1)}\right)\left(\Psi^{(0)} + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \cdots\right) =$$
$$\left(E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)} + \cdots\right)\left(\Psi^{(0)} + \lambda\Psi^{(1)} + \lambda^2\Psi^{(2)} + \cdots\right)$$

Combine w.r.t. the orders of powers of λ:

$$\left[\hat{H}^{(0)}\Psi^{(0)} - E^{(0)}\Psi^{(0)}\right] +$$
$$\lambda\left[\hat{H}^{(0)}\Psi^{(1)} + \hat{H}^{(1)}\Psi^{(0)} - E^{(0)}\Psi^{(1)} - E^{(1)}\Psi^{(0)}\right]$$
$$+ \lambda^2\left[\hat{H}^{(0)}\Psi^{(2)} + \hat{H}^{(1)}\Psi^{(1)} - E^{(0)}\Psi^{(2)}\right.$$
$$\left. - E^{(1)}\Psi^{(1)} - E^{(2)}\Psi^{(0)}\right] + \cdots = 0$$

For an arbitrary perturbation parameter λ each order has to be zero.

$$\hat{H}^{(0)}\Psi^{(0)} = E^{(0)}\Psi^{(0)}$$
$$\hat{H}^{(0)}\Psi^{(1)} + \hat{H}^{(1)}\Psi^{(0)} = E^{(0)}\Psi^{(1)} + E^{(1)}\Psi^{(0)}$$
$$\hat{H}^{(0)}\Psi^{(2)} + \hat{H}^{(1)}\Psi^{(1)} = E^{(0)}\Psi^{(2)} + E^{(1)}\Psi^{(1)} + E^{(2)}\Psi^{(0)}$$
$$\vdots$$
$$\hat{H}^{(0)}\Psi^{(p)} + \hat{H}^{(1)}\Psi^{(p-1)} = E^{(0)}\Psi^{(p)} + E^{(1)}\Psi^{(p-1)} + \cdots + E^{(p)}\Psi^{(0)}$$

The solution for the un-perturbed problem is known. For the calculation of perturbative corrections the equations above have to be solved.
Solutions: Equations are multiplied from left with $\Psi^{(0)*}$ and then integrated over all coordinates.

2.2.2 Energy corrections

First order:

$$\underbrace{\langle \Psi^{(0)} \mid \hat{H}^{(0)}\Psi^{(1)} \rangle}_{=0} + \langle \Psi^{(0)} \mid \hat{H}^{(1)}\Psi^{(0)} \rangle = E^{(0)} \underbrace{\langle \Psi^{(0)} \mid \Psi^{(1)} \rangle}_{=0} + E^{(1)} \underbrace{\langle \Psi^{(0)} \mid \Psi^{(0)} \rangle}_{=1}$$

The right hand side can be simplified using that the unperturbed wave function $\Psi^{(0)}$ is normalised and orthogonal to the corrections $\Psi^{(p)}$. The first term on the left hand side is zero because $\hat{H}^{(0)}$ is Hermitian and the orthogonality condition.

$$\langle \Psi^{(0)} \mid \hat{H}^{(0)}\Psi^{(1)} \rangle = \langle \hat{H}^{(0)}\Psi^{(0)} \mid \Psi^{(1)} \rangle = E^{(0)}\langle \Psi^{(0)} \mid \Psi^{(1)} \rangle = 0$$

Therefore we get

$$E^{(1)} = \langle \Psi^{(0)} \mid \hat{H}^{(1)}\Psi^{(0)} \rangle$$

The energy correction in first order is equal to the expectation value of the perturbation operator w.r.t. the un-perturbed wave functions.
Higher orders:
From using the same derivation we get:

$$E^{(2)} = \langle \Psi^{(0)} \mid \hat{H}^{(1)}\Psi^{(1)} \rangle$$
$$\vdots$$
$$E^{(p)} = \langle \Psi^{(0)} \mid \hat{H}^{(1)}\Psi^{(p-1)} \rangle$$

The wavefunction in $(p-1)$-th order determines the energy in p-th order within perturbation theory.
Note: A more thorough analysis shows that the wave function in p-th order is sufficient to calculate the energy up to $(2p+1)$-th order.

2.2.3 Corrections $\Psi_n^{(1)}$ and $E_n^{(2)}$

$\Psi_n^{(1)}$ is expanded in a complete set of solutions to the un-perturbed problem $\Psi_j^{(0)}$.

$$\Psi_n^{(1)} = \sum_{j \neq n} a_j \Psi_j^{(0)}$$

With this orthogonality of $\Psi_n^{(0)}$ w.r.t. $\Psi_n^{(1)}$ is guaranteed.

$$\langle \Psi_n^{(0)} \mid \Psi_n^{(1)} \rangle = \sum_{j \neq n} a_j \langle \Psi_n^{(0)} \mid \Psi_j^{(0)} \rangle = 0$$

From the general equation in first order we get:

$$\begin{aligned}
\hat{H}^{(1)} \Psi_n^{(0)} - E_n^{(1)} \Psi_n^{(0)} &= E_n^{(0)} \Psi_n^{(1)} - \hat{H}^{(0)} \Psi_n^{(1)} \\
&= \left(E_n^{(0)} - \hat{H}^{(0)} \right) \sum_{j \neq n} a_j \Psi_j^{(0)} \\
&= \sum_{j \neq n} a_j \left(E_n^{(0)} - E_j^{(0)} \right) \Psi_j^{(0)}
\end{aligned}$$

Now we use the usual procedure: Multiplication from left with $\Psi_i^{(0)*}$ $(i \neq n)$ and integration.

$$\langle \Psi_i^{(0)} \mid \hat{H}^{(1)} \Psi_n^{(0)} \rangle = a_i \left(E_n^{(0)} - E_i^{(0)} \right) = H_{in}^{(1)}$$

This defines the matrix element $H_{in}^{(1)}$ of the perturbation operator. With this definition we get:

$$a_i = \frac{H_{in}^{(1)}}{E_n^{(0)} - E_i^{(0)}}$$

$$\Psi_n^{(1)} = \sum_{i \neq n} \frac{H_{in}^{(1)}}{E_n^{(0)} - E_i^{(0)}} \Psi_i^{(0)}$$

Energy correction in second order:

$$E_n^{(2)} = \langle \Psi_i^{(0)} \mid \hat{H}^{(1)} \Psi_n^{(1)} \rangle = \sum_{i \neq n} \frac{H_{in}^{(1)}}{E_n^{(0)} - E_i^{(0)}} \langle \Psi_i^{(0)} \mid \hat{H}^{(1)} \Psi_n^{(0)} \rangle$$

$$= \sum_{i \neq n} \frac{|H_{in}^{(1)}|^2}{E_n^{(0)} - E_i^{(0)}}$$

2.2.4 Summary

The equations of perturbation theory contain the energies $E_n^{(0)}$ and wave functions $\Psi_n^{(0)}$ of the un-perturbed problem and the matrix elements $H_{in}^{(1)}$ of the perturbation operator:

$$H_{in}^{(1)} = \langle \Psi_i^{(0)} \mid \hat{H}^{(1)} \Psi_n^{(0)} \rangle$$

Perturbation series of the energy:

$$E_n = E_n^{(0)} + H_{nn}^{(1)} + \sum_{i \neq n} \frac{|H_{in}^{(1)}|^2}{E_n^{(0)} - E_i^{(0)}} + \cdots$$

Perturbation series of the wave function:

$$\Psi_n = \Psi_n^{(0)} + \sum_{i \neq n} \frac{H_{in}^{(1)}}{E_n^{(0)} - E_i^{(0)}} \Psi_i^{(0)} + \cdots$$

Higher order corrections are not used very often.

2.3 The variation method

The expectation value of the energy E_Ψ calculated with an arbitrary (valid) wave function Ψ is an upper bound for the exact energy E_0 of the ground state of the system.

$$E_\Psi = \langle \Psi \mid \hat{H} \Psi \rangle \geq E_0 \qquad \Psi \text{ normalised}$$

Proof: Expand Ψ in the complete set of the exact eigenfunctions Ψ_n of the system.

$$\hat{H} \Psi_n = E_n \Psi_n$$

$$\Psi = \sum_{n=0}^{\infty} a_n \Psi_n$$

Normalisation:

$$\langle \Psi \mid \Psi \rangle = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_n^* a_m \langle \Psi_n \mid \Psi_m \rangle$$

$$\langle \Psi \mid \Psi \rangle = \sum_{n=0}^{\infty} |a_n|^2 = 1$$

Expectation value of the energy:

$$E_\Psi = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_n^* a_m \langle \Psi_n \mid \hat{H} \Psi_m \rangle$$

$$= \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_n^* a_m E_m \langle \Psi_n \mid \Psi_m \rangle$$

$$= \sum_{n=0}^{\infty} |a_n|^2 E_n$$

In the last equation we replace all eigenvalues E_n with the lowest eigenvalue E_0 (ground state), with $E_0 < E_1 < E_2 \ldots$. Then we have

$$E_\Psi \geq E_0 \sum_{n=0}^{\infty} |a_n|^2 = E_0$$

2.3.1 The variation principle

Within all valid wave functions Ψ for the ground state we are looking for the function with the lowest expectation value for the energy E_Ψ.

$$E_\Psi \longrightarrow \text{Minimum}$$

Practical procedure: Choose an Ansatz for Ψ with free parameters α_i and determine the parameters using the variational principle. This gives the best possible solution within this Ansatz.
This is equivalent to a extremal value problem. The variation δE_Ψ has to be zero for any arbitrary variation of the parameters α_i.

$$\delta E_\Psi(\alpha_1, \alpha_2, \ldots) = \frac{\partial E_\Psi}{\partial \alpha_1} d\alpha_1 + \frac{\partial E_\Psi}{\partial \alpha_2} d\alpha_2 + \cdots = 0$$

Because the variations $d\alpha_i$ are arbitrary, we have:

$$\frac{\partial E_\Psi}{\partial \alpha_i} = 0 \qquad \text{for all } i$$

2.3.2 Linear variations

Variational function Ψ is chosen as a linear combination of n fixed basis functions Φ_i:

$$\Psi = \sum_{i=1}^{n} c_i \Phi_i$$

The coefficients c_i are varied until the minimal value of the energy expectation value E_Ψ is found. For not normalised functions Ψ we have:

$$E_\Psi = \frac{\langle \Psi \mid \hat{H}\Psi \rangle}{\langle \Psi \mid \Psi \rangle} = \frac{A}{B}$$

$$A = \sum_k \sum_j c_k^* c_j \langle \Phi_k \mid \hat{H}\Phi_j \rangle$$

$$B = \sum_k \sum_j c_k^* c_j \langle \Phi_k \mid \Phi_j \rangle$$

Notation:

$$H_{kj} = \langle \Phi_k \mid \hat{H}\Phi_j \rangle \qquad = H_{jk}^* \qquad \text{(Matrix element of } \hat{H}\text{)}$$

$$S_{kj} = \langle \Phi_k \mid \Phi_j \rangle \qquad = S_{jk}^* \qquad \text{(Overlap integral)}$$

Minimisation of E_Ψ w.r.t. the variational parameters c_i:

$$\frac{\partial E_\Psi}{\partial c_i} = 0 \qquad \text{for } i = 1, 2, \ldots, n$$

$$\frac{\partial E_\Psi}{\partial c_i} = \frac{1}{B^2}\left(B\frac{\partial A}{\partial c_i} - A\frac{\partial B}{\partial c_i} \right) = \frac{1}{B}\left(\frac{\partial A}{\partial c_i} - \frac{A}{B}\frac{\partial B}{\partial c_i} \right) = 0$$

Because $B \neq 0$ (Normalisation integral) and with $E_\Psi = A/B$ we get:

$$\frac{\partial A}{\partial c_i} - E_\Psi \frac{\partial B}{\partial c_i} = 0$$

Calculation of the derivatives:

$$A = \sum_k \sum_j c_k^* c_j H_{kj}$$

$$\frac{\partial A}{\partial c_i} = 2 \sum_j c_j H_{ij}$$

$$B = \sum_k \sum_j c_k^* c_j S_{kj}$$

$$\frac{\partial B}{\partial c_i} = 2 \sum_j c_j S_{ij}$$

Inserting the derivatives we get the generalised eigenvalue equation:

$$\sum_j \left(H_{ij} - E_\Psi S_{ij} \right) c_j = 0 \qquad \text{for } i = 1, 2, \ldots, n$$

Known: Integrals H_{ij} and S_{ij}

Unknown: Coefficients c_j and corresponding energies E_Ψ.

Linear algebra: Non-trivial solution of a homogeneous system of linear equations, for the case of vanishing determinant:

$$|H_{ij} - E_\Psi S_{ij}| = 0$$

The solution of this equation gives n energies E_k, where the lowest value is the variational energy of the ground state. For each energy E_k we get after inserting in the system of linear equations the coefficients c_{jk} and therefore the corresponding wave function Ψ_k.

Additional index to distinguish the individual solutions:

$$\Psi_k = \sum_{j=1}^{n} c_{jk} \Phi_j$$

$$\sum_j H_{ij} c_{ij} = \sum_j S_{ij} c_{jk} E_k \qquad \text{for } i, k = 1, 2, \ldots, n$$

Using matrices:

$$H = \{H_{ij}\}, \qquad S = \{S_{ij}\}, \qquad C = \{C_{ij}\}, \qquad E = \{E_i \delta_{ij}\}$$
$$HC = SCE$$

Using the $n \times n$ -matrices H, C, S and E we can write the equation in matrix form. H and S are real for real basis functions Φ_j and symmetric (hermitian otherwise), E is a real diagonal matrix.

Important special case: Orthonormal basis

$$S_{ij} = \delta_{ij}$$
$$HC = CE$$

H is the matrix representation of the Hamilton operator in the orthonormal basis. Through diagonalisation of H we get the energy eigenvalues (diagonal elements of E) and eigenvectors (columns of C). We have

$$E = C^{-1} H C = C^T H C$$

$$C^T C = I \qquad \text{or} \quad \sum_{j=1}^{n} |c_{jk}|^2 = 1 \qquad \text{for } k = 1, 2, \ldots, n$$

The inverse and transpose of the matrix C are identical. Eigenvectors are orthonormal.

Lecture 3

Molecular Hamiltonian

3.1 Time-independent non-relativistic Schrödinger equation

For a system with N electrons and K nuclei the time-independent Schrödinger equation is

$$\hat{H}(\vec{r}, \vec{R})\Psi_n(\vec{r}, \vec{s}, \vec{R}, \vec{\sigma}) = E_n\Psi_n(\vec{r}, \vec{s}, \vec{R}, \vec{\sigma}) \ .$$

Notation:

$\vec{r}_i = (x_i, y_i, z_i)$	Coordinates for electron i
$\vec{s}_i$	Spin - coordinate for electron i
$\vec{R}_A = (x_A, y_A, z_A)$	Coordinates for nuclei A
$\vec{\sigma}_A$	Spin - coordinate for nuclei A
$\vec{r} = (\vec{r}_1, \ldots, \vec{r}_N)$	Coordinates for all electrons
$\vec{s} = (\vec{s}_1, \ldots, \vec{s}_N)$	Spin - coordinates for all electrons
$\vec{R} = (\vec{R}_A, \ldots, \vec{R}_K)$	Coordinates for all nuclei
$\vec{\sigma} = (\vec{\sigma}_A, \ldots, \vec{\sigma}_K)$	Spin - coordinates for all nuclei

Conventions:

Electrons :	Number N ;	Indices $i, j, \ldots$;	Mass m
Nuclei:	Number K ;	Indices $A, B, \ldots$;	Mass M_A

Non-relativistic Hamilton operator

$$\hat{H}(\vec{r}, \vec{R}) = \hat{T}_{\mathrm{N}}(\vec{R}) + \hat{T}_{\mathrm{el}}(\vec{r}) + \hat{V}_{\mathrm{Ne}}(\vec{r}, \vec{R}) + \hat{V}_{\mathrm{ee}}(\vec{r}) + \hat{V}_{\mathrm{NN}}(\vec{R})$$

41

Kinetic energy:

$$\hat{T}_{\mathrm{N}}(\vec{R}) = -\sum_A \frac{\hbar^2}{2M_A} \left(\frac{\partial^2}{\partial x_A^2} + \frac{\partial^2}{\partial y_A^2} + \frac{\partial^2}{\partial z_A^2} \right) \qquad = -\sum_A \frac{\hbar^2}{2M_A} \nabla_A^2$$

$$\hat{T}_{\mathrm{el}}(\vec{r}) = -\frac{\hbar^2}{2m} \sum_i \left(\frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \right) \qquad = -\frac{\hbar^2}{2m} \sum_i \nabla_i^2$$

Potential energy:

$$\hat{V}_{\mathrm{Ne}}(\vec{r}, \vec{R}) = -\sum_A \sum_i Z_A \frac{e^2}{r_{Ai}}$$

$$\hat{V}_{\mathrm{ee}}(\vec{r}) = +\sum_{i<j} \frac{e^2}{r_{ij}}$$

$$\hat{V}_{\mathrm{NN}}(\vec{R}) = +\sum_{A<B} Z_A Z_B \frac{e^2}{R_{AB}}$$

Where Z_A is the charge of nuclei A and $r_{ij} = |r_i - r_j|$.
$\hat{H}$ includes the non-relativistic kinetic energy and the potential energy of the Coulomb interaction.

Not included are:

- Relativistic correction to the kinetic energy

- Interaction of magnetic moments
  (orbit/orbit, Spin/orbit, Spin/Spin)

- Interaction with external electric and magnetic fields

3.2 Born–Oppenheimer Approximation

3.2.1 Part I

Goal: Separation of nuclear motion and electron motion.

Idea: Electrons are much lighter than nuclei and move therefore much faster. They adjust to each new position of the nuclei almost instantaneously.

Ansatz: The electronic wave function $\Psi_{\mathrm{el}}(\vec{r}; \vec{R})$ is calculated for fixed position $(\vec{R})$ of the nuclei. It directly depends on $\vec{r}$ and parametrically on $\vec{R}$.

Electronic Schrödinger equation

$$\hat{H}_{\text{el}}(\vec{r}; \vec{R})\Psi_{\text{el}}(\vec{r}; \vec{R}) = E_{\text{el}}(\vec{R})\Psi_{\text{el}}(\vec{r}; \vec{R})$$

with the electronic Hamilton operator

$$\hat{H}_{\text{el}}(\vec{r}; \vec{R}) = \hat{T}_{\text{el}}(\vec{r}) + \hat{V}_{\text{Ne}}(\vec{r}; \vec{R}) + \hat{V}_{\text{ee}}(\vec{r}) \ .$$

For the derivation we have used

- $\hat{T}_{\text{N}}(\vec{R}) = 0$ non-moving nuclei

- $\hat{V}_{\text{NN}}(\vec{R})$ does not act on $\vec{r}$ (multiplicative operator) and can be neglected for the solution

Interpretation of the equations

- The calculation of Ψ_{el} takes into account the electronic interaction of nuclei and electrons, however not their dynamic coupling. There is no exchange of kinetic energy.

- Macroscopic analogy: No heat exchange, adiabaticity.

- The solution of the electronic Schrödinger equation is difficult but possible (Quantum chemistry).

- The eigenfunctions $\Psi_{\text{el}}^n(\vec{r}; \vec{R})$ are a complete orthonormal set of functions.

3.2.2   Part II

The total wave function $\Psi(\vec{r}, \vec{R})$ is expanded in the electronic wave functions $\Psi_{\text{el}}^n(\vec{r}; \vec{R})$

$$\Psi(\vec{r}, \vec{R}) = \sum_n \chi_n(\vec{R})\Psi_{\text{el}}^n(\vec{r}; \vec{R}) \ .$$

Insertion in the Schrödinger equation

$$\hat{H}(\vec{r}, \vec{R}) \sum_n \chi_n(\vec{R})\Psi_{\text{el}}^n(\vec{r}; \vec{R}) = E \sum_n \chi_n(\vec{R})\Psi_{\text{el}}^n(\vec{r}; \vec{R})$$

Multiplication from left with $\Psi_{\text{el}}^{m*}(\vec{r}; \vec{R})$ and integration over all electronic coordinates $\vec{r}$

$$\left[\hat{T}_{\text{N}}(\vec{R}) + E_{\text{el}}^m(\vec{R}) + \hat{V}_{\text{NN}}(\vec{R}) - E\right] \chi_m(\vec{R}) = -\sum_n \hat{C}_{mn}(\vec{R})\chi_n(\vec{R}) \ ,$$

and

$$\hat{C}_{mn}(\vec{R}) = -\sum_A \frac{\hbar^2}{2M_A} \left[\langle \Psi_{\mathrm{el}}^m \mid \nabla_A^2 \mid \Psi_{\mathrm{el}}^n \rangle + 2 \langle \Psi_{\mathrm{el}}^m \mid \nabla_A \mid \Psi_{\mathrm{el}}^n \rangle \cdot \nabla_A \right] \quad .$$

Note for the derivation: The coupling terms $\hat{C}_{mn}(\vec{R})$ are due to the action of $\hat{T}_{\mathrm{N}}(\vec{R})$ on $\Psi_{\mathrm{el}}^n(\vec{r}; \vec{R})$.

Comparison of relevant matrix elements shows that $\hat{C}_{mn}(\vec{R})$ is normally vanishingly small compared to $\left[\hat{T}_{\mathrm{N}}(\vec{R}) + E_{\mathrm{el}}^m(\vec{R}) + \hat{V}_{\mathrm{NN}}(\vec{R}) \right]$.

Possible approximations

$\hat{C}_{mn} = 0$ for $m \neq n$:   Adiabatic approximations

$\hat{C}_{mn} = 0$ for all m, n :   Born–Oppenheimer approximation

Both approximations lead to non-coupled systems of equations for the nuclear functions $\chi_n(\vec{R})$.

3.2.3   Part III

After the decoupling of the nuclear Schrödinger equation we get for each electronic state an effective Schrödinger equation for the nuclear motion:

$$\hat{H}_{\mathrm{N}}(\vec{R}) \chi_m(\vec{R}) = E \chi_m(\vec{R}) \quad ,$$

with

$$\hat{H}_{\mathrm{N}}(\vec{R}) = \hat{T}_{\mathrm{N}}(\vec{R}) + \hat{V}_{\mathrm{eff}}(\vec{R}) \quad ,$$

where we have been using the adiabatic approximation

$$\hat{V}_{\mathrm{eff}}(\vec{R}) = E_{\mathrm{el}}^m(\vec{R}) + \hat{V}_{\mathrm{NN}}(\vec{R}) + \hat{C}_{mm}(\vec{R})$$

and the Born–Oppenheimer approximation, respectively.

$$\hat{V}_{\mathrm{eff}}(\vec{R}) = E_{\mathrm{el}}^m(\vec{R}) + \hat{V}_{\mathrm{NN}}(\vec{R})$$

The diagonal correction term $\hat{C}_{mm}(\vec{R})$ is small and difficult to calculate. Most calculations are done using the Born–Oppenheimer potential energy surface:

$$U(\vec{R}) = E_{\mathrm{el}}^m(\vec{R}) + V_{\mathrm{NN}}(\vec{R})$$

The electronic energy $E_{\mathrm{el}}^m(\vec{R})$ and the interaction energy of the nuclei are independent of the nuclear masses, therefore we see that

Isotopes have the same potential energy surface $U(\vec{R})$.

The Born–Oppenheimer approximation is often valid to high accuracy. Exceptions occur if different electronic states are significantly coupled through nuclear motion. Special care is necessary for

- close lying potential energy surfaces (crossings and avoided crossings of electronically excited states)

- rapid movements of nuclei (impact of fast particles).

3.2.4 Order of magnitude of effects for H_2

a) Relativistic effects in H_2 (State $X^1\Sigma_g^+$) at the equilibrium geometry R_e

Change of D_0 :	-0.48 cm^{-1}
D_1 :	-0.34 cm^{-1}
E($\nu = 1$) - E($\nu = 0$):	-0.14 cm^{-1}

b) Non-adiabatic effects in H_2 (State $X^1\Sigma_g^+$) at the equilibrium geometry R_e

Approximately calculated change in energy

$$H_2 : 0.65 \text{ cm}^{-1}, \qquad HD : 0.42 \text{ cm}^{-1}, \qquad D_2 : 0.23 \text{ cm}^{-1}$$

All effects are certainly smaller than 1 cm^{-1},
$D_0 = 36118.3 \pm 0.5$ cm^{-1} (experiment)
$D_0 = 36118.0$ cm^{-1} (theory; adiabatic + relativistic)

c) Born–Oppenheimer effects, more precisely: energy change due to the diagonal adiabatic correction C_{nn}

$H_2\ X^1\Sigma_g^+$	120 cm^{-1}
$a^3\Sigma_g^-$	100 cm^{-1}
$B^3\Sigma_u^+$	200 cm^{-1}
$E, F^1\Sigma_g^+$	430 cm^{-1} (avoided crossing)

This could be important for the exact calculation of electron spectra (80 cm$^{-1} \approx 0.01$eV).

Not important for vibrational spectroscopy, Born–Oppenheimer effects on calculated transition energies ($\Delta\nu = 1$) are smaller than 1 cm^{-1} for H_2 (State $X^1\Sigma_g^+$) for $\nu = 1 - 10$.

3.3 Stationary points

N Atoms, $3N$ degrees of freedom (coordinates)
For non-linear molecules:

- 3 degrees of freedom for translation

- 3 degrees of freedom for rotation

- $3N - 6$ degrees of freedom for internal coordinates

Coordinates (e.g. Cartesian coordinates):

$$\vec{R} = (R_1, R_2, \ldots, R_{3N}) = \{R_i\}$$

Definition

$$g_i = \frac{\partial U(\vec{R})}{\partial R_i} \qquad \text{components of gradient}$$

$$f_{ij} = \frac{\partial^2 U(\vec{R})}{\partial R_i \partial R_j} \qquad \text{harmonic force constants}$$

Taylor expansion within the vicinity of a reference point with coordinates $\vec{R}_0 = \{R_{i,0}\}$:

$$U = U(0) + \sum_i g_i x_i + \frac{1}{2} \sum_i \sum_j f_{ij} x_i x_j + \cdots$$

with

$$x_i = R_i - R_{i,0}$$

or in matrix form

$$U = U(0) + g^T x + \frac{1}{2} x^T F x + \cdots$$

with

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{3N} \end{pmatrix}, \qquad g = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_{3N} \end{pmatrix}, \qquad F = \begin{pmatrix} F_{1,1} & \cdots & F_{1,3N} \\ \vdots & \ddots & \vdots \\ F_{3N,1} & \cdots & F_{3N,3N} \end{pmatrix}$$

Definition:
Points with the coordinates $\vec{R}_e = \{R_{i,e}\}$ are called stationary if all components of the gradient are zero.

$$g_i(\vec{R}_e) = 0 \qquad \text{for } i = 1, 2, \ldots, 3N$$

At a stationary point we have the Taylor expansion

$$U = U(0) + \frac{1}{2} \sum_i \sum_j f_{ij} x_i x_j + \cdots$$

with

$$x_i = R_i - R_{i,e} \ .$$

Characterisation of stationary points:
The force constant matrix $F(\vec{R}_e)$ has $3N$ real eigenvalues, whereof

- 3 for translation (equal zero)

- 3 for rotation (equal zero)

- $3N - 6$ for internal motion (different from zero)

The number of negative eigenvalues is used to characterise the stationary point

- 0 for a minimum

- 1 for a transition state

- $2, 3, \ldots$ for a saddle point of higher order

3.4   Internal coordinates

Changes in bond length and in the angles between chemical bonds provide the most significant and physically meaningful set of coordinates for the description of the potential energy. These coordinates are called "internal coordinates" since they describe (being unaffected by translation and rotations of the molecule as a whole) just the internal motions of the molecule, i.e. the molecular vibrations.
The types of internal coordinates which are generally used are the following:

- *Bond stretching coordinate* : variation of the length of a chemical bond.
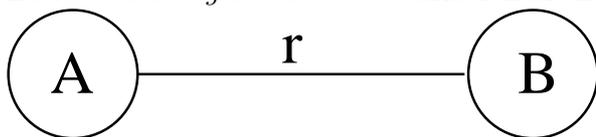
- *In-plane bending coordinate* : variation of the angle between two chemical bonds having one atom in common.
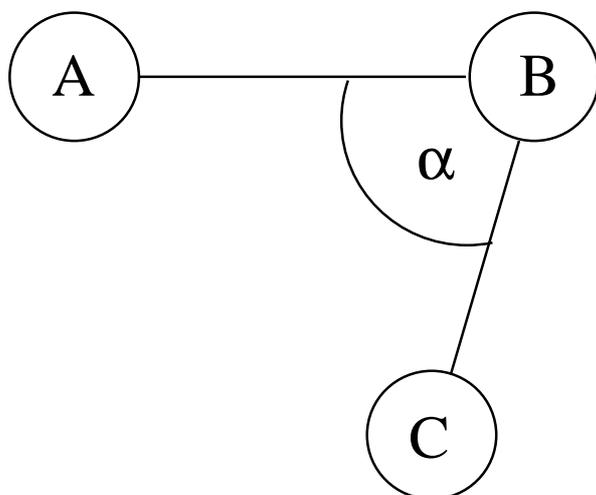
- *Out-of-plane bending coordinate* : variation of the angle between the plane defined by two bonds with one atom in common and a third bond connected to the common atom.
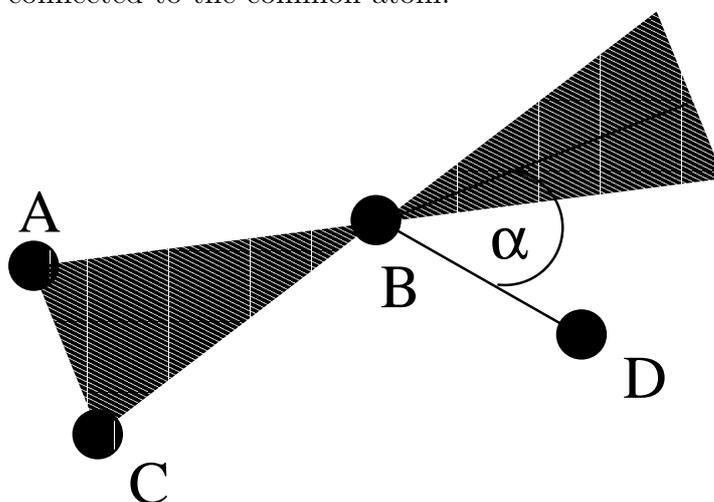
- *Torsion coordinate* : variation in the dihedral angle between the planes determined by three consecutive bonds connecting four atoms.
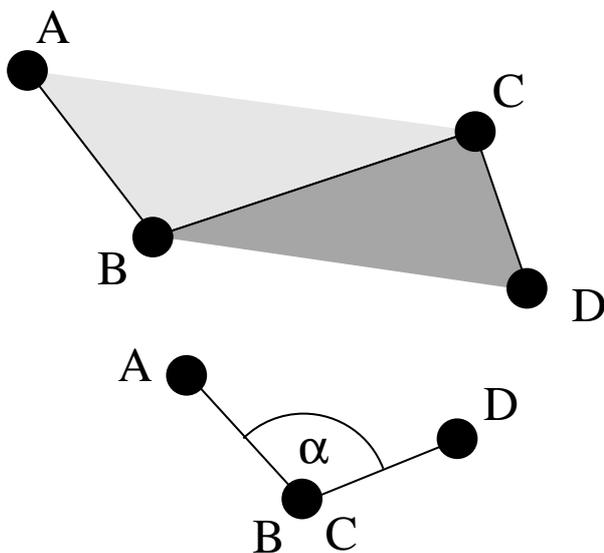
Since the variations in bond length and angles in a molecule can be written in terms of the variation in the Cartesian displacement coordinates of the atoms, the problem we have to face is that of writing the explicit form of the transformation from Cartesian coordinates to the new set of internal coordinates.

Let us call s_k a generic internal coordinate. The most general relation between s_k and the Cartesian displacement coordinates can be written in the form

$$s_k = \sum_i B_i^k x_i + \frac{1}{2} \sum_{ij} B_{ij}^k x_i x_j + \text{higher terms}$$

where the coefficients B_i^k, B_{ij}^k, etc., are determined by the molecular geometry. A drastic simplification can be achieved if we restrict our treatment to the case of infinitesimal amplitudes of vibration where we can drop all terms not linear in x.

If we call $\mathbf{s}$ and $\mathbf{x}$ the vectors whose components are the internal and the Cartesian coordinates, the linear equation in matrix notation is

$$\mathbf{s} = \mathbf{Bx}$$

where $\mathbf{B}$ is a matrix whose elements are the coefficients B_i^k.

The matrix $\mathbf{B}$ is not, in general, a square matrix. Let us consider, for example a bent triatomic molecule and choosing as internal coordinates the two bond length and the bond angle. There are also $3n - 6 = 3$ vibrational degrees of freedom and with the 9 Cartesian coordinates the $\mathbf{B}$-matrix will have 3 rows and 9 columns. The translations and rotations are not included in the set of internal coordinates which describe, by definition a molecular motion

in which the relative positions of the atoms are changed. Since $\mathbf{B}$ is not a square matrix, it cannot be inverted.

We can always include in the $\mathbf{s}$-vector six additional coordinates to describe the three translations and the three rotations. These six coordinates are most conveniently chosen in the form

$$R_1 = \frac{1}{\sqrt{M}} \sum_I \sqrt{m_I} q_{Ix} \qquad R_4 = \sum_I \left(\frac{m_I}{I_x}\right)^{1/2} \left(y_I^0 q_{Iz} - z_I^0 q_{Iy}\right)$$

$$R_2 = \frac{1}{\sqrt{M}} \sum_I \sqrt{m_I} q_{Iy} \qquad R_5 = \sum_I \left(\frac{m_I}{I_y}\right)^{1/2} \left(z_I^0 q_{Iz} - x_I^0 q_{Iy}\right)$$

$$R_3 = \frac{1}{\sqrt{M}} \sum_I \sqrt{m_I} q_{Iz} \qquad R_6 = \sum_I \left(\frac{m_I}{I_z}\right)^{1/2} \left(x_I^0 q_{Iz} - y_I^0 q_{Iy}\right)$$

where $M = \sum_I m_I$ is the total mass of the molecule and I_x, I_y, and I_z are the moments of inertia. x_I^0, y_I^0, and z_I^0 are the coordinates of the equilibrium position of the Ith atom relative to the centre of mass and

$$q_{Ix} = \sqrt{m_I} \Delta x_I \qquad q_{Iy} = \sqrt{m_I} \Delta y_I \qquad q_{Iz} = \sqrt{m_I} \Delta z_I$$

In this way $\mathbf{B}$ is a square and non-singular matrix (i.e. $|\mathbf{B}| \neq 0$) and therefore can be inverted. We can thus define the inverse transformation to get from internal to Cartesian coordinates

$$\mathbf{x} = \mathbf{B}^{-1}\mathbf{s}$$

where $\mathbf{B}^{-1}$ is the inverse matrix of $\mathbf{B}$ including translation and rotation.

3.4.1 Z-matrix

A special choice of internal coordinates often used as input to quantum chemical programs is the Z-matrix. The Z-matrix of a molecule is build from a list of the atoms. The coordinates of each new atom (A) are characterised by

1. The distance to another, already defined atom B

2. The angle between the bond to atom B and the bond of atom B to an already defined atom C.

3. The dihedral angle between A, B, C, and an already defined atom D.

The first three atoms are special cases with

- The first atom is at the origin of the Cartesian coordinate system.

- The second atom is defined by the distance to atom 1 and is oriented along a Cartesian coordinate axis.

- The third atom is defined by the distance to atom 1 or 2 and the angle (3-2-1) or the angle (3-1-2)

The Z-matrix of a molecule is not unique. There are many different possibilities to order the atoms and to define distances, angles and dihedrals.

Example: Z-matrix for staggered conformation of ethane.
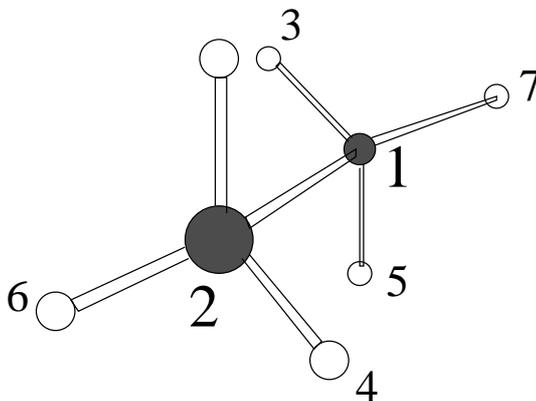

```
1  C
2  C    1.54    1
3  H    1.01    1    109.5    2
4  H    1.01    2    109.5    1    180.0    3
5  H    1.01    1    109.5    2     60.0    4
6  H    1.01    2    109.5    1    -60.0    5
7  H    1.01    1    109.5    2    180.0    6
8  H    1.01    2    109.5    1     60.0    7
```

In the first line of the Z-matrix atom 1, a carbon atom, is defined. Atom number 2 is also a carbon that is a distance of 1.54 Å from atom 1 (columns 3 and 4). Atom 3 is a hydrogen atom that is bonded to atom 1 with a bond length of 1.01 Å. The angle formed by atoms 2-1-3 is 109.5^o, information that is specified in columns 5 and 6. The fourth atom is a hydrogen, a distance of 1.01 Å from atom 2, the angle 4-2-1 is 109.5^o, and the torsion angle for atoms 4-2-1-3 is 180°. Thus for all except the first three atoms, each atom has

three internal coordinates: the distance of the atom from one of the atoms previously defined, the angle formed by the atom and two of the previous atoms, and the torsion angle defined by the atom and three of the previous atoms.

The Z-matrix can often be written in many different ways as there are many combinations of internal coordinates. As a rule internal coordinates should be chosen to have minimal coupling. However, it should be stressed that the final results of a quantum mechanical calculations are not dependent on the actual coordinate system chosen.

Lecture 4

Two-Electron Systems and Spin

4.1 Helium atom: Ground state

4.1.1 Basics

Electronic Hamilton operator within Born–Oppenheimer approximation:

$$\hat{H}_{\text{el}}(\vec{r}; \vec{R}) = \hat{T}_{\text{e}}(\vec{r}) + V_{\text{Ne}}(\vec{r}; \vec{R}) + V_{\text{ee}}(\vec{r})$$

Coordinate system: nuclei at origin, spherical coordinates for electron 1 (r_1, θ_1, ϕ_1) and electron 2 (r_2, θ_2, ϕ_2).
Electronic Hamilton operator for He (Z=2):

$$\hat{H}_{\text{el}}(\vec{r}_1, \vec{r}_2) = -\frac{\hbar^2}{2m_e}(\nabla_1^2 + \nabla_2^2) - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} + \frac{e^2}{r_{12}}$$

Corresponding Schrödinger equation:

$$\hat{H}_{\text{el}}(\vec{r}_1, \vec{r}_2)\Psi_{\text{el}}(\vec{r}_1, \vec{r}_2) = E_{\text{el}}\Psi_{\text{el}}(\vec{r}_1, \vec{r}_2)$$

For this differential equation with 6 variables there is no exact solution. We are looking for approximations, first for the ground state.

Experiment:
The energy E_0 of the ground state is equal to the sum of ionisation potentials for He (24.59 eV) and He$^+$ (54.42 eV).

$$E_0(\text{exp}) = -79.01 \text{ eV}$$

This is our reference to assess different approximations.

4.1.2 Orbital model

Ansatz:
$$\Psi_{el}(\vec{r}_1, \vec{r}_2) = \Psi_1(\vec{r}_1)\Psi_2(\vec{r}_2)$$

One electron functions are called orbitals. In the orbital model the N-electron function is approximated by a product of N orbitals.
Probability interpretation:

$$\Psi_{el}^*\Psi_{el}d\tau = \left[\Psi_1(\vec{r}_1)^*\Psi_1(\vec{r}_1)d\tau_1\right]\left[\Psi_2(\vec{r}_2)^*\Psi_2(\vec{r}_2)d\tau_2\right]$$
$$W_{12} = W_1W_2$$

The probability W_{12} to find at the same time electron 1 at $\vec{r}_1$ and electron 2 at $\vec{r}_2$ is equal to the product of the single probabilities W_1 and W_2.
In the orbital model the probability distribution of an electron is independent from the other electrons ('independent particle model').
Note:
More accurate methods that go beyond the orbital model have to include the correlation of electrons.

4.1.3 Formal separation

The Hamilton operator consists of one- and two-electron parts

$$\hat{H}_{el} = \hat{H}^{(0)} + \hat{V}_{ee} = \hat{H}^{(0)} + \frac{e^2}{r_{12}}$$

$$\hat{H}^{(0)} = \sum_{i=1}^{2} \hat{h}_i(\vec{r}_i)$$

$$\hat{h}_i = -\frac{\hbar^2}{2m_e}\nabla_i^2 - \frac{Ze^2}{r_i}$$

The one-electron operator $\hat{h}_i$ is the same as the Hamilton operator of a hydrogen-like atom with $Z = 2$ for He. Eigenvalues and eigenfunctions Ψ_i are known.

$$\hat{h}_i\Psi_k(\vec{r}_i) = \epsilon_k\Psi_k(\vec{r}_i)$$

$$\epsilon_k = -\frac{Z^2e^2}{2a}\frac{1}{n_k^2} \qquad \text{with } n_i = 1,2,3,\ldots$$

$$\Psi_k = \pi^{-1/2}(Z/a)^{3/2}e^{-\frac{Zr_1}{a}} \qquad \text{for } n_k = 1$$

Crude Approximation: With $\hat{V}_{ee} = 0$ the Hamilton operator separates like $\hat{H}^{(0)}$. The two-electron wave function can in this case be written as a product

of orbitals (the solutions for the hydrogen atom with $Z = 2$), and we get for the energy

$$E^{(0)} = \epsilon_1(n_1) + \epsilon_2(n_2)$$

Ground state of the He atom with $n_1 = n_2 = 1$:

$$E_0^{(0)} = -\frac{Z^2 e^2}{a} = -108.84 \text{ eV}$$

$$\Psi_0^{(0)} = \frac{Z^3}{\pi a^3} e^{-\frac{Z}{a}(r_1 + r_2)}$$

Because of the assumption $\hat{V}_{\text{ee}} = 0$ we get a huge discrepancy to the experimental ground state energy (-79.01 eV).

4.1.4 Perturbation theory

We take the electron-electron interaction $\hat{V}_{\text{ee}}$ as the perturbation operator $\hat{H}^{(1)}$. This is not really a valid choice as $\hat{V}_{\text{ee}}$ is not small compared to the other terms in the Hamilton operator. Nevertheless we calculate the correction to the energy to first order:

$$E_0^{(1)} = \langle \Psi_0^{(0)} \mid \hat{V}_{\text{ee}} \mid \Psi_0^{(0)} \rangle = \langle \Psi_0^{(0)} \mid \frac{e^2}{r_{12}} \mid \Psi_0^{(0)} \rangle$$

Inserting the analytic solutions for $\Psi_0^{(0)}$ leads to a difficult six-dimensional integral for the variables $(r_1, \theta_1, \phi_1, r_2, \theta_2, \phi_2)$, that can be solved analytically.

$$E_0^{(1)} = \frac{5Z}{8} \frac{e^2}{a}$$

For He with $Z = 2$ we get:

$$E_0^{(1)} = 34.01 \text{ eV}$$

$$E_0^{(0)} + E_0^{(1)} = -74.83 \text{ eV}$$

With the perturbation correction to first order the energy is off by 5.3% from the experimental value (-79.01 eV).

Corrections to higher order are possible, but much more complicated than simple variational calculations.

4.1.5 Variational calculations

Idea: In the solutions used up to know, both electrons were seeing the field of the nuclei with $Z = 2$. This neglects the partial screening of the nuclear charge due to the other electron.

Ansatz: We replace the nuclear charge $Z = 2$ in the wave function for the ground state by an effective nuclear charge $\eta < 2$.

$$\Psi(r_1, r_2) = \frac{\eta^3}{\pi a^3} e^{-\frac{\eta}{a}(r_1 + r_2)}$$

Variational calculation for the determination of the optimal value of η:

$$\frac{dE_\Psi}{d\eta} = \frac{d}{d\eta} \langle \Psi \mid \hat{H}_{\text{el}} \mid \Psi \rangle = 0$$

The calculation of the expectation value of $\hat{H}_{\text{el}}$ w.r.t. to Ψ affords the solution of several integrals (difficult but analytic solutions exist).
Solution:

$$\langle \Psi \mid \hat{H}_{\text{el}} \mid \Psi \rangle = \left[\eta^2 - 2\eta Z + \frac{5}{8}\eta \right] \frac{e^2}{a}$$

$$\frac{dE_\Psi}{d\eta} = \left[2\eta - 2Z + \frac{5}{8} \right] \frac{e^2}{a} = 0$$

$$\eta_{\text{opt}} = Z - \frac{5}{16}$$

$$E_{\text{opt}} = -\eta_{\text{opt}}^2 \frac{e^2}{a} = -\left(Z - \frac{5}{16} \right)^2 \frac{e^2}{a}$$

For He with $Z = 2$ we get:

$$\eta_{\text{opt}}(He) = 1.6875$$
$$E_{\text{opt}}(He) = -77.49 \text{ eV}$$

The nuclear charge is screened by $0.3e$. The error in the ground state energy is reduced from 5.3% with $\eta = 2$ and a perturbative calculation to 1.9% with $\eta = \eta_{\text{opt}}$.

4.1.6 Overview of ground state energies

Goal: Assess accuracy of approximative solutions by comparison to experiment.

	E_0 (a.u.)	error %
Experiment	-2.903784(3)	
Orbital model		
Zeroth order (without $\hat{V}_{\text{ee}}$), $E_0^{(0)}$	-4.00	-37.8
Perturbation theory 1. order, $E_0^{(0)} + E_0^{(1)}$	-2.75	5.3
Simple variational calculation, E_{opt}	-2.8477	1.9
Best variational calculation	-2.8617	1.4
Including electron correlation		
Simple variational calculation, 40 Parameters	-2.90276	0.04
Best variational calculation, 1078 Parameters	-2.9037462	0.001
With additional relativistic corrections	-2.9037848	

4.2 Electron spin

Experimental fact:
Electrons have a proper angular momentum (Spin), the components of which can have only two distinct values in any arbitrary direction:

$$s_z = \pm \frac{\hbar}{2}$$

Quantum mechanical description with spin operators:

$$\hat{s}^2 = \hat{s}_x^2 + \hat{s}_y^2 + \hat{s}_z^2 = \hat{s}_+\hat{s}_- - \hat{s}_z + \hat{s}_z^2 = \hat{s}_-\hat{s}_+ + \hat{s}_z + \hat{s}_z^2$$
$$\hat{s}_+ = \hat{s}_x + i\hat{s}_y$$
$$\hat{s}_- = \hat{s}_x - i\hat{s}_y$$

The linear combinations $\hat{s}_+$ and $\hat{s}_-$ are useful for derivations. Spin operators have the same commutation relations as the angular momentum operators:

$$[\hat{s}_x, \hat{s}_y] = i\hat{s}_z; \qquad [\hat{s}_y, \hat{s}_z] = i\hat{s}_x; \qquad [\hat{s}_z, \hat{s}_x] = i\hat{s}_y$$
$$[\hat{s}^2, \hat{s}_x] = [\hat{s}^2, \hat{s}_y] = [\hat{s}^2, \hat{s}_z] = 0$$

Eigenfunctions:
For a one electron system there are two orthonormal eigenfunctions $|\alpha\rangle$ and $|\beta\rangle$. Application of operators:

	$|\alpha\rangle$	$|\beta\rangle$
$\hat{s}^2$	$\frac{3}{4}|\alpha\rangle$	$\frac{3}{4}|\beta\rangle$
$\hat{s}_z$	$\frac{1}{2}|\alpha\rangle$	$-\frac{1}{2}|\beta\rangle$
$\hat{s}_x$	$\frac{1}{2}|\beta\rangle$	$\frac{1}{2}|\alpha\rangle$
$\hat{s}_y$	$\frac{i}{2}|\beta\rangle$	$-\frac{i}{2}|\alpha\rangle$
$\hat{s}_+$	0	$|\alpha\rangle$
$\hat{s}_-$	$|\beta\rangle$	0

General eigenvalue equation for spin:

$$\hat{s}_z|\sigma\rangle = m_s\hbar|\sigma\rangle$$
$$\hat{s}^2|\sigma\rangle = s(s+1)\hbar^2|\sigma\rangle$$

Spin eigenfunction $|\sigma\rangle$, quantum numbers s and m_s for the total spin and its z-component with $-s \le m_s \le s$.
For a single electron we have:

$$m_s = \pm\frac{1}{2}; \qquad s = \frac{1}{2}$$

Spin orbitals: For the complete description of an electron we need three spatial variables and one spin variable. The spin orbitals $\Psi(\vec{r})|\sigma\rangle$ are products of a spatial orbital $\Psi(\vec{r})$ and a spin eigenfunction $|\sigma\rangle$, $|\alpha\rangle$ oder $|\beta\rangle$.

Two electron systems:
Operators and eigenvalue equation:

$$\hat{S}_x = \hat{s}_{x1} + \hat{s}_{x2}; \qquad \hat{S}_y = \hat{s}_{y1} + \hat{s}_{y2}; \qquad \hat{S}_z = \hat{s}_{z1} + \hat{s}_{z2}$$
$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2$$

Because we have the general commutator relations for angular moment operators, the eigenvalue equations have the form:

$$\hat{S}_z|\sigma_1\sigma_2\rangle = M_S\hbar|\sigma_1\sigma_2\rangle \qquad (-S \le M_S \le S)$$
$$\hat{S}^2|\sigma_1\sigma_2\rangle = S(S+1)\hbar^2|\sigma_1\sigma_2\rangle$$

Spin eigenfunction $|\sigma_1 \sigma_2\rangle$, quantum numbers S and M_S for total spin and its z-component.

Relation to the Hamilton operator:
We use a non-relativistic Hamilton operator without any spin components, therefore Hamilton operator and spin operators commute.

$$[\hat{H}, \hat{S}^2] = [\hat{H}, \hat{S}_z] = 0$$

Conclusions:

- There are exact expectation values for $\hat{S}^2$ and $\hat{S}_z$. The system can be characterised by the spin quantum numbers S and M_S.

- For $|\sigma_1 \sigma_2\rangle$ we can use a product function.

4.2.1 Electrons are indistinguishable

In microscopic systems similar particles (e.g. electrons) can experimentally not be distinguished. The probability density $\Psi^*\Psi$ is in such many particle systems not allowed to be dependent on the arbitrary enumeration of the particles. A change of the numeration, exchange of particles, is not allowed to change $\Psi^*\Psi$.
Conclusion: With an exchange of two particles of the same type, the many-body wavefunction Ψ either stays the same or changes sign. Wave functions are either symmetric or anti-symmetric w.r.t. the exchange of two particles of the same kind.

Spin eigenfunctions for two-electron systems:
A naive product function leads to the following combination

$$|\sigma_1 \sigma_2\rangle = \begin{cases} \alpha(1)\alpha(2) \\ \alpha(1)\beta(2) \\ \beta(1)\alpha(2) \\ \beta(1)\beta(2) \end{cases}$$

In the second and third product we can distinguish the two electrons, this is not possible for the products one and four. In the first case symmetric wave functions can be constructed through linear combinations:

$$2^{-1/2}[\alpha(1)\beta(2) + \beta(1)\alpha(2)]$$
$$2^{-1/2}[\alpha(1)\beta(2) - \beta(1)\alpha(2)]$$

An acceptable set of product functions is

$$
\begin{aligned}
&\alpha(1)\alpha(2); \\
&2^{-1/2}[\alpha(1)\beta(2) + \beta(1)\alpha(2)]; \\
&2^{-1/2}[\alpha(1)\beta(2) - \beta(1)\alpha(2)]; \\
&\beta(1)\beta(2)
\end{aligned}
$$

Electron spin for two-electron systems: Singlets and triplets

Spin eigenfunction	Exchange	M_S	S	Notation
$\alpha(1)\alpha(2)$	symmetric	1	1	triplet
$2^{-1/2}[\alpha(1)\beta(2) + \beta(1)\alpha(2)]$	symmetric	0	1	triplet
$\beta(1)\beta(2)$	symmetric	-1	1	triplet
$2^{-1/2}[\alpha(1)\beta(2) - \beta(1)\alpha(2)]$	antisymmetric	0	0	singlet

Simple example:

$$
\begin{aligned}
\hat{S}_z \alpha(1)\alpha(2) &= (\hat{s}_{z1} + \hat{s}_{z2})\alpha(1)\alpha(2) \\
&= \hat{s}_{z1}\alpha(1)\alpha(2) + \hat{s}_{z2}\alpha(1)\alpha(2) \\
&= \frac{1}{2}\hbar\alpha(1)\alpha(2) + \frac{1}{2}\hbar\alpha(1)\alpha(2) \\
&= \hbar\alpha(1)\alpha(2) \qquad \rightarrow M_S = 1
\end{aligned}
$$

4.3 Pauli principle

All in nature realisable states of many-electron systems are described by space-spin-wave functions that are anti-symmetric w.r.t. electron exchange.

Special case of the Pauli principle: Every spatial orbital can be occupied by at most two electrons with opposite spin.

Generalisation to other particles:

Fermions	= Particles with half-integer spin
Bosons	= Particles with integer spin

All systems with several equal fermions (bosons) are described by wave functions, that are anti-symmetric (symmetric) w.r.t. exchange of two equal fermions (bosons).

Example for fermions: electron, proton, neutron
Example for bosons: H-Atom, He-nuclei

Lecture 5

Hartree–Fock Approximation

5.1   Slater-determinant

We are looking for a n-electron wave function Ψ within the orbital model that satisfies the Pauli principle. This wave function is build from orthonormal spin orbitals η.

$$\eta(\vec{r}, \sigma) = \Psi(\vec{r})|\sigma\rangle$$
$$\langle \eta_i | \eta_j \rangle = \delta_{ij}$$

Short hand for electron 1:

$$\eta(1) = \Psi(1)|\sigma(1)\rangle = \left\{ \begin{array}{ccc} \Psi\alpha(1) & = & \Psi(1) \\ \Psi\beta(1) & & \bar{\Psi}(1) \end{array} \right.$$

Ansatz

$$\Psi(1, 2, \ldots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \eta_1(1) & \eta_1(2) & \cdots & \eta_1(N) \\ \eta_2(1) & \eta_2(2) & \cdots & \eta_2(N) \\ \vdots & \vdots & & \vdots \\ \eta_N(1) & \eta_N(2) & \cdots & \eta_N(N) \end{vmatrix}$$

Exchange of two electrons corresponds to an exchange of two columns. This leads to a change of sign for the determinant. The determinantal wave functions is therefore anti-symmetric w.r.t. an exchange of electrons (Pauli principle).

A determinant vanishes if two rows are the same. In non-trivial wave functions $\Psi \neq 0$ all spin orbitals have to be different. A given spatial orbital can only appear twice in the determinant, combined with either α or β.

Pauli principle in the orbital model: A spatial orbital can at most be occupied by two electrons with opposite spin.

Writing out the determinant leads to $N!$ products of the form $\eta_1\eta_2\cdots\eta_N$. All possible permutations of the N electrons in the N spin orbitals appear. In this approach the electrons are evenly spread out over all spin orbitals and therefore indistinguishable.

Short notation: Often we omit the normalisation constant $(N!)^{-1/2}$ and we only write the diagonal elements.

$$\Psi(1, 2, \ldots, N) = |\eta_1(1)\eta_2(2)\cdots\eta_N(N)|$$

Slater determinants are always eigenfunctions of $\hat{S}_z$, however, they are not necessarily eigenfunctions of $\hat{S}^2$. For the general case there are always linear combination of determinants that are eigenfunctions of $\hat{S}_z$ and $\hat{S}^2$ at the same time. Such spin-adapted linear combination of determinants (configurations) are needed to describe open-shell systems.

Important special case: In closed-shell systems the spins of all electrons are paired.

$$\Psi(1, 2, \ldots, N) = \left|\Psi_1(1)\bar{\Psi}_1(2), \Psi_2(3)\bar{\Psi}_2(4), \ldots \Psi_{N/2}(N-1)\bar{\Psi}_{N/2}(N)\right|$$

This Slater determinant is an eigenfunction of $\hat{S}_z$ and $\hat{S}^2$ with $S = M_S = 0$ (Singlet).

5.1.1 Energy expectation value for Slater determinants

The electronic energy E_{el} is the expectation value of the non-relativistic electronic Hamilton operators $\hat{H}_{\text{el}}$ w.r.t. the N-electron wave function, written as a Slater determinant of orthonormal spin orbitals.

$$E_{\text{el}} = \langle\Psi|\hat{H}_{\text{el}}|\Psi\rangle$$
$$\hat{H}_{\text{el}} = \sum_{i=1}^{N} \hat{h}(i) + \sum_{i<j}^{N} \frac{e^2}{r_{ij}}$$
$$\hat{h}(i) = -\frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_A \frac{Z_A e^2}{r_{iA}}$$

Using the Slater determinant function we get

$$E_{\text{el}} = \sum_{i=1}^{N} h_{ii}^{(\eta)} + \sum_{i<j}^{N} \left(J_{ij}^{(\eta)} - K_{ij}^{(\eta)} \right)$$

$$h_{ii}^{(\eta)} = \int \eta_i^*(1) \hat{h}(1) \eta_i(1) d\tau_1$$

$$J_{ij}^{(\eta)} = \int \int \eta_i^*(1) \eta_i(1) \frac{e^2}{r_{12}} \eta_j^*(2) \eta_j(2) d\tau_1 d\tau_2$$

$$K_{ij}^{(\eta)} = \int \int \eta_i^*(1) \eta_j(1) \frac{e^2}{r_{12}} \eta_j^*(2) \eta_i(2) d\tau_1 d\tau_2$$

$h_{ii}^{(\eta)}$ is the one-electron integral, $J_{ij}^{(\eta)}$ Coulomb integral, and $K_{ij}^{(\eta)}$ an exchange integral. In these equations all integrals are over spin orbitals η. By integration over the spin coordinates we get to a description in terms of spatial orbitals.

Example for the transition to spatial orbitals:

$$\eta_i = \Psi_i |\alpha\rangle$$

$$h_{ii}^{(\eta)} = h_{ii}^{(\Psi)} \langle \alpha | \alpha \rangle = h_{ii}^{(\Psi)}$$

$$h_{ii}^{(\Psi)} = \int \Psi^*(\vec{r}_1) \hat{h}(1) \Psi_i(\vec{r}_1) d\vec{r}_1$$

A similar separation of spatial and spin parts occurs in all integrals. This is due to the fact that $\hat{H}_{\text{el}}$ does not contain any spin coordinates. A systematic investigation of all integrals leads to:

$$h_{ii}^{(\eta)} = h_{ii}^{(\Psi)}$$

$$J_{ij}^{(\eta)} = J_{ij}^{(\Psi)}$$

$$K_{ij}^{(\eta)} = \begin{cases} K_{ij}^{(\Psi)} \\ 0 \end{cases} \text{ if spin functions in } \eta_i, \eta_j \begin{array}{l} \text{equal} \\ \text{opposite} \end{array}$$

Interaction of two electrons in spatial orbitals Ψ_i, Ψ_j:

J_{ij} with opposite spin

$J_{ij} - K_{ij}$ with equal spin   (smaller, because $K_{ij} > 0$)

Important special case: closed shell systems

$$E_{\text{el}}^{(\Psi)} = 2 \sum_{i=1}^{N/2} h_{ii}^{(\Psi)} + \sum_{i=1}^{N/2} \sum_{j=1}^{N/2} \left(2 J_{ij}^{(\Psi)} - K_{ij}^{(\Psi)} \right)$$

5.1.2 Unitary transformation of spin orbitals

In the definition of the Slater determinant we regard the spin orbitals $\eta_i(j)$ formally as elements A_{ji} of a matrix A.

$$\Psi = (N!)^{-1/2}\det(A)$$

A unitary transformation U of the spin orbitals leaves them orthonormal:

$$\eta'_k = \sum_i \eta_i U_{ik}$$

$$U^{-1} = U^\dagger, \qquad (U^{-1})_{ik} = (U^\dagger)_{ik} = (U^*)_{ki}$$

In matrix form:

$$A' = AU$$
$$\Psi' = (N!)^{-1/2}\det(A')$$
$$= (N!)^{-1/2}\det(AU)$$
$$= (N!)^{-1/2}\det(A)\,\det(U)$$
$$\Psi' = \Psi\,\det(U)$$

Calculation of $\det(U)$ for a unitary transformation:

$$U^\dagger U = I$$
$$\det(U^\dagger U) = \det(U^\dagger)\,\det(U) = [\det(U)]^*\,\det(U) = |\det(U)|^2 = 1$$
$$\det(U) = e^{i\Phi}$$

Ψ' differs from Ψ at most by a phase factor $e^{i\Phi}$ with ± 1 for the real case. For all expectation values we get

$$\langle\Psi|\hat{O}|\Psi\rangle = \langle\Psi'|\hat{O}|\Psi'\rangle$$

In this sense determinantal wave functions are invariant w.r.t. unitary transformation of spin orbitals.

5.2 Hartree–Fock method

Goal: Calculation of optimal orbitals

Approach: Variational method, minimisation of the energy expectation value for the Slater determinant under the condition that orbitals stay orthonormal.

Derivation for spin and spatial orbitals is similar. We study the more general case of spin orbitals. For orthonormal spin orbitals we have the energy in simplified notation:

$$E_{\text{el}} = \sum_{i=1}^{N} h_{ii} + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (J_{ij} - K_{ij})$$

$$S_{ij} = \langle \eta_i | \eta_j \rangle = \delta_{ij}, \qquad S_{ij} = S_{ij}^*$$

All integrals are for spin orbitals. Spin orbitals may be complex. Optimal spin orbitals minimise the energy expression while maintaining the constraint condition. This can be achieved by the method of Lagrange multipliers:

$$\mathcal{L}[\{\eta_i\}] = E_{\text{el}}[\{\eta_i\}] - \sum_{i=1}^{N} \sum_{j=1}^{N} \epsilon_{ji} \left(S_{ij} - \delta_{ij} \right)$$

$$\epsilon_{ji} = \epsilon_{ij}^\star$$

$$\delta \mathcal{L} = 0$$

Lagrange multipliers ϵ_{ji} have to be Hermitian in order for the Lagrange-Function $\mathcal{L}$ to be real. Variation of the spin orbitals must lead to a vanishing variation of $\mathcal{L}$:

$$\delta \mathcal{L} = \delta E_{\text{el}} - \sum_{i=1}^{N} \sum_{j=1}^{N} \epsilon_{ji} \delta S_{ij} = 0$$

We investigate the variation of the integrals that is caused by the variation of the spin orbitals (c.c. = conjugate complex)

$$\delta S_{ij} = \int \delta \eta_i^*(1) \eta_j(1) d\tau_1 + \int \eta_i^*(1) \delta \eta_j(1) d\tau_1$$

$$\delta E_{\text{el}} = \sum_{i} \delta h_{ii} + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\delta J_{ij} - \delta K_{ij})$$

$$\delta h_{ii} = \int \delta \eta_i^*(1) \hat{h}(1) \eta_j(1) d\tau_1 + \int \eta_i^*(1) \hat{h}(1) \delta \eta_j(1) d\tau_1$$

$$\delta h_{ii} = \int \delta \eta_i^*(1) \hat{h}(1) \eta_j(1) d\tau_1 + \text{c.c.}$$

Because of Hermiticity of $\hat{h}$, the second term in the equation for δh_{ii} is the complex conjugate of the first term.

$$\delta J_{ij} = \int \int d\tau_1 d\tau_2 \delta\eta_i^*(1)\eta_i(1)\frac{e^2}{r_{12}}\eta_j^*(2)\eta_j(2) \qquad A$$

$$+ \int \int d\tau_1 d\tau_2 \eta_i^*(1)\delta\eta_i(1)\frac{e^2}{r_{12}}\eta_j^*(2)\eta_j(2) \qquad B=A^*$$

$$+ \int \int d\tau_1 d\tau_2 \eta_i^*(1)\eta_i(1)\frac{e^2}{r_{12}}\delta\eta_j^*(2)\eta_j(2) \qquad C$$

$$+ \int \int d\tau_1 d\tau_2 \eta_i^*(1)\eta_i(1)\frac{e^2}{r_{12}}\eta_j^*(2)\delta\eta_j(2) \qquad D=C^*$$

$$\sum_{i=1}^N \sum_{j=1}^N \delta J_{ij} = \sum_{i=1}^N \sum_{j=1}^N 2 \int \int d\tau_1 d\tau_2 \delta\eta_i^*(1)\eta_i(1)\frac{e^2}{r_{12}}\eta_j^*(2)\eta_j(2) + \text{c.c.}$$

In the double sum in the last equation the terms A and C appear in equivalent pairs and can be combined (same is true for δK_{ij}):

$$\sum_{i=1}^N \sum_{j=1}^N \delta K_{ij} = \sum_{i=1}^N \sum_{j=1}^N 2 \int \int d\tau_1 d\tau_2 \delta\eta_i^*(1)\eta_j(1)\frac{e^2}{r_{12}}\eta_j^*(2)\eta_i(2) + \text{c.c.}$$

For the remaining terms we get:

$$\sum_{i=1}^N \sum_{j=1}^N \epsilon_{ji}\delta S_{ij} = \sum_{i=1}^N \sum_{j=1}^N \epsilon_{ji} \int \delta\eta_i^*(1)\eta_j(1)d\tau_1 + \text{c.c.}$$

A further simplification can be achieved with the introduction of the Coulomb operator $\hat{J}_j$ and exchange operator $\hat{K}_j$:

$$\hat{J}_j(1)\eta_i(1) = \int d\tau_2 \eta_j^*(2)\frac{e^2}{r_{12}}\eta_j(2)\eta_i(1)$$

$$\hat{K}_j(1)\eta_i(1) = \int d\tau_2 \eta_j^*(2)\frac{e^2}{r_{12}}\eta_i(2)\eta_j(1)$$

$$J_{ij} = \int d\tau_1 \eta_i^*(1)\hat{J}_j(1)\eta_i(1)$$

$$K_{ij} = \int d\tau_1 \eta_i^*(1)\hat{K}_j(1)\eta_i(1)$$

Taking all terms together we get:

$$\sum_i \int d\tau_1 \delta\eta_i^*(1) \left[\hat{h}(1)\eta_i(1) \right.$$

$$\left. + \sum_j \left(\hat{J}_j(1) - \hat{K}_j(1) \right) \eta_i(1) - \sum_j \epsilon_{ji}\eta_j(1) \right] + \text{c.c.} = 0$$

Variations $\delta\eta_i^*$ are arbitrary and therefore every single integral to an index i has to disappear and the integrand $[\ldots]$ has to be zero.

$$\left[\hat{h}(1) + \sum_j \left(\hat{J}_j(1) - \hat{K}_j(1) \right) \right] \eta_i(1) = \sum_j \epsilon_{ji}\eta_j(1) \qquad \text{for } i = 1, \ldots, N$$

The same steps performed on the conjugate-complex part lead to the same result.

Introduction of the Fock-operator:

$$\hat{F}(1) = \hat{h}(1) + \sum_{j=1}^{N} \left(\hat{J}_j(1) - \hat{K}_j(1) \right)$$

$$\hat{F}(1)\eta_i(1) = \sum_j \epsilon_{ji}\eta_j(1)$$

The current derivation did not lead to the desired eigenvalue equation for spin orbitals. We now investigate unitary transformations of the spin orbitals. We have already shown that determinantal wave functions and their expectation values are invariant w.r.t. such unitary transformations. The same is true for the Fock operator. For the Lagrange multipliers we get ϵ_{ji}:

$$\langle \eta_k(1)|\hat{F}(1)|\eta_i(1)\rangle = \sum_j \epsilon_{ji}\langle \eta_k(1)|\eta_j(1)\rangle = \epsilon_{ki}$$

The multipliers are equivalent to the matrix elements of the Fock operator. A transformation of the spin orbitals leads to (in matrix form):

$$\epsilon' = U^\dagger \epsilon U$$

ϵ is an Hermitian matrix. Therefore there is a special unitary matrix U, which transforms ϵ to diagonal form. For this special set of transformed spin orbitals we get

$$\hat{F}(1)\eta_i'(1) = \epsilon_i'\eta_i'(1)$$

Normally one omits the special mark for these canonic spin orbitals and one writes:

$$\hat{F}\eta_i = \epsilon_i\eta_i \qquad \text{for } i = 1, \ldots, N$$

5.2.1 Solution of the Hartree-Fock equations

Problem: The Fock operator consists of Coulomb $\hat{J}_j$ and exchange operators $\hat{K}_j$. Therefore, it depends on the spin orbitals η_i, we want to calculate with the Hartree-Fock equations.

Solution: Iterative method starting from an arbitrary but reasonable initial set of orbitals $\eta_i^{(0)}$.

Hartree-Fock equations have to be solved iteratively.
Scheme

$$\{\eta_i^{(0)}\} \longrightarrow \hat{F}^{(1)} \longrightarrow \{\eta_i^{(1)}\} \longrightarrow \hat{F}^{(2)} \longrightarrow \{\eta_i^{(2)}\} \longrightarrow \text{etc.}$$

This sequence is stopped, if the solutions in two consecutive iterations don't change (within a given tolerance). This method is called SCF (Self-Consistent-Field).
For a known (self-consistent) set of occupied spin orbitals the Fock operator $\hat{F}$ is a well-defined Hermitian operator with an infinite number of eigenfunctions η_i (N occupied spin orbitals with lowest orbital energies and un-occupied (virtual) orbitals with higher orbital energies).

5.2.2 Orbital energies and Koopmans' theorem

For an N-electron system, minimisation of the energy of the determinant $|\Psi\rangle = |\eta_1\eta_2\cdots\eta_N\rangle$ leads to an eigenvalue equation $\mathcal{F}|\eta_a\rangle = \epsilon_a|\eta_a\rangle$ for the n occupied spin orbitals $\{\eta_A\}$. The Fock operator has a functional dependence on these occupied spin orbitals, but once the occupied spin orbitals are known the Fock operator becomes a well-defined Hermitian operator, which will have an infinite number of eigenfunctions

$$\mathcal{F}|\eta_j\rangle = \epsilon_j|\eta_j\rangle \qquad j = 1, 2, \ldots, \infty$$

Each of the solutions has a spin orbital energy ϵ_j. The N spin orbitals with the lowest orbital energies are just the spin orbitals occupied in $|\Psi\rangle$ for which we use the indices $a, b, \ldots$. the remaining infinite number of spin orbitals with higher energies are the virtual or un-occupied spin orbitals, which we label with the indices $r, s, \ldots$.
The matrix elements of the Fock operator in the basis of the spin orbital eigenfunctions is diagonal with diagonal elements equal to the orbital energies

$$\langle \eta_i \mid \mathcal{F} \mid \eta_j \rangle = \epsilon_j \langle \eta_i \mid \eta_j \rangle = \epsilon_j \delta_{ij}$$

The orbital energies can be expressed

$$\epsilon_i = \langle \eta_i \mid \mathcal{F} \mid \eta_i \rangle = \langle \eta_i \mid h \mid \eta_i \rangle + \sum_b \langle \eta_i \mid (\mathcal{J}_b - \mathcal{K}_b) \mid \eta_i \rangle$$

$$= \langle \eta_i \mid h \mid \eta_i \rangle + \sum_b \langle \eta_i \mid \mathcal{J}_b \mid \eta_i \rangle - \sum_b \langle \eta_i \mid \mathcal{K}_b \mid \eta_i \rangle$$

$$= \langle i \mid h \mid i \rangle + \sum_b ((ii, bb) - (ib, bi)) \ ,$$

where we have introduced the short form (ij, kl) for the two-electron integrals

$$(ij, kl) = \int \int d\tau_1 d\tau_2 \delta \eta_i^*(1) \eta_j(1) \frac{e^2}{r_{12}} \eta_k^*(2) \eta_l(2) \ .$$

In particular we get for occupied orbitals

$$\epsilon_a = \langle a \mid h \mid a \rangle + \sum_b ((aa, bb) - (ab, ba))$$

$$= \langle a \mid h \mid a \rangle + \sum_{b \neq a} ((aa, bb) - (ab, ba))$$

and for virtual orbitals

$$\epsilon_r = \langle r \mid h \mid r \rangle + \sum_b ((rr, bb) - (rb, br))$$

The orbital energy ϵ_a represents the energy of an electron in the spin orbital $|\eta_a\rangle$. This energy is the kinetic energy and the attraction to the nuclei ($\langle a \mid h \mid a \rangle$) plus a Coulomb and exchange interaction with each of the remaining $N - 1$ electrons. The result for ϵ_r has a different character. It includes the kinetic energy and the nuclear attraction of an electron in $|\eta_r\rangle$, but includes Coulomb and exchange interactions with n electrons of the Hartree–Fock ground state. It is as if an electron has been added to $|\Psi_0\rangle$ to produce a $(N + 1)$-electron state.

If we simply add up the orbital energies of the occupied states, we get

$$\sum_a^N \epsilon_a = \sum_a^N \langle a \mid h \mid a \rangle + \sum_a^N \sum_b^N ((aa, bb) - (ab, ba))$$

If we compare this with the total Hartree–Fock energy

$$E_{\mathrm{HF}} = \sum_a^N \langle a \mid h \mid a \rangle + \frac{1}{2} \sum_a^N \sum_b^N ((aa, bb) - (ab, ba))$$

we see that

$$E_{\text{HF}} \neq \sum_a^N \epsilon_a$$

and the total energy of the state $|\Psi_0\rangle$ is not just the sum of the orbital energies. The reason is as follows. The energy ϵ_a includes Coulomb and exchange interactions between an electron in orbital a and electrons in all other occupied spin orbitals, in particular orbital b. But ϵ_b includes Coulomb and exchange interactions between an electron in orbital b and electrons in all other occupied spin orbitals, in particular orbital a. Thus if we add ϵ_a and ϵ_b we include the electron-electron interactions between an electron in a and one in b twice. The sum of orbital energies counts the electron-electron interactions twice. This is the reason for the factor $\frac{1}{2}$ in the correct energy expression for the total energy relative to the sum of orbital energies.

If the total energy is not the sum of the orbital energies what physical significance can we attach to orbital energies?

Koopmans' theorem
Given an N-electron Hartree–Fock single determinant $|^N\Psi_0\rangle$ with occupied and virtual spin orbital energies ϵ_a and ϵ_r, then the ionisation potential to produce an $(N-1)$-electron single determinant $|^{N-1}\Psi_a\rangle$ with identical spin orbitals, obtained by removing an electron from spin orbital η_a, and the electron affinity to produce an $(N+1)$-electron single determinant $|^{N+1}\Psi^r\rangle$ with identical spin orbitals, obtained by adding an electron to spin orbital η_r, are just $-\epsilon_a$ and $-\epsilon_r$, respectively.

Proof: The ionisation potential is

$$\text{IP} =^{N-1} E_a -^N E_0$$

where $^{N-1}E_a$ and $^N E_0$ are the expectation values of the energy of the two relevant single determinants

$$^N E_0 = \langle ^N\Psi_0 \mid \mathcal{H} \mid ^N \Psi_0 \rangle$$
$$^{N-1}E_a = \langle ^{N-1}\Psi_a \mid \mathcal{H} \mid ^{N-1} \Psi_a \rangle$$

The ionisation potential for orbital c is therefore

$$
\begin{aligned}
\text{IP} =&^{N-1} E_c -^N E_0 \\
=& \sum_{a \neq c} \langle a \mid h \mid a \rangle + \frac{1}{2} \sum_{a \neq c} \sum_{b \neq c} ((aa, bb) - (ab, ba)) \\
& - \sum_{a} \langle a \mid h \mid a \rangle + \frac{1}{2} \sum_{a} \sum_{b} ((aa, bb) - (ab, ba)) \\
=& -\langle c \mid h \mid c \rangle - \frac{1}{2} \sum_{a} ((aa, cc) - (ac, ca)) - \frac{1}{2} \sum_{b} ((cc, bb) - (bc, cb)) \\
=& -\langle c \mid h \mid c \rangle - \sum_{b} ((cc, bb) - (bc, cb)) \\
=& -\epsilon_c
\end{aligned}
$$

Occupied orbital energies are generally negative and ionisation potentials are positive.

With a very similar calculation we can also calculate

$$
\text{EA} =^N E_0 -^{N+1} E^r = -\epsilon_r
$$

This result is consistent with the previous observation that ϵ_r included interactions with all N other electrons of the ground state and thus describes an $N + 1$th electron.

Koopmans' ionisation potentials are reasonable first approximations to experimental ionisation potentials (missing are orbital relaxation and correlation effects). Koopmans' electron affinities are unfortunately often bad. Many neutral molecules will add an electron to form a stable negative ion. Hartree–Fock calculations on neutral molecules, however, almost always give positive orbital energies for all the virtual orbitals.

5.2.3 Brillouin's theorem

The Hartree–Fock equation produces a set of spin orbitals. The single determinant $|\Psi_0\rangle$, formed from the N lowest spin orbitals is the Hartree–Fock approximation to the ground state. There are many other determinants with N functions that can be formed from the full set of spin orbitals. Brillouin's theorem makes a statement of a subset of these determinants. This subset is the set of singly excited determinants $|\Psi_a^r\rangle$ obtained from $|\Psi_0\rangle$ by a single replacement of η_a with η_r.

Brillouin's theorem
Singly excited determinants $|\Psi_a^r\rangle$ will not interact directly with a reference

Hartree–Fock determinant $|\Psi_0\rangle$, i.e., $\langle \Psi_0 \mid \mathcal{H} \mid \Psi_a^r \rangle = 0$.

Calculating the matrix element

$$\langle \Psi_0 \mid \mathcal{H} \mid \Psi_a^r \rangle = \langle a \mid h \mid r \rangle + \sum_b \left((ar, bb) - (ab, br) \right)$$

The right hand side of this equation is just the matrix element of the Fock operator in the spin orbital basis. Therefore,

$$\langle \Psi_0 \mid \mathcal{H} \mid \Psi_a^r \rangle = \langle \eta_a \mid \mathcal{F} \mid \eta_r \rangle$$

The matrix element that mixes singly excited determinants with $|\Psi_0\rangle$ is thus equal to an off-diagonal element of the Fock matrix. Now, by definition, solving the Hartree–Fock eigenvalue problem requires the off-diagonal elements to satisfy $\langle \eta_i \mid \mathcal{F} \mid \eta_j \rangle = 0$ $(i \neq j)$. One can then say that solving the Hartree–Fock eigenvalue equation is equivalent to ensuring that $|\Psi_0\rangle$ will not mix with any singly excited determinants.

Lecture 6

Molecular Orbital Theory

6.1   LCAO-Ansatz

LCAO: Linear Combination of Atomic Orbitals
In the LCAO approach the molecular orbitals (MOs) Ψ_i are written as linear combinations of atomic orbitals (AOs) Φ_μ.

$$\Psi_i = \sum_\mu c_{\mu i} \Phi_\mu$$

In the following the equations can also be applied to other types of basis sets that are not necessarily of atomic type.
Conventions:
$i, j, k, l, \ldots$ for MOs
$\mu, \nu, \lambda, \sigma, \ldots$ for AOs
We restrict the following derivations to closed-shell systems, singlet states with $S = M_S = 0$, with all orbitals doubly occupied.
Energy formula:

$$E_{\mathrm{el}} = 2 \sum_{i=1}^{N/2} h_{ii} + \sum_{i=1}^{N/2} \sum_{j=1}^{N/2} (2J_{ij} - K_{ij})$$

$$h_{ii} = \int \Psi_i^*(1)\hat{h}(1)\Psi_i(1)d\tau_1 = \langle \Psi_i | \hat{h}\Psi_i \rangle$$

$$J_{ij} = \int \int \Psi_i^*(1)\Psi_i(1)\frac{e^2}{r_{12}}\Psi_j^*(2)\Psi_j(2)d\tau_1 d\tau_2 = (ii, jj)$$

$$K_{ij} = (ij, ij)$$

We insert the LCAO equation and express all MO integrals with AO integrals. We assume that MOs and AOs are real functions.

Connection of MO and AO integrals:

$$h_{ii} = \left\langle \sum_\mu c_{\mu i} \Phi_\mu \mid \hat{h} \sum_\nu c_{\nu i} \Phi_\nu \right\rangle = \sum_\mu \sum_\nu c_{\mu i} c_{\nu i} h_{\mu\nu}$$

$$h_{\mu\nu} = \langle \Phi_\mu | \hat{h} \Phi_\nu \rangle$$

$$J_{ij} = \sum_\mu \sum_\nu \sum_\lambda \sum_\sigma c_{\mu i} c_{\nu i} c_{\lambda j} c_{\sigma j} (\mu\nu, \lambda\sigma)$$

$$K_{ij} = \sum_\mu \sum_\nu \sum_\lambda \sum_\sigma c_{\mu i} c_{\nu j} c_{\lambda i} c_{\sigma j} (\mu\nu, \lambda\sigma)$$

$$(\mu\nu, \lambda\sigma) = \int \int \Phi_\mu(1) \Phi_\nu(1) \frac{e^2}{r_{12}} \Phi_\lambda(2) \Phi_\sigma(2) d\tau_1 d\tau_2$$

Energy expression in AO basis:

$$E_{\mathrm{el}} = 2 \sum_i \sum_\mu \sum_\nu c_{\mu i} c_{\nu i} h_{\mu\nu}$$
$$+ \sum_i \sum_j \sum_\mu \sum_\nu \sum_\lambda \sum_\sigma c_{\mu i} c_{\nu i} c_{\lambda j} c_{\sigma j} [2(\mu\nu, \lambda\sigma) - (\mu\lambda, \nu\sigma)]$$

Introduction of the density matrix $P_{\mu\nu}$:

$$P_{\mu\nu} = 2 \sum_{i=1}^{N/2} c_{\mu i} c_{\nu i}$$

Simplification of the energy expression in the AO basis:

$$E_{\mathrm{el}} = \sum_\mu \sum_\nu P_{\mu\nu} h_{\mu\nu} + \frac{1}{4} \sum_\mu \sum_\nu \sum_\lambda \sum_\sigma P_{\mu\nu} P_{\lambda\sigma} [2(\mu\nu, \lambda\sigma) - (\mu\lambda, \nu\sigma)]$$

For a given basis $\{\Phi_\mu\}$ all the AO integrals $h_{\mu\nu}$ and $(\mu\nu, \lambda\sigma)$ are (in principle) known.
The LCAO coefficients $c_{\mu i}$ and the density matrix elements $P_{\mu\nu}$ are not known. They can be calculated from a linear variational calculation (Roothaan-Hall method).

6.2 Roothaan-Hall equations

Instead of a complete derivation it is easier to insert the LCAO approach in the general Hartree–Fock equations and multiply from left with Φ_μ^* and

integrate:

$$\hat{F}\left(\sum_\nu c_{\nu i}\Phi_\nu\right) = \epsilon_i\left(\sum_\nu c_{\nu i}\Phi_\nu\right)$$

$$\sum_\nu c_{\nu i}\langle\Phi_\mu|\hat{F}\Phi_\nu\rangle = \epsilon_i\sum_\nu c_{\nu i}\langle\Phi_\mu|\Phi_\nu\rangle$$

Introduction of the Fock matrix F and the overlap matrix S in AO basis:

$$F_{\mu\nu} = \langle\Phi_\mu|\hat{F}\Phi_\nu\rangle$$
$$S_{\mu\nu} = \langle\Phi_\mu|\Phi_\nu\rangle$$

Eigenvalue equation:

$$\sum_\nu F_{\mu\nu}c_{\nu i} = \epsilon_i\sum_\nu S_{\mu\nu}c_{\nu i} \qquad i = 1, 2, \ldots, K$$

$$FC = SC\epsilon \qquad \text{Matrix-Notation}$$

$$F = \{F_{\mu\nu}\}, \qquad C = \{c_{\nu i}\}, \qquad S = \{S_{\mu\nu}\}, \qquad \epsilon = \{\epsilon_i\delta_{ij}\}$$

For K basis functions Φ_μ there are K equations. The matrices have the dimension $K \times K$; F and S are symmetric, ϵ is diagonal.

Fock matrix elements in AO basis:

$$F_{\mu\nu} = \langle\Phi_\mu|\hat{h} + \sum_{i=1}^{N/2}(2\hat{J}_j - \hat{K}_j)|\Phi_\nu\rangle$$

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{j=1}^{N/2}[2(\mu\nu, jj) - (\mu j, j\nu)]$$

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{j=1}^{N/2}\sum_\lambda\sum_\sigma c_{\lambda j}c_{\sigma j}[2(\mu\nu, \lambda\sigma) - (\mu\lambda, \sigma\nu)]$$

$$F_{\mu\nu} = h_{\mu\nu} + \sum_\lambda\sum_\sigma P_{\lambda\sigma}[(\mu\nu, \lambda\sigma) - \frac{1}{2}(\mu\lambda, \sigma\nu)]$$

Solution of the Roothaan equations

As for the general Hartree–Fock equations, the Roothaan equations can only be solved iteratively: The Fock matrix elements $F_{\mu\nu}$ depend on the density matrix elements $P_{\lambda\sigma}$. The density matrix has to be calculated from the unknown LCAO coefficients $c_{\nu i}$.

A LCAO-MO-SCF calculation:

1. Choose a basis of atomic orbitals $\{\Phi_\mu\}$

2. Calculate the AO integrals $h_{\mu\nu}$, $(\mu\nu, \lambda\sigma)$ and $S_{\mu\nu}$.

3. Choose a reasonable initial density matrix $\{P_{\mu\nu}^{(0)}\}$.

4. Calculate the corresponding Fock matrix.

5. Solve the eigenvalue equation (Roothaan–Hall).

6. Calculate a new density matrix.

7. Iterate (4)-(5)-(6) until self consistency is achieved.

Solving the eigenvalue equations:
We examine a linear transformation of the non-orthogonal basis $\{\Phi_\mu\}$ into an orthogonal basis $\{\Phi'_\mu\}$.

$$\Phi'_\mu = \sum_\nu X_{\nu\mu}\Phi_\nu$$

$$\delta_{\mu\nu} = \langle \Phi'_\mu | \Phi'_\nu \rangle = \langle \sum_\lambda X_{\lambda\mu}\Phi_\lambda | \sum_\sigma X_{\sigma\nu}\Phi_\sigma \rangle$$

$$= \sum_\lambda \sum_\sigma X_{\lambda\mu}^* S_{\lambda\sigma} X_{\sigma\nu}$$

$$I = X^\dagger S X$$

The overlap matrix S in the basis $\{\Phi_\mu\}$ is Hermitian and can be diagonalised using a unitary transformation U.

$$U^\dagger S U = s$$

The diagonal matrix s contains the eigenvalues of S ($s_{ii} > 0$). For the transformation matrix X the following approaches are used:

$$X = S^{-1/2} = U s^{-1/2} U^\dagger \qquad \text{symmetric orthogonalisation}$$

$$X = U s^{-1/2} \qquad \text{canonical orthogonalisation}$$

The symmetric orthogonalisation is used most often. The canonical orthogonalisation has the advantage that linear dependencies in the basis ($s_{ii} \approx 0$) can easily be eliminated.

Connection between the different sets of LCAO coefficients:

$$\Psi_i = \sum_\mu c'_{\mu i}\Phi'_\mu = \sum_\mu c'_{\mu i} \sum_\nu X_{\nu\mu}\Phi_\nu = \sum_\nu \left[\sum_\mu X_{\nu\mu}c'_{\mu i} \right] \Phi_\nu = \sum_\nu c_{\nu i}\Phi_\nu$$

$$C = XC', \qquad C' = X^{-1}C$$

With these definitions we can rewrite the eigenvalue equations:

$$FXC' = SXC'\epsilon$$
$$(X^\dagger FX)C' = (X^\dagger SX)C'\epsilon = C'\epsilon$$

Using the Fock matrix F' in the basis $\{\Phi'_\mu\}$ we get a standard eigenvalue equation:

$$F' = X^\dagger FX$$
$$F'C' = C'\epsilon$$

Modified LCAO-MO-SCF procedure:

1. Choose a basis of atomic orbitals $\{\Phi_\mu\}$

2. Calculate AO integrals $h_{\mu\nu}, (\mu\nu, \lambda\sigma)$ and $S_{\mu\nu}$. In addition calculate the matrix X.

3. Choose a reasonable initial density matrix $\{P^{(0)}_{\mu\nu}\}$.

4. Calculate the Fock matrix.

5. Solve the Roothaan–Hall equations.

   (a) Transform the Fock matrix

   (b) Solve eigenvalue equation

   (c) Back transform coefficients

6. Calculate a new density matrix

7. Iterate (4)-(5)-(6) until self-consistency is reached

6.3   Unrestricted open-shell Hartree–Fock

Up to now we have first derived the Hartree–Fock equation for general spin orbitals η

$$\mathcal{F}\eta_i = \epsilon_i\eta_i$$

and then have had a look at the special case of closed shell orbitals. The special case lead to the restricted Hartree–Fock (RHF) equations. This special case has the advantage that the wavefunction is an eigenfunction not only of the Hartree–Fock Hamiltonian, but also of the spin operators $\mathcal{S}_z$ and

$\mathcal{S}^2$. For the case of unpaired spins there are two possible methods available. Either the electrons are partitioned in paired and non-paired sets and the paired electrons treated with equal spatial orbitals (the restricted open-shell Hartree–Fock approach, ROHF) or all electrons are treated on the same footing using general spin orbitals in an unrestricted Hartree–Fock approach (UHF). The first method has the advantage that the wavefunction is still an eigenfunction of the spin operators but the disadvantage that it needs a rather involved theory and that the solution is not variationally stable w.r.t. an unrestricted wavefunction. The UHF method is conceptually simpler and leads to the variational ground state, but might lead to wavefunctions that are no longer eigenfunctions of $\mathcal{S}^2$.

6.3.1 Spin contamination

We start from a Slater determinant

$$|\Psi_{\mathrm{UHF}}\rangle = |\eta_1 \eta_2 \cdots \eta_N\rangle$$

where the spin orbitals are

$$\eta(\mathbf{x}) = \begin{cases} \phi^\alpha(\mathbf{r})\alpha(\omega) \\ \phi^\beta(\mathbf{r})\beta(\omega) \end{cases}$$

either occupied with an α or β electron. There are N_α α and N_β β electrons in the system.

$$N = N_\alpha + N_\beta$$

If we assume that $N_\alpha > N_\beta$ then we can easily find that the expectation value of $\mathcal{S}_z$ is

$$\langle \mathcal{S}_z \rangle = \frac{1}{2} \left(N_\alpha - N_\beta \right)$$

The exact expectation value of a corresponding wavefunction for $\mathcal{S}^2$ would be

$$\langle \mathcal{S}^2_{\mathrm{exact}} \rangle = \frac{1}{4} \left(N_\alpha - N_\beta \right) \left(N_\alpha - N_\beta + 2 \right)$$

However, if we calculate the expectation value for an unrestricted Slater determinant, we get

$$\langle \mathcal{S}^2_{\mathrm{UHF}} \rangle = \langle \mathcal{S}^2_{\mathrm{exact}} \rangle + N_\beta - \sum_i^{N_\alpha} \sum_j^{N_\beta} |S_{ij}^{\alpha\beta}|^2$$

where

$$S_{ij}^{\alpha\beta} = \langle \phi_i^\alpha \mid \phi_j^\beta \rangle$$

The value for $\langle \mathcal{S}_{\mathrm{UHF}} \rangle$ should always be monitored when working with unrestricted wavefunctions.

6.3.2 Pople–Nesbet equations

The Hartree–Fock equations for a spin orbital is

$$\mathcal{F}\phi_i^\sigma \sigma(\omega_1) = \epsilon_i^\sigma \phi_i^\sigma \sigma(\omega_1)$$

where σ stands for a general spin function. If we multiply from the left with $\sigma(\omega_1)$ and integrate over the spin variable ω_1 we get

$$\mathcal{F}^\alpha \phi_i^\alpha = \epsilon_i^\alpha \phi_i^\alpha$$
$$\mathcal{F}^\beta \phi_i^\beta = \epsilon_i^\beta \phi_i^\beta$$

where the spatial Fock operators are defined by

$$\mathcal{F}^\alpha(\mathbf{r}_1) = \int d\omega_1 \alpha^\star(\omega_1) \mathcal{F}(\mathbf{r}_1, \omega_1) \alpha(\omega_1)$$

$$\mathcal{F}^\beta(\mathbf{r}_1) = \int d\omega_1 \beta^\star(\omega_1) \mathcal{F}(\mathbf{r}_1, \omega_1) \beta(\omega_1)$$

The operator $\mathcal{F}^\alpha(1)$ is the kinetic energy, nuclear attraction, and effective potential of an electron of α spin. The effective interactions of an electron of α spin include a Coulomb and exchange interaction with all other electrons of α spin plus only a Coulomb interaction with electrons of β spin. Thus

$$\mathcal{F}^\alpha(1) = h(1) + \sum_a^{N_\alpha} [J_a^\alpha(1) - K_a^\alpha(1)] + \sum_a^{N_\beta} J_a^\beta(1)$$

The sum over the N_α orbitals formally includes the interaction of an α electron with itself. However, since

$$[J_a^\alpha(1) - K_a^\alpha(1)] \phi_a^\alpha(1) = 0$$

this self-interaction is eliminated. The corresponding Fock operator for electrons of β spin is

$$\mathcal{F}^\beta(1) = h(1) + \sum_a^{N_\beta} \left[J_a^\beta(1) - K_a^\beta(1) \right] + \sum_a^{N_\alpha} J_a^\alpha(1)$$

The kinetic energy and nuclear attraction of an electron in one of the unrestricted orbitals ϕ_i^α or ϕ_i^β is the expectation value

$$h_{ii}^\alpha = \langle \phi_i^\alpha \mid h \mid \phi_i^\alpha \rangle \quad \text{or} \quad h_{ii}^\beta = \langle \phi_i^\beta \mid h \mid \phi_i^\beta \rangle$$

The Coulomb interaction of an electron in ϕ_i^α with one of opposite spin in ϕ_j^β is

$$J_{ij}^{\alpha\beta} = J_{ji}^{\beta\alpha} = \langle \phi_i^\alpha \mid J_j^\beta \mid \phi_i^\alpha \rangle = \langle \phi_j^\beta \mid J_i^\alpha \mid \phi_j^\beta \rangle = (\phi_i^\alpha \phi_i^\alpha, \phi_j^\beta \phi_j^\beta)$$

The corresponding Coulomb interactions between electrons of the same spin are

$$J_{ij}^{\alpha\alpha} = \langle \phi_i^\alpha \mid J_j^\alpha \mid \phi_i^\alpha \rangle = \langle \phi_j^\alpha \mid J_i^\alpha \mid \phi_j^\alpha \rangle = (\phi_i^\alpha \phi_i^\alpha, \phi_j^\alpha \phi_j^\alpha)$$

and

$$J_{ij}^{\beta\beta} = \langle \phi_i^\beta \mid J_j^\beta \mid \phi_i^\beta \rangle = \langle \phi_j^\beta \mid J_i^\beta \mid \phi_j^\beta \rangle = (\phi_i^\beta \phi_i^\beta, \phi_j^\beta \phi_j^\beta)$$

The exchange interactions between electrons of parallel spin are

$$K_{ij}^{\alpha\alpha} = \langle \phi_i^\alpha \mid K_j^\alpha \mid \phi_i^\alpha \rangle = \langle \phi_j^\alpha \mid K_i^\alpha \mid \phi_j^\alpha \rangle = (\phi_i^\alpha \phi_j^\alpha, \phi_j^\alpha \phi_i^\alpha)$$

and

$$K_{ij}^{\beta\beta} = \langle \phi_i^\beta \mid K_j^\beta \mid \phi_i^\beta \rangle = \langle \phi_j^\beta \mid K_i^\beta \mid \phi_j^\beta \rangle = (\phi_i^\beta \phi_j^\beta, \phi_j^\beta \phi_i^\beta)$$

There is no exchange interaction between electrons of opposite spin.
The total energy for a unrestricted determinantal wavefunction is

$$E_0 = \sum_a^{N_\alpha} h_{aa}^\alpha + \sum_a^{N_\beta} h_{aa}^\beta + \frac{1}{2} \sum_a^{N_\alpha} \sum_b^{N_\alpha} (J_{ab}^{\alpha\alpha} - K_{ab}^{\alpha\alpha})$$

$$+ \frac{1}{2} \sum_a^{N_\beta} \sum_b^{N_\beta} \left(J_{ab}^{\beta\beta} - K_{ab}^{\beta\beta} \right) + \sum_a^{N_\alpha} \sum_b^{N_\beta} J_{ab}^{\alpha\beta}$$

The factor of $1/2$ in the third and fourth terms eliminates the double counting in the free summation. The self-interaction disappears since $J_{aa}^{\alpha\alpha} - K_{aa}^{\alpha\alpha} = J_{aa}^{\beta\beta} - K_{aa}^{\beta\beta} = 0$.
Basis set expansion of the orbitals (LCAO)

$$\psi_i^\alpha = \sum_{\mu=1}^{K} C_{\mu i}^\alpha \chi_\mu \qquad\qquad i = 1, 2, \ldots, K$$

$$\psi_i^\beta = \sum_{\mu=1}^{K} C_{\mu i}^\beta \chi_\mu \qquad\qquad i = 1, 2, \ldots, K$$

Notice that both sets of orbitals are expanded in only one set of basis function. The two eigenvalue equations for the orbital sets guarantee that the eigenfunctions for α and β sets individually form orthogonal sets. There is no reason, however, that a member of one set need be orthogonal to a member

of the other set. Even though the two sets of spatial orbitals overlap with each other, the set of $2K$ spin orbitals will form an orthonormal set, either from spatial orthogonality ($\alpha\alpha$ and $\beta\beta$ case) or spin orthogonality ($\alpha\beta$ case). Substituting the expansions of the orbitals into the Hartree–Fock equations gives

$$\sum_\nu C^\sigma_{\nu j} \mathcal{F}^\sigma(1) \chi_\nu(1) = \epsilon^\sigma_j \sum_\nu C^\sigma_{\nu j} \chi_\nu(1)$$

If we multiply this equation by $\chi^\star_\mu(1)$ and integrate we get

$$\sum_\nu F^\sigma_{\mu\nu} C^\sigma_{\nu j} = \epsilon^\sigma_j \sum_\nu S_{\mu\nu} C^\sigma_{\nu j}$$

where $\mathbf{S}$ is the overlap matrix of the basis functions and $\mathbf{F}^\sigma$ is the matrix representation of $\mathcal{F}^\sigma$

$$F^\sigma_{\mu\nu} = \int d\mathbf{r}_1 \chi^\star_\mu(1) \mathcal{F}^\sigma(1) \chi_\nu(1)$$

Therefore we end up with two coupled matrix equations, the Pople–Nesbet equations

$$\mathbf{F}^\alpha \mathbf{C}^\alpha = \mathbf{S} \mathbf{C}^\alpha \epsilon^\alpha$$
$$\mathbf{F}^\beta \mathbf{C}^\beta = \mathbf{S} \mathbf{C}^\beta \epsilon^\beta$$

Unrestricted density matrices

The charge density of the system is

$$\rho^T(\mathbf{r}) = \sum_{a=1}^{N_\alpha} |\psi^\alpha_a(\mathbf{r})|^2 + \sum_{a=1}^{N_\beta} |\psi^\beta_a(\mathbf{r})|^2 = \rho^\alpha(\mathbf{r}) + \rho^\beta(\mathbf{r})$$

The corresponding spin density is defined by

$$\rho^S(\mathbf{r}) = \rho^\alpha(\mathbf{r}) - \rho^\beta(\mathbf{r})$$

By substituting the basis set expansion for the orbitals we get

$$\rho^\alpha(\mathbf{r}) = \sum_{a=1}^{N_\alpha} |\psi^\alpha_a(\mathbf{r})|^2 = \sum_\mu \sum_\nu P^\alpha_{\mu\nu} \chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r})$$

and

$$\rho^\beta(\mathbf{r}) = \sum_{a=1}^{N_\beta} |\psi^\beta_a(\mathbf{r})|^2 = \sum_\mu \sum_\nu P^\beta_{\mu\nu} \chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r})$$

where the α and β electron density matrices are used

$$P^\alpha_{\mu\nu} = \sum_a^{N_\alpha} C^\alpha_{\mu a}(C^\alpha_{\nu a})^\star$$

$$P^\beta_{\mu\nu} = \sum_a^{N_\beta} C^\beta_{\mu a}(C^\beta_{\nu a})^\star$$

In addition to these two density matrices we can also define a total and spin density matrix

$$\mathbf{P}^T = \mathbf{P}^\alpha + \mathbf{P}^\beta$$

$$\mathbf{P}^S = \mathbf{P}^\alpha - \mathbf{P}^\beta$$

With these definitions we can write the Fock matrices in compact form

$$F^\alpha_{\mu\nu} = h^{\text{core}}_{\mu\nu} + \sum_\lambda \sum_\sigma P^T_{\lambda\sigma}(\mu\nu, \sigma\lambda) - P^\alpha_{\lambda\sigma}(\mu\lambda, \sigma\nu)$$

$$F^\beta_{\mu\nu} = h^{\text{core}}_{\mu\nu} + \sum_\lambda \sum_\sigma P^T_{\lambda\sigma}(\mu\nu, \sigma\lambda) - P^\beta_{\lambda\sigma}(\mu\lambda, \sigma\nu)$$

and the total energy is

$$E_0 = \frac{1}{2} \sum_\mu \sum_\nu \left[P^T_{\mu\nu} h^{\text{core}}_{\mu\nu} + P^\alpha_{\mu\nu} F^\alpha_{\mu\nu} + P^\beta_{\mu\nu} F^\beta_{\mu\nu} \right]$$

Lecture 7

Correlation Energy

7.1 Basics

For many systems the Hartree–Fock method accounts for about 99 % of the total energy. However, often the remaining 1 % are important for physical and chemical properies of the system.

Correlation:

Reduced probability to find an electron with different spin close to another electron: Coulomb hole.

Reduced probability to find an electron with the same spin close to another electron: Fermi hole.

Hartree–Fock: Variational optimal one determinant wave function.

Improvements are only possible with multi determinant wave functions.

$$\Psi = a_0 \Phi_{\mathrm{HF}} + \sum_i a_i \Phi_i$$

Excited Slater determinants:

Hartree–Fock determinant

$$\Phi_{\mathrm{HF}} = |\Phi_1 \bar{\Phi}_1 \cdots \Phi_{N/2} \bar{\Phi}_{N/2}|$$

Not used MOs (virtual, unoccupied MOs)

$$\Phi_{N/2+1}, \cdots, \Phi_K$$

Produce new determinants by exchange of MOs in Φ_{HF} with one, two, three, ... virtual MOs. Singly, doubly, triply, ... excited determinants.

$$\Phi_0, \Phi_a^r, \Phi_{ab}^{rs}, \Phi_{abc}^{rst}, \ldots$$

This procedure generates determinants with all possible eigenvalues of $\hat{S}_z$, linear combination of these determinants that are also eigenfunctions of $\hat{S}^2$ are called CSF (Configuration State Functions).
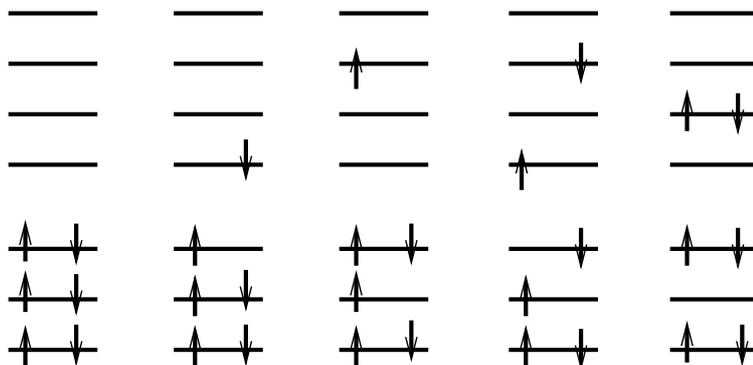
The complete optimization problem to find an exact wave function with this method is two dimensional. Enlargement of the one-electron basis to get a converged Hartree–Fock wave function also increases the number of virtual MOs. This means an increased number of excited determinants. The complete solution for a given basis set (with K functions) is called Full Configuration Interaction (FCI).

86
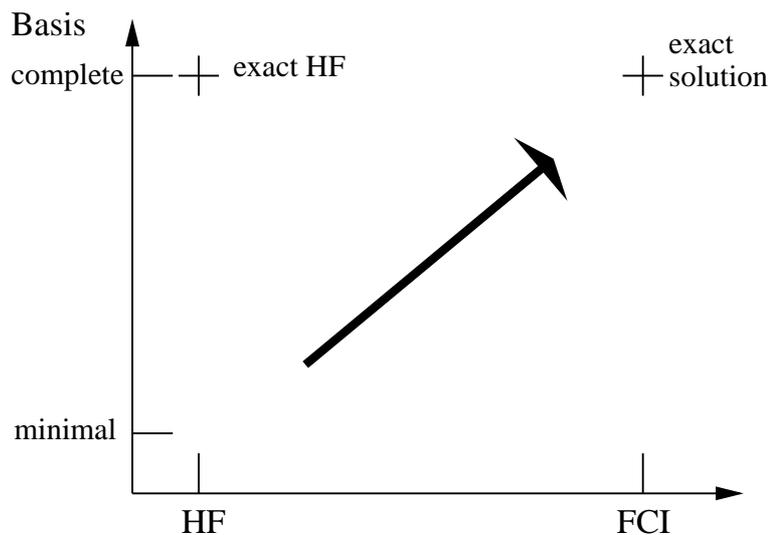
7.2 Matrixelements

Examples of determinants with two orbitals

$$\Phi_0 : \qquad \frac{1}{\sqrt{2}}[a(1)b(2) - b(1)a(2)]$$

$$\Phi_a^r : \qquad \frac{1}{\sqrt{2}}[r(1)b(2) - b(1)r(2)]$$

$$\Phi_{ab}^{rs} : \qquad \frac{1}{\sqrt{2}}[r(1)s(2) - s(1)r(2)]$$

$$\langle \Phi_0 \mid \hat{O} \mid \Phi_0 \rangle = \frac{1}{2}\langle a(1)b(2) - b(1)a(2) \mid \hat{O} \mid a(1)b(2) - b(1)a(2) \rangle$$

$$= \frac{1}{2}\langle a(1)b(2) \mid \hat{O} \mid a(1)b(2) \rangle - \frac{1}{2}\langle b(1)a(2) \mid \hat{O} \mid b(1)a(2) \rangle$$

$$- \frac{1}{2}\langle a(1)b(2) \mid \hat{O} \mid b(1)a(2) \rangle + \frac{1}{2}\langle b(1)a(2) \mid \hat{O} \mid a(1)b(2) \rangle$$

Operators:

$$\hat{O}_0 = 1$$

$$\hat{O}_1 = \sum_{i=1}^{N} \hat{h}(i)$$

$$\hat{O}_2 = \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{e^2}{r_{ij}}$$

Case 1:

$$\langle \Phi | \Phi \rangle = 1$$

$$\langle \Phi | \hat{O}_1 | \Phi \rangle = \sum_i \langle i | \hat{h} | i \rangle$$

$$\langle \Phi | \hat{O}_2 | \Phi \rangle = \frac{1}{2} \sum_i \sum_j (ii|jj) - (ij|ji)$$

Case 2:

$$\langle \Phi | \Phi_a^r \rangle = 0$$

$$\langle \Phi | \hat{O}_1 | \Phi_a^r \rangle = \langle a | \hat{h} | r \rangle$$

$$\langle \Phi | \hat{O}_2 | \Phi_a^r \rangle = \sum_j (ar|jj) - (aj|jr)$$

Case 3:

$$\langle\Phi|\hat{O}_1|\Phi_{ab}^{rs}\rangle = 0$$

$$\langle\Phi|\hat{O}_2|\Phi_{ab}^{rs}\rangle = (ar|bs) - (as|br)$$

Case 4:

$$\langle\Phi|\hat{O}_2|\Phi_{abc}^{rst}\rangle = 0$$

Energies of determinants:

$$\langle\Phi \mid \hat{H} \mid \Phi\rangle = \sum_i H_{ii} + E_J + E_K$$

$h_i i$: contribution from each electron

E_J : contribution from each unique electron pair

E_K : contribution from each unique electron pair with same spin

Examples
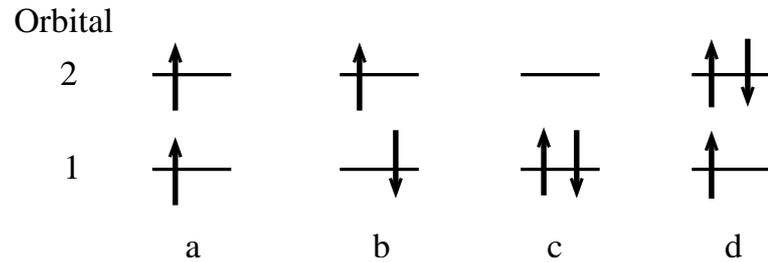
a) $h_{11} + h_{22} + J_{12} - K_{12}$
b) $h_{11} + h_{22} + J_{12}$
c) $2h_{11} + J_{11}$
d) $h_{11} + 2h_{22} + 2J_{12} + J_{22} - K_{12}$

7.3 Configuration interaction method (CI)

Ansatz

$$\Psi_{\mathrm{CI}} = a_0\Phi_{\mathrm{HF}} + \sum_S a_S\Phi_s + \sum_D a_D\Phi_D + \cdots$$

Variational calculation with normalisation constraint

$$\langle\Psi_{\mathrm{CI}}|\Psi_{\mathrm{CI}}\rangle = 1$$

88

Lagrange function

$$\mathcal{L} = \langle \Psi_{\mathrm{CI}}|\hat{H}|\Psi_{\mathrm{CI}}\rangle - \lambda\left[\langle\Psi_{\mathrm{CI}}|\Psi_{\mathrm{CI}}\rangle - 1\right]$$

$$\langle\Psi_{\mathrm{CI}}|\hat{H}|\Psi_{\mathrm{CI}}\rangle = \sum_{i=0}\sum_{j=0} a_i a_j \langle\Phi_i|\hat{H}|\Phi_j\rangle$$
$$= \sum_{i=0} a_i^2 E_i + \sum_{i=0}\sum_{j\neq j} a_i a_j \langle\Phi_i|\hat{H}|\Phi_j\rangle$$

$$\langle\Psi_{\mathrm{CI}}|\Psi_{\mathrm{CI}}\rangle = \sum_{i=0}\sum_{j=0} a_i a_j \langle\Phi_i|\Phi_j\rangle$$
$$= \sum_{i=0} a_i^2 \langle\Phi_i|\Phi_i\rangle$$
$$= \sum_{i=0} a_i^2$$

Variation:

$$\frac{\partial\mathcal{L}}{\partial a_i} = 2\sum_j a_j \langle\Phi_i|\hat{H}|\Phi_j\rangle - 2\lambda a_i = 0$$

$$a_i\left(\langle\Phi_i|\hat{H}|\Phi_j\rangle - \lambda\right) + \sum_{j\neq i} a_j \langle\Phi_i|\hat{H}|\Phi_j\rangle = 0$$

$$a_i\left(E_i - \lambda\right) + \sum_{j\neq i} a_j \langle\Phi_i|\hat{H}|\Phi_j\rangle = 0$$

Multiplication of the last equation with a_i and summation over all equations leads to

$$\sum_i a_i^2 E_i - \sum_i a_i^2 \lambda + \sum_i \sum_{j\neq i} a_i a_j \langle\Phi_i|\hat{H}|\Phi_j\rangle = 0$$

and therefore

$$\lambda = \langle\Psi_{\mathrm{CI}}|\hat{H}|\Psi_{\mathrm{CI}}\rangle \ .$$

The Lagrange multiplier is the CI energy ($\lambda = E_{\mathrm{CI}}$). We have a CI eigenvalue equation in matrix form

$$\mathbf{H}\mathbf{a} = E\mathbf{a}$$

Size of the CI matrix:
N electorns (indistinguishable) in K MOs.

$$\text{Number of Determinants} = \binom{K}{N} = \frac{K!}{N!(K-N)!}$$

89

Many determinants have a wrong (not desired) M_S and can be neglected. In addition the CI matrix is reduced in the CSF basis

$$\text{Number of CSF(Singlet)} = \frac{K!(K+1)!}{(N/2)!(N/2+1)!(K-N/2)!(K-N/2+1)!}$$

Examplel:
H_2O 10 electrons in 38 AOs (= 38 MOs)

Excitation	CSF	total CSF
0	1	1
1	71	72
2	2485	2557
3	40040	42597
4	348530	391127
5	1723540	2114667
6	5033210	7147877
7	8688080	15836557
8	8653645	24490202
9	4554550	29044752
10	1002001	30046752

Sizes:
CI-vector $\approx 30 \cdot 10^6$ elements ($\approx$ 240 Mbytes)
CI-matrix $\approx (30 \cdot 10^6)^2 = 900 \cdot 10^{12}$ elements ($\approx$ 7200 Tbytes)
but most $H_{ij} = 0$, number of non-zero matrix elements is proportional to the matrix size.

Solution of the CI eigenvalue equation:
Only lowest eigenvalues are of interest.
Iterative diagonalization for some of the lowest eigenvalues

- Lanczos method

- Davidson method

- Residuum methode

The CI matrix is sparely occupied. The number of matrix elements $H_{ij} \neq 0$ is proportional to the dimension of the matrix. **H** is not explicitly calculated but only **Ha**.

Reduced CI methods:

Idea: Not all determinants are equally important.

Ansatz: Only allow excitations from a subset of MOs into a subset of virtual MOs (active space).
Allow only a maximal number of excitations.

Ne-Atom: Weight of CSF

Excitation	Weight
0	0.9644945573
1	0.0009864929
2	0.0336865893
3	0.0003662339
4	0.0004517826
5	0.0000185090
6	0.0000017447
7	0.0000001393
8	0.0000000011

Active Space: restriction of orbitals for excitations

active space
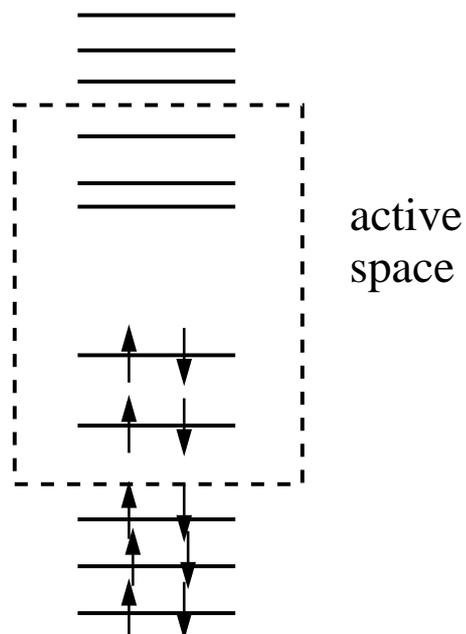
7.4 Dynamical and non-dynamical electron correlation

For the majority of molecules, for example all moleculaes that can be assigned a single Lewis structure, the main error from in the Hartree–Fock approximation comes from ignoring the correlated motion of each electron with all the other electrons. This is called dynamical correlation because it refers to the dynamical character of the electron-electron interactions. Even though the correlation energy may be large, the Hartree–Fock determinant dominates the configuration interaction expansion in these cases. The correction is made up from many small contributions from excited state determinants. In certain cases, however, it happens that a set of nearly degenerate orbitals occure in a molecule that can only be partially occupied. In a restricted Hartree–Fock calculation one or several of these orbitals cannot be included into the wavefunction. An equally good wavefunction could be generated by including these orbitals and exculding the others. In a subsequent configuration interaction calculation we expect that these determinants will have similar weights. The correlation energy gained for these cases is called non-dynamical correlation.

Lecture 8

Coupled Cluster Approaches

8.1   Size Consistency
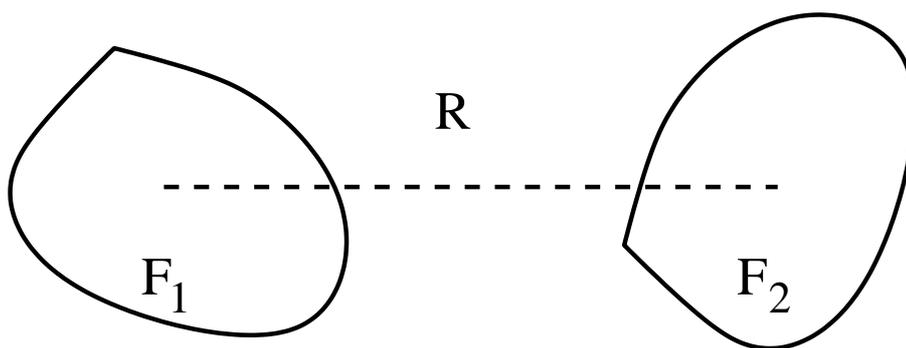
Method X is size consistent if

$$E_X(R = \infty) = 2E_X(F)$$

Example: Hartree-Fock

$$\Psi_{r=\infty} = \Psi_{F_1} \cdot \Psi_{F_2}$$

therefore $E(R = \infty) = E(F_1) + E(F_2) = 2E(F)$

Example: SDCI

Minimal basis set $(H_2)_2$

93
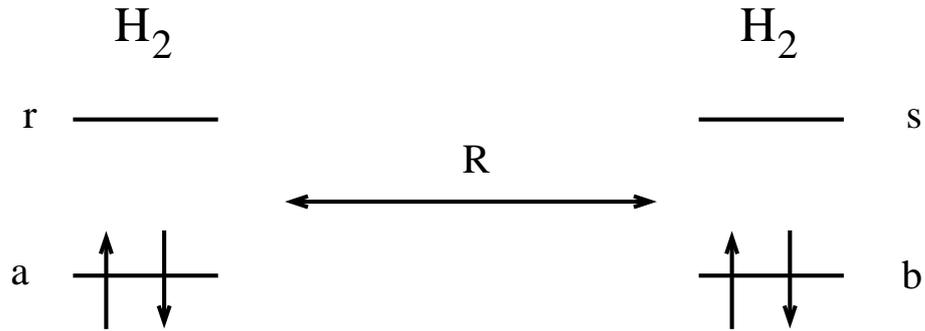
Determinants for the full calculation:

HF	$a\bar{a}b\bar{b}$
singly	$a\bar{r}b\bar{b}$
singly	$a\bar{a}b\bar{r}$
CT	$a\bar{s}b\bar{b}$
CT	$a\bar{a}b\bar{r}$
doubly	$r\bar{r}b\bar{b}$
doubly	$a\bar{a}s\bar{s}$
CT	$a\bar{s}b\bar{r}$

For $R = \infty$ the CT determinants are zero and we are left with 5 determinants.

For the calculation of the separated systems we have three determinants per system.

HF	$a\bar{a}$
singly	$a\bar{r}$
doubly	$r\bar{r}$
HF	$b\bar{b}$
singly	$b\bar{s}$
doubly	$s\bar{s}$

The total wavefunction will be a linear combination of all the products. We get 9 determinants, including triply and quadruply excited ones. The energy of these two calculations will certainly differ. Truncated CI methods are not size consistent.

$$2E(F) < E(R = \infty)$$

The bigger the system the smaller the percentage of correlation energy is included in a truncated CI calculation.

$$E_{\text{Corr}}(SDCI)/E_{\text{Corr}}(FCI) \longrightarrow 0 \qquad \text{for} \quad N_{\text{el}} \to \infty$$

A related concept is size extensivity. A method is size extensive if

$$E_{\text{Corr}}(F_N) \propto N \cdot E_{\text{Corr}}(F)$$

94

8.2 Coupled Cluster Methods

Perturbation theory based methods add all corrections (singly, doubly, triply
... excitations) up to a given order to the reference function. In Coupled
Cluster (CC) methods for a given type of correction all order will be included.
Ansatz for the CC wave function

$$\Psi_{CC} = e^{\hat{T}} \Phi_0$$

$$e^{\hat{T}} = \sum_{k=0}^{\infty} \frac{1}{k!} \hat{T}^k$$

The operator $\hat{T}$ is defined as

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots + \hat{T}_N \ ,$$

where the operator $\hat{T}_i$ applied to the Hartree–Fock wave function generates
all i-th excited Slater determinants.

$$\hat{T}_1 \Phi_0 = \sum_a^{occ} \sum_r^{vir} t_a^r \Phi_a^r$$

$$\hat{T}_2 \Phi_0 = \sum_{a<b}^{occ} \sum_{r<s}^{vir} t_{ab}^{rs} \Phi_{ab}^{rs}$$

The expansion coefficients t are called amplitudes. They are our new varia-
tional parameters.

$$e^{\hat{T}} = 1 + \hat{T}_1 + (\hat{T}_2 + \frac{1}{2}\hat{T}_1^2) + (\hat{T}_3 + \hat{T}_2\hat{T}_1 + \frac{1}{6}\hat{T}_1^3)$$
$$+ (\hat{T}_4 + \hat{T}_3\hat{T}_1 + \frac{1}{2}\hat{T}_2^2 + \frac{1}{2}\hat{T}_2\hat{T}_1^2 + \frac{1}{24}\hat{T}_4^4) + \cdots$$

Insertion into the Schrödinger equation:

$$\hat{H}e^{\hat{T}}\Phi_0 = Ee^{\hat{T}}\Phi_0$$

multiplication from left with Φ_0^* and integration

$$\langle \Phi_0 | \hat{H} e^{\hat{T}} \Phi_0 \rangle = E_{CC} \langle \Phi_0 | e^{\hat{T}} \Phi_0 \rangle$$
$$= E_{CC} \langle \Phi_0 | (1 + \hat{T}_1 + \hat{T}_2 + \cdots) \Phi_0 \rangle$$
$$= E_{CC}$$

We make use of the fact that the Hamilton operator only contains one- and two-electron terms:

$$\begin{aligned}
E_{\text{CC}} &= \langle \Phi_0 | \hat{H}(1 + \hat{T}_1 + \hat{T}_2 + \frac{1}{2}\hat{T}_1^2)\Phi_0 \rangle \\
&= \langle \Phi_0 | \hat{H}\Phi_0 \rangle + \langle \Phi_0 | \hat{H}\hat{T}_1\Phi_0 \rangle \\
&\quad + \langle \Phi_0 | \hat{H}\hat{T}_2\Phi_0 \rangle + \frac{1}{2}\langle \Phi_0 | \hat{H}\hat{T}_1^2\Phi_0 \rangle \\
&= E_0 + \sum_a^{occ}\sum_r^{vir} t_a^r \langle \Phi_0 | \hat{H}\Phi_a^r \rangle \\
&\quad + \sum_{a<b}^{occ}\sum_{r<s}^{vir} (t_{ab}^{rs} + t_a^r t_b^s - t_a^s t_b^r)\,\langle \Phi_0 | \hat{H}\Phi_{ab}^{rs} \rangle
\end{aligned}$$

For $\Phi_0 = \Phi_{\text{HF}}$ we have $\langle \Phi_0 | \hat{H}\Phi_a^r \rangle = 0$ (Brillouins theorem) and the second integral can easily be written with MOs:

$$E_{\text{CC}} = E_0 + \sum_{a<b}^{occ}\sum_{r<s}^{vir} (t_{ab}^{rs} + t_a^r t_b^s - t_a^s t_b^r)\,[(ar|bs) - (as|br)]$$

The CC energy is completely determined by the Hartree–Fock energy and the one and two electron amplitudes.

Coupled (nonlinear) equations for the amplitude can be generated from the Schrödinger equation by multiplication from left with $\Phi_a^{r*}, \Phi_{ab}^{rs*}, \ldots$ and subsequent integration.

$$\begin{aligned}
\langle \Phi_a^r | &\longrightarrow & N_{occ} \cdot N_{virt} \text{ equations} \\
\langle \Phi_{ab}^{rs} | &\longrightarrow & N_{occ}^2 \cdot N_{virt}^2 \text{ equations} \\
\langle \Phi_{abc}^{rst} | &\longrightarrow & \cdots
\end{aligned}$$

8.3 Approximate Coupled Cluster methods

For the exact CC method all operators $\hat{T}_i(i \leq N)$ have to be used.

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots \hat{T}_N$$

The method scales like M^{2N+2} for M basis functions and N electrons. The only practical method is to use approximations of the form $\hat{T} \approx \hat{T}_1 + \hat{T}_2$ (CCSD).

$\hat{T} = \hat{T}_1 + \hat{T}_2$	CCSD	M^6
$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3$	CCSDT	M^8
$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \hat{T}_4$	CCSDTQ	M^{10}

The case $\hat{T} = \hat{T}_1$ does not improve on the Hartree–Fock energy. Multiplication of the Schrödinger equation from left with $\Phi_0^*, \Phi_a^{r*}, \Phi_{ab}^{rs*}$ and integration leads for CCSD to a system of coupled, non-linear equations for the amplitudes t_a^r und t_{ab}^{rs}. These equations have to be solved iteratively. The most common improvement over the CCSD method is achieved by a perturbative correction for $\hat{T}_3$. Inclusion of $\hat{T}_3$ with perturbation theory leads to the CCSD(T) method with M^7 scaling.

$$e^{\hat{T}_1 + \hat{T}_2} = 1 + \hat{T}_1 + \left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2 \right)$$
$$+ \left(\hat{T}_2\hat{T}_1 + \frac{1}{6}\hat{T}_1^3 \right) + \left(\frac{1}{2}\hat{T}_2^2 + \frac{1}{2}\hat{T}_2\hat{T}_1^2 + \frac{1}{24}T_1^4 \right) + \cdots$$

CCSD equations

$$E_{CCSD} = E_0 + \langle \Phi_0 \mid \hat{H} \left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2 \right) \Phi_0 \rangle$$

$$\langle \Phi_a^r \mid \hat{H} \left(1 + \hat{T}_1 + \left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2 \right) + \left(\hat{T}_2\hat{T}_1 + \frac{1}{6}\hat{T}_1^3 \right) \right) \Phi_0 \rangle = E_{CCSD} \langle \Phi_a^r \mid \hat{T}_1 \Phi_0 \rangle$$

$$\langle \Phi_{ab}^{rs} \mid \hat{H} \left(\cdots \right) \Phi_0 \rangle = E_{CCSD} \langle \Phi_{ab}^{rs} \mid \left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2 \right) \Phi_0 \rangle$$

CCSD is non-variational. CCSD is size-consistent; higher excitations are included through products $(t_{ab}^{rs} t_{cd}^{tu})$.

Table 8.1: H_2O at the basis set limit

	Exp.	CCSD(T)	HF
r_{OH} [pm]	95.72	95.64	94.0
α(HOH)	104.52	104.2	106.2
μ_e [D]	1.8473	1.853	1.936
Atomization energy [kJ/mol]	975.3	975.5	652.3

Table 8.2: Basis set error: Statistics over many small molecules

Basis	HF	CCSD - HF	CCSD(T) - CCSD
cc-pVDZ	5.8	23.8	1.3
cc-pVTZ	3.9	6.3	0.5
cc-pVQZ	1.5	1.5	0.2
cc-pV5Z	0.2	0.5	0.0
cc-pV6Z	0.0	0.4	0.0

8.4 Accuracy in Coupled Cluster calculations
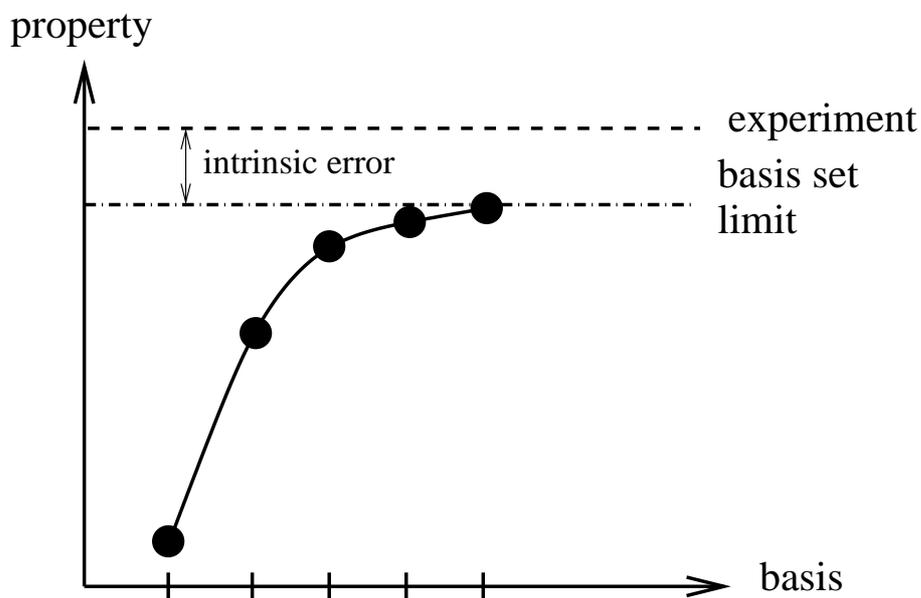
Lecture 9

Møller–Plesset Perturbation Theory

9.1 Electron correlation with perturbation theory

Problem: What is $\hat{H}^{(0)}$ and what is $\hat{H}^{(1)}$?

Perturbation theory
$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)}$$

Basis $\Psi_n^{(0)}$
$$\hat{H}^{(0)}\Psi_n^{(0)} = E_n^{(0)}\Psi_n^{(0)}$$

First order energy correction to ground state $\Psi_g^{(0)}$
$$E_g^{(1)} = \langle \Psi_g^{(0)} \mid \hat{H}^{(1)} \mid \Psi_g^{(0)} \rangle$$

Second order energy correction
$$E_g^{(2)} = \langle \Psi_g^{(0)} \mid \hat{H}^{(1)} \mid \Psi_g^{(1)} \rangle$$

and first order correction to wavefunction
$$\Psi_n^{(1)} = \sum_{i \neq g} \frac{H_{ig}^{(1)}}{E_g^{(0)} - E_i^{(0)}} \Psi_i^{(0)}$$

and we have
$$E_g^{(2)} = \sum_{i \neq g} \frac{|H_{ig}^{(1)}|^2}{E_g^{(0)} - E_i^{(0)}}$$

with
$$H_{ig}^{(1)} = \langle \Psi_i^{(0)} \mid \hat{H}^{(1)} \mid \Psi_j^{(0)} \rangle$$

Definition of Hamiltonian for the unperturbed system

$$\hat{H}^{(0)} = \sum_i \hat{F}(i)$$

The unperturbed Hamiltonian is chosen to be the sum of all Fock-operators of all electrons of the system. This is a sum of one-electron operators without coupling. Therefore the solution will be the sum of all energies of the independent systems and the wavefunction will be the product of the one-electron wavefunctions (orbitals). Because of the Pauli-principle we have to anti-symmetrize this product and get a determinantal wavefunction.

$$\Psi_g^{(0)} = |\Phi_1 \Phi_2 \cdots \Phi_N| = \Phi_{HF}$$

The complete basis of all zeroth-order solution are all possible determinants generated from the orbitals.

$$\Psi_i^{(0)} = |\Phi_1 \Phi_2 \cdots \Phi_r \cdots \Phi_N| = \Phi_a^r$$

and all other excited determinants (Φ_{ab}^{rs}, Φ_{abc}^{rst}, ...).
Energy

$$E_g^{(0)} = \sum_i \epsilon_i \neq E_{HF}$$

The energy of the other determinants (basis functions) is also just the sum of the orbital energies.

$$E_i^{(0)} = \epsilon_1 + \epsilon_2 + \epsilon_3 + \cdots$$

Perturbation Hamiltonian

$$\hat{H}^{(0)} = \sum_i \hat{F}(i)$$
$$\hat{H}^{(1)} = \hat{H} - \hat{H}^{(0)} = \sum_i \hat{h}(i) + \hat{V}_{\text{ee}} - \sum_i \hat{F}(i)$$
$$= \hat{V}_{\text{ee}} - \sum_j (\hat{J}_j - \hat{K}_j)$$

Møller–Plesset perturbation theory

$$\text{MP0} = E(\text{MP0}) = \langle\Phi_{\text{HF}}|\hat{H}^{(0)}|\Phi_{\text{HF}}\rangle = \sum_i \epsilon_i$$

$$\text{MP1} = E(\text{MP0}) + E(\text{MP1}) = \sum_i \epsilon_i + \langle\Phi_{\text{HF}}|\hat{H}^{(1)}|\Phi_{\text{HF}}\rangle$$

$$= \sum_i \epsilon_i + \frac{1}{2}\sum_{ij}\langle\Phi_{\text{HF}}|\frac{e^2}{r_{ij}}|\Phi_{\text{HF}}\rangle - \sum_{ij}\left((ii|jj) - (ij|ij)\right)$$

$$= \sum_i \epsilon_i + \frac{1}{2}\sum_{ij}\left(J_{ij} - K_{ij}\right) - \sum_{ij}\left(J_{ij} - K_{ij}\right)$$

$$= \sum_i \epsilon_i - \frac{1}{2}\sum_{ij}\left(J_{ij} - K_{ij}\right)$$

$$= E(HF)$$
$$\text{MP2} = E(\text{HF}) + E(\text{MP2})$$

From general perturbation theory:

$$E(\text{MP2}) = \sum_{i\neq 0}\frac{\langle\Phi_0|\hat{H}^{(1)}|\Phi_i\rangle\langle\Phi_i|\hat{H}^{(1)}|\Phi_0\rangle}{E_0 - E_i}$$

With

$$\langle\Phi_0|\hat{H}^{(1)}|\Phi_a^r\rangle = 0 \qquad \text{(Brillouin Theorem)}$$
$$\langle\Phi_0|\hat{H}^{(1)}|\Phi_{ab}^{rs}\rangle = (ab|rs) - (ar|bs)$$
$$\langle\Phi_0|\hat{H}^{(1)}|\Phi_{abc}^{rst}\rangle = 0$$

and the energy denominator

$$E_0 - E(\Phi_{ab}^{rs}) = \epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s$$

we get

$$E(\text{MP2}) = \sum_{a<b}^{occ}\sum_{r<s}^{vir}\frac{[(ab|rs) - (ar|bs)]^2}{\epsilon_a + \epsilon_b - \epsilon_r - \epsilon_s}$$

MP perturbation theory is size-consistent and size-extensive.

Order	Scaling	% correlation
1	M^4	0
2	M^5	≈ 90
3	M^6	≈ 95
4	M^7	≈ 98
5	M^8	-
6	M^9	-

Table 9.1: Water, DZ basis set, RHF calculation

Method	r_e	$1.5r_e$	$1.5r_e$
MP2	92.9	86.9	80.6
MP3	95.5	88.0	78.5
MP4	99.2	96.8	94.0
MP5	99.6	97.5	95.4
MP6	99.9	99.3	99.6
MP7	99.95	99.5	99.9
MP8	99.98	99.9	101.3
FCI	-0.130085	-0.191996	-0.299864

MP2	efficient ($\approx$ HF) implementation possible
	recovers $\approx$ 90 % of correlation energy
MP3	often not more accurate than MP2
	$\rightarrow$ not a standard method
MP4	improves on MP2
	similar cost, but inferior to CCSD(T)

MP5$\cdots$ only for benchmarking

Only MP2 is a commonly used method in quantum chemistry.

9.2   Convergence of MP series

No smooth convergence, or no convergence at all is possible.

Detection of difficult cases: HF determinant is not a good starting point. Weight of Φ_0 in perturbation wavefunction is less than 90 %. Example: H_2O at a $2.0r_e$ geometry. The HF determinant has only a 58% weight in MP2. The wavefunction has multi-configuration character.

MP based on unrestricted Hartree–Fock

- Slow convergence due to spin contamination problems

- To be used with care! Better use spin-projection methods.

MP based on multi-configurational wavefunctions (e.g. CAS-PT2)

- expensive

- often very accurate

- needs expert knowledge

9.3 The dissociation problem

Restricted vs. unrestricted solutions.

Example: $H_2 \longrightarrow H + H$, homolytic bond breaking
Minimal basis Φ_1, Φ_2 with overlap $S_{12} = \langle \Phi_1 \mid \Phi_2 \rangle$

$$\Psi_1 = [2(1 + S_{12})]^{-1/2} (\Phi_1 + \Phi_2)$$
$$\Psi_2 = [2(1 - S_{12})]^{-1/2} (\Phi_1 - \Phi_2)$$

Symmetry therefore fully defines the restricted solution $\Psi_0 = |\Psi_1 \bar{\Psi}_1|$.
Unrestricted case

$$\Psi_1^\alpha = \cos \Theta \ \Psi_1 + \sin \Theta \ \Psi_2$$
$$\Psi_1^\beta = \cos \Theta \ \Psi_1 - \sin \Theta \ \Psi_2$$

if $\Theta = 0$ then RHF=UHF.

$\Psi^{\alpha} = \Psi^{\beta}$

small r

Ψ^{α} Ψ^{β}

large r

Energy

RHF

UHF

Instability point
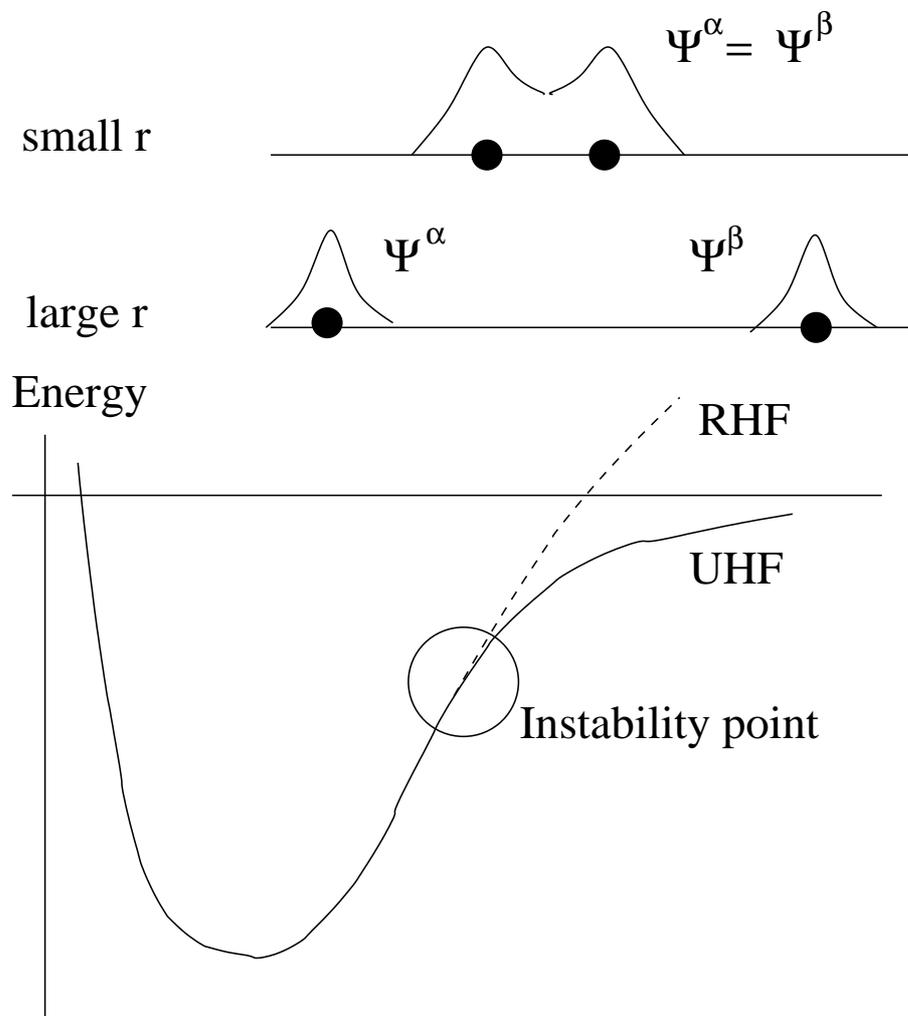
Bond breaking (homolytic)

- restricted wavefunctions fail

- unrestricted wavefunctions may have spin-contamination

- MP2 is highly affected

- CCSD(T) might solve problems partially

Ultimate solution: Multi-configurational SCF (MC-SCF)

$$|\Psi_0\rangle = a_0\Phi_0 + a_1\Phi_1 + \cdots$$

The wavefunction is approximated by a small CI expansion. Variationally optimize orbitals and CI-coefficients at the same time.

Lecture 10

Density Functional Theory: Part I

10.1 Basics

In the following we investigate N-electron systems within the Born–Oppenheimer approximation. The system is described by a Hamilton operator $\hat{H}$

$$\hat{H} = \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{ee}} \ ,$$

where $\hat{T}$ is the operator of the kinetic energy, $\hat{V}_{\text{ee}}$ the electron-electron interaction, and $\hat{V}_{\text{ext}}$ the external potential.

10.1.1 Hohenberg–Kohn Theorems

Theorem 1
The external potential $\hat{V}_{\text{ext}}$ is, except for a trivial constant, uniquely defined by the electron density $\rho(\vec{r})$.

Proof:
Assumption: There are two potentials $\hat{V}$ and $\hat{V}'$ compatible with the electron density $\rho(\vec{r})$.

$$1) \qquad \rho \longrightarrow \hat{V} \longrightarrow \hat{H} \longrightarrow \Psi$$

$$2) \qquad \rho \longrightarrow \hat{V}' \longrightarrow \hat{H}' \longrightarrow \Psi'$$

Using the variation principle for wavefunctions we get

$$
\begin{aligned}
E_0 < \langle \Psi'|\hat{H}|\Psi'\rangle \ &= \ \langle \Psi'|\hat{H}'|\Psi'\rangle + \langle \Psi'|\hat{H} - \hat{H}'|\Psi'\rangle \\
&= \ E_0' + \int \rho(r)\left(\hat{V}(r) - \hat{V}'(r)\right) dr
\end{aligned}
$$

and

$$E_0' < \langle \Psi | \hat{H}' | \Psi \rangle = \langle \Psi | \hat{H} | \Psi \rangle + \langle \Psi | \hat{H}' - \hat{H} | \Psi \rangle$$
$$= E_0 - \int \rho(r) \left(\hat{V}(r) - \hat{V}'(r) \right) dr \ .$$

We have used $|\Psi\rangle$ and $|\Psi'\rangle$ as test functions with $\hat{H}'$ and $\hat{H}$ and have made use of the variation principle. After adding up the two equations we get the following contradiction

$$E_0 + E_0' < E_0 + E_0' \ ,$$

and therefore our initial assumption must be wrong. Since we have used the variation principle this proof holds only for the ground state.

Total energy

$$E_0 = E_V(\rho_0) = T[\rho_0] + \int V_{\text{ext}}(r)\rho_0(r)dr + J[\rho_0] + E_{\text{NC}}[\rho_0]$$

Classical Coulomb-energy:

$$J[\rho_0] = \frac{1}{2} \int\int \frac{\rho_0(r_1)\rho_0(r_2)}{|r_1 - r_2|} \, dr_1 dr_2$$

Kinetic energy: $T[\rho_0]$
Non-classical electron-electron interaction energy: $E_{\text{NC}}[\rho_0]$

Only V_{ext} directly depends on the system, therefore we can write

$$E_0 = \int V_{\text{ext}}(r)\rho_0(r)dr + F_{\text{HK}}[\rho_0]$$

where F_{HK} is a universal functional of the electron density:

$$F_{\text{HK}}[\rho] = T[\rho] + J[\rho] + E_{\text{NC}}[\rho] \ .$$

Theorem 2:
For a trial density $\tilde{\rho}(r)$, with the properties $\tilde{\rho}(r) \geq 0$ and $\int \tilde{\rho}(r)dr = N$ (number of electrons) we always have $E_0 \leq E_V(\tilde{\rho})$.

Proof:

$$\tilde{\rho} \xrightarrow[HK1]{} V' \to \hat{H}' \to \Psi'$$

106

We use $|\Psi'\rangle$ as a trial function with the Hamilton operator $\hat{H}$. We have (variation principal):

$$E_0 \leq \langle \Psi' | \hat{H} | \Psi' \rangle = E_V[\tilde{\rho}]$$

and therefore

$$E_0 \leq E_V[\tilde{\rho}] \ .$$

This proof is based on the assumption that $\tilde{\rho} \to V'$ is valid. But not all $\tilde{\rho}(r) \geq 0$ with $\int \tilde{\rho}(r)dr = N$ have a corresponding potential.

The set of all electron densities with a corresponding potential are called v-representable.

The variation principle

$$E_0 = \text{Min}_{\tilde{\rho}}\{E_V[\tilde{\rho}]\}$$

is only valid for v-representable electron densities.

The proof can be expanded to N-representable densities (Levi constraint search method) A N-representable density is defined by:

- calculated from an anti-symmetric wavefunction

- fulfils

$$\rho(r) \geq 0; \quad \int \rho(r)dr = N; \quad \int |\nabla \rho^{1/2}|^2 dr < \infty$$

Plausibility arguments by Wilson

1. Starting from $\rho(r)$

2. Calculate the number of electrons N in the system

$$N = \int \rho(r)dr$$

3. Find the position and charge of the nuclei (cusp-condition)

$$\frac{\partial}{\partial r_A}\bar{\rho}(r_A)\bigg|_{r_A=0} = -2Z_A\bar{\rho}(0) \ .$$

where $\bar{\rho}(r_A)$ is the over all space directions averaged density.
This defines the external potential $V_{\text{ext}}(r) = \sum_A \frac{Z_A e^2}{|r-R_A|}$

4. N and V_{ext} define the Hamilton operator.

5. With $\hat{H}$ we know Ψ and therfore all properties of the system can be calculated.

The electron density $\rho(r)$ defines all properties of an N-electron system.
The Hohenberg-Kohn theorems give an existence proof for the functionals
$T[\rho]$ and $E_{\mathrm{NC}}[\rho]$, as well as a variation principle for the case of the known
functionals. What is missing is the concrete form of the functionals. Many
approximate functionals have been proposed. The accuracy, especially for
the kinetic energy is for applications in chemistry not sufficient. In real
application we will use another method.

10.1.2 Kohn–Sham Theory

We discuss a system with N non-interacting electrons in a local potential $\hat{V}_s$:

$$\hat{H}_s = \hat{T} + \hat{V}_s \ .$$

From the 1. Hohenberg–Kohn theorem we get

$$E_s[\rho] = T_s[\rho] + \int V_s(r)\rho(r)dr$$

and with the 2. Hohenberg–Kohn theorem it follows that

$$\delta E_s[\rho_s] = 0 \ ,$$

This allows us to calculate the ground state density ρ_s.
The solution of the Schrödinger equation

$$\hat{H}_s \Psi_s = E_s \Psi_s$$

is a one-determinant wave function. This is due to the non-interacting elec-
trons that allow for a simple product wavefunction for the solution. At the
same time we also want the Pauli-principle to hold

$$\Psi_s = |\Phi_1(r) \cdots \Phi_N(r)|$$

The one-particle density matrix for this system is

$$\rho_1^s(r, r') = \sum_{i=1}^{N} \Phi_i(r)\Phi_i^*(r') \ ,$$

the density

$$\rho_s(r) = \sum_{i=1}^{N} |\Phi_i(r)|^2 \ ,$$

and the kinetic energy

$$T_s[\rho] = \int \Phi_i^*(r) \left(-\frac{1}{2}\nabla^2 \right) \Phi_i(r) dr \ .$$

The functions Φ_i are orthonormal. With this is the functional $T_s[\rho]$ for all densities of the form

$$\rho(r) = \sum_{i=1}^{N} |\Phi_i(r)|^2$$

defined.

Let's now turn to a system of N interacting electrons and an external potential $\hat{V}_0$. The corresponding energy functional is

$$E_{V_0}[\rho] = T[\rho] + \int V_0(r)\rho(r)dr + J[\rho] + E_{\mathrm{NC}}[\rho] \ .$$

Adding and subtracting $T_s[\rho]$ we get

$$E_{V_0}[\rho] = T_s[\rho] + \int V_0(r)\rho(r)dr + J[\rho] + T[\rho] - T_s[\rho] + E_{\mathrm{NC}}[\rho]$$

$$= T_s[\rho] + \int V_0(r)\rho(r)dr + J[\rho] + E_{\mathrm{xc}}[\rho]$$

This defines the exchange–correlation functional $E_{\mathrm{xc}}[\rho]$

$$E_{\mathrm{xc}}[\rho] = T[\rho] - T_s[\rho] + E_{\mathrm{NC}}[\rho] \ .$$

The corresponding potential is defined by the functional derivative

$$V_{\mathrm{xc}}[\rho](r) = \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho(r)} \ .$$

We assume that the optimal density ρ_0 defined by

$$\delta E_{V_0}[\rho]\big|_{\rho=\rho_0} = 0$$

is non-interacting v-representable. Experience shows, that this is a reasonable assumption for physical systems. With that we have a unique potential V_s with

$$\delta E_{V_s}[\rho]\big|_{\rho=\rho_0} = 0 \ .$$

Comparing the results we get

$$V_s(r) = V_0(r) + \int \frac{\rho_0(r')}{|r - r'|}dr' + V_{\mathrm{xc}}[\rho_0](r) \ ,$$

the definition of the Kohn–Sham potential.

From the variation principle for the non-interacting system we get the eigen-value equation

$$\left(\hat{T}_s + \hat{V}_s\right)\Phi_i = \epsilon_i\Phi_i \ ,$$

where the Lagrange multipliers ϵ_i have been introduced for the orthogonality constraint of the Kohn–Sham orbitals Φ_i. In analogy to the Hartree–Fock equation the potential V_s depends on the solution of the Kohn–Sham equations and the Kohn–Sham orbitals have to be calculated using an iterative procedure.

Exactly like the Hartree–Fock orbitals the Kohn–Sham orbitals can be expanded in basis functions. We get a matrix representation of the Kohn–Sham operator

$$\langle\mu|\hat{H}_{\mathrm{KS}}|\nu\rangle = h_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma}(\mu\nu,\lambda\sigma) + \langle\mu|\hat{V}_{\mathrm{xc}}|\nu\rangle \ ,$$

where P is the one-particle density matrix of the Kohn–Sham orbitals. Formally we can get the Kohn–Sham equations from the Hartree–Fock equations by replacing the matrix representation of the exchange operator by the matrix representation of the exchange–correlation potentials.

$$-\frac{1}{2}\sum_{\lambda\sigma} P_{\lambda\sigma}(\mu\lambda,\nu\sigma) \longrightarrow \langle\mu|\hat{V}_{\mathrm{xc}}|\nu\rangle$$

Lecture 11

Density Functional Theory: Part II

11.1 Exchange and correlation energy

11.1.1 Density operators

For a system with the wave function $\Psi_N(x_1, x_2, \ldots, x_N)$ (x_i are combined spatial and spin coordinates) we define a density operator γ_N as

$$\gamma_N(x_1', x_2', \ldots, x_N', x_1, x_2, \ldots, x_N) = \Psi_N(x_1', x_2', \ldots, x_N')\Psi_N^\star(x_1, x_2, \ldots, x_N)$$

or in Dirac notation as a projection operator in position representation

$$\gamma_N = |\Psi_N\rangle\langle\Psi_N| \ .$$

The trace of γ_N is one

$$\mathrm{Tr}(\gamma_N) = \int \cdots \int \Psi_N \Psi_N^\star \, dx_1 \cdots dx_N = 1$$

and expectation values of operators $\hat{A}$ can be calculated from

$$\langle A \rangle = \mathrm{Tr}(\gamma_N \hat{A}) = \mathrm{Tr}(\hat{A}\gamma_N) \ .$$

For fermions we define reduced density operators of the form

$$\gamma_p(x_1', x_2', \ldots, x_p', x_1, x_2, \ldots, x_p) = \binom{N}{p} \int \cdots \int \gamma_N \, dx_{p+1} \ldots dx_N \ .$$

Specially we will need the two-particle density operator

$$\gamma_2(x_1', x_2', x_1, x_2) = \frac{N(N-1)}{2} \int \cdots \int \gamma_N \, dx_3 \ldots dx_N$$

and the one-particle density operator

$$\gamma_1(x_1', x_1) = N \int \cdots \int \gamma_N \, dx_2 \ldots dx_N \ .$$

Properties of γ_1 and γ_2 are

- $\text{Tr}(\gamma_2) = \frac{N(N-1)}{2}$

- $\text{Tr}(\gamma_1) = N$

- γ_1 and γ_2 are positive semi-definite and Hermitian and can be diagonalised.

$$\int \gamma_1(x_1', x_1)\phi_i(x_1) \, dx_1 = n_i\phi_i(x_1')$$

$$\int\int \gamma_2(x_1', x_2', x_1, x_2)\theta_i(x_1, x_2) \, dx_1 dx_2 = g_i\theta_i(x_1', x_2')$$

ϕ_i are the natural spin orbitals and θ_i the natural spin geminals of the system. The eigenvalues n_i and g_i are called occupation numbers of the orbitals and geminals, respectively. We have

$$0 \leq n_i, g_i \leq 1 \ .$$

and

$$\sum_i n_i = N$$

$$\sum_i g_i = \frac{N(N-1)}{2}$$

Remember:

Orbital one-electron function

Geminal two-electron function

In the basis of the eigenfunctions we can write the density operators as

$$\hat{\gamma}_1 = \sum_i n_i|\phi_i\rangle\langle\phi_i|$$

$$\hat{\gamma}_2 = \sum_i g_i|\theta_i\rangle\langle\theta_i| \ .$$

Integration over spin variables:

$$\rho_1(r_1', r_1) = \int \gamma_1(x_1', x_1)\Big|_{s_1=s_1'} ds_1$$

$$\rho_2(r_1', r_2', r_1, r_2) = \int \int \gamma_2(x_1', x_2', x_1, x_2)\Big|_{\substack{s_1=s_1' \\ s_2=s_2'}} ds_1 ds_2$$

ρ_1 and ρ_2 are spin-free density operators. Information about spin has been lost in ρ:

$$\rho_1^\sigma(r_1', r_1) = \gamma_1(r_1'\sigma, r_1\sigma)$$

$$\rho_1(r_1', r_1) = \rho_1^\alpha(r_1', r_1) + \rho_1^\beta(r_1', r_1)$$

Expectation values of operators that are not dependent on spin variables can be calculated by

$$\langle \mathcal{O}_1 \rangle = \mathrm{Tr}(\mathcal{O}_1\rho_1) = \int \mathcal{O}_1(r_1)\rho_1(r_1', r_1)\Big|_{r_1=r_1'} dr_1'$$

$$\langle \mathcal{O}_2 \rangle = \mathrm{Tr}(\mathcal{O}_2\rho_2) = \int \int \mathcal{O}_2(r_1, r_2)\rho_2(r_1', r_2', r_1, r_2)\Big|_{\substack{r_1=r_1' \\ r_2=r_2'}} dr_1 dr_2$$

The electronic Hamilton operator is a two-particle operator, therefore we have

$$E = \langle H \rangle = \mathrm{Tr}(H\rho_2) = E(\rho_2)$$

i.e. the total energy is a functional of the two-particle density operator. Electron density:

$$\rho(r) = \rho_1(r, r)$$

$$\rho(r) = \frac{2}{N-1} \int \rho_2(r, r') dr'$$

$$\rho_2(r, r') = \rho_2(r, r', r, r')$$

11.1.2   Electron-electron interaction

We write the diagonal part of the two-particle density operator in the form

$$\rho_2(r_1, r_2) = \frac{1}{2}\rho(r_1)\rho(r_2)[1 + h(r_1, r_2)] \ ,$$

where $h(r_1, r_2)$ is the pair correlation function. The electron-electron interaction energy is

$$E_{\text{ee}} = \int\int \frac{1}{|r_1 - r_2|} \frac{1}{2} \rho(r_1)\rho(r_2)[1 + h(r_1, r_2)] \, dr_1 dr_2$$

$$= J[\rho] + \frac{1}{2} \int\int \frac{1}{|r_1 - r_2|} \rho(r_1)\rho(r_2)h(r_1, r_2) \, dr_1 dr_2$$

$$= J[\rho] + \frac{1}{2} \int\int \frac{1}{|r_1 - r_2|} \rho(r_1)\rho_{\text{xc}}(r_1, r_2) \, dr_1 dr_2 \ .$$

where $\rho_{\text{xc}}(r_1, r_2) = \rho(r_2)h(r_1, r_2)$ is the exchange–correlation hole of an electron at position r_1.

Sum rule:

$$\rho(r_1) = \frac{2}{N-1} \int \rho_2(r_1, r_2) dr_2$$

$$= \frac{2}{N-1} \int \frac{1}{2}\rho(r_1)\rho(r_2)[1 + h(r_1, r_2)] dr_2$$

$$= \frac{1}{N-1}\rho(r_1) \left[\int \rho(r_2)dr_2 + \int \rho(r_2)h(r_1, r_2)dr_2 \right]$$

$$(N-1)\rho(r_1) = \rho(r_1) \left[N + \int \rho_{\text{xc}}(r_1, r_2)dr_2 \right]$$

From this we get

$$\int \rho_{\text{xc}}(r_1, r_2)dr_2 = -1$$

11.1.3 One-determinant wave functions

$$\Psi^S = |\chi_1, \chi_2, \ldots, \chi_N\rangle$$

Reduced density operators

$$\gamma_1(x_1', x_1) = \int \Psi^S(x_1', \ldots)\Psi^{S\star}(x_1, \ldots)dx_2 \ldots dx_N$$

$$\gamma_1(x_1', x_1) = \sum_{i=1}^{N} \chi_i(x_1')\chi_i^\star(x_1)$$

$$\gamma_2(x_1', x_2', x_1, x_2) = \frac{1}{2} \left[\gamma_1(x_1', x_1)\gamma_1(x_2', x_2) - \gamma_1(x_1', x_2)\gamma_1(x_2', x_1) \right]$$

γ_2 is a simple function of γ_1.

Electron density:

$$\rho(r) = \sum_{i=1}^{N} |\chi_i(r)|^2$$

Pair correlation function:

$$\rho_2(r_1, r_2, r_1, r_2) = \frac{1}{2} \left[\rho(r_1)\rho(r_2) - \rho_1(r_1, r_2)\rho_1(r_2, r_1) \right]$$

Be careful with spin!

$$\rho_1(r_1, r_2)\rho_1(r_2, r_1) = \rho_1^\alpha(r_1, r_2)\rho_1^\alpha(r_2, r_1) + \rho_1^\beta(r_1, r_2)\rho_1^\beta(r_2, r_1)$$

$$h(r_1, r_2) = -\frac{\rho_1(r_1, r_2)\rho_1(r_2, r_1)}{\rho(r_1)\rho(r_2)}$$

Sum rule:

$$
\begin{aligned}
\int \rho(r_2) h(r_1, r_2) dr_2 &= -\int \rho(r_2) \frac{\rho_1(r_1, r_2)\rho_1(r_2, r_1)}{\rho(r_1)\rho(r_2)} dr_2 \\
&= -\frac{1}{\rho(r_1)} \int \rho_1(r_1, r_2)\rho_1(r_2, r_1) dr_2 \\
&= -\frac{1}{\rho(r_1)} \int \sum_{ij} \chi_i(r_1)\chi_i^\star(r_2)\chi_j(r_2)\chi_j^\star(r_1) dr_2 \\
&= -\frac{1}{\rho(r_1)} \sum_i \chi_i(r_1)\chi_i^\star(r_1) \\
&= -\frac{1}{\rho(r_1)} \rho(r_1) \\
&= -1
\end{aligned}
$$

The exchange-correlation hole of a Slater determinant fulfils the sum rule exactly.

Invariance to unitary transformations:

$$\bar{\chi}_j^\sigma = \sum_i \chi_i^\sigma U_{ij}^\sigma$$

U^σ is a unitary matrix ($UU^\star = 1$). Spin symmetry can not be broken by the transformation.

Example:

$$\rho_1^\sigma(r_1^{'}, r_1) = \sum_i \bar{\chi}_i^\sigma(r_1^{'}) \sum_i \bar{\chi}_i^{\sigma\star}(r_1)$$

$$= \sum_{ijk} \chi_j^\sigma(r_1^{'}) U_{ji}^\sigma \{\chi_k^\sigma(r_1) U_{ki}^\sigma\}^\star$$

$$= \sum_{ijk} \chi_j^\sigma(r_1^{'}) U_{ji}^\sigma U_{ki}^{\sigma\star} \chi_k^{\sigma\star}(r_1)$$

$$= \sum_j \chi_j^\sigma(r_1^{'}) \chi_j^{\sigma\star}(r_1)$$

With ρ_1 is also ρ_2 invariant and therefore the expectation values of the Hamilton operator and the total energy.

Natural orbitals:

$$\gamma_1 = \sum_i n_i |\phi_i\rangle\langle\phi_i|$$

and

$$\gamma_1 = \sum_i |\chi_i\rangle\langle\chi_i|$$

and therefore $n_i = 1$ and $\phi_i = \chi_i$.

Idem-potency

$$\gamma_1\gamma_1 = \left\{\sum_i n_i |\phi_i\rangle\langle\phi_i|\right\} \left\{\sum_i n_i |\phi_i\rangle\langle\phi_i|\right\}$$

$$= \sum_i n_i^2 |\phi_i\rangle\langle\phi_i|$$

$$= \sum_i |\phi_i\rangle\langle\phi_i|$$

$$= \gamma_1$$

11.1.4 Exchange and correlation energy in the Kohn–Sham method

Kohn–Sham exchange and correlation functionals:

$$E_{\text{XC}}[\rho] = T[\rho] - T_s[\rho] + V_{\text{ee}}[\rho] - J[\rho]$$

We want to make a connection of $E_{\text{XC}}[\rho]$ with the pair correlation function from the electron-electron interaction energy. We consider Hamilton operators where the operator $\hat{V}_{\text{ee}}$ has been multiplied by a scaling factor λ. For

$\lambda = 1$ we get a system of fully interacting electrons and for $\lambda = 0$ a system of non-interacting electrons (similar to the Kohn–Sham reference system). The wave function to these operators will be denoted by Ψ^λ. From the Hohenberg–Kohn theorems we get for non-interactive v-representable ρ

$$F_{\lambda=1}[\rho] = F[\rho] = T[\rho] + V_{ee}[\rho]$$
$$F_{\lambda=0}[\rho] = T_s[\rho] \ .$$

And therefore we have

$$E_{XC}[\rho] = T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho]$$
$$= F_{\lambda=1}[\rho] - F_{\lambda=0}[\rho] - J[\rho]$$
$$= \int_0^1 d\lambda \frac{\partial F_\lambda[\rho]}{\partial \lambda} - J[\rho]$$

From a non-trivial derivation one gets finally

$$\frac{\partial F_\lambda[\rho]}{\partial \lambda} = \langle \Psi^\lambda \mid V_{ee} \mid \Psi^\lambda \rangle$$

and

$$E_{XC}[\rho] = \int_0^1 d\lambda \langle \Psi^\lambda \mid V_{ee} \mid \Psi^\lambda \rangle - J[\rho]$$
$$= \int \int \frac{1}{r_{12}} \bar{\rho}_2(r_1, r_2) \, dr_1 dr_2 - J[\rho]$$
$$= \frac{1}{2} \int \int \frac{1}{r_{12}} \rho(r_1)\rho(r_2)\bar{h}(r_1, r_2) \, dr_1 dr_2$$
$$= \frac{1}{2} \int \int \frac{1}{r_{12}} \rho(r_1)\bar{\rho}_{xc}(r_1, r_2) \, dr_1 dr_2$$

where the averaged exchange–correlation hole $\bar{\rho}_{xc}$ is defined by

$$\bar{\rho}_{xc}(r_1, r_2) = \rho(r_2)\bar{h}(r_1, r_2)$$

and the averaged pair correlation function $\bar{h}$

$$\int_0^1 d\lambda \, \rho_2^\lambda(r_1, r_2) = \frac{1}{2}\rho(r_1)\rho(r_2)[1 + \bar{h}(r_1, r_2)]$$

$\rho_2^\lambda(r_1, r_2)$ is the diagonal part of the two-particle density matrix calculated from the wave function Ψ^λ.

The averaged pair correlation function has the property
$$\bar{h}(r_1, r_2) = \bar{h}(r_2, r_1)$$
and the sum rule
$$\int \bar{\rho}_{\mathrm{xc}}(r_1, r_2) dr_2 = \int \rho(r_2)\bar{h}(r_1, r_2)dr_2 = -1 \ .$$

The exchange-correlation energy depends only on a spherically averaged $\bar{\rho}_{\mathrm{xc}}(r_1, r_2)$
$$E_{\mathrm{XC}}[\rho] = \frac{1}{2}\int dr\rho(r) \int_0^\infty 4\pi s ds \rho_{\mathrm{xc}}^{\mathrm{SA}}(r, s) \ ,$$
where the spherically averaged exchange–correlation hole is defined by
$$\rho_{\mathrm{xc}}^{\mathrm{SA}}(r, s) = \frac{1}{4\pi}\int_\Omega \bar{\rho}_{\mathrm{xc}}(r, r')dr'$$
with the integration domain $\Omega : |r - r'| = s$. The sum rule gets
$$4\pi \int s^2 ds \rho_{\mathrm{xc}}^{\mathrm{SA}}(r, s) = -1 \ .$$

Definition of exchange and correlation energy in quantum chemistry:
$$E_{\mathrm{XC}}[\rho] = E_{\mathrm{X}}^{\mathrm{HF}}[\rho_{\mathrm{HF}}] + E_{\mathrm{C}}^{\mathrm{HF}}[\rho]$$
$E_{\mathrm{C}}^{\mathrm{HF}}$: Exact energy - Hartree–Fock energy.
$E_{\mathrm{X}}^{\mathrm{HF}}$: Hartree–Fock exchange energy calculated with Hartree–Fock orbitals.
For the Kohn–Sham theory we will use a slightly different separation:
$$E_{\mathrm{XC}}[\rho] = E_{\mathrm{X}}^{\mathrm{KS}}[\rho] + E_{\mathrm{C}}^{\mathrm{KS}}[\rho]$$
$E_{\mathrm{X}}^{\mathrm{KS}}$: Hartree–Fock exchange energy calculated with Kohn–Sham orbitals.
The exchange hole for spin compensated systems can be calculated from
$$\rho_x(r_1, r_2) = -\frac{1}{2}\frac{|\rho_1(r_1, r_2)|}{\rho(r_1)}$$
and is
$$\int \rho_x(r_1, r_2) dr_2 = -1 \ .$$
The correlation hole is defined as the difference to ρ_x
$$\bar{\rho}_{\mathrm{xc}}(r_1, r_2) = \rho_x(r_1, r_2) + \rho_c(r_1, r_2)$$
and with that we have
$$\int \rho_c(r_1, r_2) dr_2 = 0 \ .$$

The exchange energy is identical to the Coulomb interaction of the electrons with a charge distribution with total charge one. The correlation energy is a Coulomb interaction of the electrons with a neutral charge distribution.

Lecture 12

Density Functional Theory: Part III

12.1 Local density approximation

$$E_{\mathrm{XC}}[\rho] = \int dr \rho(r) \epsilon_{\mathrm{XC}}[\rho(r)]$$

where $\epsilon_{\mathrm{XC}}(r) = f(\rho(r))$ is a function (not functional) of the density.
Potential:

$$V_{\mathrm{XC}}(r) = \epsilon_{\mathrm{XC}}(r) + \rho(r)\frac{\delta\epsilon_{\mathrm{XC}}[\rho(r)]}{\delta\rho(r)}$$

The approximation is to have at each point in space the energy density with the value of the homogenous electron gas of the local density.
Separation of XC energy density

$$\epsilon_{\mathrm{XC}}(r) = \epsilon_{\mathrm{X}}(r) + \epsilon_{\mathrm{C}}(r)$$

Hartree–Fock exchange for the homogenous electron gas (Dirac 1930)

$$\epsilon_{\mathrm{X}}(\rho(r)) = -\frac{3}{4}\left(\frac{3}{\pi}\rho\right)^{1/3}$$

Correlation energy for the homogenous electron gas:
Quantum-Monte-Carlo calculations (Ceperley, Alder 1980, Ortiz, Ballone 1994)
Analytic function fitted to QMC values:

$$r_s = \left[\frac{3}{4\pi\rho}\right]^{1/3}$$

$$\epsilon_{\mathrm{C}} = \frac{\gamma_c}{1 + \beta_1 \sqrt{r_s} + \beta_2 r_s} \qquad\qquad r_s > 1$$

$$\epsilon_{\mathrm{C}} = A \ln r_s + B + C r_s \ln r_s + D r_s \qquad\qquad r_s < 1$$

$$\gamma_c = -0.1423, \qquad \beta_1 = 1.0529, \qquad \beta_2 = 0.3334$$
$$A = 0.0311, \qquad B = -0.048, \qquad C = 0.0020, \qquad D = -0.0116$$

The formula for E_{XC} in the local density approximation gives

$$\rho_{\mathrm{xc}}^{\mathrm{LDA}}(r_1, r_2) = \rho(r_1)\bar{h}_0(|r_1 - r_2|; \rho(r_1)) \ ,$$

where $\bar{h}_0(|r_1 - r_2|; \rho(r_1))$ is the averaged pair correlation function for the homogenous electron gas. Contrary to the exact formula we have here the density $\rho(r_1)$ as a factor and not $\rho(r_2)$. The LDA exchange correlation density is

$$\epsilon_{\mathrm{xc}}(\rho(r_1)) = \frac{1}{2} \int \frac{1}{r_{12}} \rho_{\mathrm{xc}}^{\mathrm{LDA}}(r_1, r_2) dr_2 \ .$$

$\rho_{\mathrm{xc}}^{\mathrm{LDA}}$ is spherically symmetric and fulfils the sum rule

$$\int \rho_{\mathrm{xc}}^{\mathrm{LDA}}(r_1, r_2) dr_2 = -1 \ ,$$

but $h_0(|r_1 - r_2|; \rho(r_1))$ is not symmetric in r_1 and r_2. The exact function ρ_{xc} is not spherical symmetric, but only the symmetric part is used in an exchange correlation energy calculation. Therefore, a compensation of errors can be expected from the LDA approximation.

12.2 Gradient corrections

Simple Taylor expansion don't work, but 'empirical' functionals of the type

$$E_{\mathrm{XC}} = \int F_{\mathrm{XC}}[\rho, \nabla\rho] dr$$

give clearly improved results.
Potential:

$$V_{\mathrm{XC}}(r) = \frac{\delta F_{\mathrm{XC}}}{\delta\rho(r)} - \sum_{\alpha=1}^{3} \frac{\partial}{\partial r_\alpha} \left[\frac{\delta F_{\mathrm{XC}}}{\delta(\nabla_\alpha \rho)} \right]$$

12.2.1 Exchange functional of Becke (1988)

$$E_{\mathrm{X}} = E_{\mathrm{X}}^{\mathrm{LDA}}[\rho(r)] - \int F_{\mathrm{X}}[s]\rho^{4/3}dr$$

$$s = \frac{|\nabla\rho|}{\rho^{4/3}}$$

$$F_{\mathrm{X}}[s] = \beta\,2^{-1/3}\frac{s^2}{1 + 6\beta s \sinh^{-1}s}$$

β optimised using Hartree–Fock exchange energies of noble gas atoms.

$$\beta = 0.0042$$

Typical error in the exchange energy $E_{\mathrm{X}}^{\mathrm{LDA}} \approx 10\%$ and $E_{\mathrm{X}}(\text{Becke}) \approx 0.1\%$.

12.2.2 Correlation functionals of Perdew et al.

(Perdew 1986, Perdew-Wang 1991, Perdew-Bourke-Ernzerhof (PBE) 1996)
These functionals model the correlation hole based on exact properties without using empirical parameters.
The PBE functional

$$E_c^{\mathrm{PBE}} = \int \rho(r)\left[\epsilon_c + H(r_s, t)\right]dr$$

$$t = \frac{|\nabla\rho|}{2\rho k_s}, \qquad k_s = \sqrt{\frac{4k_f}{\pi}}, \qquad k_f = \left(\frac{9\pi}{4}\right)^{1/3}r_s \qquad r_s = \left(\frac{3}{4\pi\rho}\right)^{1/3}$$

$$H(r_s, t) = \gamma \log\left(1 + \frac{\beta}{\gamma}t^2\left[\frac{1 + At^2}{1 + At^2 + A^2 t^4}\right]\right)$$

$$\text{with}$$

$$A = \frac{\beta}{\gamma}\left[\exp(-\epsilon_c/\gamma) - 1\right]^{-1}$$

$$\gamma = 0.031090690869, \qquad \beta = 0.066724550$$

12.2.3 Lee-Yang-Parr (LYP) correlation functional (1988)

This functional is based on an approximative formula for the correlation energy of the Helium atom by Colle and Salvetti (1975). The basic assumption is

$$\rho_2(r_1, r_2) = C\rho_2^{\mathrm{HF}}(r_1, r_2)\ .$$

Then they replaced ρ_2^{HF} with ρ_1 and after a partial integration one gets

$$E_c^{\mathrm{LYP}} = \int \rho \left[\frac{-a}{1 + dx} \left[1 + bC_F e^{-cx} \right] \right] dr$$
$$+ \int ab|\nabla\rho|^2 x^5 \frac{e^{-cx}}{1 + dx} \frac{1}{24} \left[1 + \frac{7}{3} \left(cx + \frac{dx}{1 + dx} \right) \right] dr$$

where $x = \rho^{-1/3}$ and

$$C_F = \frac{3}{10}(3\pi^2)^{2/3} \qquad a = 0.04918 \qquad b = 0.132 \qquad c = 0.2533 \qquad d = 0.349$$

For $|\nabla\rho| = 0$ (i.e. a homogenous system) the correct LDA value is not recovered. We have approximately

$$E_c^{\mathrm{LDA}} \approx 2E_c^{\mathrm{LYP}}(|\nabla\rho| = 0)$$

The LYP functional has the special form

$$F_c(\rho, \nabla\rho) = |\nabla\rho|^2 \, G(\rho)$$

12.3 Hybrid functionals

Let's go back to the definition of the exchange–correlation energy in the Kohn–Sham method

$$E_{\mathrm{XC}}[\rho] = T[\rho] - T_s[\rho] + V_{\mathrm{ee}}[\rho] - J[\rho]$$
$$= F_{\lambda=1}[\rho] - F_{\lambda=0}[\rho] - J[\rho]$$
$$= \int_0^1 d\lambda \, \frac{\partial F_\lambda[\rho]}{\partial \lambda} - J[\rho]$$

and define the function E_{NC}^λ as the non-classical Coulomb energy at the value λ. With this we get

$$E_{\mathrm{XC}}[\rho] = \int_0^1 E_{NC}^\lambda \, d\lambda \ .$$

For a value of $\lambda = 0$ (a non-interacting system) this energy is exactly known. It's the Hartree-Fock exchange energy, because the wavefunction for this system is a single Slater determinant. If we now approximate the end point of integration $\lambda = 1$ with a common density functional, we can estimate the total integral with a weighted sum of the two end points. How to choose the weights is again an open question. There are many possibilities, some of them purely empirical and others based on theoretical considerations.
The two most often used hybrid functionals are

Table 12.1: Dependence on basis set for some bond length [Å]. LDA values

Bond	6-31G(d,p)	6-311++G(d,p)	Limit	Experiment
$H-H$	$-/0.765$	$-/0.765$	$-/0.765$	$-/0.741$
H_3C-CH_3	$1.513/1.105$	$1.510/1.101$	$1.508/1.100$	$1.526/1.088$
H_2C-CH_2	$1.330/1.098$	$1.325/1.094$	$1.323/1.093$	$1.339/1.085$
$HC-CH$	$1.212/1.078$	$1.203/1.073$	$1.203/1.074$	$1.203/1.061$

- B3LYP

$$E_{\mathrm{XC}}^{\mathrm{B3LYP}} = (1-a)E_x^{\mathrm{LDA}} + aE_x^{\mathrm{HF}} + bE_x^{\mathrm{B88}} + cE_c^{\mathrm{LYP}} + (1-c)E_c^{\mathrm{LDA}}$$

where E_x^{B88} is the Becke exchange functional. The values for the parameters are

$$a = 0.20, \qquad b = 0.72, \qquad c = 0.81$$

- PBE0

$$E_{\mathrm{XC}}^{\mathrm{PBE0}} = E_{\mathrm{XC}}^{\mathrm{PBE}} + 0.25(E_x^{\mathrm{HF}} - E_x^{\mathrm{PBE}})$$

This is a parameter free functional. Neither the PBE functional nor the weight 0.25 are fit to any experimental data.

12.4 Basis sets

In density functional theory the same kind of basis sets can be used as in Hartree–Fock calculations. In principal the parameters would have to be newly optimised (e.g. for a 6-31G* type basis). In practice even individual parameters for each density functional would have to be generated. However, it turned out that such optimised parameters only have a minor effect on results and the expense needed for the optimisation is not justifiable. It is therefore common to use the same basis sets in density functional theory that are also used in wavefunction calculations.

In Kohn–Sham density functional theory only simple orbitals have to be described. Therefore, the demand on the basis is comparable to other independent particle models like the Hartree–Fock model. Basis sets of double-ζ quality with one set of polarisation functions are often sufficient to get good results. Triple-ζ basis sets with two sets of polarisation functions are for most cases close to converged results.

Table 12.2: Averaged and maximal absolute error of the atomisation energy for a test set (G2) of molecules [kcal/mol]. Basis: 6-311+G(3df,2p)

Functional	RMS	max error
LDA	36.4	84
PBE	8.6	26
BLYP	4.7	15
B3LYP	2.4	10
PBE0	3.5	10

12.5 Model calculations

Comparing different functional for all kind of molecular properties is a favourite past time of all quantum chemists. There is a plethora of research papers on this topic.

12.6 Coulomb energy

Density functional programs are build exactly in the same way as Hartree–Fock programs. The only difference is that exchange integrals are replace by integrals with the potential from the XC functionals. Many DFT programs differ in another point from Hartree–Fock programs. In Hartree–Fock program the Fock matrix elements for Coulomb and exchange are calculated together.

$$F_{\mu\nu}^c = \sum_{\lambda\sigma} P_{\lambda\sigma}(\mu\nu, \lambda\sigma)$$

$$F_{\mu\nu}^x = -\frac{1}{2} \sum_{\lambda\sigma} P_{\lambda\sigma}(\mu\lambda, \sigma\nu)$$

Computational time is dominated by the calculation of the two-electron integrals. Therefore one uses

$$F_{\mu\nu}^c \longleftarrow 4P_{\lambda\sigma}(\mu\nu, \lambda\sigma)$$
$$F_{\lambda\sigma}^c \longleftarrow 4P_{\mu\nu}(\mu\nu, \lambda\sigma)$$
$$F_{\nu\sigma}^x \longleftarrow -P_{\mu\lambda}(\mu\nu, \lambda\sigma)$$
$$F_{\mu\sigma}^x \longleftarrow -P_{\nu\lambda}(\mu\nu, \lambda\sigma)$$
$$F_{\nu\lambda}^x \longleftarrow -P_{\mu\sigma}(\mu\nu, \lambda\sigma)$$
$$F_{\mu\lambda}^x \longleftarrow -P_{\nu\sigma}(\mu\nu, \lambda\sigma)$$

to achieve a minimal number of integral calculations. In DFT calculation only the Coulomb part is needed, except for hybrid functionals. The Coulomb contribution has a special structure that can be used together with further approximations to increase the efficiency of the program.

For the calculation of the Coulomb part of the Kohn–Sham matrix we need all integrals $(\mu\nu, \lambda\sigma)$ over all basis functions. This method scales like M^4 where M is the number of basis functions. For large systems the number of integrals can be reduced drastically if only integrals with a significant contribution to the energy are calculated. All modern programs are using such screening procedures. It can be shown that asymptotically only M^2 integrals have to be calculated. However, this procedure has a rather large pre-factor, i.e. the asymptotic scaling is only reached for very large systems. For this reason other approximate method for the calculation of the Coulomb energy are used.

12.6.1 Resolution of identity method

The electron density is defined as

$$\rho(r) = \sum_{\mu\nu} P_{\mu\nu} \phi_\mu(r) \phi_\nu(r) \ .$$

The product of two atom-centred basis functions is a new function

$$\chi_{(\mu\nu)}(r) = \phi_\mu(r)\phi_\nu(r) \ .$$

The density is therefore described exactly in a basis of the size M^2.

$$\rho(r) = \sum_{(\mu\nu)} P_{\mu\nu} \chi_{(\mu\nu)}(r)$$

The basic idea is now to replace this exact description of the density by an approximate description in a smaller basis, where the new basis is only slightly bigger than the basis used for the orbitals.

$$\rho(r) \approx \sum_{\alpha} c_\alpha \chi_\alpha(r)$$

With this we get for the contribution to the Kohn–Sham matrix

$$F_{\mu\nu}^c = \sum_{\alpha} c_\alpha \int \int \frac{\phi_\mu(r)\phi_\nu(r)\chi_\alpha(r')}{|r - r'|} \, dr dr'$$
$$= \sum_{\alpha} c_\alpha (\mu\nu, \alpha)$$

This method scales like M^3 and asymptotically again like M^2, but it has a much smaller pre-factor. Special basis sets for the electron density have to be used that match the basis sets for the wavefunctions. Typically these basis sets have three times more functions than the wave function basis.

We haven't yet discussed how the expansion coefficients c_α are calculated. We can transform this problem into an optimisation. We are looking for those expansion coefficients which minimise the difference to the exact electron density with respect to a given metric.

$$\tilde{\rho}(r) = \sum_\alpha c_\alpha \chi_\alpha(r)$$

$$\Delta(r) = \rho(r) - \tilde{\rho}(r)$$

$$\Omega(\{c\}) = \int\int \Delta(r)\mathcal{O}(r,r')\Delta(r')drdr'$$

$$\mathbf{c} \leftarrow \text{MIN}_{\{c\}}\Omega(\{c\})$$

where $\mathcal{O}(r,r')$ defines a distance metric. The solution of this problem is

$$c_\alpha = \sum_\beta S_{\alpha\beta}^{-1}\,T_\beta$$

with

$$S_{\alpha\beta} = \int\int \chi_\alpha(r)\mathcal{O}(r,r')\chi_\beta(r')drdr'$$

$$T_\beta = \int\int \chi_\beta(r)\mathcal{O}(r,r')\rho(r')drdr'$$

The operators $\mathcal{O}(r,r') = \delta(r-r')$, $\mathcal{O}(r,r') = |r-r'|$ and $\mathcal{O}(r,r') = \frac{1}{|r-r'|}$ have been investigated. Most of the programs use the Coulomb operator. These methods are called RI (resolution of identity) method, because they can be derived using the identity operator (for orthogonal basis sets).

$$1 = \sum_\alpha |\alpha\rangle\langle\alpha|$$

This method is available in many quantum chemistry codes z.B. `Turbomole, Gaussian 2003, DeMON`.

12.6.2 Local expansion of density

The local expansion of density are based on similar ideas as the RI methods, but instead of looking for a global approximation to the density separate

approximations to all pair densities are used.

$$\rho_{AB}(r) = \sum_{\substack{\alpha\beta \\ \alpha\in A,\beta\in B}} P_{\alpha\beta}\Phi_\alpha(r)\Phi_\beta(r)$$

$$\rho_{AB}(r) \approx \sum_{\mu\in A} c_\mu\chi_\mu^A(r) + \sum_{\nu\in B} c_\nu\chi_\nu^B(r)$$

The pair density of atoms A and B is expanded in a basis located at the atoms A and B. The advantage of this method is that the calculation of the expansion coefficients is done locally and is therefore independent of the size of the molecule. The disadvantage lies in the limitation of the expansion to the functions on the same atoms. To achieve good results large basis sets (with high angular momenta) have to be used. This method is used in the program ADF.

12.7   Numerical integration

With Gaussian functions as basis sets, all integrals needed in standard quantum chemistry calculations can be derived analytically. Integrals for the exchange and correlation functionals needed in density functional theory cannot be calculated analytically even for special types of basis functions. They have to be calculated using numerical procedures.

12.7.1   Voronoi polyhedra

We investigate a general integral over all space

$$I = \int F(r)dr \ .$$

In a first step we divide the integral into different parts using a weight function $w_A(r)$. The function $w_A(r)$ should be close to unity near nucleus A and close to zero near all other nuclei. The function should be continuous and differentiable and defined in all space. We require further that

$$\sum_A w_A(r) = 1 \ .$$

With this we get

$$I = \int F(r)dr$$

$$= \int \left(\sum_A w_A(r) \right) F(r)dr$$

$$= \sum_A \int w_A(r) F(r)dr$$

$$= \sum_A I_A$$

To calculate the weight functions we introduce confocal elliptic coordinates (λ, μ, ϕ). These coordinates are defined by

$$\lambda = \frac{r_1 + r_2}{R_{12}}$$

$$\mu = \frac{r_1 - r_2}{R_{12}} \quad .$$

Dabei ist r_1 die Distanz zum Zentrum 1 und r_2 die Distanz zum Zentrum 2, R_{12} der Abstand der beiden Zentren und ϕ der Winkel um die Verbindungslinie der Zentren. Der Definitionsbereich der Koordinaten ist

$$\phi : \quad [0, 2\pi]$$

$$\lambda : \quad [1, \infty]$$

$$\mu : \quad [-1, 1]$$

Die wichtige Koordinate ist μ. Für zwei Zentren A und B haben wir

$$\mu_{AB} = \frac{r_A - r_B}{R_{AB}}$$

und wir sehen, dass μ_{AB} den Wert -1 am Zentrum A und auf der Achse von A weg besitzt. μ_{AB} hat den Wert 1 am Zentrum B und auf der Achse von B weg. Definieren wir eine neue Funktion $s(\mu_{AB})$

$$s(\mu_{AB}) = \begin{cases} 1 & \text{for } -1 \leq \mu_{AB} \leq 0 \\ 0 & \text{for } 0 < \mu_{AB} \leq 1 \end{cases} \quad .$$

With this we can define the weigth functions

$$P_A(r) = \prod_{B \neq A} s(\mu_{AB})$$

$$w_A(r) = \frac{P_A(r)}{\sum_A P_A(r)} \quad .$$

128

The whole space is devided into Voronoi polyhedra with respect to the nuclei. The discontinuity of the function $s(\mu_{AB})$ at $\mu_{AB} = 0$ is not suited for numerical calculations. Therefore we choose

$$\frac{ds}{d\mu} = A_m(1 - \mu^2)^m \ ,$$

where A_m is taken such that $s(-1) = 1$ and $s(1) = 0$. A typical value for m is 10.

It is advantageous to use also the relative size of the atoms, i.e. we make a variable transformation such that different atoms get different weights.

$$\nu_{AB} = \mu_{AB} + a_{AB}(1 - \mu_{AB}^2)$$

The parameter a_{AB} is calcuated from the Bragg-Slater radii (R_A, R_B) of the atoms.

$$a_{AB} = \frac{u_{AB}}{u_{AB}^2 - 1}$$
$$u_{AB} = \frac{\chi - 1}{\chi + 1}$$
$$\chi = \frac{R_A}{R_B}$$

12.7.2 Radial integration

We calculate the integrals I_A using numerical integration, where the integration is split into a radial part and an integration over the unit sphere.

$$I_A \approx \sum_i w_r(i) \sum_j w_\Omega(j) \ w_A(r(i), \Omega(j)) \ F_A(r(i), \Omega(j))$$

There are many different integration methods for radial grids. The best methods for this problem work in two steps. First the integral $[0, \infty]$ is mapped onto $[-1, 1]$ using a variable transformation. Possible transformations are

$$r = \eta \frac{a + x}{\sqrt{1 - x^2}} \ ,$$
$$r = \eta \frac{a + x}{1 - x} \ ,$$
$$r = \frac{\eta}{\ln 2} \ \ln \frac{a + 1}{1 - x} \ ,$$

where the transformations are normalised such that the midpoint of the x-interval is maped onto $r = \eta$. The value of η depends on the type of atom.

H 0.8	He 0.9	Li 1.8	Be 1.4	B 1.3	C 1.1	N 0.9	O 0.9
F 0.9	Ne 0.9						
Na 1.4	Mg 1.3	Al 1.3	Si 1.2	P 1.1	S 1.0	CL 1.0	Ar 1.0
K 1.5	Ca 1.4	Sc 1.3	Ti 1.2	V 1.2	Cr 1.2	Mn 1.2	Fe 1.2
Co 1.2	Ni 1.1						
Cu 1.1	Zn 1.1	Ga 1.1	Ge 1.0	As 0.9	Se 0.9	Br 0.9	Kr 0.9

Finally we use a Gauss quadrature for the integral on the interval $[-1, 1]$.

$$\int_{-1}^{1} f(x)dx \approx \sum_{j=1}^{N} \frac{v_j}{W(x_j)} f(x_j)$$

The form of the function $W(x)$ depends on the orthogonal polynome that was used as a base for the integration method.

<div align="center">

Gauss-Legendre
</div>

$$W(x) = 1, \qquad -1 < x < 1$$
$$(j+1)P_{j+1} = (2j+1)xP_j - jP_{j-1}$$

<div align="center">

Gauss-Chebyshev
</div>

$$W(x) = \frac{1}{\sqrt{1-x^2}}, \qquad -1 < x < 1$$
$$T_{j+1} = 2xT_j - T_{j-1}$$

12.7.3 Angular integration

Two-dimensional integrals over the unit sphere

$$\int_0^\pi \int_0^{2\pi} f(\theta, \phi) sin\theta d\theta d\phi$$

can be efficiently performed using Lebedev grids. Lebedev grids with a certain number of points allow for the exact integration of all spherical harmonics up to a certain value of $L_{\max}$. These grids are based on octahedral symmetry. The position of the special points and their integration weights have to be taken from a table.

Gitter	$L_{\max}$	I_{nred}	I_{tot}
1	3	1	6
2	5	2	14
3	7	3	26
4	9	3	38
5	11	4	50
6	15	5	86
7	17	6	110
8	19	7	146
9	23	9	194
10	29	12	302
11	35	16	434
12	41	20	590
13	47	25	770
14	53	30	974

Lecture 13

Molecular Properties

13.1 Wavefunction analysis

$$\text{Wavefunction } \Psi \quad \longrightarrow \quad \begin{array}{l} \text{chemical concepts} \\ \text{atomic charges} \\ \text{oxidation states} \\ \text{reactivity} \\ \text{bonding} \end{array}$$

13.1.1 Basisfunction based population analysis

For $\Psi = |\Phi_1 \cdots \Phi_N|$ or $\gamma(r, r') = \sum_i f_i \Phi_i(r)\Phi_i(r')$ (single determinant wavefunctions) we have

$$\rho(r) = \sum_i \rho_i(r) = \sum_i f_i |\Phi_i(r)|^2$$

Basis set expansion

$$\Phi_i(r) = \sum_\alpha^{\text{AO}} c_{\alpha i} \phi_\alpha(r)$$

Inserting this into the density we get

$$\rho_i(r) = f_i |\Phi_i(r)|^2 = f_i \sum_{\alpha\beta} c_{\alpha i} c_{\beta i} \phi_\alpha(r)\phi_\beta(r)$$

133

Calulating the total number of electrons

$$N = \sum_i^{\text{MO}} \int \rho_i(r) \, dr = \sum_i^{\text{MO}} f_i \sum_{\alpha\beta} c_{\alpha i} c_{\beta i} \int \phi_\alpha(r)\phi_\beta(r) \, dr$$

$$= \sum_i^{\text{MO}} f_i \sum_{\alpha\beta} c_{\alpha i} c_{\beta i} S_{\alpha\beta}$$

$$= \sum_{\alpha\beta} P_{\alpha\beta} S_{\alpha\beta}$$

$$= \text{Tr}\,(PS)$$

where we have used the density matrix

$$P_{\alpha\beta} = \sum_i^{\text{MO}} f_i c_{\alpha i} c_{\beta i}$$

We will now look into schemes to partition the N electrons using (PS).

Mulliken population analysis

$P_{\alpha\alpha}S_{\alpha\alpha}$: fraction of electron in AO α
$P_{\alpha\beta}S_{\alpha\beta}$: put $\frac{1}{2}$ electron in α and $\frac{1}{2}$ electron in β Now we sum over all α
associated with atom A. This only works with basis sets where the functions can be easily subdivided using atoms.

$$\rho_A = \sum_{\alpha \in A}^{\text{AO}} \sum_{\beta}^{\text{AO}} P_{\alpha\beta} S_{\alpha\beta}$$

The total charge on atom A is then

$$q_A = Z_A - \rho_A$$

Löwdin population analysis

This population analysis is based on a slightly different calculation of the total number of electrons based on the invariance of the trace.

$$N = \text{Tr}\,(PS) = \text{Tr}\left(PS^{\frac{1}{2}}S^{\frac{1}{2}}\right)$$

$$= \text{Tr}\left(S^{\frac{1}{2}}PS^{\frac{1}{2}}\right) = \text{Tr}\left(\bar{P}\right)$$

$$= \sum_\alpha \bar{P}_{\alpha\alpha}$$

The number of electrons on atom A is

$$\rho_A = \sum_{\alpha \in A}^{\text{AO}} \bar{P}_{\alpha\alpha}$$

The total charge on atom A is again

$$q_A = Z_A - \rho_A$$

These types of charges are

- strongly depend on basis functions

- can only be used for trends, not absolute values

- use only with small basis sets (no diffuse functions)

13.1.2   Electrostatic potential derived charges

Die electrostatic potential of an electronic charge distribution $\rho(r)$ and a collection of atomic cores with charge Z_A is

$$V_{\text{ESP}}(r) = \sum_A \frac{Z_A}{|r - R_A|} - - \int \frac{\rho(r')}{|r - r'|} \, dr'$$

The potential of a collection of atomic charges q_A at the same positions R_A as the atomic cores is

$$V_{\text{ESP}}^q(r) = \sum_A \frac{q_A}{|r - R_A|}$$

We now use the electrostatic potential to define the optimal atomic charges q_A with

- outside the charge distribution $\rho(r)$ we require

$$V_{\text{ESP}}(r) = V_{\text{ESP}}^q(r)$$

- charge conservation

$$\sum_A q_A = \text{total charge of molecule}$$

- fix higher multipoles (dipole etc.)

We achieve this using a least square fit with constraints

$$\text{Min}_q \sum_i \left(V_{\text{ESP}}(r_i) - V_{\text{ESP}}^q(r_i) \right)^2 + \text{constraints}$$

These type of charges are often used in force fields and QM/MM calculations.

13.1.3 Density based charges

Hirschfeld method

Definition

$$q_A = Z_A - \int_{\text{Atom}} \rho(r) \, dr$$

This requires a definition of an atomic volume. We start with a weighting function

$$w_A(r) = \frac{\bar{\rho}_A(r)}{\bar{\rho}(r)}$$

where $\bar{\rho}_A(r)$ is the electron density of an isolated atom and $\bar{\rho}(r) = \sum_A \bar{\rho}_A(r)$. The weighting function has the property

$$\sum_A w_A(r) = 1$$

for any choice of r. Therefore we can write

$$N = \int \rho(r) \, dr = \int \rho(r) \sum_A w_A(r) \, dr$$

$$= \sum_A \int \rho(r) w_A(r) \, dr = \sum_A \rho_A$$

with the definition of ρ_A

$$\rho_A = \int \rho(r) w_A(r) \, dr$$

Atoms in molecules (AIM)

Bader analysis: based on the topology of $\rho(r)$ (maxima, minima, saddle points).
Maxima of $\rho(r)$ are at nuclei = attractors of electronic charge. The gradient of density at each point in space points to strongest attractor (basin of attraction = atom).

13.2 Response properties

Response of a molecule to external (internal) perturbation

- external electric field (F)

Table 13.1: Example: Atomic charge of oxygen in water

Basis set	Mulliken	Löwdin	ESP	AIM
STO-3G	-0.39	-0.27	-0.65	-0.89
3-21G	-0.74	-0.46	-0.90	-0.93
6-31G**	-0.67	-0.44	-0.81	-1.24
6-311G(2d2p)	-0.52	0.00	-0.74	-1.24
6-311++G(2d2p)	-0.47	-0.12	-0.76	-1.25

- external magnetic field (B)

- nuclear magnetic moment (I)

- nuclear geometry (R)

time-independent perturbation : static property
time-dependent perturbation : dynamic property
Total property: contribution from electrons (QM) and nuclei

Taylor series expansion of energy

$$E(\lambda) = E(0) + \frac{\partial E}{\partial \lambda}\lambda + \frac{1}{2}\frac{\partial^2 E}{\partial \lambda^2}\lambda^2 + \cdots$$

Energy derivatives = "Properties"

Example: External electric field

$$E_{\text{int}} = \int \rho(r)V_{\text{ext}}(r)\ dr$$

Assume electric field $F = -\frac{\partial V_{\text{ext}}}{\partial r}$ is uniform over molecule (Multipole expansion)

$$E_{\text{int}} = qV - \mu \cdot F - \frac{1}{2}Q : F' - \cdots$$

$$q = \text{net charge}$$
$$\mu = \text{dipole moment}$$
$$Q = \text{quadrupole moment}$$
$$F' = \text{field gradient}$$

Field $\longrightarrow$ wavefunction $\longrightarrow$ dipole

$$\mu = \mu_0 + \alpha F + \frac{1}{2}\beta F^2 + \frac{1}{6}\gamma F^3 + \cdots$$
$$\mu_0 = \text{permanent dipole}$$
$$\alpha = \text{polarisability}$$
$$\beta, \gamma = \text{hyper-polarisability}$$

The total energy is
$$E(F) = \langle \Psi \mid H_0 - \mu F \mid \Psi \rangle$$

with the total Hamiltonian

$$H = H_0 - \mu_0 F - \alpha F^2 - \frac{1}{2}\beta F^3 - \cdots$$

Series expansion

$$E(F) = E(0) + \underbrace{\frac{\partial E}{\partial F}}_{-\mu_0} F + \underbrace{\frac{1}{2}\frac{\partial^2 E}{\partial F^2}}_{-\alpha} F^2 + \underbrace{\frac{1}{6}\frac{\partial^3 E}{\partial F^3}}_{-\frac{1}{2}\beta} F^3 + \cdots$$

General:
$$\text{Property} \propto \frac{\partial^{n+m} E}{\partial F^n \partial R^m}$$

n	m	Property
0	0	Energy
1	0	Electric dipole moment
0	1	Energy gradient
2	0	Polarisability
0	2	Harmonic vibrational frequency
1	1	Infrared absorption
2	1	Raman intensity

13.2.1 Derivative techniques

Hamiltonian
$$H = H_0 + \lambda P_1 + \lambda^2 P_2$$

Expectation value

$$E(\lambda) = \langle \Psi \mid H_0 + \lambda P_1 + \lambda^2 P_2 \mid \Psi \rangle$$

First derivative

$$\frac{\partial E(\lambda)}{\partial \lambda} = \langle \frac{\partial \Psi}{\partial \lambda} \mid H \mid \Psi \rangle + \langle \Psi \mid P_1 + \lambda P_2 \mid \Psi \rangle + \langle \Psi \mid H \mid \frac{\partial \Psi}{\partial \lambda} \rangle$$

For real wavefunctions we have $\lambda \to 0$ then $\Psi \to \Psi_0$.

$$\frac{\partial E(\lambda)}{\partial \lambda} \bigg|_{\lambda=0} = \langle \Psi_0 \mid P_1 \mid \Psi_0 \rangle + 2 \langle \frac{\partial \Psi}{\partial \lambda} \bigg|_{\lambda=0} \mid H_0 \mid \Psi_0 \rangle$$

The wavefunction Ψ depends on two different set of parameters, the wavefunction expansion coefficients (c) and the basis functions (ψ). The derivative wrt λ is therefore

$$\frac{\partial \Psi}{\partial \lambda} = \frac{\partial \Psi}{\partial \psi} \frac{\partial \psi}{\partial \lambda} + \frac{\partial \Psi}{\partial c} \frac{\partial c}{\partial \lambda}$$

For fully variational wavefunctions we have

$$\frac{\partial \Psi(\alpha_1, \alpha_2, \ldots, \alpha_n)}{\partial \alpha_i} = 0$$

and we get the Hellmann-Feynman theorem

$$\frac{\partial}{\partial \lambda} \langle \Psi \mid H \mid \Psi \rangle = \langle \Psi \mid \frac{\partial H}{\partial \lambda} \mid \Psi \rangle$$

For variationally optimised wavefunctions (HF, DFT, MCSCF, but not MPx, CI, CCSD) we have

$$\frac{\partial \Psi}{\partial c} = 0 \quad \text{and} \quad \frac{\partial \Psi}{\partial \psi} \neq 0$$

and therefore

$$\frac{\partial E(\lambda)}{\partial \lambda} \bigg|_{\lambda=0} = \underbrace{\langle \Psi_0 \mid P_1 \mid \Psi_0 \rangle}_{\text{Hellmann-Feynman}} + 2 \underbrace{\langle \frac{\partial \Psi}{\partial \lambda} \bigg|_{\lambda=0} \mid H_0 \mid \Psi_0 \rangle \frac{\partial \psi}{\partial \lambda}}_{\text{Pulay term}}$$

139

Lecture 14

NMR Chemical Shielding

14.1 General

Maxwell equations

$$\mathrm{div}\vec{E} = 4\pi\rho$$

$$\mathrm{div}\vec{B} = 0$$

$$\mathrm{rot}\vec{E} = -\frac{1}{c}\frac{\partial\vec{B}}{\partial t}$$

$$\mathrm{rot}\vec{B} = \frac{4\pi}{c}\vec{J} + \frac{1}{c}\frac{\partial\vec{E}}{\partial t}$$

Partial solution by introducing a vector field $\vec{A}$ with the property

$$\vec{B} = \mathrm{rot}\vec{A}$$

$\vec{A}$ determines $\vec{B}$ uniquely, but $\vec{A}(\vec{B})$ is not unique. We can add a gradient field to $\vec{A}$ without change of the physical properties.

$$\vec{A}' = \vec{A} + \mathrm{grad}\lambda$$

$$\mathrm{rot}\vec{A}' = \mathrm{rot}\vec{A} + \mathrm{rot}(\mathrm{grad}\lambda) = \mathrm{rot}\vec{A}$$

λ: gauge function
Insert into Maxwell equation

$$\mathrm{rot}\vec{E} = -\frac{1}{c}\frac{\partial\vec{B}}{\partial t}$$

$$\vec{E} + \frac{1}{c}\frac{\partial\vec{A}}{\partial t} = -\mathrm{grad}\Phi$$

Φ : scalar potential with

$$\Phi' = \Phi + \text{const.}$$

Special form of $\vec{A}$: Coulomb gauge

$$\text{div}\vec{A} = 0$$

Poisson equation

$$\Delta\Phi = -4\pi\rho$$

$$-\Delta\vec{A} + \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} + \frac{1}{c^2}\text{grad}\frac{\partial\Phi}{\partial t} = \frac{4\pi}{c}\vec{J}$$

In the non-relativistic limit $(c \to \infty)$ these equations simplify to

$$\Delta\Phi = -4\pi\rho$$

$$\Delta\vec{A} = 0$$

However, there is still a change in the gauge possible

$$\vec{A}' = \vec{A} + \text{grad}\lambda$$

as we get for

$$\text{div}\vec{A} = \text{div}(\vec{A} + \text{grad}\lambda)$$
$$= \text{divgrad}\lambda$$
$$= \Delta\lambda$$

This means that a change of gauge is possible for λ that fulfill the Laplace equation

$$\Delta\lambda = 0$$

Special case of a homogeous external magnetic field $\vec{B}(r) = \vec{B}_0$

$$\vec{A}_0 = \frac{1}{2}\vec{B}_0 \times (\vec{r} - \vec{R})$$
$$= \frac{1}{2}\vec{B}_0 \times \vec{r} - \frac{1}{2}\vec{B}_0 \times \vec{R}$$
$$= \frac{1}{2}\vec{B}_0 \times \vec{r} + \text{grad}\lambda$$
$$\lambda = -\frac{1}{2}(\vec{B}_0 \times \vec{R}) \cdot \vec{r}$$

The position $\vec{R}$ is called the gauge origin.

14.2 Chemical shifts

Nuclear spins couple to the local magnetic field $\vec{B}_{\text{loc}}(\vec{R}_A)$

- level split

- transitions between levels can be induced by radiofrequency fields

- measure energy difference $\propto \vec{B}_{\text{loc}}(\vec{R}_A)$

Local magnetic field

$$\vec{B}_{\text{loc}}(\vec{r}) = \vec{B}_{\text{ext}}(\vec{r}) + \vec{B}_{\text{ind}}(\vec{r})$$
$$\vec{B}_{\text{ext}}(\vec{r}) = \vec{B}_0$$

Biot-Savart law

$$\vec{B}_{\text{ind}}(\vec{s}) = -\frac{1}{c} \int \frac{\vec{j}(\vec{r}) \times (\vec{r} - \vec{s})}{|\vec{r} - \vec{s}|} \, d\vec{r}$$

where $\vec{j}(\vec{r})$ is the current density. We introduce the magnetic shielding constant

$$\vec{B}_{\text{ind}}(\vec{s}) = -\sigma(\vec{s})\vec{B}_0$$

and the chemical shift

$$\delta = \sigma_{\text{std}} - \sigma(\vec{s})$$

The chemical shift is independent on the external magnetic filed $\vec{B}_0$ and the magnetic moment I of the nuclei.

Calculation: For molecule and standard calculate $\sigma(\vec{s})$ at atomic positions, this needs the calculation of the current density $\vec{j}(\vec{r})$.

Hamiltonian:

$$H_0 \xrightarrow{\vec{B}_{\text{ext}}} H$$
$$\vec{p} \to \vec{p} + \vec{A}_{\text{ext}}$$

For the Coulomb gauge and gauge origin $\vec{R}$ we have

$$\hat{H} = H_0 + \vec{A}_{\text{ext}}\vec{p} + \frac{1}{2}\vec{A}_{\text{ext}}^2$$
$$\vec{A}_{\text{ext}} = \frac{1}{2}\vec{B}_{\text{ext}} \times (\vec{r} - \vec{R})$$
$$\hat{H}(\vec{B}_{\text{ext}}, \vec{R}) = \hat{H}_0 + \frac{1}{2}\left((\vec{r} - \vec{R}) \times \vec{p}\right)\vec{B}_{\text{ext}} + \frac{1}{8}\left(\vec{B}_{\text{ext}} \times (\vec{r} - \vec{R})\right)^2$$

For one-determinant wavefunctions and closed shells we have

$$\Psi = |\Phi_1 \cdots \Phi_n|$$

and for the current density

$$\vec{j} = i \sum_{k=1}^{N} (\Phi_k^* \vec{\nabla} \Phi_k - (\vec{\nabla} \Phi_k^*) \Phi_k) - 2\vec{A}_{\text{ext}} \sum_{k=1}^{N} \Phi_k^* \Phi_k$$

14.3   Linear response for Hartree–Fock

Fock operator

$$\hat{h} = \hat{h}_0 + i\hat{h}_1 + \mathcal{O}(B_{\text{ext}}^2)$$

Hartree–Fock orbitals

$$\Phi_k = \Phi_k^0 + i\Phi_k^1 + \mathcal{O}(B_{\text{ext}}^2)$$

The current density vanishes for $B_{\text{ext}} = 0$

$$\vec{j} = \vec{j}^1 + \mathcal{O}(B_{\text{ext}}^2)$$

We get for the first order current density

$$\vec{j}^1 = 2 \underbrace{\sum_{k=1}^{N} \left(\Phi_k^1 \nabla \Phi_k^0 - \Phi_k^0 \nabla \Phi_k^1 \right)}_{\text{paramagnetic part}} - 2\underbrace{\vec{A}_{\text{ext}} \sum_{k=1}^{N} \Phi_k^0 \Phi_k^0}_{\text{diamagnetic part}}$$

$\vec{j}^1$ is independent of the gauge origin $\vec{R}$. The paramagnetic and diamagnetic parts depend on the gauge origin. This causes numerical problems if the two terms are not calculated with the same accuracy. This is called the gauge invariance problem of chemical shifts. It is effectively related to the problem of non-complete basis sets. The ground state and linear response orbitals have different needs for basis sets.

Solution: distributed gauge methods

IGLO individual gauge origins $\vec{R}_k$ for every localised orbital

GIAO gauge including atomic orbitals

CSGT continous set of gauge transformations

LORG localised orbital/local origin

Table 14.1: Example: chemical shifts for C calculated with different methods (CH_4)

Basis	common gauge origin		IGLO	GIAO
	C	H		
TZ	25.58	62.01	32.48	32.57
TZ+3p	30.97	41.22	31.29	31.35
uncontracted	30.98	32.07	31.20	31.34
extended	31.23	31.26	31.24	31.25

14.3.1 Calculation of response orbitals

We expand the linear response orbitals in the set of ground state orbitals.

$$\Phi_p^1 = \sum_{q=1}^N Y_{qp}\Phi_q^0$$

with $Y_{qp} = Y_{pq}$. Stationarity conditions for Hartree-Fock

$$F_{ia} = 0 \qquad \begin{cases} \text{i occupied} \\ \text{o unoccupied} \end{cases}$$

Expansion of Fock matrix

$$F_{ia} = F_{ia}^0 + iF_{ia}^1 + \mathcal{O}(B_{\text{ext}}^2)$$

with

$$F_{ia}^1 = \langle \Phi_1^0 \mid \hat{h}^1 \mid \Phi_a^0 + \sum_{jb} H_{ia,jb}Y_{jb}$$

and

$$H_{ia,jb} = \delta_{ij}F_{ab}^0 - \delta_{ab}F_{ij}^0 + (ab,ij) - (aj,bi)$$

The condition $F_{ia}^1 = 0$ leads to a set of linear equations for Y of the form

$$\mathcal{H}Y = h$$

$\mathcal{H}$ depends on Φ^0 and h depends on Φ^0 and $\hat{h}^1$.

$$\hat{h}^1 = -i\frac{1}{2}\left((\vec{r} - \vec{R}) \times \vec{p}\right)B_{\text{ext}}$$

$$= -\frac{1}{2}\left((\vec{r} - \vec{R}) \times \vec{\nabla}\right)B_{\text{ext}}$$

For canonical orbitals $F_{pq} = \epsilon_p \delta_{pq}$ and we get

$$Y_{ia} = -\frac{\langle \Phi_i^0 \mid \hat{h}^1 \mid \Phi_a^0 \rangle}{\epsilon_i - \epsilon_a}$$

The induced current density is

$$j^1 = 2 \sum_{ia} Y_{ia} \left(\Phi_i^0 \nabla \Phi_a^0 - \Phi_a^0 \nabla \Phi_i^0 \right) - \left(\vec{B}_{\text{ext}} \times (\vec{r} - \vec{R}) \right) \sum_i \Phi_i^0 \Phi_i^0$$

Calculations of chemical shifts have been performed with Hartree–Fock, MP2, CCSD(T), and DFT electronic structure methods. For heavy elements relativistic corrections are needed.