Course SS22

Stefan.Goedecker@unibas.ch

Organization

- Lecture with script: http://comphys.unibas.ch/teaching.htm
  Simple exercises: Traditional analytic problems and numerical problems (on your own laptop?)

Lecture together with exercises give 4 credit points if exercises are done well and the oral exam is passed. Prerequisite for being admitted to oral examination on the material of the course: successful solution of at least half of the number of points for each exercise set. (Easy exercises give 1 point, medium 2 and difficult 3 points)

Ab-initio electronic structure calculations

Matter is composed of electrons and atomic nuclei. Within the framework of condensed matter physics, chemistry, biology and materials science one is not interested in the structure of the nuclei. The nuclei are just considered to be charged point particles. The basic laws governing systems of electrons and nuclei are the many body equations of quantum mechanics. These equations govern therefore all phenomena of matter at the atomistic scale. In principle any condensed matter problem can therefore be solved by using quantum mechanics.

Ab-initio electronic structure calculations solve numerically the quantum mechanical equations. Ab initio means that the theory behind such calculations is based on ”first principles”. i.e. on the basic physical laws. The only input for an ab-initio electronic structure calculation are the atomic charges of the constituent atoms. Ab-initio method can be subdivided into two categories, wave-function and density functional methods. In wave-function methods one calculates directly the many-body wave-function. These methods are also called quantum-chemistry methods or post Hartree-Fock methods and they can solve the quantum mechanical equations with arbitrary precision. Unfortunately the computer time grows tremendously when very high precision is required. For this reason wave-function methods can essentially only be applied to small molecules, but not to larger systems. Because of these limitations, density functional methods have become very popular. Because they are computationally much cheaper they can be applied to systems containing up to a few hundred atoms. In contrast to wave-function methods they can not systematically be improved in accuracy, but they give reasonable accuracy for a wide range of properties. The bond lengths and angles of molecular geometries and crystalline structures can be predicted with errors of a few percent. Vibrational frequencies and the bulk modulus with errors of 5 to 10 percent. However, there are also properties whose accuracy is insufficient in density functional theory. When one compares different structures very small energy differences can be important. Experimentally one can measure energy differences with an accuracy of 1 kcal/mole (about 0.04 eV/molecule). The best density functional schemes give only an accuracy of .2 per atom in molecules. Errors in transition state energies can be up to a 100 percent. Density functional methods can well describe covalent, metallic and ionic bonds. Hydrogen and short range van der Waals interactions are described with moderate accuracy, long range van der Waals interactions are entirely missing in standard density functional schemes.

The virtual chemistry/physics laboratory

Simulation methods are replacing traditional experimental methods in academic as well as in industrial research. Instead of doing research in a traditional lab one does it in the 'virtual' lab. Such a virtual chemistry/physics laboratory needs the following ingredients:

- Models of the physical reality, e.g. density functional theory for the description of interacting electronic systems or elasticity theory for the description of macroscopic bodies
- Algorithms that allow us to solve the fundamental equations of these models numerically
- Fast computers
- Efficient implementations of the algorithms on modern computers

Due to large research efforts worldwide considerable progress is to be expected in all these fields and simulation will become the workhorse of research and development in various industries such as in the pharmaceutical and in the chemical industry

Topics

- The Born-Oppenheimer approximation
- The Hartree-Fock method
- Configuration Interaction methods
- Quantum Monte Carlo
- Density Functional methods
- Tight binding methods
- Linear scaling algorithms for electronic structure calculations
- Pseudopotentials
- Basis sets for electronic structure calculations
- Diagonalization and minimization algorithms
- Quantum molecular dynamics

- Calculation of various properties such as forces, stress, phonons, excitations and magnetic properties
- Overview over existing electronic structure codes

Atomic units

All equations in this course will us atomic units. These units are formally obtained by setting $\hbar = m_e = e = \kappa_0 = 1$, where $\kappa_0 = 4\pi\varepsilon_0$. In this way important atomic properties have unit values

- charge of an electron = 1 (instead of 1.60 $\times 10^{-19}$ C)
- mass of an electron = 1 (instead of 9.11 $\times 10^{-31}$ kg)
- Angular momentum, $\hbar$ = 1 (instead of 1.05 $\times 10^{-34}$ J s
- Bohr radius of hydrogen atom $a_0 = \frac{\hbar^2}{m_e e^2}$ = 1 (instead of .529 $\times 10^{-10}$ m)
- Ground state energy of hydrogen atom $-\frac{1}{2}\frac{m_e e^4}{\hbar^2} = -\frac{1}{2}$ (instead of -2.18 $\times 10^{-18}$ J)

Because many other atomic and molecular properties are related to the above quantities, they will also have numerical values that are of the order of unity. For instance

- Bond lengths are of the order of the Bohr radius
- The binding energy of a molecule is typically a fraction of the ground state energy of the hydrogen atom

- The electric dipole moment of a molecule is typically of the order of $ea_0 = 1$ (instead of 8.45×10^{-30} C m)

A brief repetition of the single particle Schrödinger equation

The time independent single particle Schrödinger equation is given by

$$\mathcal{H}\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}) \tag{1}$$

where

$$\mathcal{H} = -\frac{1}{2}\nabla^2 + V(\mathbf{r}) \tag{2}$$

The eigenvalue ε_i gives the energy of the i-th state To solve this equation on a computer, we have to discretize it. One expresses in most cases the eigen-functions as a linear combination of a set of so-called basis functions $U_k(\mathbf{r})$.

$$\phi_i(\mathbf{r}) = \sum_k u_{k,i}U_k(\mathbf{r}) \tag{3}$$

We will assume that the basis functions are orthogonal, i.e that $\int U_l(\mathbf{r})U_k(\mathbf{r})d\mathbf{r} = \delta_{k,l}$. Substituting Eq. 3 into Eq. 1 one obtains

$$\mathcal{H}\sum_k u_{k,i}U_k(\mathbf{r}) = \varepsilon_i \sum_k u_{k,i}U_k(\mathbf{r})$$

Multiplying from the left by $U_l(\mathbf{r})$ and integrating, one obtains

$$\sum_k u_{k,i} \int U_l(\mathbf{r}) \mathcal{H} U_k(\mathbf{r}) d\mathbf{r} = \varepsilon_i \sum_k u_{k,i} \int U_l(\mathbf{r}) U_k(\mathbf{r}) d\mathbf{r} = \varepsilon_i u_{l,i}$$

Introducing the Hamiltonian matrix $H_{l,k}$

$$H_{l,k} = \int U_l(\mathbf{r}) \mathcal{H} U_k(\mathbf{r}) d\mathbf{r} \tag{4}$$

we obtain the following eigenvalue problem

$$\sum_k H_{l,k} u_{k,i} = \varepsilon_i u_{l,i} \tag{5}$$

Eq. 5 is the discretized version of the continuous Schrödinger equation Eq. 1. Introducing a vector $\vec{u}_i$ that contains the expansion coefficients of the i-th eigenfunction Eq. 5 can be rewritten in matrix vector notation

$$H \vec{u}_i = \varepsilon_i \vec{u}_i \tag{6}$$

Exercise [1pt]: *Prove for the 1-dim case Schrödinger equation that H (Eq. 4) is symmetric*

The Born Oppenheimer approximation

A condensed matter system is in essence just a collection of N_{at} atomic nuclei and N electrons. The Born Oppenheimer approximation justifies treating the nuclei as classical particles in the electronic Schrödinger equation. A priori both the nuclei and the electrons have to be treated quantum-mechanically and are therefore described by the combined electron-nuclei ground state wave-function $\Psi^{en}(\mathbf{R}_1,...,\mathbf{R}_{N_{at}},\mathbf{r}_1,...,\mathbf{r}_N)$ that is an eigenfunction

$$\mathcal{H}^{en}(\mathbf{R}_1,...,\mathbf{R}_{N_{at}},\mathbf{r}_1,...,\mathbf{r}_N)\Psi^{en}(\mathbf{R}_1,...,\mathbf{R}_{N_{at}},\mathbf{r}_1,...,\mathbf{r}_N) = E^{en}\,\Psi^{en}(\mathbf{R}_1,...,\mathbf{R}_{N_{at}},\mathbf{r}_1,...,\mathbf{r}_N) \tag{7}$$

of the Hamiltonian

$$\mathcal{H}^{en}(\mathbf{R}_1,...,\mathbf{R}_{N_{at}},\mathbf{r}_1,...,\mathbf{r}_N) = \underbrace{\frac{1}{M}\sum_{i=1}^{N_{at}} -\frac{1}{2}\nabla^2_{\mathbf{N}_i}}_{\mathcal{T}^n} \tag{8}$$

$$+\underbrace{\sum_{i=1}^{N_{at}}\sum_{j=1}^{i-1}\frac{Z_iZ_j}{|\mathbf{R}_i-\mathbf{R}_j|}+\sum_{i=1}^{N}-\frac{1}{2}\nabla^2_{\mathbf{r}_i}+\sum_{i=1}^{N}\sum_{j=1}^{i-1}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}-\sum_{i=1}^{N}\sum_{j=1}^{N_{at}}\frac{Z_j}{|\mathbf{r}_i-\mathbf{R}_j|}}_{\mathcal{H}}$$

where $\nabla^2_{\mathbf{N}_i} = \frac{M}{M_i}\left(\frac{\partial^2}{\partial X_i^2}+\frac{\partial^2}{\partial Y_i^2}+\frac{\partial^2}{\partial Z_i^2}\right)$ and $\nabla^2_{\mathbf{r}_i} = \frac{\partial^2}{\partial x_i^2}+\frac{\partial^2}{\partial y_i^2}+\frac{\partial^2}{\partial z_i^2}$. M is the mass of a proton in atomic units which has the numerical value of 1836.1. $\frac{M_i}{M}$ is thus the atomic mass in units of the proton mass. In addition to the combined electron-nuclei wave-function Ψ^{en} there exist the electronic wave-functions $\Psi_k(\mathbf{R}_1, ..., \mathbf{R}_{N_{at}}, \mathbf{r}_1, ..., \mathbf{r}_N)$ that are by definition eigenfunctions of the electronic part of the Hamiltonian

$$\begin{aligned} \mathcal{H}(\mathbf{R}_1, ..., \mathbf{R}_{N_{at}}, \mathbf{r}_1, ..., \mathbf{r}_N) \quad \Psi_k(\mathbf{R}_1, ..., \mathbf{R}_{N_{at}}, \mathbf{r}_1, ..., \mathbf{r}_N) \quad &= \\ E_k(\mathbf{R}_1, ..., \mathbf{R}_{N_{at}}) \quad \Psi_k(\mathbf{R}_1, ..., \mathbf{R}_{N_{at}}, \mathbf{r}_1, ..., \mathbf{r}_N) & \end{aligned} \tag{9}$$

Since $\mathcal{H}$ is a hermitian operator the eigenfunctions Ψ_k form a complete orthonormal set with respect to the space of the electronic coordinates and the electron-nuclei ground state wave-function can therefore be expanded in terms of these eigenfunctions

$$\Psi^{en}(\mathbf{R}_1, ..., \mathbf{R}_{N_{at}}, \mathbf{r}_1, ..., \mathbf{r}_N) = \sum_k \Psi_k(\mathbf{R}_1, ..., \mathbf{R}_{N_{at}}, \mathbf{r}_1, ..., \mathbf{r}_N)\Psi_k^n(\mathbf{R}_1, ..., \mathbf{R}_{N_{at}}) \tag{10}$$

Inserting Eq. 10 into Eq. 7 we obtain

$$\begin{aligned} \sum_k \left(\frac{-1}{2M} \nabla_{\mathbf{N}}^2 + \mathcal{H} \right) \Psi_k \Psi_k^n \quad &= \quad E^{en} \sum_k \Psi_k \Psi_k^n \\ \frac{-1}{2M} \sum_k \nabla_{\mathbf{N}}^2 \Psi_k \Psi_k^n + \sum_k E_k \Psi_k \Psi_k^n \quad &= \quad E^{en} \sum_k \Psi_k \Psi_k^n \end{aligned} \tag{11}$$

where $\nabla_{\mathbf{N}}^2 = \sum_{i=1}^{N_{at}} \nabla_{\mathbf{N}_i}^2$

Introducing the vector

$$\nabla_{\mathbf{N}} = \begin{pmatrix} \sqrt{\frac{M}{M_1}}\frac{\partial}{\partial X_1} \\ \sqrt{\frac{M}{M_1}}\frac{\partial}{\partial Y_1} \\ \sqrt{\frac{M}{M_1}}\frac{\partial}{\partial Z_1} \\ \ldots\ldots \\ \sqrt{\frac{M}{M_{N_{at}}}}\frac{\partial}{\partial X_{N_{at}}} \\ \sqrt{\frac{M}{M_{N_{at}}}}\frac{\partial}{\partial Y_{N_{at}}} \\ \sqrt{\frac{M}{M_{N_{at}}}}\frac{\partial}{\partial Z_{N_{at}}} \end{pmatrix}$$

we apply the product rule for differentiation to Eq. 11

$$\begin{aligned} \frac{-1}{2M}\sum_k \nabla_{\mathbf{N}}^2 \Psi_k \Psi_k^n + \sum_k E_k \Psi_k \Psi_k^n &= E^{en}\sum_k \Psi_k \Psi_k^n \\ \frac{-1}{2M}\sum_k \nabla_{\mathbf{N}}\left(\Psi_k \nabla_{\mathbf{N}} \Psi_k^n + \Psi_k^n \nabla_{\mathbf{N}} \Psi_k\right) + \sum_k E_k \Psi_k \Psi_k^n &= E^{en}\sum_k \Psi_k \Psi_k^n \\ \frac{-1}{2M}\sum_k \left[\Psi_k \nabla_{\mathbf{N}}^2 \Psi_k^n + 2(\nabla_{\mathbf{N}}\Psi_k)(\nabla_{\mathbf{N}}\Psi_k^n) + \Psi_k^n \nabla_{\mathbf{N}}^2 \Psi_k\right] + \sum_k E_k \Psi_k \Psi_k^n &= E^{en}\sum_k \Psi_k \Psi_k^n \quad (12) \end{aligned}$$

To obtain an equation for the nuclear wave-functions Ψ_k^n we multiply Eq. 12 from the left by Ψ_i and integrate over $d\mathbf{r}_1, ...d\mathbf{r}_N$. Using the orthogonality of Ψ_k

$$\langle \Psi_i | \Psi_j \rangle = \int d\mathbf{r}_1, ...d\mathbf{r}_N \Psi_i(\mathbf{r}_1, ...\mathbf{r}_N) \Psi_j(\mathbf{r}_1, ...\mathbf{r}_N) = \delta_{i,j}$$

we obtain

$$\frac{-1}{2M} \nabla_{\mathbf{N}}^2 \Psi_i^n + \underbrace{\frac{-1}{2M} \sum_k \left[\langle \Psi_i | \nabla_{\mathbf{N}} | \Psi_k \rangle (\nabla_{\mathbf{N}} \Psi_k^n) + \langle \Psi_i | \nabla_{\mathbf{N}}^2 | \Psi_k \rangle \Psi_k^n \right]}_{\text{non-adiabatic coupling elements}} + E_i \Psi_i^n = E^{en} \Psi_i^n \qquad (13)$$

In the adiabatic approximation the non-adiabatic terms coupling different electronic states are neglected and only the electronic ground state (E_0, Ψ_0) is used for describing the ground state of the combined electron nucleon system:

$$\frac{-1}{2M} \nabla_{\mathbf{N}}^2 \Psi_0^n + \frac{-1}{2M} \left[2 \langle \Psi_0 | \nabla_{\mathbf{N}} | \Psi_0 \rangle (\nabla_{\mathbf{N}} \Psi_0^n) + \langle \Psi_0 | \nabla_{\mathbf{N}}^2 | \Psi_0 \rangle \Psi_0^n \right] + E_0 \Psi_0^n = E^{en} \Psi_0^n \qquad (14)$$

The wavefunction Ψ^{en} in Eq. 10 is then simply the product of the electronic ground state wavefunction Ψ_0 and a nucleonic part Ψ^n. Since the non-adiabatic coupling terms in Eq. 13 are small due to the presence of the factor $\frac{1}{M}$ this approximation is justified unless there are several electronic energies E_i that are nearly degenerate. This is for instance the case in photochemical reactions that involve more than one electronic surface

$E_i(\mathbf{R}_1,...,\mathbf{R}_{N_{at}})$. In the Born-Oppenheimer approximation the second factor in Eq. 14 is neglected as well. The justification is again that it is much smaller than E_i due to the presence of the factor $\frac{1}{M}$.

$$\frac{-1}{2M}\nabla^2_{\mathbf{N}}\Psi^n_0 + E_0\Psi^n_0 = E^{en}\,\Psi^n_0 \tag{15}$$

The interpretation of the Born-Oppenheimer equation 15 is the following. The nucleonic wave-function is moving in the potential $E_0(\mathbf{R}_1,...,\mathbf{R}_{N_{at}})$ generated by the eigenvalues of the electronic Schrödinger equation. Because of this interpretation $E_0(\mathbf{R}_1,...,\mathbf{R}_{N_{at}})$ is also called the ground state potential energy surface or ground state Born-Oppenheimer surface. Solving Eq. 15 gives the ground state energy E^{en} of the combined electron-nucleon system. In the limit of an infinitely large mass the nucleonic wave-function would be a delta function that is located at the minimum of the potential energy surface. Because the mass is large but not infinite the nucleonic wave-function is localized in a small region around this minimum. Since the kinetic energy $\frac{-1}{2M}\int\Psi^n_0\nabla^2_{\mathbf{N}}\Psi^n_0 = \frac{1}{2M}\int\left|\nabla_{\mathbf{N}}\Psi^n_0\right|^2$ is always positive, E^{en} is always greater than the minimum of the potential energy surface.

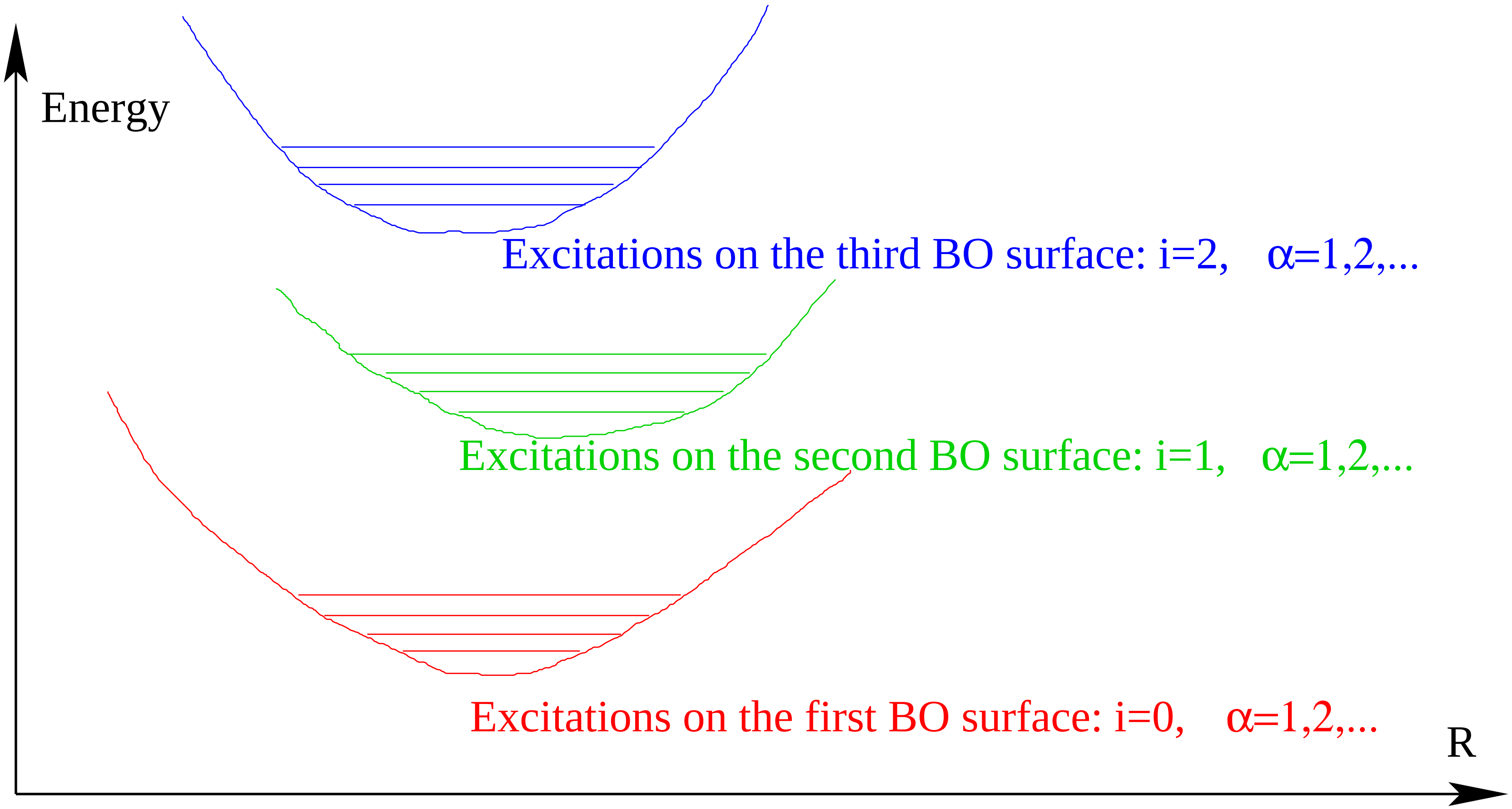
Up to now we have only considered the ground state of the combined electron-nucleus system. Lets now look at the exited states. Being an eigenvalue equation, Eq. 15 allows to solve not only for the ground state but also for excited states:

$$\frac{-1}{2M}\nabla^2_{\mathbf{N}}\Psi^n_{0,\alpha} + E_0\Psi^n_{0,\alpha} = E^{en}_{0,\alpha}\,\Psi^n_{0,\alpha} \tag{16}$$

The index α denotes here the different vibrational levels. The energy of vibrational excitations is small, of the order of a meV. If we are also interested in higher energy excitations (of the order of an eV), we have to consider also the different vibrational levels associated to the higher Born-Oppenheimer surfaces, associated to electronic excitations i:

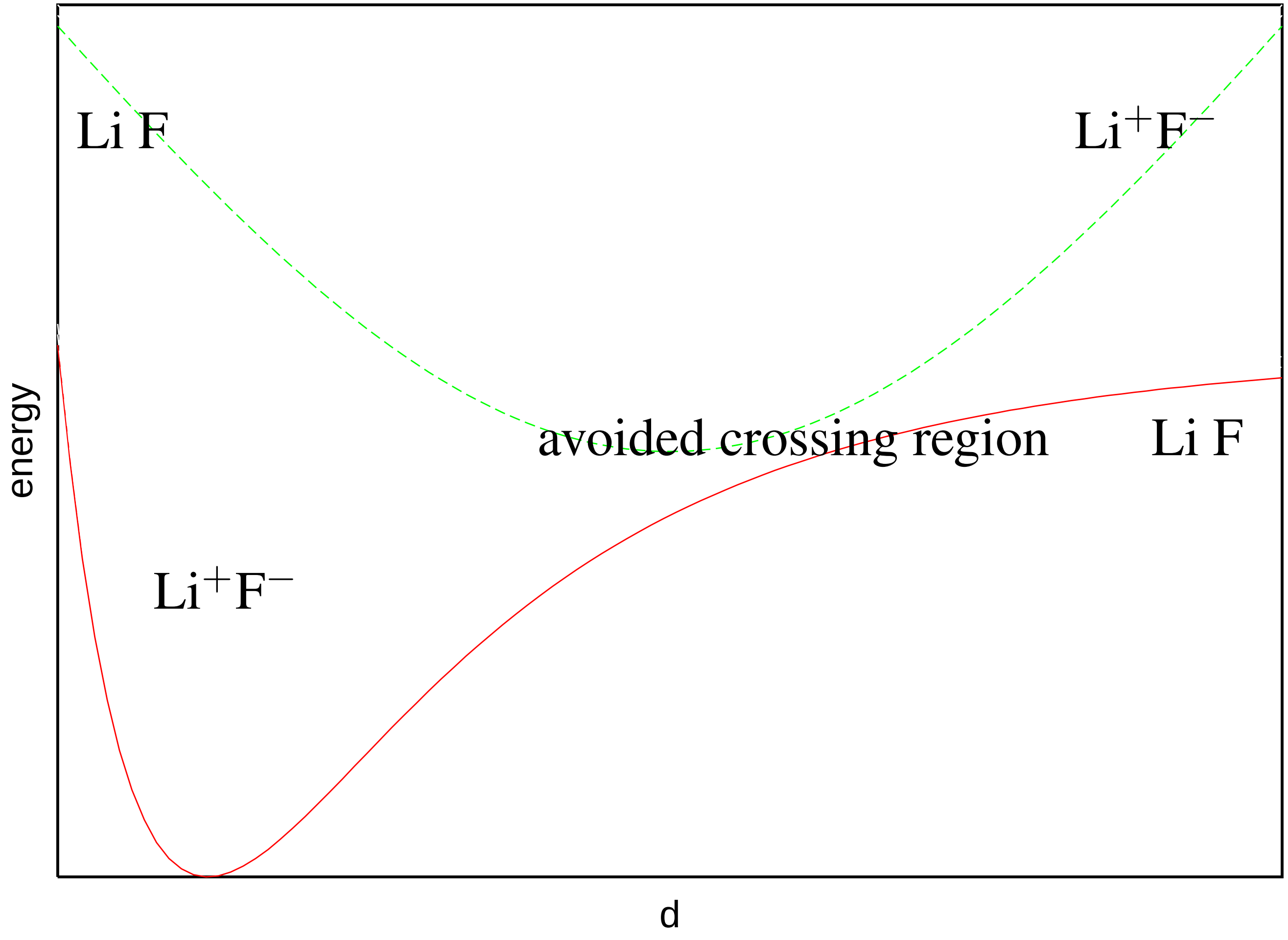
$$\frac{-1}{2M}\nabla_{\mathbf{N}}^{2}\Psi_{i,\alpha}^{n}+E_{i}\Psi_{i,\alpha}^{n}=E_{i,\alpha}^{en}\,\Psi_{i,\alpha}^{n} \tag{17}$$

A schematic view of the entire excitation spectrum of a molecule is shown below.



The difference between $E^{en}_{i,0}$ and the minimum on the Born Oppenheimer surface is called the zero point energy. It represents the smallest possible energy of a nuclear wavefunction.

Whenever two electronic levels come close to each other, the adiabatic approximation fails. An example of such a failure is given below. The figure below shows the potential energy surface of the LiF molecule as a function of the distance d between the two atoms. In the region of the avoided crossing the adiabatic approximation fails.



Exercise [2pts]: The nuclear Schrödinger equation for H_2

Let us consider a H_2 *molecule whose nuclei, both of mass* M *are constrained to move along a line with coordinates* x_1 *and* x_2 *respectively. Let us further assume the Born-Oppenheimer surface is purely quadratic around the equilibrium bond length* d*. The equation for the nuclear wavefunction in the Born-Oppenheimer approximation eq. 15 then becomes:*

$$\left[-\frac{1}{2M}\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}\right)+k(x_2-x_1-d)^2\right]\Psi_n(x_1,x_2)=E^{en}\,\Psi_n(x_1,x_2) \tag{18}$$

Transform the equation first to a new system of coordinates:

$$\begin{aligned} x &= x_1-x_2 \\ X &= \frac{1}{2}(x_1+x_2) \end{aligned}$$

Use the fact that the ground state solution of the Schrödinger equation in a quadratic potential is a Gaussian $\exp(-\frac{1}{2}(\frac{x}{\alpha})^2)$*. Determine the value of* α *and the discrete eigenvalues of equation (18). Compare* α *and the energy of the eigenvalue with typical electronic length and energy scales for the case where* $k=1$*.*

Multiplying the room temperature with the Boltzmann constant gives 25 meV. Only excitations whose energy is less than or comparable to 25 meV will be populated at room temperature. The room temperature energy is very small compared to the typical variation of the Born Oppenheimer surface which is of the order of a few eV's. Hence the nucleonic wave-functions will only sample small regions around the minimum $E^0 = E(\mathbf{R}_1^0, ..., \mathbf{R}_{N_{at}}^0)$ of the Born Oppenheimer surface. Translating this statement into a picture where the nuclei are particles, means that the amplitude of vibration of the nuclei at room temperature is small compared to the inter-atomic distance. The reason why the inter-atomic distance is the right length scale to compare with follows from the fact that a displacement that is of the order of an inter-atomic distance would give rise to an energy of a few eV. In this small region around the minimum of the Born Oppenheimer surface, the surface can well be approximated by a quadratic function. Eq. 16 thus becomes.

$$\left[\frac{-1}{2M}\nabla_{\mathbf{N}}^2 + E^0 + \frac{1}{2}\sum_{I,J}(\mathbf{R}_I - \mathbf{R}_I^0)D_{I,J}(\mathbf{R}_J - \mathbf{R}_J^0)\right]\Psi_{0,\alpha}^n(\mathbf{R}_1, ..., \mathbf{R}_{N_{at}}) = E_{0,\alpha}^{en}\Psi_{0,\alpha}^n(\mathbf{R}_1, ..., \mathbf{R}_{N_{at}}) \tag{19}$$

The elements of the 3 by 3 matrices $D_{I,J}$ are called the inter-atomic force constants.

$$D_{I,J} = \begin{pmatrix} \left.\frac{\partial^2 E}{\partial X_I \partial X_J}\right|_{\mathbf{R}_I=\mathbf{R}_I^0,\mathbf{R}_J=\mathbf{R}_J^0} & \left.\frac{\partial^2 E}{\partial X_I \partial Y_J}\right|_{\mathbf{R}_I=\mathbf{R}_I^0,\mathbf{R}_J=\mathbf{R}_J^0} & \left.\frac{\partial^2 E}{\partial X_I \partial Z_J}\right|_{\mathbf{R}_I=\mathbf{R}_I^0,\mathbf{R}_J=\mathbf{R}_J^0} \\ \left.\frac{\partial^2 E}{\partial Y_I \partial X_J}\right|_{\mathbf{R}_I=\mathbf{R}_I^0,\mathbf{R}_J=\mathbf{R}_J^0} & \left.\frac{\partial^2 E}{\partial Y_I \partial Y_J}\right|_{\mathbf{R}_I=\mathbf{R}_I^0,\mathbf{R}_J=\mathbf{R}_J^0} & \left.\frac{\partial^2 E}{\partial Y_I \partial Z_J}\right|_{\mathbf{R}_I=\mathbf{R}_I^0,\mathbf{R}_J=\mathbf{R}_J^0} \\ \left.\frac{\partial^2 E}{\partial Z_I \partial X_J}\right|_{\mathbf{R}_I=\mathbf{R}_I^0,\mathbf{R}_J=\mathbf{R}_J^0} & \left.\frac{\partial^2 E}{\partial Z_I \partial Y_J}\right|_{\mathbf{R}_I=\mathbf{R}_I^0,\mathbf{R}_J=\mathbf{R}_J^0} & \left.\frac{\partial^2 E}{\partial Z_I \partial Z_J}\right|_{\mathbf{R}_I=\mathbf{R}_I^0,\mathbf{R}_J=\mathbf{R}_J^0} \end{pmatrix}$$

One can now find the normal modes $\mathbf{U}^i$ of the corresponding classical system of harmonic oscillators. The normal modes satisfy the generalized eigenvalue problem

$$\sum_J D_{I,J} \mathbf{U}_J^i = M_I \omega_i^2 \mathbf{U}_I^i$$

and they are orthogonal in the sense that

$$\sum_I M_I \mathbf{U}_I^i \mathbf{U}_I^j = \delta_{i,j}$$

Note that in this notation $\mathbf{U}^i$ is the i-th eigenvector of length $3N_{at}$ that consists of N_{at} subvectors $\mathbf{U}_I^i$ of length 3. In the same way the $3N_{at}$ times $3N_{at}$ matrix D is made of of N_{at}^2 3 by 3 matrices $D_{I,J}$. Since the $\mathbf{U}^i$'s form a complete orthonormal basis any vector can be expanded in terms of the $\mathbf{U}^i$'s. This is in particular true for the displacements of the nuclei around the minimum:

$$\mathbf{R} - \mathbf{R}^0 = \sum_i c_i \mathbf{U}^i$$

It can be shown that Eq. 19 becomes in the new coordinate systems spanned by the $\mathbf{U}^i$'s

$$\left[E^0 + \sum_i \left(\frac{1}{2}\omega_i^2 c_i^2 - \frac{1}{2}\frac{\partial^2}{\partial c_i^2}\right)\right] \Psi^n_{0,\alpha}(\mathbf{c}_1, \ldots, \mathbf{c}_{N_{at}}) = E^{en}_{0,\alpha}\, \Psi^n_{0,\alpha}(\mathbf{c}_1, \ldots, \mathbf{c}_{N_{at}}) \tag{20}$$

The solutions of Eq. 20 are products of 1-dim functions

$$\Psi^n_{0,\alpha}(\mathbf{c}_1, \ldots, \mathbf{c}_{N_{at}}) = \chi_{n_1^\alpha}(c_1) \ldots \chi_{n^\alpha_{3N_{at}}}(c_{3N_{at}})$$

and their energy is given by

$$E^{en}_{0,\alpha} = E^0 + \sum_{i=1}^{3N_{at}} \left(\frac{1}{2} + n_i^\alpha\right) \omega_i \tag{21}$$

Each vibrational excitation is thus characterized by a set of occupation numbers n_i^α. Even if all the occupation numbers are zero, the energy of the system where the nuclei are described by wave-functions is larger than of the system where the nuclei are classical particles. This additional energy is called the zero point motion energy. Since the potential energy $E(\mathbf{R}_1^0,...,\mathbf{R}_{N_{at}}^0)$ does not change under translations of the entire system, there are 3 translation vectors that are eigenvectors of D with eigenvalue ω_i equal to zero. Since these modes give no energy contribution we can take them out of the sum and Eq. 21 becomes

$$E_{0,\alpha}^{en} = E^0 + \sum_{i=1}^{3N_{at}-3} \left(\frac{1}{2} + n_i^\alpha\right)\omega_i \tag{22}$$

A further analysis of the vibrational frequencies of a periodic solid shows that the frequencies form bands that have as quantum numbers the wave-vector **k** and the band index i: $\omega = \omega_i(\mathbf{k})$

In the independent electron picture of a metal, electronic excitations of arbitrarily small energy exist. From what was said before one might suspect that the adiabatic approximation therefore fails for a metal. This is fortunately not true. The phonon spectrum of metals is determined with great accuracy within the adiabatic framework.

Quantum vibrations in molecules

The theory of quantum vibrations in molecules is somewhat more complicated than the theory of phonons in solids. In analogy to the classical case the vibrational excitations are classified into translation, rotations and proper vibrations. Thus $E^{en} - E^0$ is a sum of a translational contribution E^{trans}, a rotational contribution E^{rot} and a vibrational contribution E^{vib}

- Translations
  In the classical context, the motion of a molecule can be split up into a part describing the motion of the center of mass of the system and a part describing the motion of the individual atoms relative to this center of mass. Such a distinction can not be carried over straightforwardly into the quantum mechanical context since the positions of the nuclei (and consequently also the center of mass) are not any more well defined. One can however see what happens to the nucleonic Schrödinger equation if one introduces a new set of coordinates. Introducing the center of mass $\mathbf{R}_{CM}$

  $$\mathbf{R}_{CM} = \sum_{i=1}^{N_{at}} \frac{M_i}{M_T} \mathbf{R}_i$$

  where $M_T = \sum_{i=1}^{N_{at}} M_i$, the positions of the first $N_{at} - 1$ nuclei are defined with respect to this center of mass by

  $$\mathbf{R}'_i = \mathbf{R}_i - \mathbf{R}_{CM}$$

and the position of the N_{at}-th nuclei is then consequently given by

$$\mathbf{R}'_{N_{at}} = -\sum_{i=1}^{N_{at}-1} \frac{M_i}{M_{N_{at}}} (\mathbf{R}_i - \mathbf{R}_{CM})$$

It is then easy to show that the Hamiltonian of Eq. 15 is transformed into

$$\frac{-1}{2M_T}\nabla^2_{\mathbf{R}_{CM}} + \sum_{i=1}^{N_{at}-1} \frac{-1}{2M_i}\nabla^2_{\mathbf{R}'_i} + \sum_{i=1}^{N_{at}-1}\sum_{j=1}^{N_{at}-1} \frac{1}{2M_T}\nabla_{\mathbf{R}'_i}\nabla_{\mathbf{R}'_j} + E_0(\mathbf{R}'_1, ..., \mathbf{R}'_{N_{at}-1}) \tag{23}$$

For a large molecule the total mass M_T is much larger than any individual mass M_i and the 3rd term in Eq. 23 can be neglected. Then we are indeed left with the first term that describes the kinetic energy of the center of mass and the second term that describes the kinetic energy of the motion relative to the center of mass. This second term will later on be split into a rotational and vibrational term.
Exercise [3pts]: *Derive Eq. 23 from Eq. 15*

The eigenfunctions of the center of mass part of the Hamiltonian of Eq. 23

$$\frac{-1}{2M_T}\nabla^2_{\mathbf{R}_{CM}}$$

are plane waves and the energy levels are given by

$$E^{trans} = \frac{\pi^2}{2}\frac{1}{M_T L^2}(n_x^2 + n_y^2 + n_z^2)$$

L is the length of the periodic box that is used in solving the eigenvalue problem and n_x, n_y and n_z are the quantum numbers. Note that the density of states becomes continuous in the limit where L tends to infinity.

- Rotations
  For the calculation of the rotational levels, it is generally assumed that the molecule is a rigid rotator, i.e. one neglects the fact that the molecules stretches due to centrifugal forces that are particularly strong for high rotational excitations. Then the quantum mechanical levels depend only on the moments of inertia of the molecule. In the simplest case of a diatomic molecule they are given by

  $$E^{rot} = \frac{J(J+1)}{2I}$$

  where I is the moment of inertia $I = M_1 R_1^2 + M_2 R_2^2$ and where the origin is chosen to be the center of mass.

- Vibrations
  The treatment of proper vibrations is analogous to the treatment of phonons in

solids. One does a Taylor expansion of the potential energy around the equilibrium positions including the constant and quadratic term. The energy is then expressed in terms of the eigenvalues ω_i^2 of the matrix D. In contrast to the case of periodic boundary condition where the matrix D had 3 zero eigenvalues due to the 3 translations, the matrix D has now 6 zero eigenvalues (unless the molecule is diatomic in which case only 2 rotations exist). 3 are again due to translations and 3 are due to the rotations that leave the energy invariant as well. Hence the vibrational energy is given by

$$E^{vib} = \sum_{i=1}^{3N_{at}-6} \left(\frac{1}{2} + n_i^{\alpha} \right) \omega_i \tag{24}$$

The many electron Schrödinger equation: the physical background

The chemical and physical properties of substances such as atoms, molecules, solids or liquids are given by the laws of many body quantum mechanics. The central equation is the non-relativistic many-electron Schrödinger equation. As has been shown before the corrections due to the quantum nature of the nucleons are relatively small and can be added in an approximative way relatively easily after the potential energy surface has been determined. Up to now we have ignored the fact that electrons have a spin. Spin will from now on be included in our considerations and we introduce therefore the combined spatial and spin variables $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N$ for the N electron system, where $\mathbf{x}_i = (\mathbf{r}_i s_i)$. The spatial part $\mathbf{r}_i = (x_i, y_i, z_i)$ is a continuous variable that can take on any value, s_i is a formal discrete variable. We will assume that it takes on the value $1/2$ for an electron with spin up and $-1/2$ for spin down. The electronic Schrödinger equation takes then the form:

$$\mathcal{H}\Psi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) = E\Psi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) \tag{25}$$

The Operator $\mathcal{H}$ is the many electron Hamiltonian for an N electron system, Ψ the many electron wave-function and E the energy of the system. From a mathematical point of view , Schrödinger's equation is an eigenvalue problem where E is the eigenvalue and Ψ the eigenvector. Unless specified differently we are always interested in the ground-state wave-function Ψ_0 which is associated to the lowest energy eigenvalue E_0. Higher energy eigenstates E_1, Ψ_1 correspond to excited states.

$|\Psi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N)|^2 dr_1 ... dr_N$ is the probability of finding a configuration where electron 1 with spin s_1 is within $\mathbf{r}_1$ and $\mathbf{r}_1 + d\mathbf{r}_1$, electron 2 with spin s_2 is within $\mathbf{r}_2$ and $\mathbf{r}_2 + d\mathbf{r}_2$, etc. $d\mathbf{r}_i$ is a short hand notation for $dx_1\, dy_i\, dz_i$.
Since electrons are indistinguishable particles the probability of finding electron 1 in the neighborhood of $\mathbf{r}_1$ and electron 2 in the neighborhood of $\mathbf{r}_2$ must equal the probability of finding electron 1 in the neighborhood of $\mathbf{r}_2$ and electron 2 in the neighborhood of $\mathbf{r}_1$. Since the probability must be equal not only under the exchange of electron 1 and 2, but under exchange of any pair i, j we have

$$|\Psi(..., \mathbf{x}_i, ..., \mathbf{x}_j, ...)|^2 = |\Psi(..., \mathbf{x}_j, ..., \mathbf{x}_i, ...)|^2$$

This constraint can be satisfied by a wave-function that is either symmetric or antisymmetric under the exchange of two coordinates. It turns out that a symmetric wave-function corresponds to a boson, whereas an antisymmetric wave-function corresponds to a fermion. Electrons are fermions and hence we have

$$\Psi(..., \mathbf{x}_i, ..., \mathbf{x}_j, ...) = -\Psi(..., \mathbf{x}_j, ..., \mathbf{x}_i, ...)$$

Exercise [2pts]: *One can not only exchange one electron pair, but many electron pairs, one after the other. How many possibilities do exist to arrange the indices of the coordinates without changing the probability?*

Because of its interpretation as a probability we can normalize the eigenvector of the Schrödinger equation such that

$$\int d\mathbf{x}_1 ... \int d\mathbf{x}_N \Psi^*(\mathbf{x}_1, ..., \mathbf{x}_N)\Psi(\mathbf{x}_1, ..., \mathbf{x}_N) = 1 \tag{26}$$

The non-relativistic many electron Hamiltonian is spin independent and has the form (Eq. 8)

$$\mathcal{H} = \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^{N}\sum_{j=1}^{i-1} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i=1}^{N}\sum_{j=1}^{N_{at}} \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} + \sum_{i=1}^{N_{at}}\sum_{j=1}^{i-1} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \tag{27}$$

The first term containing the Laplacian represents the kinetic energy of the electrons, the second term the electrostatic repulsion between the electrons, the third term the attraction of the electrons to the nuclei and the fourth term the classical electrostatic repulsion between the nuclei. In the Born Oppenheimer approximation that we have adopted here, the nuclei are treated as classical point particles with charge Z_j, whose position is given by R_j.

Exercise [2pts]: *Show that the Hamiltonian (Eq. 27) is a hermitian operator*

$$\int d\mathbf{x}_1 ... \int d\mathbf{x}_N \Psi_a^*(\mathbf{x}_1, ..., \mathbf{x}_N)\mathcal{H}\Psi_b(\mathbf{x}_1, ..., \mathbf{x}_N) \tag{28}$$

$$= \left[\int d\mathbf{x}_1 ... \int d\mathbf{x}_N \Psi_b^*(\mathbf{x}_1,...,\mathbf{x}_N)\mathcal{H}\Psi_a(\mathbf{x}_1,...,\mathbf{x}_N)\right]^* \tag{29}$$

Exercise [1pts]: *Show that different eigenstates (Ψ_0, Ψ_1, Ψ_2, ...) are orthogonal to each other*

$$\int d\mathbf{x}_1 ... \int d\mathbf{x}_N \Psi_i^*(\mathbf{x}_1,...,\mathbf{x}_N)\Psi_j(\mathbf{x}_1,...,\mathbf{x}_N) = \delta_{i,j}$$

Exercise [1pts]: *Show that the energy can be written as the following expectation value*

$$E = \int d\mathbf{x}_1 ... \int d\mathbf{x}_N \Psi^*(\mathbf{x}_1,...,\mathbf{x}_N)\mathcal{H}\Psi(\mathbf{x}_1,...,\mathbf{x}_N)$$

Exercise [1pts]: *Show that the expectation value of the kinetic energy term is always positive*

$$E_{kin} = \int d\mathbf{x}_1 ... \int d\mathbf{x}_N \Psi^*(\mathbf{x}_1,...,\mathbf{x}_N)\left(\sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2\right)\Psi(\mathbf{x}_1,...,\mathbf{x}_N)$$

Exercise [2pts]: *Show that the expectation value of the term representing the interaction* *between the nuclei gives simply the classical expression*

$$\begin{aligned} E_{n-n} &= \int d\mathbf{x}_1 ... \int d\mathbf{x}_N \Psi^*(\mathbf{x}_1, ..., \mathbf{x}_N) \left(\sum_{i=1}^{N_{at}} \sum_{j=1}^{i-1} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \right) \Psi(\mathbf{x}_1, ..., \mathbf{x}_N) \\ &= \sum_{i=1}^{N_{at}} \sum_{j=1}^{i-1} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \end{aligned} \tag{30}$$

Additional terms need to be added to the Hamiltonian if external fields are present. If for instance an external electric field is applied that is associated with a potential $V_{ext}(\mathbf{r})$ then the following term has to be added:

$$\sum_i V_{ext}(\mathbf{r}_i)$$

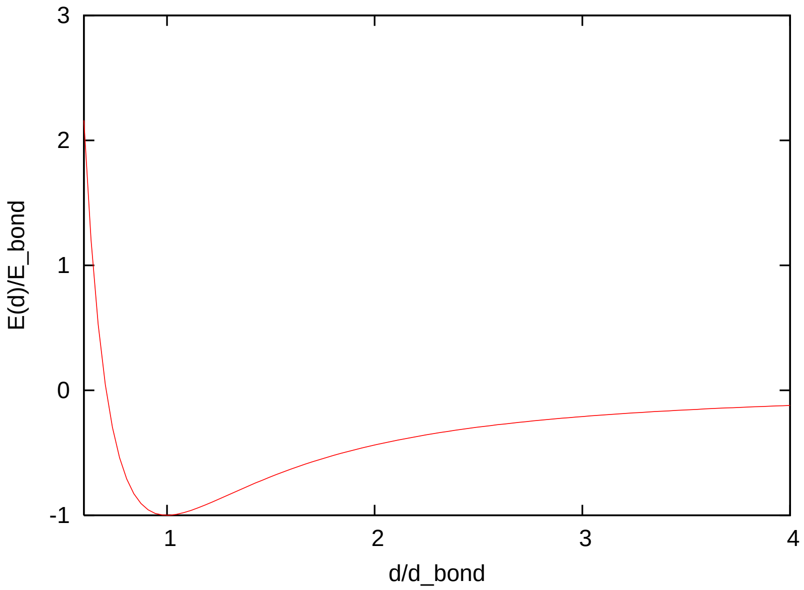
Electronic structure calculations solve the many electron Schrödinger equation in its exact form (Eq. 25) or in approximate forms. From a formal point of view, the input for an electronic structure calculation specifies the positions $\mathbf{R_j}$ for the different kind of atoms characterized only by their charge Z. The output is the total energy E and the wave-function Ψ. Evidently the Born-Oppenheimer surface is the total energy as a function of the atomic coordinates. Many other physical observables can be derived from this output.

Example: The hydrogen molecule H_2

Without restriction we can assume that one nucleus is located at the origin and the other at a distance d along the z-axis. The resulting Hamiltonian is then given by

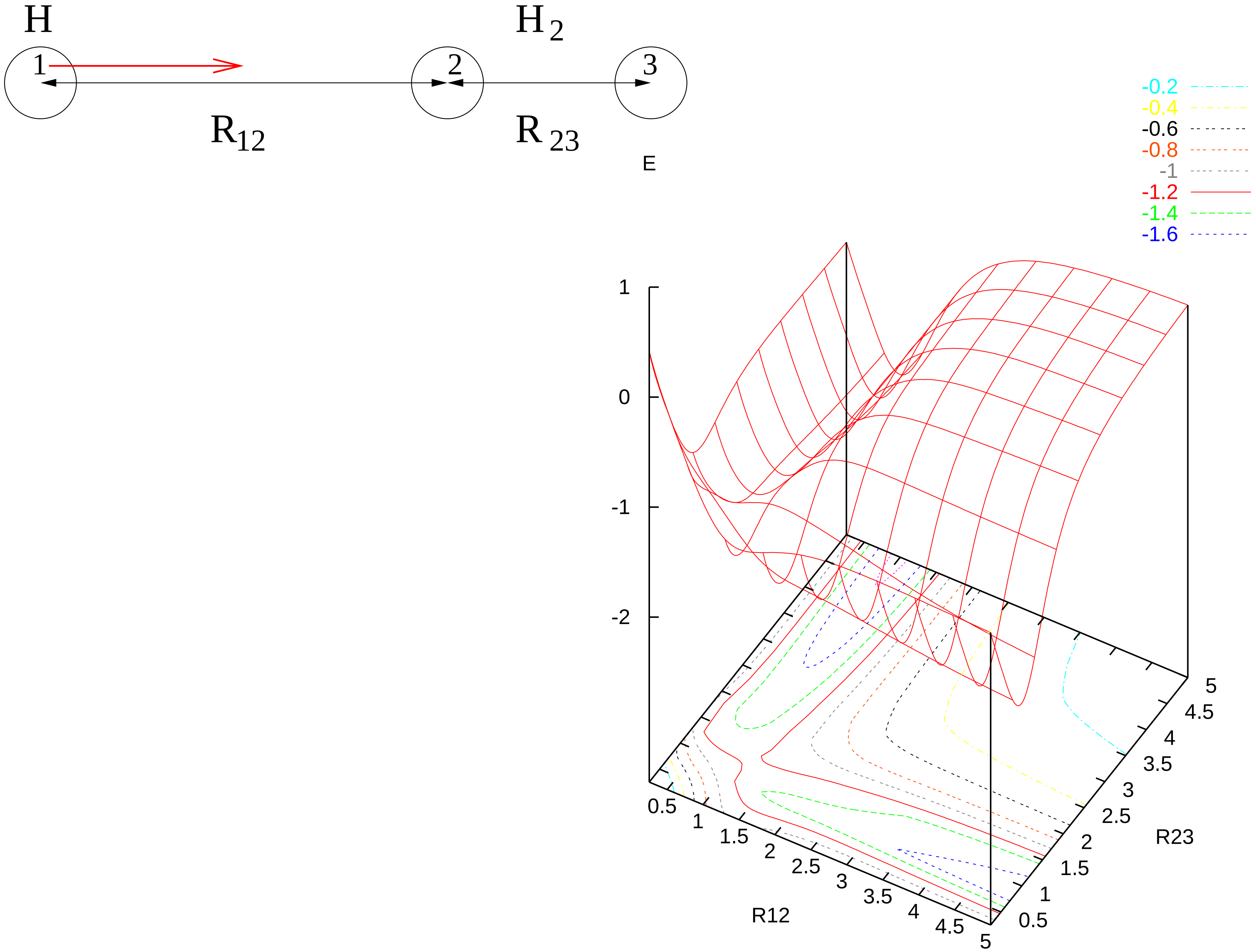
$$\mathcal{H} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} - \frac{1}{|\mathbf{r}_1|} - \frac{1}{|\mathbf{r}_1 - \mathbf{d}\hat{\mathbf{z}}|} - \frac{1}{|\mathbf{r}_2|} - \frac{1}{|\mathbf{r}_2 - \mathbf{d}\hat{\mathbf{z}}|} + \frac{1}{d} \qquad (31)$$

One can solve the many electron Schrödinger equation for various distances d and obtain for each distance a total energy $E(d)$ and the associated wave-function. The potential energy surface (or Born Oppenheimer surface) $E(d)$ of the electronic ground state has the following form



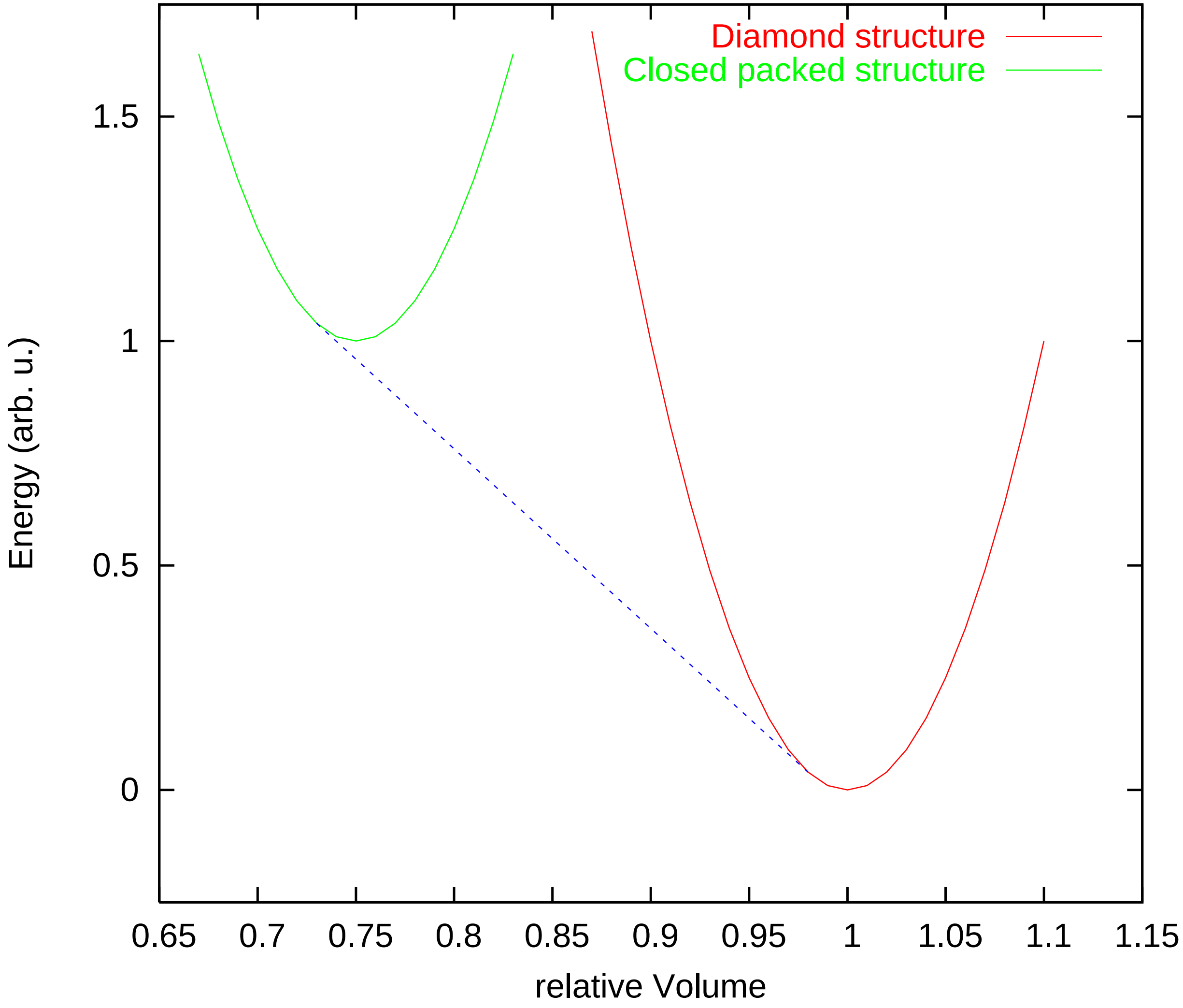
The minimum of this curve give the bond-length d_{bond} and the binding energy E_{bond}

Example: The chemical reaction $H + H_2 \to H_2 + H$



Example: Equilibrium lattice constant, pressure, phase transitions of crystalline materials

In a crystalline (periodic) material one can scale all the atomic coordinates and thus compress or expand the crystal. In this way one can calculate the total energy as a function of the volume of the crystal. Such a plot is schematically given below for the case of silicon in 2 crystalline phases.



Such a plot contains a lot of information. The volume for which each curve takes on a minimum determines the equilibrium volume at zero pressure determining the lattice constant of the material. The curve with the lower minimum determines which structure is the thermodynamically stable one at zero temperature. This does however not necessarily mean that the other energetically higher structure can not be found in nature. It can be a meta-stable structure. Such a case is encountered in the case of carbon. The graphite structure is lower in energy than the diamond structure, but the time scale for the spontaneous transformation from diamond into graphite is so long at ambient pressure and temperature that nobody has to worry about the stability of his/her diamond. The slope of the energy versus volume (Ω) curve gives the pressure P of the two different phases at a certain volume since

$$P = -\frac{\partial E}{\partial \Omega}$$

The blue tangent line gives the pressure at which a phase transformation takes place.

The variational principle

The ground state wave function Ψ_0 is the wave-function Ψ which minimizes

$$E = \int d\mathbf{x}_1 ... \int d\mathbf{x}_N \Psi^*(\mathbf{x}_1, ..., \mathbf{x}_N) \mathcal{H} \Psi(\mathbf{x}_1, ..., \mathbf{x}_N)$$

Proof: Since the eigenvectors of a Hermitian matrix form a complete set we can expand Ψ in terms of the eigenstates Ψ_i

$$\Psi(\mathbf{x}_1, ..., \mathbf{x}_N) = \sum_i c_i \Psi_i(\mathbf{x}_1, ..., \mathbf{x}_N) \tag{32}$$

The normalization condition (Eq. 26) implies

$$\sum_i c_i^2 = 1$$

The expectation value for the energy is given by

$$E = \sum_i E_i c_i^2$$

Exercise [2pts]: *Prove the 2 above equations*

Evidently the minimum is obtained if $c_0 = 1$ with all other coefficients being zero. Hence $E = E_0$ and $\Psi = \Psi_0$

Generalizations of the variational principle

- The variational principle does not only hold if the wave-function is written as a linear combination of basis functions (Eq. 32) but also for any nonlinear parameterization of a wave-function.

- The variational principle can be generalized to excited states. It can be shown that the M-th excited state Ψ_M minimizes the energy expectation values

  $$E_M = \int d\mathbf{x}_1 ... \int d\mathbf{x}_N \Psi^*_M(\mathbf{x}_1, ..., \mathbf{x}_N) \mathcal{H} \Psi_M(\mathbf{x}_1, ..., \mathbf{x}_N)$$

  under the normalization constraint

  $$\int d\mathbf{x}_1 ... \int d\mathbf{x}_N \Psi^*_M(\mathbf{x}_1, ..., \mathbf{x}_N) \Psi_M(\mathbf{x}_1, ..., \mathbf{x}_N) = 1$$

  and the additional constraint that Ψ_M is orthogonal to all $M-1$ lower eigenstates Ψ_m, $m = 1, .., M-1$

  $$\int d\mathbf{x}_1 ... \int d\mathbf{x}_N \Psi^*_M(\mathbf{x}_1, ..., \mathbf{x}_N) \Psi_m(\mathbf{x}_1, ..., \mathbf{x}_N) = 0$$

Density matrices

The density matrix of a many electron quantum state is defined as

$$\gamma(\mathbf{x}_1, ..., \mathbf{x}_N; \mathbf{x}'_1, ..., \mathbf{x}'_N) = \Psi(\mathbf{x}_1, ..., \mathbf{x}_N)\Psi^*(\mathbf{x}'_1, ..., \mathbf{x}'_N)$$

Since a many electron wave-function is a high dimensional and therefore highly complicated quantity, it is difficult to use it for a physical interpretation of numerical results. Reduced density matrices are most helpful in this context since they are lower dimensional. Particularly useful are the first and second order reduced density matrices γ_1 and γ_2 often also called one-particle and two-particle density matrices.

$$\gamma_1(\mathbf{x}_1; \mathbf{x}'_1) = N \int d\mathbf{x}_2 ... \int d\mathbf{x}_N \Psi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N)\Psi^*(\mathbf{x}'_1, \mathbf{x}_2, ..., \mathbf{x}_N)$$

$$\gamma_2(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) = N(N-1) \int d\mathbf{x}_3 ... \int d\mathbf{x}_N \Psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, ..., \mathbf{x}_N)\Psi^*(\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}_3, ..., \mathbf{x}_N)$$

The spin charge density γ is the diagonal part of γ_1

$$\gamma(\mathbf{x}) = \gamma_1(\mathbf{x}; \mathbf{x})$$

It represents the probability of finding some electron at $\mathbf{x}$ and is thus the expectation value of $\sum_i \delta(\mathbf{x} - \mathbf{x}_i)$

Exercise [2pts]: *Demonstrate the above statement*

Exercise [1pts]: *Show that*

$$\int d\mathbf{x}_1 \gamma(\mathbf{x}_1) = N$$

Exercise [2pts]: *Express* γ_1 *in terms of* γ_2

The diagonal of the two-particle density matrix

$$\gamma_2(\mathbf{x},\mathbf{y}) = \gamma_2(\mathbf{x},\mathbf{y};\mathbf{x},\mathbf{y})$$

represents the conditional probability of finding an electron at $\mathbf{x}$ if there is another electron at $\mathbf{y}$. It is thus the expectation value of $\sum_i \sum_{i \neq j} \delta(\mathbf{x}-\mathbf{x}_i)\delta(\mathbf{y}-\mathbf{x}_j)$

The pair correlation function h is defined by the equation

$$\gamma_2(\mathbf{x},\mathbf{y}) = \rho(\mathbf{x})\rho(\mathbf{y})\left[1 + h(\mathbf{x},\mathbf{y})\right] \tag{33}$$

Spin-less density matrices

Even though the spin variable is a discrete variable it is common practice to denote a sum over the two realizations of this quantity by an integral

Two particle spin-less density matrix

$$\rho_2(\mathbf{r}_1,\mathbf{r}_2;\mathbf{r}'_1,\mathbf{r}'_2) = \int ds_1 ds_2 \gamma_2(\mathbf{r}_1 s_1,\mathbf{r}_2 s_2;\mathbf{r}'_1 s_1,\mathbf{r}'_2 s_2)$$

One particle spin-less density matrix

$$\rho_1(\mathbf{r}_1;\mathbf{r}'_1) = \int ds_1 \gamma_1(\mathbf{r}_1 s_1;\mathbf{r}'_1 s_1) \tag{34}$$

Charge density

$$\rho(\mathbf{r}_1) = \rho_1(\mathbf{r}_1;\mathbf{r}_1) \tag{35}$$

Natural orbitals and natural occupation numbers

The natural occupation numbers n_i and natural orbitals ψ_i are the eigenvalues and eigenvectors of the one-particle density matrix

$$\int d\mathbf{r}' \rho_1(\mathbf{r},\mathbf{r}')\psi_i(\mathbf{r}') = n_i \psi_i(\mathbf{r}) \tag{36}$$

The total energy in terms of density matrices

$$\begin{aligned}
E_{kin} &= \int d\mathbf{x}_1 ... \int d\mathbf{x}_N \, \Psi^*(\mathbf{x}_1, ..., \mathbf{x}_N) \left(\sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 \right) \Psi(\mathbf{x}_1, ..., \mathbf{x}_N) \\
&= \int \left[-\frac{1}{2}\nabla_1^2 \gamma_1(\mathbf{x}_1; \mathbf{x}'_1) \right]_{\mathbf{x}_1 = \mathbf{x}'_1} d\mathbf{x}_1 = \int \left[-\frac{1}{2}\nabla_1^2 \rho_1(\mathbf{r}_1; \mathbf{r}'_1) \right]_{\mathbf{r}_1 = \mathbf{r}'_1} d\mathbf{r}'_1 \qquad (37) \\
E_{e-e} &= \int d\mathbf{x}_1 ... \int d\mathbf{x}_N \, \Psi^*(\mathbf{x}_1, ..., \mathbf{x}_N) \left(\sum_{i=1}^{N} \sum_{j=1}^{i-1} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right) \Psi(\mathbf{x}_1, ..., \mathbf{x}_N) \\
&= \frac{1}{2} \int \int \frac{\gamma_2(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}_1, \mathbf{x}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{x}_1 d\mathbf{x}_2 = \frac{1}{2} \int \int \frac{\rho_2(\mathbf{r}_1, \mathbf{r}_2; \mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \qquad (38) \\
E_{e-n} &= \int d\mathbf{x}_1 ... \int d\mathbf{x}_N \, \Psi^*(\mathbf{x}_1, ..., \mathbf{x}_N) \left(\sum_{i=1}^{N} \sum_{j=1}^{N_{at}} \frac{-Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} \right) \Psi(\mathbf{x}_1, ..., \mathbf{x}_N) \\
&= \int \gamma_1(\mathbf{x}_1; \mathbf{x}_1) \sum_{j=1}^{N_{at}} \frac{-Z_j}{|\mathbf{r}_1 - \mathbf{R}_j|} d\mathbf{x}_1 = \int \rho_1(\mathbf{r}_1; \mathbf{r}_1) \sum_{j=1}^{N_{at}} \frac{-Z_j}{|\mathbf{r}_1 - \mathbf{R}_j|} d\mathbf{r}_1 \qquad (39)
\end{aligned}$$

Exercise [3pts]: *Derive the 3 above equations*

The curse of dimensionality

What makes the many electron problem so difficult is its high dimensionality. The many electron wave-function for an N electron system is a $3N$ dimensional object taking into account only the spatial part. Discretizing the wave-function along each dimension by K points would require to store the values of the wave-function on K^{3N} grid point. Even for moderate values such as $K = 10, N = 100$ the storage requirements would by far exceed the memory of the largest computers. As we will see the antisymmetry constraint will slightly reduce the storage requirements, but nevertheless the dimensionality problem remains the main obstacle for electronic structure calculations.

The N-representability problem

The fact that the total energy can be written in terms of the two particle density matrix, which is only a 6-dimensional quantity, and quantities that can be derived from it (one particle density matrix and density) suggests that one might solve the electronic structure problem by minimizing the energy expression with respect to the two particle density matrix. Unfortunately such a minimization gives a too low ground state energy. The reason for this is that there is some hidden constraint, namely that the two particle density matrix can be obtained from a N-electron wave-function. A two particle density matrix that can be obtained from a wave-function is called N-representable. No simple sufficient conditions for a two particle density matrix to be N-representable are known. For the one particle density matrix on the other hand the conditions are known. A spin-less one particle density matrix is N-representable if and only if all the eigenvalues n_i are in the interval [0:2]. If the one particle density matrix contains spin the interval is [0:1].

The Hellmann Feynman theorem

Let us consider an Hamiltonian H that depends on a parameter λ. Typically this parameter is the position of a nucleus or the strength of an electric field. We then have also λ dependent eigenvalues and vectors that satisfy

$$H(\lambda)\Psi(\lambda,x) = E(\lambda)\Psi(\lambda,x)$$

This theorem is valid both for many electron and single electron wave-functions and the variable x can therefore denote any set of coordinates of a single or many electron system. The Hellmann Feynman theorem then states that

$$\frac{\partial E(\lambda)}{\partial\lambda} = \int dx\,\Psi^*(\lambda,x)\frac{\partial H(\lambda)}{\partial\lambda}\Psi(\lambda,x) \tag{40}$$

Proof:

$$\begin{aligned}
\frac{\partial E(\lambda)}{\partial\lambda} &= \frac{\partial}{\partial\lambda}\int dx\,\Psi^*(\lambda,x)H(\lambda)\Psi(\lambda,x) \\
&= \int dx\,\Psi^*(\lambda,x)\frac{\partial H(\lambda)}{\partial\lambda}\Psi(\lambda,x) + \int dx\,\frac{\partial\Psi^*(\lambda,x)}{\partial\lambda}H(\lambda)\Psi(\lambda,x) + \int dx\,\Psi^*(\lambda,x)H(\lambda)\frac{\partial\Psi(\lambda,x)}{\partial\lambda} \\
&= \int dx\,\Psi^*(\lambda,x)\frac{\partial H(\lambda)}{\partial\lambda}\Psi(\lambda,x) + E(\lambda)\int dx\,\Psi^*(\lambda,x)\frac{\partial\Psi(\lambda,x)}{\partial\lambda} + E(\lambda)\int dx\,\Psi(\lambda,x)\frac{\partial\Psi^*(\lambda,x)}{\partial\lambda}
\end{aligned}$$

$$
\begin{aligned}
&= \int dx\,\Psi^*(\lambda,x)\frac{\partial H(\lambda)}{\partial\lambda}\Psi(\lambda,x)+E(\lambda)\frac{\partial}{\partial\lambda}\int dx\,\Psi^*(\lambda,x)\Psi(\lambda,x)\\
&= \int dx\,\Psi^*(\lambda,x)\frac{\partial H(\lambda)}{\partial\lambda}\Psi(\lambda,x)
\end{aligned}
$$

Let us now apply the Hellmann Feynman theorem to the case where λ is a component of an atomic position $\mathbf{R}_j$. Without restriction we can consider the first component of the atom number one: $\lambda = X_1$ From Eqs. 40, 27, 39, 30 we obtain

$$
\begin{aligned}
f_1^x = -\frac{\partial E}{\partial X_1} &= \frac{\partial}{\partial X_1}\left(\int \rho(\mathbf{r}_1)\sum_{j=1}^{N_{at}}\frac{Z_j}{|\mathbf{R}_j-\mathbf{r}_1|}d\mathbf{r}_1 - \sum_{i=1}^{N_{at}}\sum_{j=1}^{i-1}\frac{Z_iZ_j}{|\mathbf{R}_i-\mathbf{R}_j|}\right)\\
&= \int \rho(\mathbf{r})\frac{\partial}{\partial X_1}\left(\sum_{j=1}^{N_{at}}\frac{Z_j}{|\mathbf{R}_j-\mathbf{r}|}\right)d\mathbf{r} - \frac{\partial}{\partial X_1}\left(\frac{1}{2}\sum_{i=1}^{N_{at}}\sum_{j=1}^{N_{at}}{}'\frac{Z_iZ_j}{|\mathbf{R}_i-\mathbf{R}_j|}\right)\\
&= -Z_1\int \rho(\mathbf{r})\frac{X_1-x}{|\mathbf{R}_1-\mathbf{r}|^3}d\mathbf{r}+\sum_{j=2}^{N_{at}}\frac{Z_1Z_j(X_1-X_j)}{|\mathbf{R}_1-\mathbf{R}_j|^3} \qquad (41)
\end{aligned}
$$

The forces acting on the nuclei due to the influence of the electrons described as quantum mechanical particles depends only on the probability distribution ρ and are identical to the forces that would arise from a classical charge distribution ρ.

Exercise [1pts]: *Write down the formula for the y component of the 2nd nucleus*

The Hellmann Feynman theorem is not only valid for the exact many electron Hamiltonian but it is also valid in the case where approximate Hamiltonians are used such as the Hartree Fock or Density Functional Hamiltonian.

The total energy of the electronic ground state as a function of coordinates of the atoms in a molecule or solid $E(\mathbf{R}_1, ..., \mathbf{R}_{N_{at}})$ is called the potential energy surface or Born-Oppenheimer surface. Points where the forces vanish i.e. $\mathbf{f}_i = 0$ are stationary points of this surface and have a physical importance. The global minimum of the potential energy surface gives the geometry of the configurational ground state of the molecule. Other local minima give other stable geometries. Saddle points are transition states in chemical reactions. For our example of the H_2 molecule we have one ground state geometry, for the $H_2 + H$ system we have one transition state in addition to the two configurational ground states.

Finding a configurational ground state with the Hellmann Feynman theorem is done in the following way. One starts with an initial guess for the atomic positions. For these atomic positions one solves the Schrödinger equation (Eq. 25). From the electronic charge density $\rho(\mathbf{r})$ one then obtains via the Hellmann Feynamn theorem the forces acting on the nuclei. The nuclei are then moved in the direction of the forces to obtain new atomic positions. The above steps are then repeated for each new atomic configuration until the forces on the nuclei vanish. Finding the ground state geometry of a molecule or solid requires thus solving the electronic Schrödinger equation (Eq. 25) for several non-equilibrium atomic configurations until the equilibrium configuration is found.

The virial theorem

Let us consider a molecule that is in its electronic and configurational ground state. As a consequence its energy will increase if either its wave-function is modified and/or if the atomic coordinates are modified. One such particular modification is a scaling of all the electronic and configurational coordinates, $\mathbf{r}_i \to \lambda \mathbf{r}_i$ and $\mathbf{R}_i \to \lambda^{-1}\mathbf{R}_i$. Since the wave-function has to remain normalized we have

$$\Psi(\mathbf{r}_1 s_1, \mathbf{r}_2 s_2, ...) \to \lambda^{3N/2}\Psi(\lambda\mathbf{r}_1 s_1, \lambda\mathbf{r}_2 s_2, ...)$$

Exercise [1pts]: *Prove that the wave-function remains normalized under this transformation*

Hence the energy becomes

$$\begin{aligned} E(\lambda) = \quad & - \quad \lambda^3 \int \left[\frac{1}{2}\nabla_1^2 \rho_1(\lambda\mathbf{r}_1; \lambda\mathbf{r}'_1)\right]_{\mathbf{r}_1=\mathbf{r}'_1} d\mathbf{r}_1 + \lambda^6 \int\int \frac{1}{2}\frac{\rho_2(\lambda\mathbf{r}_1, \lambda\mathbf{r}_2; \lambda\mathbf{r}_1, \lambda\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \\ & - \quad \lambda^3 \int \rho_1(\lambda\mathbf{r}_1; \lambda\mathbf{r}_1) \sum_{j=1}^{N_{at}} \frac{Z_j}{|\mathbf{r}_1 - \lambda^{-1}\mathbf{R}_j|} d\mathbf{r}_1 + \sum_{i=1}^{N_{at}} \sum_{j=1}^{i-1} \frac{Z_i Z_j}{\lambda^{-1}|\mathbf{R}_i - \mathbf{R}_j|} \end{aligned}$$

Introducing a new variable $\mathbf{t} = \lambda \mathbf{r}$ we get

$$
\begin{aligned}
E(\lambda) = & - \lambda^2 \int \left[\frac{1}{2} \nabla_1^2 \rho_1(\mathbf{t}_1; \mathbf{t}'_1) \right]_{\mathbf{t}_1 = \mathbf{t}'_1} d\mathbf{t}_1 + \lambda \int \int \frac{1}{2} \frac{\rho_2(\mathbf{t}_1, \mathbf{t}_2; \mathbf{t}_1, \mathbf{t}_2)}{|\mathbf{t}_1 - \mathbf{t}_2|} d\mathbf{t}_1 d\mathbf{t}_2 \\
& - \lambda \int \rho_1(\mathbf{t}_1; \mathbf{t}_1) \sum_{j=1}^{N_{at}} \frac{Z_j}{|\mathbf{t}_1 - \mathbf{R}_j|} d\mathbf{t}_1 + \lambda \sum_{i=1}^{N_{at}} \sum_{j=1}^{i-1} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \\
= & \quad \lambda^2 E_{kin} + \lambda E_{pot}
\end{aligned}
$$

The stationary point condition gives

$$
0 = \left. \frac{\partial E}{\partial \lambda} \right|_{\lambda=1} = 2E_{kin} + E_{pot}
$$

Together with the fact that $E = E_{kin} + E_{pot}$ one thus obtains the virial theorem

$$
E = -E_{kin} = \frac{1}{2} E_{pot} \tag{42}
$$

Slater determinant wave-functions

Writing the many electron wave function Ψ as a sum of Slater determinants Ψ_I is the easiest way to impose antisymmetry. Such an ansatz is called configuration interaction ansatz, since a determinant can be considered as an electronic configuration.

$$\Psi(\mathbf{x}_1, ..., \mathbf{x}_N) = \sum_I C_I \underbrace{\frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{i_1}(\mathbf{x}_1) & ... & \psi_{i_N}(\mathbf{x}_1) \\ ... & ... & ... \\ \psi_{i_1}(\mathbf{x}_N) & ... & \psi_{i_N}(\mathbf{x}_N) \end{vmatrix}}_{\Psi_I(\mathbf{x}_1,...,\mathbf{x}_i,...,\mathbf{x}_j,...,\mathbf{x}_N)} \tag{43}$$

$\psi_i(\mathbf{x})$ are orthonormal single particle spin orbitals, i.e. $\int \psi_i(\mathbf{x})\psi_j(\mathbf{x})d\mathbf{x} = \delta_{i,j}$ and I is a composite index $I = (i_1, ...i_N)$. The most general form of a spin orbital is

$$\psi_j(\mathbf{x}) = \psi_j^{\alpha}(\mathbf{r})\alpha(s) + \psi_j^{\beta}(\mathbf{r})\beta(s)$$

where $\alpha(s)$ and $\beta(s)$ are the two realizations of the spin function $\sigma(s)$ representing a spin up and spin down electron.

A Slater determinant wave-function is obviously antisymmetric

$$
\begin{aligned}
\Psi_I(\mathbf{x}_1,...,\mathbf{x}_i,...,\mathbf{x}_j,...,\mathbf{x}_N) &= \begin{vmatrix} \psi_{i_1}(\mathbf{x}_1) & ... & \psi_{i_N}(\mathbf{x}_1) \\ ... & ... & ... \\ \psi_{i_1}(\mathbf{x}_i) & ... & \psi_{i_N}(\mathbf{x}_i) \\ ... & ... & ... \\ \psi_{i_1}(\mathbf{x}_j) & ... & \psi_{i_N}(\mathbf{x}_j) \\ ... & ... & ... \\ \psi_{i_1}(\mathbf{x}_N) & ... & \psi_{i_N}(\mathbf{x}_N) \end{vmatrix} \\
&= - \begin{vmatrix} \psi_{i_1}(\mathbf{x}_1) & ... & \psi_{i_N}(\mathbf{x}_1) \\ ... & ... & ... \\ \psi_{i_1}(\mathbf{x}_j) & ... & \psi_{i_N}(\mathbf{x}_j) \\ ... & ... & ... \\ \psi_{i_1}(\mathbf{x}_i) & ... & \psi_{i_N}(\mathbf{x}_i) \\ ... & ... & ... \\ \psi_{i_1}(\mathbf{x}_N) & ... & \psi_{i_N}(\mathbf{x}_N) \end{vmatrix} = -\Psi_I(\mathbf{x}_1,...,\mathbf{x}_j,...,\mathbf{x}_i,...,\mathbf{x}_N)
\end{aligned}
$$

Using a basis set of M orbitals ψ_i, there are $M(M-1)(M-2)...(M-(N-1))$ possibilities to arrange the indices $i_1,...,i_N$. Since the determinant does not depend on the ordering (except for the sign), we can constrain the sum $\sum_I = \sum_{i_1,...,i_N}$ to $\sum'_I = \sum_{i_1<i_2<...<i_N}$ and we get $\frac{M!}{(M-N)!N!}$ possible configurations. The curse of dimensionality has not disappeared!

The Ψ_I's form an orthonormal basis set for $i_1 < i_2 < ... < i_N$. For simplicity we consider only real (not complex) wave-functions

$$\begin{aligned}
\int d\mathbf{x}_1,...d\mathbf{x}_N\,\Psi_I(\mathbf{x}_1,...\mathbf{x}_N)\Psi_J(\mathbf{x}_1,...\mathbf{x}_N) &= \frac{1}{N!}\int d\mathbf{x}_1,...d\mathbf{x}_N \\
\Big(\psi_{i_1}(\mathbf{x}_1)\psi_{i_2}(\mathbf{x}_2)\psi_{i_3}(\mathbf{x}_3)...\psi_{i_N}(\mathbf{x}_N) &- \psi_{i_2}(\mathbf{x}_1)\psi_{i_1}(\mathbf{x}_2)\psi_{i_3}(\mathbf{x}_3)...\psi_{i_N}(\mathbf{x}_N)+....\Big) \\
\Big(\psi_{j_1}(\mathbf{x}_1)\psi_{j_2}(\mathbf{x}_2)\psi_{j_3}(\mathbf{x}_3)...\psi_{j_N}(\mathbf{x}_N) &- \psi_{j_2}(\mathbf{x}_1)\psi_{j_1}(\mathbf{x}_2)\psi_{j_3}(\mathbf{x}_3)...\psi_{j_N}(\mathbf{x}_N)+....\Big) = \delta_{I,J}
\end{aligned}$$

Calculation of the Configuration Interaction Hamiltonian

The Schrödinger equation

$$\mathcal{H}\Psi = E\,\Psi$$

together with the ansatz

$$\Psi = \sum_J C_J \Psi_J \tag{44}$$

gives

$$\mathcal{H}\sum_J C_J \Psi_J = E\sum_J C_J \Psi_J$$

Multiplying from the left by Ψ_I and integrating we obtain

$$\sum_J H^{CI}_{I,J} C_J = E\,C_I \tag{45}$$

where

$$H^{CI}_{I,J} = \int d\mathbf{x}_1, \ldots d\mathbf{x}_N \Psi^*_I(\mathbf{x}_1, \ldots \mathbf{x}_N)\mathcal{H}\Psi_J(\mathbf{x}_1, \ldots \mathbf{x}_N)$$

$H^{CI}_{I,J}$ is given by the Slater Condon rules:

- if $I = J$

$$H^{CI}_{I,J} = \sum_{i=i_1,\ldots,i_N} [i|h|i] + \frac{1}{2} \sum_{i=i_1,\ldots,i_N} \sum_{j=i_1,\ldots,i_N} [i,i|j,j] - [i,j|j,i] \tag{46}$$

- if I and J differ by one index $i_k \neq j_k$

$$H^{CI}_{I,J} = [i_k|h|j_k] + \sum_{j=i_1,\ldots,i_N} [i_k, j_k|j,j] - [i_k, j|j, j_k] \tag{47}$$

- if I and J differ by two indices $i_k \neq j_k$, $i_l \neq j_l$

$$H^{CI}_{I,J} = [i_k, j_k|i_l, j_l] - [i_k, j_l|i_l, j_k] \tag{48}$$

- else

$$H^{CI}_{I,J} = 0 \tag{49}$$

where

$$[i|h|j] = \int d\mathbf{x}\, \psi^*_i(\mathbf{x}) \left(-\frac{1}{2}\nabla^2 + \sum_{l=1}^{N_{at}} \frac{-Z_l}{|\mathbf{r} - \mathbf{R}_l|}) \right) \psi_j(\mathbf{x})$$

$$[i,j|k,l] = \int\int d\mathbf{x}\, d\mathbf{x}' \, \frac{\psi^*_i(\mathbf{x})\, \psi_j(\mathbf{x})\, \psi^*_k(\mathbf{x}')\, \psi_l(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} \tag{50}$$

Configuration interaction calculations

There are two conceptually important versions of Configuration Interaction (CI) calculations

- In the full CI method one uses M given spin-orbitals and finds the set of expansion coefficients C_J of Eq. 44 that minimize the energy. This minimization problem gives rise to the eigenvalue problem of Eq. 45. The dimension of the eigenvalue is typically several billions.

- In the multi-reference self-consistent field method (MRSCF) one finds not only the optimal coefficients C_J but also the optimal set of orbitals $\psi_i(\mathbf{x})$. In addition to the eigenvalue problem one also has a set of differential equations for the spin-orbitals. The MRSCF is in principle the most accurate method for solving the many electron Schrödinger equation. To obtain the same accuracy with a full CI methods one has to use a much larger set of fixed orbitals $\psi_i(\mathbf{x})$. A widely used method is the ordinary self-consistent field method (SCF) method where one has only one determinant. It can thus be considered as a special case of the MRSCF method, even though it is not really a configuration interaction method since there is only one configuration. The SCF method will be discussed in detail later.

Spin properties of many electron wave-functions

The non-relativistic many electron Hamiltonian 27 does not depend on any spin variable. Hence the Hamiltonian commutes with the total spin operators

$$S_z = \sum_{i=1}^{N} s_z(i)$$

and

$$S^2 = S_x^2 + S_y^2 + S_z^2$$

where S_x and S_y are defined analogously to S_z. As a consequence any eigenstate and in particular the ground state Ψ_0 of $\mathcal{H}$ has to be an eigenfunction of S_z and S^2.

Reminder of single particle spin properties

Any spin function or spinor $\sigma(s)$ is a linear combination of the two spinors $\alpha(s)$ and $\beta(s)$ which represent a spin up and spin down electron. We have

$$\begin{aligned} \alpha(+1/2) &= 1 \qquad & \alpha(-1/2) &= 0 \\ \beta(+1/2) &= 0 \qquad & \beta(-1/2) &= 1 \end{aligned} \tag{51}$$

and we can therefore represent α and β as 2 component vectors

$$\alpha \quad = \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The Pauli spin operators are consequently 2 by 2 matrices

$$\begin{aligned}
s_x &= \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
s_y &= \frac{1}{2}\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \\
s_z &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\end{aligned}$$

And their action on α and β is

$$\begin{array}{ll}
s_x\alpha = \frac{1}{2}\beta & s_x\beta = \frac{1}{2}\alpha \\
s_y\alpha = I\frac{1}{2}\beta & s_y\beta = -I\frac{1}{2}\alpha \\
s_z\alpha = \frac{1}{2}\alpha & s_z\beta = -\frac{1}{2}\beta
\end{array}$$

In the future we will frequently encounter the 'integral'

$$\int \sigma_i(s)\sigma_j(s)ds$$

As mentioned before the integral is a short hand notation for a sum:

$$\sigma_i(-1/2)\sigma_j(-1/2)+\sigma_i(1/2)\sigma_j(1/2)$$

We can now consider the following 4 cases. The result always follows from Eq. 51

- $\sigma_i = \alpha$ and $\sigma_j = \beta$. Then $\alpha(-1/2)\beta(-1/2)+\alpha(1/2)\beta(1/2) = 0$
- $\sigma_i = \beta$ and $\sigma_j = \alpha$. Then $\beta(-1/2)\alpha(-1/2)+\beta(1/2)\alpha(1/2) = 0$
- $\sigma_i = \alpha$ and $\sigma_j = \alpha$. Then $\alpha(-1/2)\alpha(-1/2)+\alpha(1/2)\alpha(1/2) = 1$
- $\sigma_i = \beta$ and $\sigma_j = \beta$. Then $\beta(-1/2)\beta(-1/2)+\beta(1/2)\beta(1/2) = 1$

Hence the σ's behave like a set of orthogonal functions:

$$\int \sigma_i(s)\sigma_j(s)ds = \delta_{i,j}$$

Theorem: A Slater determinant wave-function is an eigenfunction of S_z if all of its spin orbitals are of the form $\psi(\mathbf{r})\alpha(s)$ or $\psi(\mathbf{r})\beta(s)$. So we will now assume that σ is either equal to $\alpha(s)$ or $\beta(s)$. Consequently $s_z\,\sigma = \lambda\,\sigma$, where λ is either $\frac{1}{2}$ or $-\frac{1}{2}$.
Proof:

$$\left(\sum_{i=1}^{N} s_z(i)\right)\Psi_I(\mathbf{x}_1,...,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}$$

$$\left(\sum_{i=1}^{N} s_z(i)\right)\Big[\psi_{i_1}(\mathbf{r}_1)\sigma_{i_1}(s_1)\;\psi_{i_2}(\mathbf{r}_2)\sigma_{i_2}(s_2)\;\psi_{i_3}(\mathbf{r}_3)\sigma_{i_3}(s_3)...\psi_{i_N}(\mathbf{r}_N)\sigma_{i_N}(s_N)$$

$$-\psi_{i_2}(\mathbf{r}_1)\sigma_{i_2}(s_1)\;\psi_{i_1}(\mathbf{r}_2)\sigma_{i_1}(s_2)\;\psi_{i_3}(\mathbf{r}_3)\sigma_{i_3}(s_3)...\psi_{i_N}(\mathbf{r}_N)\sigma_{i_N}(s_N) + ...\Big]$$

$$= \frac{1}{\sqrt{N!}}\Big[(\lambda_{i_1}+...+\lambda_{i_N})\Big(\psi_{i_1}(\mathbf{r}_1)\sigma_{i_1}(s_1)\;\psi_{i_2}(\mathbf{r}_2)\sigma_{i_2}(s_2)\;\psi_{i_3}(\mathbf{r}_3)\sigma_{i_3}(s_3)...\psi_{i_N}(\mathbf{r}_N)\sigma_{i_N}(s_N)\Big)$$

$$-(\lambda_{i_1}+...+\lambda_{i_N})\Big(\psi_{i_2}(\mathbf{r}_1)\sigma_{i_2}(s_1)\;\psi_{i_1}(\mathbf{r}_2)\sigma_{i_1}(s_2)\;\psi_{i_3}(\mathbf{r}_3)\sigma_{i_3}(s_3)...\psi_{i_N}(\mathbf{r}_N)\sigma_{i_N}(s_N)\Big) + ...\Big]$$

$$= (\lambda_{i_1}+...+\lambda_{i_N})\Psi_I(\mathbf{x}_1,...,\mathbf{x}_N)$$

A single determinant wave-function of this type is in general not an eigenfunction of the S^2 operator. There is however one important exception, closed shell systems

Closed shell systems

A closed shell systems consists of an even number N of electrons. The electrons form pairs that share the same spatial orbital, but differ in their spin function being either α or β. A closed system is represented by a single or a linear combination of Slater determinants of the following form:

$$\Psi_I(\mathbf{x}_1, ..., \mathbf{x}_N) =$$

$$\frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{i_1}(\mathbf{r}_1)\alpha(s_1) & \psi_{i_1}(\mathbf{r}_1)\beta(s_1) & ... & \psi_{i_{N/2}}(\mathbf{r}_1)\alpha(s_1) & \psi_{i_{N/2}}(\mathbf{r}_1)\beta(s_1) \\ ... & ... & ... & ... & ... \\ \psi_{i_1}(\mathbf{r}_i)\alpha(s_i) & \psi_{i_1}(\mathbf{r}_i)\beta(s_i) & ... & \psi_{i_{N/2}}(\mathbf{r}_i)\alpha(s_i) & \psi_{i_{N/2}}(\mathbf{r}_i)\beta(s_i) \\ ... & ... & ... & ... & ... \\ \psi_{i_1}(\mathbf{r}_N)\alpha(s_N) & \psi_{i_1}(\mathbf{r}_N)\beta(s_N) & ... & \psi_{i_{N/2}}(\mathbf{r}_N)\alpha(s_N) & \psi_{i_{N/2}}(\mathbf{r}_N)\beta(s_N) \end{vmatrix}$$

Such a Slater determinant is an eigenfunction of both S_z and S^2 with eigenvalues 0. (see exercise below). By definition it is thus a singlet state. Closed shell systems represent a very important class of systems in chemistry and physics. Most systems formed by covalent bonds are closed shell systems.

Exercise [3pts]: *Using the previous theorem show that a closed shell determinant is an eigenfunction of* S_z *with eigenvalue 0. Next show that it is also a eigenfunction of* $S^2 = \sum_{i,j} s_x(i)s_x(j) + s_y(i)s_y(j) + s_z(i)s_z(j)$ *with eigenvalue 0. Hint: Show first that a closed shell determinant is an eigenfunction of* $\sum_{i,j} s_z(i)s_z(j)$ *with eigenvalue zero. Next examine* $\sum_{i,j} s_x(i)s_x(j)$ *and* $\sum_{i,j} s_y(i)s_y(j)$*. Consider the cases where* $i = j$, $i = 2k-1$ *and* $j = 2k$ *(*$k = 1,...,N/2$*) and the remaining cases.*

The Hartree Fock energy expression

The Hartree Fock (HF) method is a standard method for medium accuracy electronic structure calculations. In the chemistry community the method is also frequently called self-consistent field (SCF) method. In this method the wave-function is given by a single determinant. This approximation is better than one might think. It turns out that in a MRSCF calculation expansion (Eq. 44) only one coefficient C_J of the expansion of Eq. 44 is significant, the majority of the coefficients are in general very small. In the HF method one has to solve a system of differential equations for the orbitals. This system of equations is obtained by minimizing the expectation value of the total energy for a single determinant wave-function. According to Eq. 46 the energy is given by

$$E = \sum_{i=1}^{N} [i|h|i] + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left([i,i|j,j] - [i,j|j,i] \right) \tag{52}$$

- The first term is the ordinary one particle (kinetic plus potential) energy
- The second term is the classical electrostatic repulsion
- The last term is the non-classical exchange term. This term ensures that the functional is self-interaction free. Because the term $[i,i|i,i]$ from the exchange part cancels the same term from the classical Hartree part, an electron in an orbital i does not repel itself.

The Hartree Fock equations will be derived in the following for the important case of a closed shell system. Since we require the spatial part of the orbitals to be identical for pairs of spin up and spin down electrons this variety of Hartree Fock is called restricted Hartree Fock. Since the spin part of our spin orbitals is either α or β we can integrate out the spin part:

$$
\begin{aligned}
[i|h|i] &= \int d\mathbf{r}ds\, \psi_i^*(\mathbf{r})\sigma_i(s) \left(-\frac{1}{2}\nabla^2 + V_{en}(\mathbf{r})\right) \psi_i(\mathbf{r})\sigma_i(s) \\
&= \int d\mathbf{r}\, \psi_i^*(\mathbf{r}) \left(-\frac{1}{2}\nabla^2 + V_{en}(\mathbf{r})\right) \psi_i(\mathbf{r})
\end{aligned}
$$

where

$$
V_{en}(\mathbf{r}) = \sum_{j=1}^{N_{at}} \frac{-Z_j}{|\mathbf{r} - \mathbf{R}_j|}
$$

$$
\begin{aligned}
[i,i,j,j] &= \int\int d\mathbf{r}ds\, d\mathbf{r}'ds'\, \frac{\psi_i^*(\mathbf{r})\sigma_i(s)\, \psi_i(\mathbf{r})\sigma_i(s)\, \psi_j^*(\mathbf{r}')\sigma_j(s')\, \psi_j(\mathbf{r}')\sigma_j(s')}{|\mathbf{r} - \mathbf{r}'|} \\
&= \int\int d\mathbf{r}\, d\mathbf{r}'\, \frac{\psi_i^*(\mathbf{r})\, \psi_i(\mathbf{r})\, \psi_j^*(\mathbf{r}')\, \psi_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}
\end{aligned}
$$

$$
\begin{aligned}
[i,j,i,j] &= \int\int d\mathbf{r} ds\, d\mathbf{r}' ds'\, \frac{\psi_i^*(\mathbf{r})\sigma_i(s)\,\psi_j(\mathbf{r})\sigma_j(s)\,\psi_i^*(\mathbf{r}')\sigma_i(s')\,\psi_j(\mathbf{r}')\sigma_j(s')}{|\mathbf{r}-\mathbf{r}'|} \\
&= \int\int d\mathbf{r}\, d\mathbf{r}'\, \frac{\psi_i^*(\mathbf{r})\,\psi_j(\mathbf{r})\,\psi_i^*(\mathbf{r}')\,\psi_j(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\delta_{\sigma_i,\sigma_j}
\end{aligned}
$$

Introducing spin free one electron integrals and electron repulsion integrals

$$
(i|h|j) = \int d\mathbf{r}\,\psi_i^*(\mathbf{r})\left(-\frac{1}{2}\nabla^2 + V_{en}(\mathbf{r}))\right)\psi_j(\mathbf{r})
$$

$$
(i,j|k,l) = \int\int d\mathbf{r}\, d\mathbf{r}'\, \frac{\psi_i^*(\mathbf{r})\,\psi_j(\mathbf{r})\,\psi_k^*(\mathbf{r}')\,\psi_l(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}
$$

the Hartree Fock energy (eq. 52) thus becomes for a closed shell system

$$
E = \sum_{i=1}^{N/2} 2(i|h|i) + \frac{1}{2}\sum_{i=1}^{N/2}\sum_{j=1}^{N/2} 4(i,i|j,j) - 2(i,j|j,i) \tag{53}
$$

Exercise [2pts]: *Show that the one-particle density matrix for a general single determinant wave-function is given by*

$$\gamma_1(\mathbf{x}_1;\mathbf{x'}_1) = \sum_{j=1}^{N} \psi_j(\mathbf{x})\psi_j^*(\mathbf{x'}) \tag{54}$$

and that the spin-less one-particle density matrix becomes

$$\rho_1(\mathbf{r}_1;\mathbf{r'}_1) = 2\sum_{j=1}^{N/2} \psi_j(\mathbf{r})\psi_j^*(\mathbf{r'}) \tag{55}$$

for a closed shell single determinant wave-function

Exercise [2pts]: *Verify that* $\sum_i \sum_j [i,i|j,j]$ *represents indeed the classical electrostatic interaction, i.e. show that*

$$\sum_{i=1}^{N}\sum_{j=1}^{N}[i,i|j,j] = \int d\mathbf{r}d\mathbf{r'}\frac{\rho(\mathbf{r})\rho(\mathbf{r'})}{|\mathbf{r}-\mathbf{r'}|}$$

hint: use Eq. 34, 35

Exercise [2pts]: *Show that the HF energy of Eq. 53 and one-particle density matrix of Eq. 55 is invariant under unitary transformations among the single particle orbitals*

$$\tilde{\psi}_i(\mathbf{r}) = \sum_j U_{i,j}\psi_j(\mathbf{r}) \tag{56}$$

The HF energy in terms of density matrices

The total HF energy can entirely be written in terms of the one particle density matrix. The relevant part of the energy is obviously the electrostatic interaction among the electrons that is in general a function of the two particle density matrix (Eq. 38). From Eq. 52 and the definition of γ_1 (Eq. 54) it follows that the HF electron-electron repulsion energy can be written as

$$E_{e-e} = \frac{1}{2}\left(\int d\mathbf{x}\int d\mathbf{x}'\frac{\gamma_1(\mathbf{x};\mathbf{x})\gamma_1(\mathbf{x}';\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|} - \int d\mathbf{x}\int d\mathbf{x}'\frac{\gamma_1(\mathbf{x};\mathbf{x}')\gamma_1(\mathbf{x}';\mathbf{x})}{|\mathbf{r}-\mathbf{r}'|}\right) \tag{57}$$

This shows that the HF method neglects certain correlations among the electrons because the entire correlations are only visible in the two particle density matrix.
It can be demonstrated that in an insulator $\gamma_1(\mathbf{x};\mathbf{x}')$ decays exponentially fast to zero for large separations of $\mathbf{r}$ and $\mathbf{r}'$. This implies that all the quantum mechanical effects which are represented by the second term in Eq. 57, are short range in the HF approximation. The only long range term is the electrostatic interaction, i.e the first term in Eq. 57. This is not correct. Even if the above electrostatic term from the the electronic charge distribution is exactly canceled by ionic contributions (which is for instance the case in a noble gas dimer with a large interatomic distance), there is a long range quantum mechanical effect which is the van-der Waals interaction. Hence van-der Waals interactions are not described by HF.

Orthogonalizing orbitals

A set of orbitals ψ_i is said to be orthonormal (= orthogonal and normalized) if the overlap matrix $\mathcal{S}$,

$$\mathcal{S}_{i,j} = \int d\mathbf{r}\, \psi_i^*(\mathbf{r})\psi_j(\mathbf{r})$$

is the identity matrix. There are three major algorithms to generate an orthonormal set $\tilde{\psi}_i$ from a non-orthonormal set.

- Loewdin orthogonalization

$$\tilde{\psi}_i = \sum_j \left(\mathcal{S}^{-1/2}\right)^*_{i,j} \psi_j$$

- Gram Schmid orthogonalization
  For i=1 to N do

$$\bar{\psi}_i(\mathbf{r}) = \psi_i(\mathbf{r}) - \sum_{j=1}^{i-1} \int \psi_j^*(\mathbf{r}')\psi_i(\mathbf{r}')d\mathbf{r}'\, \psi_j(\mathbf{r})$$

$$\tilde{\psi}_i(\mathbf{r}) = \bar{\psi}_i(\mathbf{r}) / \sqrt{\int d\mathbf{r}\, \bar{\psi}_i^*(\mathbf{r})\bar{\psi}_i(\mathbf{r})}$$

enddo

The Gram Schmid procedure is usually done in place

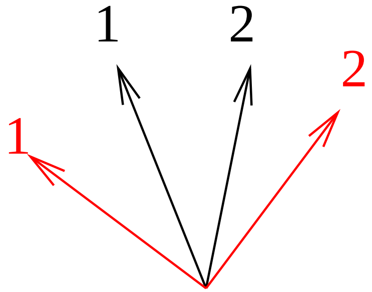
$$\begin{aligned}
\psi_1 &\Leftarrow \psi_1/|\psi_1| \\
\psi_2 &\Leftarrow \psi_2-\langle\psi_1|\psi_2\rangle\psi_1 \;;\; \psi_2 \Leftarrow \psi_2/|\psi_2| \\
\psi_3 &\Leftarrow \psi_3-\langle\psi_1|\psi_3\rangle\psi_1-\langle\psi_2|\psi_3\rangle\psi_2 \;;\; \psi_3 \Leftarrow \psi_3/|\psi_3| \\
\ldots &\Leftarrow \ldots
\end{aligned}$$

Exercise [2pts]: *Show that both methods give orthonormal orbitals*

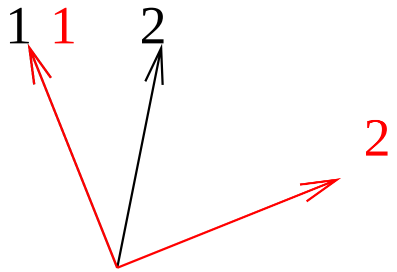
Exercise [1pts]: *How does one have to modify the formulas for Loewdin and Gram-Schmid orthogonalization if one deals with discrete vectors instead of continuous functions?*

In the Gram Schmid method the first vector is not turned (only scaled) whereas the last one is heavily turned. In the Loewdin method all vectors are roughly equally turned.

LOEWDIN



GRAM −SCHMID



- Stabilized Gram Schmidt: It turns out that the classical Gram Schmidt procedure becomes nummerically unstable in high dimensional spaces. In the stabilized Gram Schmidt the orthogonalization steps are done in a different order which is mathematically equivalent but numericallty more stable.

$$
\begin{aligned}
\psi_1 \quad &\Leftarrow \quad \psi_1/|\psi_1| \\
&\psi_2 \quad \Leftarrow \psi_2 - \langle\psi_1|\psi_2\rangle\psi_1 \\
&\psi_3 \quad \Leftarrow \psi_3 - \langle\psi_1|\psi_3\rangle\psi_1 \\
&\psi_4 \quad \Leftarrow \psi_4 - \langle\psi_1|\psi_4\rangle\psi_1 \\
&\ldots \\
\psi_2 \quad &\Leftarrow \quad \psi_2/|\psi_2| \\
&\psi_3 \quad \Leftarrow \psi_3 - \langle\psi_2|\psi_3\rangle\psi_2 \\
&\psi_4 \quad \Leftarrow \psi_4 - \langle\psi_2|\psi_4\rangle\psi_2 \\
&\psi_5 \quad \Leftarrow \psi_5 - \langle\psi_2|\psi_5\rangle\psi_2 \\
&\ldots \\
\psi_3 \quad &\Leftarrow \quad \psi_3/|\psi_3| \\
\ldots \quad & \qquad \ldots
\end{aligned}
$$

- Orthogonalization with importance ranking This method will be demonstrated for discrete vectors rather than for continous functions. However it can easily be adapted to continuous functions if necessary. The central quantity is again the overlap matrix S whose elements are in the discrete case the scalarproducts of the initial non-orthogonal vectors v_i

  $$S_{i,j} = \langle v_i | v_j \rangle = \sum_k v_i(k) v_j(k)$$

  where the elements of the vector v_i are denoted by $v_i(k)$. Since the overlap matrix is a positive definite symmetric matrix its eigenvalues λ_i are real and its eigenvectors ω_i diagonalize the original overlap matrix, i.e.

  $$\sum_{l,k} \omega_i(l) S_{l,k} \omega_j(k) = \delta_{i,j} \lambda_i$$

  The new orthogonal set of vectors w_i can now be obtained by the following linear combination of the original v_i's

  $$w_i = \sum_k \omega_i(k) v_k$$

  The orthogonality can easily be seen from the relation

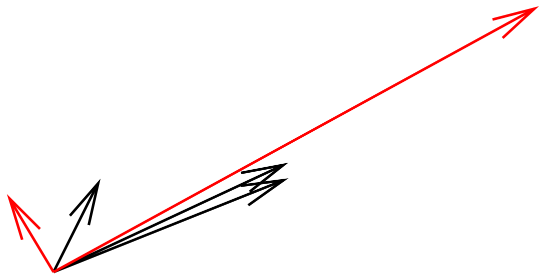
  $$\langle w_i | w_j \rangle = \sum_\mu w_i(\mu) w_j(\mu) = \sum_{k,l} \sum_\mu \omega_i(k) v_k(\mu) \omega_j(l) v_l(\mu) = \sum_{k,l} \omega_i(k) S_{k,l} \omega_j(l) = \delta_{i,j} \lambda_i$$

The length $\sqrt{\lambda_i}$ of the vectors w_i is a measure of how many of the original vectors v_i point in directions similar to w_i and how long they are. This can easily be deduced for the first eigenvalue vector pair from the variational principle. w_1 is the longest vector that can be constructed as a linear combination of the v_i by a set of coefficients $\omega_1(l)$ that are normalized to one, since the eigenvector ω_1 maximizes the expression

$$\frac{\langle w|w\rangle}{\langle\omega_1|\omega_1\rangle} = \frac{\langle\omega_1|S|\omega_1\rangle}{\langle\omega_1|\omega_1\rangle}$$

Obviously a long vector can be constructed if many of the original v_i's point in the same direction and if they are long. Eigenvalues that are zero, or very small in numerical applications, indicate that the dimension of the space spanned is less than the number of original vectors v_i. The eigenvectors correponding to these small eigenvalues indicate the degrees of freedom arising from numerical noise. In this way large eigenvalues λ_i assign a larger importance to a direction w_i, whereas small eigenvalues indicate that the corresponding directions w_i are not significant or even purely numerical noise. The situation is illustrated below.

There are three input vectors shown in black. Two of them are very similar to each other and they probably differ only by some noise. We assume that they are not lying exactly in the plane. Two output vectors are shown in red. The longest vector shows the most important direction, and the second longest the second most important direction. Not shown is the third vector which is very short and nearly perpendicular to the plane. This vector just represents unimportant noise.



The Euler-Lagrange equations of a functional

A functional is a mathematical object that takes as input a function and gives as output a number such as the total energy E. Frequently in physics we seek a function that will make the functional stationary. Such a function has the property that to first order the output of the functional will not change under any conceivable infinitesimal change of the input function. The calculus of variations allows us to calculate the functional derivative of such an object, which then in turn can be set equal to zero. The resulting differential equation is called the Euler Lagrange equation of the functional. Since the derivative is not with respect to a variable but with respect to a function it is denoted by $\frac{\delta F[f]}{\delta f(x)}$ and it is a function of x. For the moment we will assume that f is real. Later the generalization to complex functions f will be considered. Formally the Euler Lagrange equation of a functional $F[f(x)]$ can be obtained by putting the following ordinary partial derivative to zero.

$$\frac{\partial}{\partial a} F[f(x) + a\delta f(x)]|_{a=0} = 0$$

Example: Let's calculate the functional derivative of

$$E[\psi] = \frac{\int dx\ \psi(x)(-\frac{1}{2}\frac{\partial^2}{\partial x^2})\psi(x) + \psi(x)V(x)\psi(x)}{\int dx \psi(x)\psi(x)} = \frac{P[\psi]}{Q[\psi]}$$

$$\frac{\partial}{\partial a}\frac{P[\psi]}{Q[\psi]} = \frac{\frac{\partial P[\psi]}{\partial a}}{Q[\psi]} - \frac{\frac{\partial Q[\psi]}{\partial a}}{Q[\psi]^2}P[\psi]$$

$$\begin{aligned}
\left.\frac{\partial Q[\psi]}{\partial a}\right|_{a=0} &= \left.\frac{\partial}{\partial a}\int dx\,(\psi+a\delta f)(\psi+a\delta f)\right|_{a=0} = 2\int dx\,(\psi+a\delta f)\delta f\Big|_{a=0} = \int dx\,[2\psi]\,\delta f \\
\left.\frac{\partial P[\psi]}{\partial a}\right|_{a=0} &= \left.\frac{\partial}{\partial a}\int dx\;(\psi+a\delta f)(-\frac{1}{2}\frac{\partial^2}{\partial x^2})(\psi+a\delta f)+(\psi+a\delta f)V(x)(\psi+a\delta f)\right|_{a=0} \\
&= \int dx\,\psi(-\frac{1}{2}\frac{\partial^2}{\partial x^2})\delta f+\delta f(-\frac{1}{2}\frac{\partial^2}{\partial x^2})\psi+2\psi V\delta f \\
&= \int dx\,\left[(-\frac{\partial^2}{\partial x^2})\psi(x)+2V(x)\psi(x)\right]\delta f(x)
\end{aligned}$$

The terms in the square brackets are the functional derivatives of P and Q.
Putting $Q[\psi]=1$ we finally obtain

$$\left.\frac{\partial}{\partial a}\frac{P[\psi]}{Q[\psi]}\right|_{a=0} = 2\int dx\,\delta f(x)\left(-\frac{1}{2}\frac{\partial^2}{\partial x^2}+V(x)-E\right)\psi(x)=0$$

If this holds true for any δf it follows that

$$\left(-\frac{1}{2}\frac{\partial^2}{\partial x^2}+V(x)-E\right)\psi(x)=0 \tag{58}$$

which is just the 1-dim single particle Schrödinger equation.

Frequently the energy expression is written in terms of complex orbitals. The Euler-Lagrange equations of a real functional F that depends on the complex function $f(x)=f_r(x)+If_i(x)$ are

$$\begin{aligned} \frac{\delta F[f]}{\delta f_r(x)} &= 0 & (59)\\ \frac{\delta F[f]}{\delta f_i(x)} &= 0 & (60) \end{aligned}$$

The same set of equations can be obtained in a more elegant way by taking the derivative with respect to f^*. Using the chain rule and the fact that $f_r=(f+f^*)/2$ and $f_i=(f-f^*)/(2I)$ we obtain

$$\begin{aligned} \frac{\delta F[f]}{\delta f^*(x)} &= \frac{\delta F[f]}{\delta f_r(x)}\frac{\delta f_r(x)}{\delta f^*(x)}+\frac{\delta F[f]}{\delta f_i(x)}\frac{\delta f_i(x)}{\delta f^*(x)}\\ &= \frac{\delta F[f]}{\delta f_r(x)}\frac{1}{2}+\frac{\delta F[f]}{\delta f_i(x)}I\frac{1}{2} \end{aligned}$$

Since both the real and the imaginary part have to be zero, the condition $\frac{\delta F[f]}{\delta f^*(x)} = 0$ is equivalent to the set of Equations 59 and 60.
Exercise [2pts]: *Show that the Euler Lagrange equation of*

$$E[\psi] = \frac{\int dx\ \psi^*(x)(-\frac{1}{2}\frac{\partial^2}{\partial x^2})\psi(x) + \psi^*(x)V(x)\psi(x)}{\int dx \psi^*(x)\psi(x)}$$

gives as well the Schrödinger equation 58.

In the complex formalism one obtains immediately the correct functional derivative of the kinetic energy term, without doing two integrations by parts. That an integration by parts can be done is explicitly assumed. If it could not be done the operator would not be hermitian and could therefore be complex.
If no differential operators are present the following shortcut will give the correct result. Replace all integrals by sums over grid points $\mathbf{r}_i$, take the partial derivative with respect to the quantities $f(\mathbf{r}_i)$ and divide by the integration weight. Finally drop the subscript $_i$ to obtain a continuous function from the discrete partial derivatives.

As an example let's recalculate the functional derivative of the normalization integral using this trick.

$$\frac{\delta}{\delta\psi^*(x)}\int\psi^*(x)\psi(x)dx$$

the discrete version becomes

$$\left(\frac{\partial}{\partial\psi^*(x_i)}\sum_j\psi^*(x_j)\psi(x_j)\Delta x\right)/\Delta x=\psi(x_i)$$

Hence the functional derivative is $\psi(x)$

In the previous example we were looking for the lowest eigenstate of the single particle Schrödinger equation. Including the orthonormality constraint was easy in this case. Since there is only one orbital we had just to make sure that it is normalized. This was done by explicitly dividing in the energy expression by the norm $Q[\psi]$. In case where we have an energy expression that depends on several orbitals imposing orthonormality is done in the following way.

Minimizing a functional under orthonormality constraints

Orthonormality of the orbitals that are used to construct Slater determinant wave-functions leads to significant simplifications in the calculation of expectation values. If one is ready to accept more complicated expressions one can however also use non-orthonormal orbitals. In order to calculate the gradient under orthonormality constraints we have to abandon temporarily the orthonormality assumption. Only after we have calculated the gradient expression we can again simplify it by restricting ourselves to orthonormal orbitals.

In order to calculate the energy expectation value from a set of real non-orthonormal orbitals we consider some composite functional. The first part of the functional constructs a set of orthonormal orbitals $\tilde{\psi}_i$ by a symmetric Löwdin orthogonalization of the non-orthogonal set ψ_i

$$\tilde{\psi}_i = \sum_l \left(\mathcal{S}_{i,l}^{-1/2} \right)^* \psi_l \qquad ; \qquad \tilde{\psi}_i^* = \sum_l \mathcal{S}_{i,l}^{-1/2} \psi_l^* \tag{61}$$

where $\mathcal{S}_{i,j} = \int \psi_i(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r}$ is the overlap matrix among the occupied orbitals. Since in our actual calculation we use orthogonal orbitals, we have to consider only infinitesimally nonorthogonal orbitals . Then $\mathcal{S}^{-1/2} = (I + (\mathcal{S} - I))^{-1/2} \approx I - (1/2)(\mathcal{S} - I)$ and Eq. 61

becomes

$$\tilde{\psi}_j^* = \sum_l (\frac{3}{2}\delta_{j,l} - \frac{1}{2}\mathcal{S}_{j,l})\psi_l^* \qquad ; \qquad \tilde{\psi}_j = \sum_l (\frac{3}{2}\delta_{j,l} - \frac{1}{2}\mathcal{S}_{j,l}^*)\psi_l \tag{62}$$

The gradient of the total functional is then obtained by applying the chain rule:

$$\frac{\delta E}{\delta \psi_i^*(\mathbf{r})} = \sum_j \int \frac{\delta E}{\delta \tilde{\psi}_j^*(\mathbf{r}')} \frac{\delta \tilde{\psi}_j^*(\mathbf{r}')}{\delta \psi_i^*(\mathbf{r})} d\mathbf{r}' + \sum_j \int \frac{\delta E}{\delta \tilde{\psi}_j(\mathbf{r}')} \frac{\delta \tilde{\psi}_j(\mathbf{r}')}{\delta \psi_i^*(\mathbf{r})} d\mathbf{r}' \tag{63}$$

The first two parts of the two products in Eq. 63 depend on the specific energy functional and we will just denote this unconstrained gradient by $d_j(\mathbf{r})$

$$d_j(\mathbf{r}) = \frac{1}{2}\frac{\delta E}{\delta \tilde{\psi}_j^*(\mathbf{r})} \qquad ; \qquad d_j^*(\mathbf{r}) = \frac{1}{2}\frac{\delta E}{\delta \tilde{\psi}_j(\mathbf{r})} \tag{64}$$

The second parts in the two products on the right hand side of Eq. 63 give

$$\begin{aligned} \frac{\delta \tilde{\psi}_j(\mathbf{r}')}{\delta \psi_i^*(\mathbf{r})} &= -\frac{1}{2}\sum_l \psi_l(\mathbf{r}') \frac{\delta}{\delta \psi_i^*(\mathbf{r})} \int \psi_j(\mathbf{r}'')\psi_l^*(\mathbf{r}'') d\mathbf{r}'' \\ &= -\frac{1}{2}\psi_i(\mathbf{r}')\psi_j(\mathbf{r}) \end{aligned} \tag{65}$$

$$
\begin{aligned}
\frac{\delta\tilde{\psi}_j^*(\mathbf{r}')}{\delta\psi_i^*(\mathbf{r})} &= \frac{3}{2}\delta_{j,i}\delta(\mathbf{r}-\mathbf{r}') - \frac{1}{2}\mathcal{S}_{j,i}\delta(\mathbf{r}-\mathbf{r}') - \frac{1}{2}\sum_l \psi_l^*(\mathbf{r}')\frac{\delta}{\delta\psi_i^*(\mathbf{r})}\int \psi_j^*(\mathbf{r}'')\psi_l(\mathbf{r}'')d\mathbf{r}'' \\
&= \frac{3}{2}\delta_{i,j}\delta(\mathbf{r}-\mathbf{r}') - \frac{1}{2}\mathcal{S}_{i,j}\delta(\mathbf{r}-\mathbf{r}') - \frac{1}{2}\sum_l \psi_l^*(\mathbf{r}')\delta_{i,j}\psi_l(\mathbf{r}) \\
&= \delta_{i,j}\delta(\mathbf{r}-\mathbf{r}') - \frac{1}{2}\delta_{i,j}\sum_l \psi_l^*(\mathbf{r}')\psi_l(\mathbf{r})\,. \qquad (66)
\end{aligned}
$$

In the last transformation step, we have used the fact that we calculate the derivative for a set of orthonormal orbitals and therefore $\mathcal{S} = I$. Because of the orthogonality constraint we are of course allowed to put $\mathcal{S} = I$ only after calculating the derivative. Finally, we obtain the expression for the constrained gradient

$$
\begin{aligned}
\frac{1}{2}\frac{\delta E}{\delta\psi_i^*(\mathbf{r})} = d_i(\mathbf{r}) &- \frac{1}{2}\sum_j\left(\int d_j^*(\mathbf{r}')\psi_i(\mathbf{r}')d\mathbf{r}'\right)\psi_j(\mathbf{r}) \\
&- \frac{1}{2}\sum_j\left(\int d_i(\mathbf{r}')\psi_j^*(\mathbf{r}')d\mathbf{r}'\right)\psi_j(\mathbf{r}) \\
&= d_i(\mathbf{r}) - \sum_j \Lambda_{i,j}\psi_j(\mathbf{r}). \qquad (67)
\end{aligned}
$$

where Λ is a Lagrange multiplier matrix for the orthogonality constraints

The Hartree Fock equations

The Euler-Lagrange equation for the HF energy expression (Eq. 52) is

$$d_i(\mathbf{x}) = \left(-\frac{1}{2}\nabla^2 + V_{en}(\mathbf{r}) + V_H(\mathbf{r})\right)\psi_i(\mathbf{x}) - \sum_j \int d\mathbf{x}' \frac{\psi_i(\mathbf{x}')\psi_j^*(\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|}\psi_j(\mathbf{x}) \tag{68}$$

The Euler-Lagrange equation of the closed shell system HF energy expression (Eq. 53) is

$$\frac{1}{2}d_i(\mathbf{r}) = \left(-\frac{1}{2}\nabla^2 + V_{en}(\mathbf{r}) + V_H(\mathbf{r})\right)\psi_i(\mathbf{r}) - \sum_j \int d\mathbf{r}' \frac{\psi_i(\mathbf{r}')\psi_j^*(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\psi_j(\mathbf{r}) \tag{69}$$

The Hartree potential $V_H(\mathbf{r})$ is the classical electrostatic potential arising from the electronic charge density $\rho(\mathbf{r})$

$$V_H(\mathbf{r}) = \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{70}$$

For a closed shell system the charge density (Eq. 35, 55) is given by

$$\rho(\mathbf{r}) = 2\sum_{j=1}^{N/2}\psi_j(\mathbf{r})\psi_j^*(\mathbf{r})$$

The simplest way to solve the HF equations numerically consists thus of a repetition until convergence of the following steps

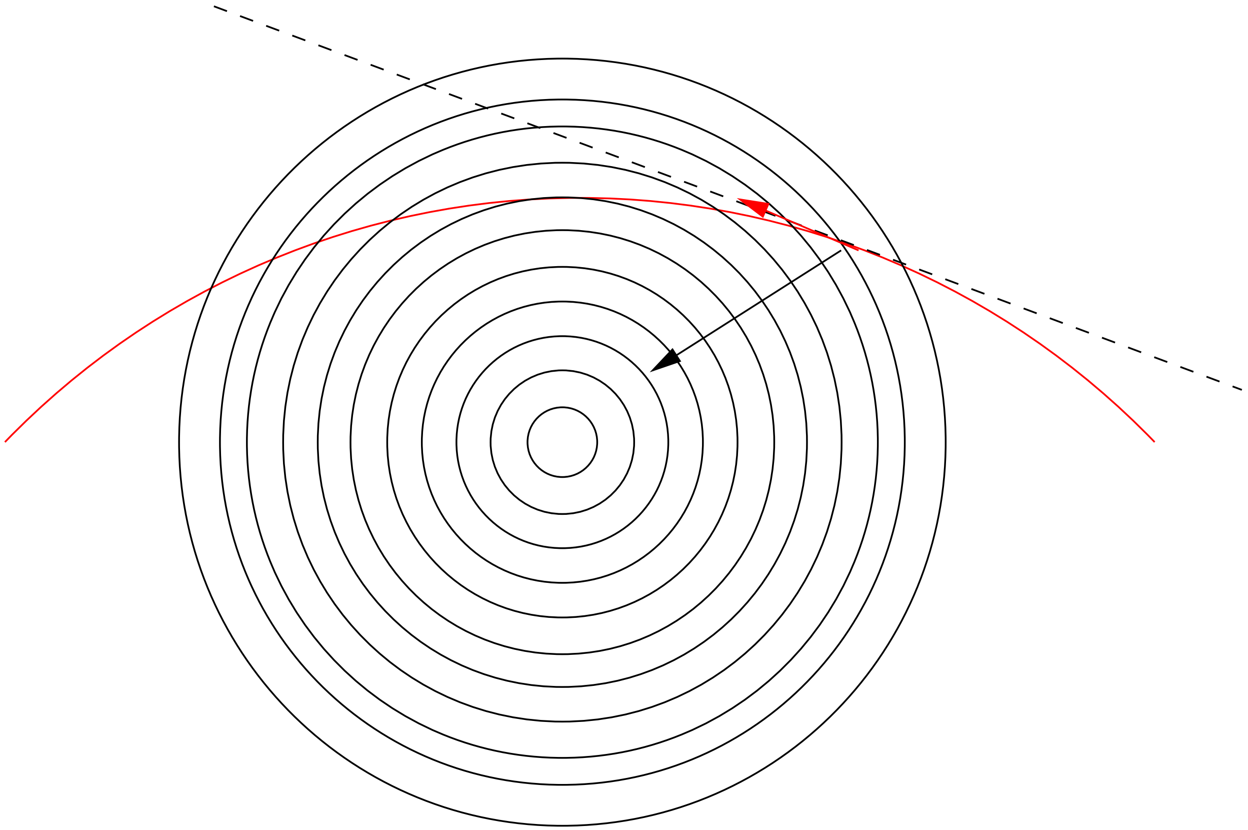
- Use Eq. 69 (or Eq. 68) to calculate the unconstrained gradient
- Use Eq. 67 to form the effective gradient $\frac{\delta E}{\delta \psi_i(\mathbf{r})}$ that contains the orthogonality constraint
- Move downhill along the gradient. Make only small moves with step-size $step$ to ensure convergence

$$\psi_i(\mathbf{r}) \leftarrow \psi_i(\mathbf{r}) - step \frac{\delta E}{\delta \psi_i(\mathbf{r})}$$

- Orthogonalize $\psi_i(\mathbf{r})$

Since the gradient incorporates already the orthogonality constraints, one might wonder why an explicit orthogonalization step is still necessary. Things are illustrated in the figure below.

Let us assume that we search the minimum of the function whose equipotential lines are shown in black under the constraint that we remain on the red curve. The unconstrained gradient is represented by the black vector, the gradient incorporating the constraints by the red vector. Moving along the red gradient will conserve the constraint only to first order but not to higher order. For this reason the constraint has to be imposed explicitly after any finite move in the direction of the red gradient.



It can be shown that

$$\int d_j^*(\mathbf{x}')\psi_i(\mathbf{x}')d\mathbf{x}' = \int d_i(\mathbf{x}')\psi_j^*(\mathbf{x}')d\mathbf{x}'$$

Exercise [2pts]: *Show that for the closed shell HF case (d_i is given by Eq. 69)*

$$\int d_j^*(\mathbf{r}')\psi_i(\mathbf{r}')d\mathbf{r}' = \int d_i(\mathbf{r}')\psi_j^*(\mathbf{r}')d\mathbf{r}'$$

Hence Eq. 67 has for the HF case the structure

$$\frac{1}{2}\frac{\delta E}{\delta\psi_i(\mathbf{x})} = d_i(\mathbf{x}) - \sum_j \Lambda_{i,j}\psi_j(\mathbf{x}) \tag{71}$$

where $\Lambda_{i,j}$ is the symmetric matrix $\int d_j^*(\mathbf{x}')\psi_i(\mathbf{x}')d\mathbf{x}'$. We can always choose as the unitary matrix U of Eq. 56 the matrix that diagonalizes Λ. The HF equations have then the structure

$$d_i(\mathbf{x}) = \varepsilon_i\psi_i(\mathbf{x}) \tag{72}$$

where ε_i are the eigenvalues of Λ. The particular orbitals that satisfy Eq. 72 are called canonical orbitals. The matrix $\Lambda_{i,j}$ is the discretization of the Fock operator discussed in the following and is called the Fock matrix.

The canonical orbitals can be obtained in two ways. Once one has solved the HF equations as a minimization problem the canonical orbitals can be obtained from the non-unique minimizing orbitals by a unitary transformation. The other possibility is to consider the solution of the HF equations as a self-consistent eigenvalue problem. In this alternative approach one finds the eigenvectors of the Fock matrix (Eq. 74). For a given Fock matrix this is an obviously an ordinary eigenvalue problem. The problem is that the Fock matrix is not fixed, but it depends on the orbitals which are the solution of the eigenvalue problem. For this reason Eq. 74 gives rise to a so-called self-consistent eigenvalue problem. The self-consistency condition is fulfilled if the orbitals that were used for the construction of the Fock matrix are identical to the eigenorbitals.

In density functional theory the situation is similar. The equations can either be solved as a minimization problem or as a self-consistent eigenvalue problem. The algorithmic aspects of the mixing approach will be discussed in more detail in the section on density functional theory.

The Fock operator and Koopman's theorem

The Fock operator $\mathcal{F}$ is given by

$$\mathcal{F}(\mathbf{x},\mathbf{x}') = \delta(\mathbf{x}-\mathbf{x}')\left(-\frac{1}{2}\nabla^2_{\mathbf{r}'} + V_{en}(\mathbf{r}) + V_H(\mathbf{r})\right) - \sum_{j=1}^{N} \frac{\psi_j^*(\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|}\psi_j(\mathbf{x}) \tag{73}$$

The gradient of Eq. 68 can thus be obtained by applying the Fock operator onto the orbitals

$$d_i(\mathbf{x}) = \int \mathcal{F}(\mathbf{x},\mathbf{x}')\psi_i(\mathbf{x}')d\mathbf{x}'$$

The canonical orbitals are thus the eigenorbitals of the Fock operator since they satisfy

$$\int \mathcal{F}(\mathbf{x},\mathbf{x}')\psi_i(\mathbf{x}')d\mathbf{x}' = \varepsilon_i\psi_i(\mathbf{x}) \tag{74}$$

In principle the Hartree Fock method allows to determine only the occupied orbitals that are used for the construction of the determinantal wave-function. If one considers however the Hartree Fock orbitals as the eigenfunctions of the Fock operator one can calculate both occupied and virtual orbitals. The Fock operator has an infinite number of eigenfunctions. The lowest N orbitals are called the occupied orbitals. All the other orbitals are called virtual orbitals. Since the virtual orbitals do not enter into the HF wave-function, they do not influence the energy or any other physical property of the Hartree Fock ground state. For a system of noninteracting electrons the energy is the sum over all occupied eigenvalues. In the Hartree Fock method we have what is called independent particles that are interacting through some mean field. The energy E can therefore be written as the sum over the eigenvalues of Eq. 74 plus a correction term that arises from averaging out the interactions as can be seen by comparing Eq. 52 and Eq. 68.

$$E = \sum_{i=1}^{N} \varepsilon_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left([ii|jj] - [ij|ji] \right) \tag{75}$$

The eigenvalues of the Fock operator ε_i are the expectation values of this operator with respect to its eigenfunctions and hence they are given by

$$\varepsilon_i = [i|h|i] + \sum_{j=1}^{N} \left([ii|jj] - [ij|ji] \right) \tag{76}$$

The physical importance of the Fock operator and its eigenvalues comes from the fact that the eigenvalues are approximations to the ionization energy and electron affinity. The Hartree Fock ionization energy is defined as the difference between the energy E^N of the N-electron determinant and the energy E^{N-1} of the $N-1$-electron determinant. It can thus in principle be determined by doing 2 separate HF calculations for the N and $N-1$ electron system. If one does 2 such separate calculations one has of course to relax separately the 2 sets of (canonical) orbitals for both systems, since both sets are different. It turns out that relaxations effects are usually not very important and one can therefore obtain a decent $N-1$-electron HF wave-function by building up the determinant using $N-1$ canonical orbitals of the N electron determinant. Let us assume that we take out in the construction of the $N-1$ electron determinant the canonical orbital m. The ionization potential is then given by (Eq. 75)

$$\begin{aligned}
E^{N-1}-E^{N} &\approx \sum_{i\neq m}[i|h|i]+\frac{1}{2}\sum_{i\neq m}\sum_{j\neq m}[i,i|j,j]-[i,j|j,i] \\
&- \left(\sum_{i=1}^{N}[i|h|i]+\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}[i,i|j,j]-[i,j|j,i]\right) \\
&= -\left([m|h|m]+\sum_{j=1}^{N}[m,m|j,j]-[m,j|j,m]\right) \\
&= -\varepsilon_m
\end{aligned}$$

In the same way it can be shown that the HF electron affinity $E^{N}-E^{N+1}$ is approximated by $-\varepsilon_l$, where $-\varepsilon_l$ is the l-th (virtual) eigenvalue of the Fock operator calculated from the Fock operator of the N electron system. These results are called Koopman's theorem.

The helium atom in HF

The Hamiltonian for the He atom is

$$\mathcal{H} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} - \frac{2}{|\mathbf{r}_1|} - \frac{2}{|\mathbf{r}_2|} \tag{77}$$

The HF ground state is obtained by putting a spin up and a spin down electron into the same 1s-like orbital ψ_{1s}. Antisymmetry is achieved through the spin part of the wavefunction:

$$\Psi(\mathbf{r}_1, s_1, \mathbf{r}_2, s_2) = \psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2)\frac{1}{\sqrt{2}}\left(\alpha(s_1)\beta(s_2) - \alpha(s_2)\beta(s_1)\right) \tag{78}$$

Exercise [1pts]: *Show that $\Psi(\mathbf{r}_1, s_1, \mathbf{r}_2, s_2)$ of Eq. 78 can be written as a closed shell determinant*

The corresponding one-particle density matrix is

$$\begin{aligned}
\rho_1(\mathbf{x}_1, \mathbf{x}_1') &= \psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_1')\left(\alpha(s_1)\alpha(s_1') + \beta(s_1)\beta(s_1')\right) \\
\rho_1(\mathbf{r}_1, \mathbf{r}_1') &= 2\psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_1') \\
\rho_1(\mathbf{r}_1) &= 2\psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_1) = 2\psi_{1s}(r_1)\psi_{1s}(r_1)
\end{aligned} \tag{79}$$

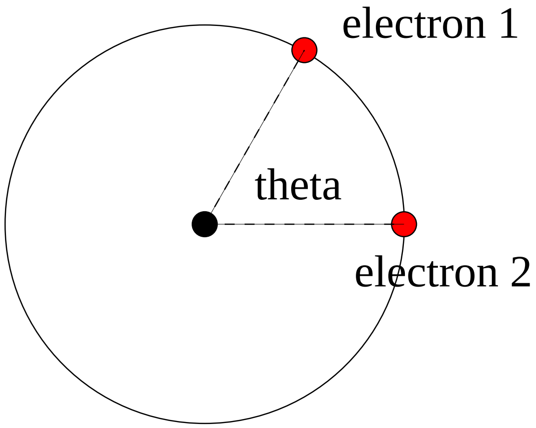
The corresponding two-particle density matrix is

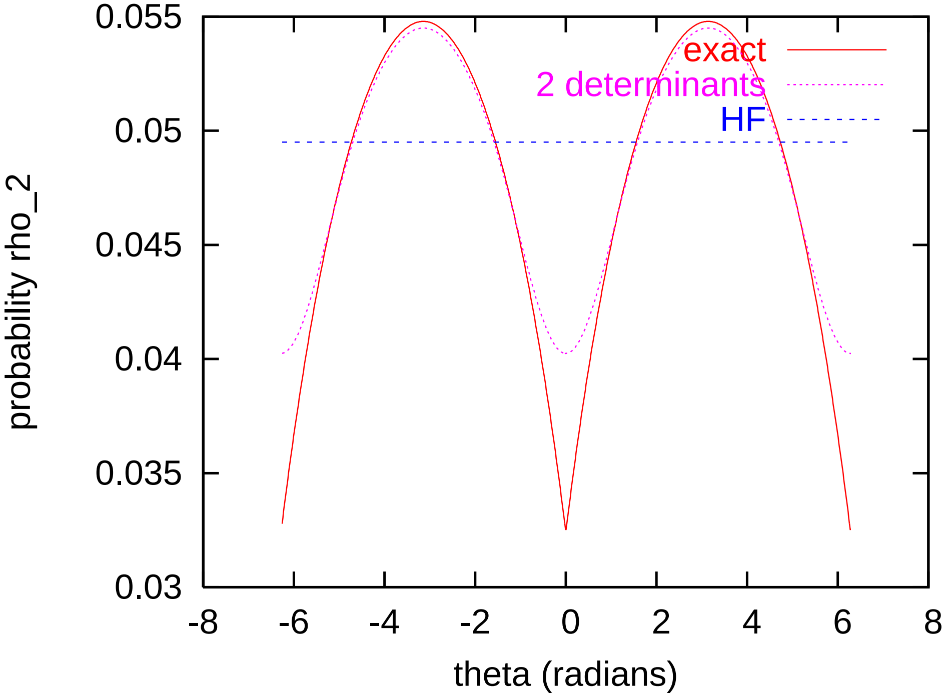
$$\rho_2(\mathbf{r}_1,\mathbf{r}_2;\mathbf{r}'_1,\mathbf{r}'_2) = 2\psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}'_1)\psi_{1s}(\mathbf{r}_2)\psi_{1s}(\mathbf{r}'_2)$$

$$\rho_2(\mathbf{r}_1,\mathbf{r}_2) = 2\psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2)\psi_{1s}(\mathbf{r}_2) = \frac{1}{2}\rho_1(\mathbf{r}_1)\rho_1(\mathbf{r}_2) = \frac{1}{2}\rho_1(r_1)\rho_1(r_2) \quad (80)$$

Exercise [2pts]: *Derive Eq. 79 and Eq. 80 from Eq. 78*

The fact that for the He atom $\rho_2(\mathbf{r}_1,\mathbf{r}_2)$ is in HF a simple product of the probabilities for finding one electron at $\mathbf{r}_1$ and another at $\mathbf{r}_2$ means that there is no correlation between the two electrons. As we will see later, HF gives only some sort of correlation called exchange correlation between electrons of the same spin, but no correlation between electrons of opposite spin





Hartree Fock treatment of the homogeneous electron gas

The homogeneous electron gas is a model system that is popular in investigations of many electron effects in interacting electron systems. It consists of many electrons in a periodic volume. A positive homogeneous background charge is added that exactly cancels the charge of the electrons. One can imagine this background charge as arising from nuclei whose nuclear charge is smeared out over the whole volume. Under normal conditions it will turn out that the electronic density is a constant. Hence the total charge density at any point in space is exactly zero and there is thus no classical electrostatic interaction energy in the system. An analytical solution for the homogeneous electron gas can be found in the HF approximation. It is not difficult to see that the normalized plane wave $Vol^{-1/2}\exp(I\mathbf{k}\cdot\mathbf{r})$ is a solution of the canonical HF equation (Eq. 72) in a periodic cell of volume Vol.

- Obviously a plane wave is an eigenfunction of the kinetic energy operator
- The electronic Hartree term is canceled by the ”nucleonic” potential arising from the uniform background.
- The only non-trivial term is the exchange term of the Fock operator

$$\sum_{\mathbf{k}'}\int d\mathbf{r}'\,\frac{\exp(I\mathbf{k}\cdot\mathbf{r}')\exp(-I\mathbf{k}'\cdot\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\exp(I\mathbf{k}'\cdot\mathbf{r})$$

Since the kernel $\frac{1}{4\pi}\frac{1}{|\mathbf{r}-\mathbf{r}'|}$ is the inverse of ∇^2 and a plane wave is a eigenfunction of the Laplacian ∇^2 we have

$$\int d\mathbf{r}' \frac{\exp(I\mathbf{k}\cdot\mathbf{r}' - I\mathbf{k}'\cdot\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \propto \frac{1}{k^2}\exp(I\mathbf{k}\cdot\mathbf{r} - I\mathbf{k}'\cdot\mathbf{r})$$

and a plane wave is indeed an eigenfunction of the exchange operator.

The lowest total energy is obtained by occupying all plane waves with a wave-vector **k** whose modulus is less or equal to the Fermi wave-vector k_F. Calculating the total energy one finds

$$E = 2\sum_{k<k_F}\frac{k^2}{2} - \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\int d\mathbf{r}\,\rho^{4/3} \tag{81}$$

where ρ is the constant charge density
Exercise [1pts]: *Show that the electron density is a constant for plane wave orbitals*

The fact the the exchange part of the energy density for the homogeneous electron gas is proportional to $\rho^{4/3}$ will later on be the starting point for the density functional treatment of real systems.

What is missing in Hartree Fock

- The true many electron kinetic energy that can be expressed in terms of the natural occupation numbers n_i and natural orbitals $\phi_i^{NO}(\mathbf{r})$ (Eq. 36) as

$$E_{kin} = \sum_i n_i \int \phi_i^{*NO}(\mathbf{r}) \frac{-1}{2} \nabla^2 \phi_i^{NO}(\mathbf{r}) d\mathbf{r} \tag{82}$$

  is approximated by a finite sum over the occupied orbitals. For a closed shell system this gives

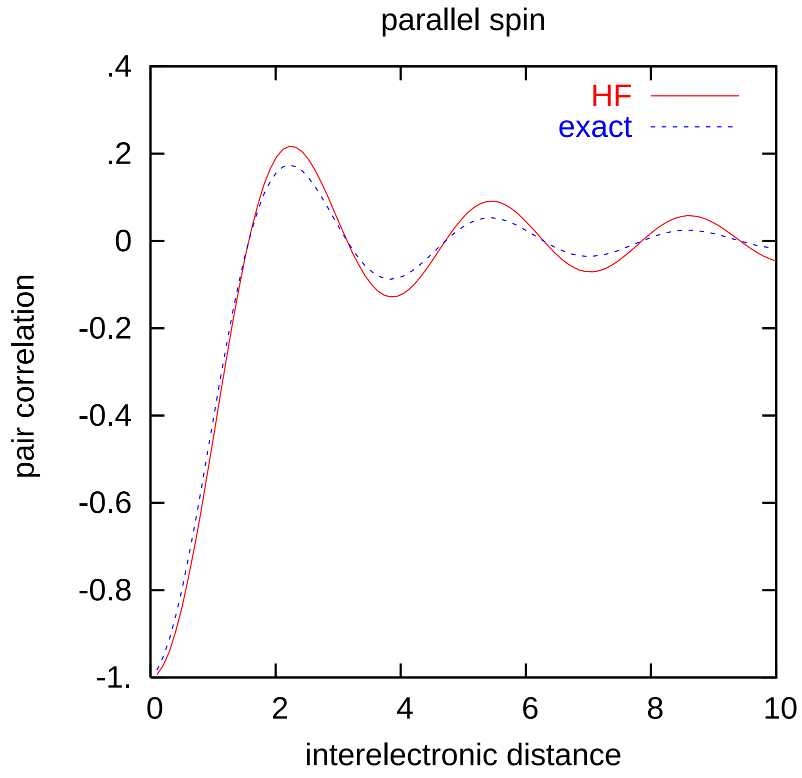
$$E_{kin} = 2 \sum_{i=1}^{N/2} \int \phi_i^{*HF}(\mathbf{r}) \frac{-1}{2} \nabla^2 \phi_i^{HF}(\mathbf{r}) d\mathbf{r}$$

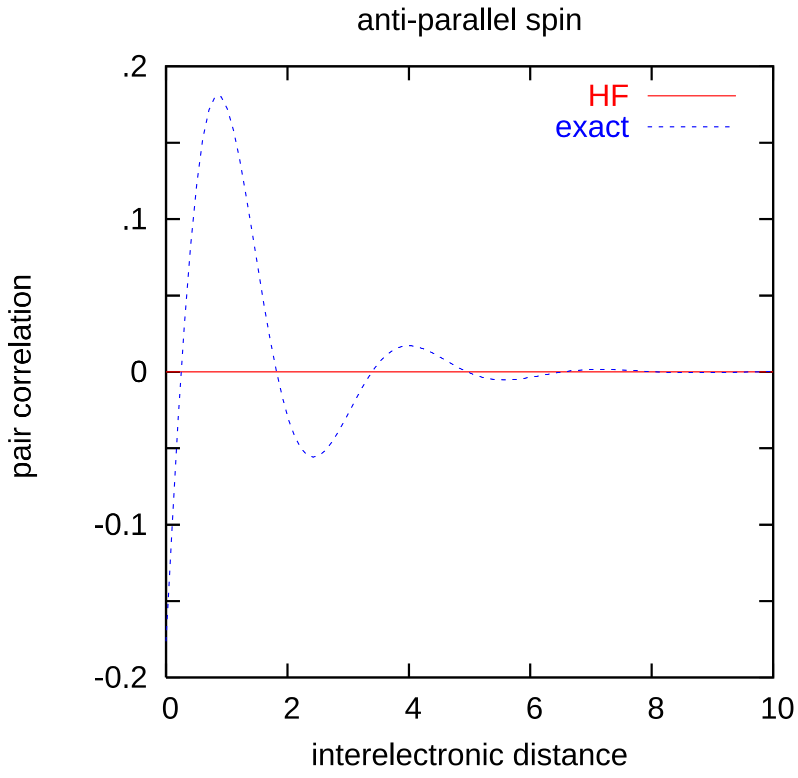
  where $\phi_i^{HF}(\mathbf{r})$ are the HF orbitals. The lowest N natural orbitals are very similar to the canonical HF functions. The critical approximation is putting all the occupation numbers for the lowest N spin orbitals to 1 (or 2 for a close shell system) and all the other ones to 0.

- It follows from Eq. 57 that γ_2 is approximated in the HF method in terms of γ_1 by

$$\gamma_2(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) = \gamma_1(\mathbf{x}_1; \mathbf{x}'_1)\gamma_1(\mathbf{x}_2; \mathbf{x}'_2) - \gamma_1(\mathbf{x}_1; \mathbf{x}'_2)\gamma_1(\mathbf{x}_2; \mathbf{x}'_1)$$

This form neglects correlations among anti-parallel spin electrons and takes only partially into account correlations among parallel spin electrons. The effect is best demonstrated for the homogeneous electron gas. If one calculates analytically the HF pair correlation function (Eq. 33) for parallel and anti-parallel spin electrons and compares with quasi exact numerical results one obtains plots with the following behavior:





Exercise [2pts]: *Derive Eq. 82 from Eq. 37 and Eq. 36*

Exercise [2pts]: *Is the HF total energy lower or higher than the exact total energy? Is the HF kinetic energy higher or lower than the exact kinetic energy? Remember that the virial theorem is valid both for the exact energy and the HF energy.*

Truncated CI and coupled cluster methods

As has been stated several times full CI and MRSCF calculations are not feasible computationally except for very small molecules, since the number of determinants that can be constructed from a given set of orbitals increases exponentially. So-called truncated CI methods select out of this very large number of determinants only a small subset that hopefully contains the most important contributions. The classification of the included determinants is according to how many electrons are excited from occupied to virtual orbitals. The occupied and virtual orbitals are the canonical orbitals that are eigenfunctions of the Fock matrix (Eq. 74). Remember that the occupied orbitals are by definition the lowest N eigenfunctions that are used to construct the HF determinant. The virtual orbitals are all the remaining eigenfunctions that have larger (generally positive) eigenvalues. Because the various determinants are constructed from the canonical HF orbitals, truncated CI and coupled cluster methods are also frequently called post Hartree-Fock methods. Because post-HF methods add additional variational parameters in form of the coefficients of the additional determinant wave-functions that are included, the energy is always lower than the energy obtained with HF. By definition the difference between the full CI energy and the HF energy is called the correlation energy. This name is misleading since, as we have seen, the HF energy already includes correlations between electrons of the same spin. But since the term correlation energy is universally used in the above described way we will therefore follow this practice. Since truncated CI methods have less degrees of freedom than the full CI method they can recover only a certain fraction of the correlation energy.

Let us now discuss in more detail the classification of determinants. In the following occupied orbitals will be denoted by the indices i, j and virtual orbitals by a, b. A post-HF wave-function can then have the following terms

- The HF ground state determinant Ψ_0
- Single excitation determinants Ψ_i^a where the occupied orbital i is replaced by the virtual orbital a.
- Double excitation determinants $\Psi_{i,j}^{a,b}$ where the occupied orbitals i and j are replaced by the virtual orbitals a and b.
- In the same way higher excitation determinants can be set up

Excitations higher than double excitations are rarely included in an exact way. The CI method where excitations up to doubles are included is called CISD. If triple excitations are included by perturbation theory the method is called CISD(T). The CISD wave-function has the following form.

$$\Psi = \Psi_0 + \sum_{i,a} C_i^a \Psi_i^a + \sum_{i<j,a<b} C_{i,j}^{a,b} \Psi_{i,j}^{a,b}$$

The number of terms in the above wave function increases as the 4th power with respect to system size because both the number of occupied and virtual orbitals is proportional to the system size. The size of the CI matrix is thus reduced in truncated CI methods compared to the full CI matrix but the cost of calculating a matrix element is the same. In the worst case it scales quadratically. The total numerical effort for the CISD method increases therefore as the 6th power. This is clearly better than the exponential increase of the full CI method but it nevertheless remains very costly.
The CISD as well as all other truncated CI methods have several severe shortcomings.

- The method is not size consistent. This means that the energy of a composite system that consists of two subsystems that are separated by a very large distance is not equal to twice the energy of the subsystem as it should be. This can be understood in the following way. If we treat with CISD only a single subsystem, two electrons of this subsystem can be excited. In the composite system we also can excite at most 2 electrons and we can therefore not simultaneously excite 2 electrons in both subsystems as it would be necessary to obtain a description that is of equal quality as the CISD description of the single subsystem.

- It can be shown, that the method is not size extensive which means that the fraction of the correlation energy recovered for a big system is less than for a small system. It would therefore not make sense to apply CISD to extended systems such as solids.

The above mentioned problems can be overcome by the so-called coupled cluster (CC) wave function method. Through its action on the HF determinant Ψ_0 we can define a single excitation operator

$$T_1\Psi_0 = \sum_i^{occ}\sum_a^{vir} C_i^a \Psi_i^a$$

a double excitation operator

$$T_2\Psi_0 = \sum_{i<j}^{occ}\sum_{a<b}^{vir} C_{i,j}^{a,b} \Psi_{i,j}^{a,b}$$

and higher order excitation operators. The coupled cluster wave-function Ψ_{cc} is now given by

$$\Psi_{cc} = \exp(T)\Psi_0$$

where the cluster operator T is given by

$$T = T_1 + T_2 + ...$$

If the cluster operator includes single and double excitations the method is called CCSD. Through the exponential ansatz higher excitations are however built into the wave-function as one can easily see by expanding the exponential

$$\exp(T_1+T_2) = 1+T_1+\left(T_2+\frac{1}{2}T_1^2\right)+\left(T_2T_1+\frac{1}{6}T_1^3\right)+\left(\frac{1}{2}T_2^2+\frac{1}{2}T_2T_1^2+\frac{1}{24}T_1^4\right)+\ldots$$

The last term in brackets represents for instance quadruple excitations since all-together 4 occupied orbitals are replaced by virtual orbitals.

Calculating the CCSD wave-function means determining the amplitudes C_i^a and $C_{i,j}^{a,b}$ of the single and double excitation operators. Determining these amplitudes is more difficult than determining for instance the coefficients of the single and doubly excited determinants in a CISD calculation. The CISD wave-function is linear in the expansion coefficients and therefore the energy is a quadratic form. As a consequence the coefficients can be found by solving an eigenvalue problem. In the CCSD method the ansatz for the wave-function contains higher powers of the excitation amplitudes. The energy expectation value is therefore not any more a quadratic form and the CC method does allow for a solution in terms of an eigenvalue problem. As a matter of fact, the calculation of the energy expectation value is too complicated. The CC wave-function is therefore determined by projection methods. This leads to a set of coupled nonlinear equations for the excitation amplitudes. Because of this solution method the CC method is not variational, i.e. one can obtain an energy that is lower than the true energy. Nevertheless the CC method is considered to be one of the best post HF methods, especially because it is size consistent and extensive. In addition to the CCSD method the CCSD(T) method is popular where triple excitation are included in perturbation theory. The scaling of the traditional implementations of the CCSD and CCSD(T) methods is the same as of the CISD and CISD(T) method, namely M^6 and M^7 where M is the size of the basis set.

Many body perturbation theory

Hartree Fock calculations frequently do not give the desired accuracy, whereas full or even truncated configuration interaction calculations are computationally not feasible. Under such circumstances a popular compromise between cost and accuracy is many body perturbation theory as derived by Möller and Plesset which gives a M^5 scaling with a small prefactor. In this scheme the zero-th order Hamiltonian $\mathcal{H}_0$ is based on the Fock operator.

$$\mathcal{H}_0 = \sum_i \mathcal{F}(\mathbf{x}_i, \mathbf{x}'_i)$$

Hence the perturbation $\mathcal{V}$ is the difference between the exact coulombic electron-electron interaction and the approximation of this interaction in the Fock operator:

$$\mathcal{V} = \sum_{i<j<N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_i \mathcal{F}^{e-e}(\mathbf{x}_i, \mathbf{x}'_i) \tag{83}$$

where

$$\mathcal{F}^{e-e}(\mathbf{x}, \mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}') V_H(\mathbf{r}) - \sum_j \frac{\psi_j^*(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{x})$$

The orbitals used for the construction of the Fock operator are the canonical Hartree Fock orbitals. It is easy to see that the HF wave function is an eigenfunction of $\mathcal{H}_0$ with eigenvalue $E^{(0)} = \sum_{i=1}^{N} \varepsilon_i$.

Exercise [2pts]: *Prove the above statement*

Hence the HF wave-function Ψ_0 is also the zero-th order wave-function for the ground state. Note that the zero-th order energy is not equal to the HF energy

The first order energy $E^1 = E^0 + \Delta E^1$ in perturbation theory is given by

$$\Delta E^{(1)} = \langle\Psi_0|\mathcal{V}|\Psi_0\rangle = \int d\mathbf{x}_1, ..., \int d\mathbf{x}_N \Psi_0^*(\mathbf{x}_1, ..., \mathbf{x}_N)\, \mathcal{V}\, \Psi_0(\mathbf{x}_1, ..., \mathbf{x}_N)$$

Since the second part of $\mathcal{V}$ in Eq. 83, $\mathcal{F}^{e-e}$ is a one body operator, its energy expectation value can be obtained in terms of the one particle density matrix which is given for the HF wave-function by Eq. 54.

$$\begin{aligned}
& \int d\mathbf{x}_1, ..., \int d\mathbf{x}_N \Psi_0^*(\mathbf{x}_1, ..., \mathbf{x}_N)\, \mathcal{F}^{e-e}\, \Psi_0(\mathbf{x}_1, ..., \mathbf{x}_N) \\
= \; & \int\int d\mathbf{x} d\mathbf{x}' \mathcal{F}^{e-e}(\mathbf{x}, \mathbf{x}')\gamma_1(\mathbf{x}', \mathbf{x}) \\
= \; & \int\int d\mathbf{x} d\mathbf{x}' \left(\frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}\delta(\mathbf{x}-\mathbf{x}')\gamma_1(\mathbf{x}, \mathbf{x}') + \frac{\sum_j \phi_j(\mathbf{x})\phi_j(\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|}\gamma_1(\mathbf{x}, \mathbf{x}') \right) \\
= \; & \sum_{i,j} [ii|jj] - [ij|ji]
\end{aligned}$$

Evaluating the first part of $\mathcal{V}$ in Eq. 83, $\sum \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$, gives obviously the HF expression for the electron electron interaction which is identical to the above expression up to a constant.

Hence

$$E^{(1)}=\sum_{i=1}^{N}\varepsilon_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{i=j}^{N}([ii|jj]-[ij|ji])$$

Thus E_1 is identical to the HF total energy as expressed in Eq. 75.
Let us next calculate the second order perturbation correction to the energy which is given by

$$\Delta E^{(2)}=\sum_{l=1,\ldots}\frac{|\langle\Psi_0|\mathcal{V}|\Psi_l\rangle|^2}{E_0^{(0)}-E_l^{(0)}} \tag{84}$$

The Slater determinant Ψ_l denotes the l-th excited state of $\mathcal{H}_0$. Such states can be formed by constructing single determinant wave-functions out of the occupied and virtual canonical HF orbitals. All these determinants are by construction eigenfunctions of $\mathcal{H}_0$. Many possibilities exist to form such determinants. We will classify them by how many of the occupied HF orbitals are replaced by virtual orbitals. In single excitations we replace one occupied HF orbital by a virtual one, in double excitation we replace two and so on. In the following the indices a and b will denote occupied orbitals whereas r and s denote virtual orbitals. Ψ_a^r will thus denote a determinant where the occupied orbital a was replaced by the virtual orbital r and Ψ_{ab}^{rs} a determinant where the occupied orbitals a and b were replaced by the virtual orbitals r and s.

Let us first consider the single excitations. The relevant matrix elements in Eq. 84 are

$$\langle\Psi_0|\mathcal{V}|\Psi_a^r\rangle = \langle\Psi_0|\mathcal{H}-\mathcal{H}_0|\Psi_a^r\rangle$$

Since Ψ_0 is a eigenfunction of $\mathcal{H}_0$ we have

$$\langle\Psi_0|\mathcal{H}_0|\Psi_a^r\rangle = E\langle\Psi_0|\Psi_a^r\rangle = 0$$

From Eq. 47 and Eq. 73 it follows that

$$\langle\Psi_0|\mathcal{H}|\Psi_a^r\rangle = \int d\mathbf{x}\int d\mathbf{x}'\psi_a(\mathbf{x})\mathcal{F}(\mathbf{x},\mathbf{x}')\psi_r(\mathbf{x}') = 0$$

$\langle\Psi_0|\mathcal{H}|\Psi_a^r\rangle$ is thus an off-diagonal element of the Fock matrix. Since the Fock matrix is diagonal for canonical HF orbitals the term vanishes as well. Consequently there are no contributions to $\Delta E^{(2)}$ from single excitations. Triple and higher excitations do not give a contribution either. This follows from the Slater Condon rule (Eq. 48).
Contributions arise only from double excitations. From Eq. 47 we get

$$\langle\Psi_0|\mathcal{H}|\Psi_{ab}^{rs}\rangle = [ar|bs]-[as|br]$$

$\langle\Psi_0|\mathcal{H}_0|\Psi_{ab}^{rs}\rangle$ is zero for the same reasons as in the case of the single excitations. Since the zero-th order energy of any determinant is equal to the sum over all the eigenvalues associated to the orbitals out of which it is build we get $E_0^{(0)} - E_{ab}^{rs\,(0)} = \varepsilon_r - \varepsilon_a + \varepsilon_s - \varepsilon_b$ and the final result for the second order energy, which is denoted in the quantum chemistry literature as MP2, is:

$$\Delta E^{(2)} = \sum_{a<b} \sum_{r<s} \frac{|[ar|bs] - [as|br]|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s} \tag{85}$$

The MP2 energy expression includes van der Waals interactions. This can easily be seen by considering two helium atoms A and B that are far away, i.e. r_{AB} is very large. In this case a set of canonical HF orbitals of the entire system consists of the union of the canonical HF atoms for the two isolated atoms. Let us assume that orbital a is a 1s like orbital centered at atom A and orbital r a 2p like orbital also centered at A. b and s are orbitals of the same type as a and r but centered on the other atom B. Since we have a closed shell system, the matrix element $[ar|bs]$ of Eq. 85 is then given by

$$[ar|bs] = 4 \int \int d\mathbf{r}\, d\mathbf{r}' \, \frac{\psi_{1s}(\mathbf{r} - \mathbf{R}_A)\, \psi_{2p}(\mathbf{r} - \mathbf{R}_A)\, \psi_{1s}(\mathbf{r}' - \mathbf{R}_B)\, \psi_{2p}(\mathbf{r}' - \mathbf{R}_B)}{|\mathbf{r} - \mathbf{r}'|}$$

The charge distribution $\rho_A(\mathbf{r}) = \psi_{1s}(\mathbf{r} - \mathbf{R}_A)\, \psi_{2p}(\mathbf{r} - \mathbf{R}_A)$ does not have a monopole because the two canonical orbitals ψ_{1s} and ψ_{2p} centered on A are orthogonal. In the same way $\rho_B(\mathbf{r}) = \psi_{1s}(\mathbf{r} - \mathbf{R}_B)\, \psi_{2p}(\mathbf{r} - \mathbf{R}_B)$ does not have a monopole. However both charge distributions have a dipole moment. Hence $[ar|bs]$ represents a dipole-dipole interaction which decays like r_{AB}^{-3}. Since the matrix elements are squared in Eq. 85, $\Delta E^{(2)} \propto r_{AB}^{-6}$ which is the well know decay behavior of van der Waals interactions. Van der Waals interactions are also included in truncated CI and coupled cluster methods as soon as double excitations are included since in this case one has CI matrix elements that contain the same type of $[ar|bs]$ terms.

Möller Plesset perturbation theory can be carried on to higher order. In addition to MP2, MP3 and MP4 are used. Higher order terms are too expensive to calculate. In addition the perturbation series frequently diverges.

Exercise [2pts]: *How rapidly decay the interactions induced by the term* $[as|br]$ *in Eq. 85. Use the fact that the canonical orbitals decay exponentially far away from their centers* $\mathbf{R}_A$ *and* $\mathbf{R}_B$

Density functional theory

Density functional theory is at present the most popular method for electronic structure calculations. It is faster than HF and the best functionals give usually higher accuracy than HF. The Hohenberg Kohn theorem is the central result upon which density functional theory is built. It demonstrates that there exists an universal functional of the density $F[\rho]$ from which the ground state can be obtained by minimizing

$$F[\rho]+\int V_{en}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$$

over all $\rho(\mathbf{r})$. The charge density that gives the minimum is the charge density of the ground state $\rho_0(\mathbf{r})$ and the total energy of the ground state E_0 is given by

$$E_0=F[\rho_0]+\int V_{en}(\mathbf{r})\rho_0(\mathbf{r})d\mathbf{r}$$

$V_{en}(\mathbf{r})$ is the external potential due to the nuclei

$$V_{en}(\mathbf{r})=\sum_{j=1}^{N_{at}}\frac{-Z_j}{|\mathbf{r}-\mathbf{R}_j|} \tag{86}$$

Before reproducing the original proof by Hohenberg and Kohn, a proof that gives an in-principle construction of the exact functional will be presented. It is called the Levy constrained search formalism. From the variational principle we know

$$E = \min_{\Psi} \langle\Psi|\mathcal{H}|\Psi\rangle$$

where $\langle\Psi|\mathcal{H}|\Psi\rangle$ equals $\int d\mathbf{x}_1 ... \int d\mathbf{x}_N \Psi(\mathbf{x}_1, ..., \mathbf{x}_N)\mathcal{H}\Psi(\mathbf{x}_1, ..., \mathbf{x}_N)$ and the search is over all normalized antisymmetric wave-functions Ψ. We can now split up this search into two parts or two 'loops'. In the inner loop we search over all wave-functions that give a certain charge density ρ and in the outer loop we search over all possible charge densities ρ

$$E = \min_{\rho} \min_{\Psi \to \rho} \langle\Psi|\mathcal{H}|\Psi\rangle$$

$\mathcal{H}$ (Eq. 27) consists of a kinetic energy part $\mathcal{H}_{kin}$, an electron-electron interaction part $\mathcal{H}_{ee}$ and an electron-nucleus interaction part $\mathcal{H}_{en}$. Whereas the contributions from $\mathcal{H}_{kin}$ and $\mathcal{H}_{ee}$ depend implicitly on ρ the $\mathcal{H}_{en}$ depends explicitly on ρ and can therefore be taken outside the constrained minimization over all the wave-functions. The inner minimization loop therefore becomes

$$\min_{\Psi \to \rho} \quad \langle \Psi | \mathcal{H}_{kin} + \mathcal{H}_{ee} | \Psi \rangle + \int V_{en}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}$$

We now define the universal functional $F[\rho]$ as

$$F[\rho] = \min_{\Psi \to \rho} \quad \langle \Psi | \mathcal{H}_{kin} + \mathcal{H}_{ee} | \Psi \rangle \tag{87}$$

It is universal because $\mathcal{H}_{kin} + \mathcal{H}_{ee}$ is independent of the atomic positions $\mathbf{R}_j$ and therefore the same for any chemical system. Hence

$$E = \min_{\rho} \quad F[\rho] + \int V_{en}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}$$

Let us now go back to the original Hohenberg-Kohn theorem. The Hohenberg-Kohn theorem states that the external potential (Eq. 86) is determined uniquely, except for a constant, by the ground state density $\rho_0(\mathbf{r})$. The proof is straightforward. Suppose there were two different external potentials, $V_1(\mathbf{r})$ and $V_2(\mathbf{r})$ which differ by more than a constant and which lead to the same ground state density $\rho_0(\mathbf{r})$. The two external potentials lead to two different hamiltonians $\mathcal{H}_1$ and $\mathcal{H}_2$. The two hamiltonians in turn give rise to two different many-electron wave-functions Ψ_1 and Ψ_2 which are hypothesized to have the same $\rho_0(\mathbf{r})$. Since Ψ_2 is not the ground state of $\mathcal{H}_1$ it follows from the variational principle that

$$E_1 = \langle\Psi_1|\mathcal{H}_1|\Psi_1\rangle < \langle\Psi_2|\mathcal{H}_1|\Psi_2\rangle$$

The strict inequality is justified by the assumption that the ground state is non-degenerate. The last term can be written

$$\begin{aligned} \langle\Psi_2|\mathcal{H}_1|\Psi_2\rangle &= \langle\Psi_2|\mathcal{H}_2|\Psi_2\rangle + \langle\Psi_2|(\mathcal{H}_1 - \mathcal{H}_2|\Psi_2\rangle && (88) \\ &= E_2 + \int (V_1(\mathbf{r}) - V_2(\mathbf{r}))\,\rho_0(\mathbf{r})d\mathbf{r} && (89) \end{aligned}$$

so that

$$E_1 < E_2 + \int (V_1(\mathbf{r}) - V_2(\mathbf{r}))\,\rho_0(\mathbf{r})d\mathbf{r}$$

Since we can exchange the role of the two superscripts 1 and 2 we can also obtain

$$E_2 < E_1 + \int (V_2(\mathbf{r}) - V_1(\mathbf{r}))\,\rho_0(\mathbf{r})d\mathbf{r}$$

Adding the two equations we arrive at the contradictory inequality

$$E_1 + E_2 < E_1 + E_2$$

This establishes the desired result. There cannot be two different external potentials differing by more than a constant which give rise to the same non-degenerate ground state density.

The Hohenberg-Kohn theorem has far reaching consequences. Once we know the potential we can in principle calculate the many-electron wave function and from it any physical properties. Hence any physical property is already determined by the density. In particular the energy is determined by the density and the functional $F[\rho_0]$ of Eq. 87 has to exist.

Both the Levy constrained search formalism and the Hohenberg-Kohn theorem do not tell us what the explicit form of the functional is. Finding such a functional is actually a rather difficult task and as a matter of fact no fully satisfactory form has been found up to now. By 'looking' at the charge density it is already very difficult to determine very fundamental things about a system such as whether it is an insulator or metal or where the chemical bonds are. So finding an accurate energy only from the density is still much harder. The most difficult part in $F[\rho]$ is the kinetic energy part. Practically all density functional calculations are therefore done in the Kohn-Sham scheme where the many electron kinetic energy is written as the kinetic energy of N independent Kohn-Sham orbitals $\phi_i^{KS}(\mathbf{r})$

$$E_{kin} = \sum_{i=1}^{N} \int \phi_i^{*\,KS}(\mathbf{r}) \frac{-1}{2} \nabla^2 \phi_i^{KS}(\mathbf{r}) d\mathbf{r} \tag{90}$$

plus a correction term that is supposed to account for the difference between the many electron kinetic energy of Eq.82 and the independent particle kinetic energy of Eq.90. The canonical Kohn-Sham orbitals are like their HF counterpart very similar to the lowest N natural orbitals. This kinetic energy correction term plus the non-classical part of the electron-electron interaction are combined into the so-called exchange correlation energy E_{xc}. The most important term in the exchange correlation energy is the exchange term that was found in the HF energy expression of the homogeneous electron gas Eq. 81 and is proportional to $\rho(\mathbf{r})^{4/3}$.

$$E_{xc}[\rho(\mathbf{r})] = \int d\mathbf{r}\, const \times \rho(\mathbf{r})^{4/3} + other\ terms$$

Depending on the nature of the 'other terms' we have the following classification:

- Local Density Approximation (LDA): the 'other terms' in the exchange correlation functional depend only on the charge density $\rho(\mathbf{r})$. The numerical form of E_{xc} was obtained by a fit to highly accurate numerical many electron calculations based on the Quantum Monte Carlo method. This approximation is thus quasi exact for the homogeneous electron gas and is supposed to be accurate if the variation of the charge density is weak. For this reason it works best for solids. In the case of atoms or molecules the density decays to zero on the surface of the atom or molecule and it can not any more be considered as slowly varying. The LDA approximation gives typically a too high energy for such finite systems. As a consequence the bonding energies of solids are in general too large. The bonding energy is the difference in energy between the energy of a system containing many individuals atom and the energy of the solid that these individual atoms can form.

- Generalized Gradient Approximation (GGA): the exchange correlation functional depends on the charge density $\rho(\mathbf{r})$ and its gradient. Systems with rapidly varying charge densities are better described. Since GGA functionals have to coincide with the LDA exchange correlation functional for the case where the density is constant, the functional form of these GGA's is typically the following. The LDA functionals times an enhancement factor that goes to one for a constant density, but grows to values larger than one if the gradient is non-vanishing. The enhancement factor depends on the reduced dimensionless gradient s

$$s(\mathbf{r}) = \frac{|\nabla\rho(\mathbf{r})|}{2(3\pi^2)^{1/3}\rho(\mathbf{r})^{4/3}}$$

- meta GGA: depends in addition on the kinetic energy density $\sum_{i=1}^{N} \phi_i^{*\,KS}(\mathbf{r})\nabla^2\phi_i^{KS}(\mathbf{r})$
- hybrid functionals: contain in addition some fraction of the HF exchange energy

Within each class of exchange correlation functionals there are many proposed forms. This zoo of functionals is so large that it is difficult to keep track of all functionals on the market.

The Kohn-Sham equations in the LDA approximation

For simplicity we will consider a closed shell system. The LDA total energy is given by

$$E = -\sum_{i=1}^{N/2} \int \phi_i^*(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}) d\mathbf{r} + \int V_{en}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' d\mathbf{r} + E_{xc}^{LDA}[\rho(\mathbf{r})] \tag{91}$$

where the charge density $\rho(\mathbf{r})$ is the sum over the square of all the occupied Kohn-Sham orbitals

$$\rho(\mathbf{r}) = 2 \sum_{i=1}^{N/2} \phi_i^*(\mathbf{r}) \phi_i(\mathbf{r}) \tag{92}$$

The first term in Eq. 91 is the kinetic energy of N independent electrons, the second the interaction of the electrons with the nuclei and potentially other external potentials, the third the classical electron-electron repulsion and the last the above described exchange correlation energy. The Kohn-Sham equations are obtained by minimizing the total energy expression Eq. 91 under the constraint that the orbitals ϕ_i are orthonormal. The procedure is analogous to the HF case. Applying the rules for functional derivatives, we obtain for the unconstrained gradient $d_i(\mathbf{r}) = \frac{1}{2} \frac{\delta E}{\delta \phi_i^*(\mathbf{r})}$

$$d_i(\mathbf{r}) = -\frac{1}{2} \nabla^2 \phi_i(\mathbf{r}) + V_{en}(\mathbf{r})\, \phi_i(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}'\, \phi_i(\mathbf{r}) + v_{xc}^{LDA}(\rho(\mathbf{r}))\, \phi_i(\mathbf{r}) \tag{93}$$

where the exchange correlation potential is defined as

$$v_{xc}^{LDA}(\rho(\mathbf{r})) = \frac{\delta E_{xc}^{LDA}(\rho(\mathbf{r}))}{\delta\rho(\mathbf{r})}$$

Introducing the Kohn-Sham Hamiltonian $\mathcal{H}^{KS}$

$$\mathcal{H}^{KS} = -\frac{1}{2}\nabla^2 + V(\mathbf{r}) \tag{94}$$

where the Kohn-Sham potential V is the sum of the external potential, the Hartree potential and the exchange correlation potential

$$V(\mathbf{r}) = V_{en}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{xc}^{LDA}(\rho(\mathbf{r}))$$

the unconstrained gradient can simply be written as

$$d_i(\mathbf{r}) = \mathcal{H}^{KS}\phi_i(\mathbf{r}) \tag{95}$$

The condition that the constrained gradient vanishes becomes according to Eq. 67

$$\mathcal{H}^{KS}\phi_i(\mathbf{r}) - \sum_{j=1}^{N/2} \Lambda_{i,j}\phi_j(\mathbf{r}) = 0 \tag{96}$$

where

$$\Lambda_{i,j} = \int \phi_j^*(\mathbf{r}) d_i(\mathbf{r}) d\mathbf{r} = \int \phi_i(\mathbf{r}) d_j^*(\mathbf{r}) d\mathbf{r} = \int \phi_j^*(\mathbf{r}) \mathcal{H}^{KS} \phi_i(\mathbf{r}) d\mathbf{r} \tag{97}$$

Exercise [3pts]: *Using functional derivatives verify Eq. 93*

As in the HF case the total energy (Eq. 91) is invariant under unitary transformations among the occupied orbitals. We may therefore again choose canonical orbitals which diagonalize the matrix Λ in Eq. 97 and the condition that the constrained gradient vanishes (Eq. 96) results in the eigenvalue problem

$$\mathcal{H}^{KS} \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}) \tag{98}$$

The canonical orbitals satisfying the above equation are called the Kohn-Sham orbitals and the eigenvalues are called the Kohn-Sham energies. Like in HF only the occupied orbitals enter into the energy expression of Eq. 91. The virtual levels have no physical meaning except that the virtual ε_i's give like in HF approximative electron affinities. The occupied ε_i's give also like in HF approximate ionization energies.

Spin polarization

Up to now we have ignored spin in our treatment of density functional theory. For a closed shell system where the density of spin up electrons equals the density of the spin down electrons, it is indeed only the density that matters. For other systems such as magnetic systems we have a spin polarization which means that the up and down electron densities are not equal. In such a case we have to use an exchange correlation functional that depends both on the density of the spin up and spin down electrons and, and in the case of a GGA on the gradients of these two densities. There exist versions of all current exchange correlation functionals that allow for the inclusion of spin polarization.

The spin polarization is determined by the spin quantum number S_z. The energy of a true many electron system does not only depend on S_z but also on the spin quantum number S^2. This later dependence is not taken into account by the current exchange correlation functionals. Wave-function methods such as CI have to be used to describe the spin multiplicity of atoms and molecules.

Treatment of system with degeneracies at the Fermi level

For systems with a degeneracy or near degeneracy at the Fermi level, i.e. for systeams where the Homo-Lumo gap either vanishes or is very small, so called fractional occupation numbers have to be introduced to obtain convergence in electronic structure calculations. In those schemes virtual orbitals get assigend some small occupation numbers, whereas the occupation number of the highest occupied orbitals are somewhat reduced from one. Based on the fractional occupation numbers we next define some "free energy" functional $F = E - TS$ that consists of an energy term E very similar to the non-degenerate case plus some entropy term S. We consider again for simplicity a closed shell system.

$$\begin{aligned} F &= -\sum_{i=1}^{\infty} f_i \int \phi_i^*(\mathbf{r}) \nabla^2 \phi_i(\mathbf{r}) d\mathbf{r} + \int V_{en}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' d\mathbf{r} + E_{xc}^{LDA}[\rho(\mathbf{r})] \\ &+ 2 k_B T \sum_{i=1}^{\infty} f_i \ln(f_i) + (1-f_i) \ln(1-f_i) = E - TS \end{aligned} \tag{99}$$

where the charge density $\rho(\mathbf{r})$ is now also given by a slightly modified formula:

$$\rho(\mathbf{r}) = 2 \sum_{i=1}^{\infty} f_i \phi_i^*(\mathbf{r}) \phi_i(\mathbf{r}) \tag{100}$$

In practice it is not really necessary to extend the summation in the above formula up to infinity since the occupation numbers f_i tend rapidly towards zero when the eigenvalues ε_i

get large. The term $S = -2k_B \sum_i f_i \ln(f_i) + (1 - f_i) \ln(1 - f_i)$ in Eq. 99 can be interpreted as some kind of electronic entropy term of an electron-hole system. The main practical advantage of this term is however that it allows for a fully variational treatment with simple derivative formulas. If we consider the f_i's as additional variational degrees of freedom, we can calculate in the Lagrange multiplier formalism the derivative of F with respect to the occupation numbers under the constraint that all the occupation numbers sum up to the total number of electrons N

$$\frac{dF}{df_i} = \frac{\partial}{\partial f_i}\left[F - \mu\left(\sum_j f_j - N\right)\right] = \frac{\partial F}{\partial f_i} - \mu = 0 \qquad (101)$$

where μ is the Lagrange multiplier. In the chosen notation, $\frac{dF}{df_i}$ is meant to be the total gradient including the constraint. Now

$$\frac{\partial E}{\partial f_i} = \langle \varepsilon_i \rangle$$

where $\langle \varepsilon_i \rangle$ is the energy expectation value of the orbital ϕ_i. In addition

$$\frac{1}{k_B}\frac{\partial S}{\partial f_i} = -\left(\ln(f_i) + 1 - \ln(1 - f_i) - 1\right) = \ln\left(\frac{1 - f_i}{f_i}\right)$$

Zeroing the total gradient of Eq. 101 therefore gives

$$\langle \varepsilon_i \rangle - \frac{1}{\beta} \ln \left(\frac{1 - f_i}{f_i} \right) - \mu = 0$$

where $\beta = 1/(k_B T)$. Solving the above equation for f_i gives

$$f_i = \frac{1}{1 + exp(\beta(\langle \varepsilon_i \rangle - \mu))} \tag{102}$$

where μ is determined by the constraint $\sum_i f_i = N$.

Let us next calculate the unconstrained gradient $d_i(\mathbf{r})$ of F with respect to the orbitals. We have to take into account that F depends not only directly on the orbitals ϕ_i but also indirectly through the f_i's.

$$d_i(\mathbf{r}) = \frac{\partial F}{\partial \phi_i^*(\mathbf{r})} + \sum_j \frac{dF}{df_j} \frac{\partial f_j}{\partial \phi_i^*(\mathbf{r})} \tag{103}$$

If we now require that the f_i's satisfy Eq. 102, i.e. that they minimize F for the present given set of orbitals, then $\frac{dF}{df_j} = 0$ and the second term disappears. Hence

$$d_i(\mathbf{r}) = \frac{\partial F}{\partial \phi_i^*(\mathbf{r})} = f_i \mathcal{H}^{KS} \phi_i(\mathbf{r}) \tag{104}$$

The Lagrange multipliers imposing the orthogonality constraints of Eq 67 are then given by

$$\Lambda_{i,j} = \frac{1}{2}\left(f_i h_{i,j} + f_j h_{i,j}\right) \tag{105}$$

where $h_{i,j} =< \phi_i|\mathcal{H}^{KS}|\phi_j >$ Hence the constrained gradient is given by

$$\frac{\delta F}{\delta \phi_i^*(\mathbf{r})} = f_i \mathcal{H}^{KS}\phi_i(\mathbf{r}) - \sum_j \frac{1}{2}\left(f_i + f_j\right) h_{i,j}\phi_j(\mathbf{r}) \tag{106}$$

It can immeadiatley be seen that this gradient expression is zero if the $\phi_j(\mathbf{r})$'s are eigenfunctions of $\mathcal{H}^{KS}$, in which case $h_{i,j} = \varepsilon_i \delta_{i,j}$. In this case the expectation value in Eq. 102 also becomes an eigenvalue and we recover the well established Fermi distribution f_i. So also in this case the equivalence of a variational minization of the free energy expression and an approach based on a selfconsistent diagonalization of the Kohn-Sham Hamiltonian is guaranteed.

The temperature in a free energy functional also mimics in a certain sense many body effects. As we have seen introducing electronic correlation lowers the energy, even though the kinetic energy increases. The free energy also decreases with increasing temperature even though the kinetic energy increases.

Exercise [2pts]: *Prove that* $\frac{dF_0}{dT} < 0$ *where* F_0 *is a minimum of the free energy functional* F *with respect to the occupation numbers and orbitals*

The ionic forces are given by the same expression (Eq. 41) with respect to ρ independently of whether the energy or free energy functional is used. When we calculate the derivative of the F or E with respect to the atomic positions, we have to consider in principle also the variation of the wave functions ϕ_i or the occupation number f_i with respect to the atomic positions. Since by Eq. 102 the occupation numbers depend on the orbitals this later dependence can be included in the ϕ_i dependence,

$$\frac{dF}{d\mathbf{X}_1} = \frac{\partial F}{\partial \mathbf{X}_1} + \sum_i \frac{dF}{d\phi_i} \frac{\partial \phi_i}{\partial \mathbf{X}_1} \tag{107}$$

Here $\frac{dF}{d\phi_i}$ denotes again the total derivative, i.e the derivative that includes all the constraints, This derivative is zero since F is minimized with respect to the ϕ_i's under the constraints. Hence the forces are given by the standard Hellmann-Feynman formulas:

$$\frac{dF}{d\mathbf{X}_1} = \frac{\partial F}{\partial \mathbf{X}_1} = -Z_1 \int \rho(\mathbf{r}) \frac{X_1 - x}{|\mathbf{R}_1 - \mathbf{r}|^3} d\mathbf{r} + \sum_{j=2}^{N_{at}} \frac{Z_1 Z_j (X_1 - X_j)}{|\mathbf{R}_1 - \mathbf{R}_j|^3} \tag{108}$$

where ρ is given by Eq. 100.

Gaussian smearing schemes

The previous approach for handling degeneracies is appealing since it leads to a Fermi distribution of the fractional occupation numbers. It has however the disadvantage that the Fermi distribution does not decay very rapidly to zero and that is is therefore necessary to include a substantial number of virtual orbitals in the calculation. For this reason so-called Gaussian smearing schemes are frequently used. In this case the relation between f_i and $\langle \varepsilon_i \rangle$ is given by

$$f_i^{Gaus} = \frac{1}{2}(1 - erf(\beta(\varepsilon_i - \mu))$$

For this distribution the f_i's have an extremely fast Gaussian like decay for large $\langle \varepsilon_i \rangle$'s. If one scales the temperature by $4/\sqrt{\pi}$ the resuling distribution function

$$f_i^{scaled} = \frac{1}{2}(1 - erf(\beta \frac{4}{\sqrt{\pi}}(\langle \varepsilon_i \rangle - \mu))$$

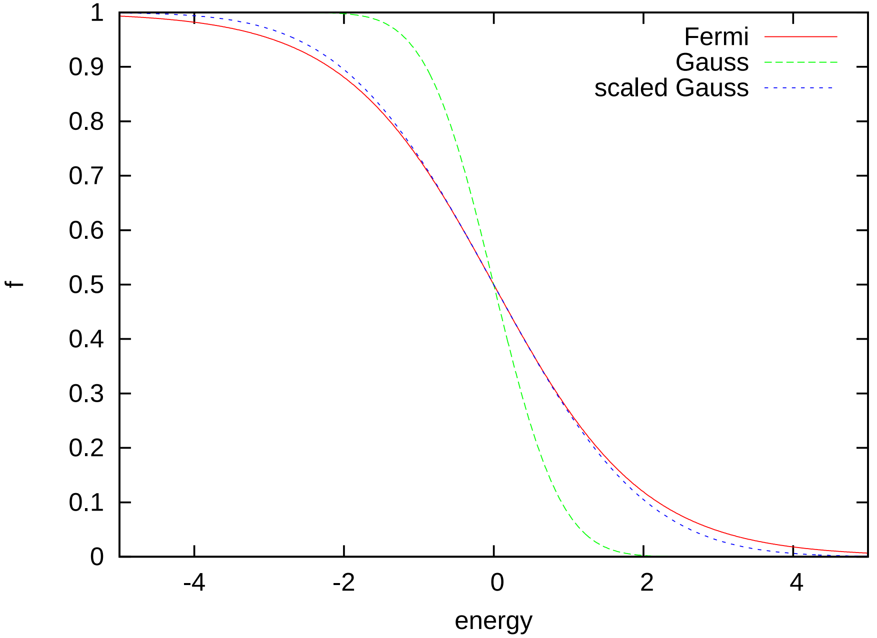
has the same slope at the Fermi level and is therefore very similar to the Fermi distribution in the region around the Fermi level, which is set to zero in the Figure below. β is set to one.

The entropy like term has be chosen in a Gaussian smearing scheme as

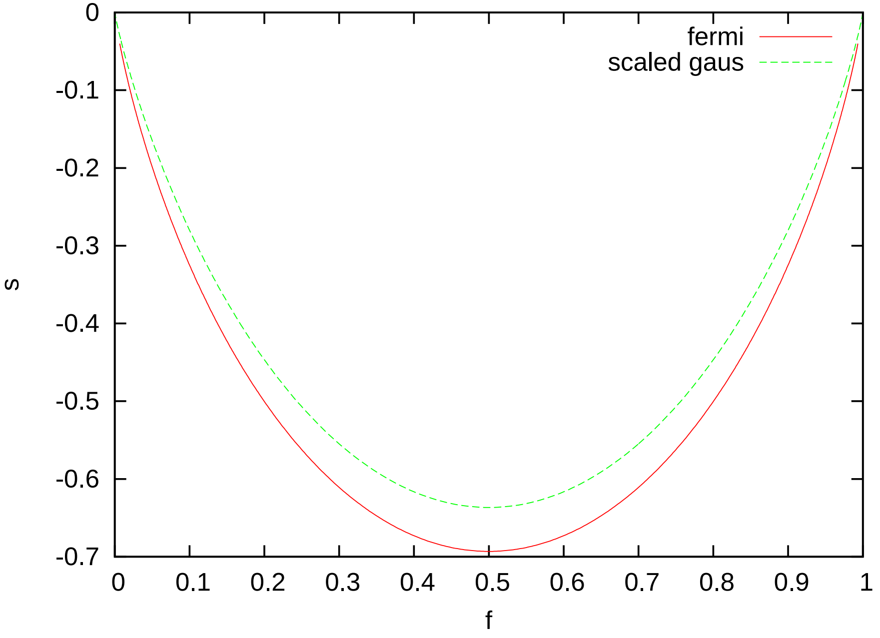
$$S = \sum_i \frac{1}{2\sqrt{\pi}} exp(-\beta(\varepsilon_i - \mu))$$

The entropy functions look then also rather similar to the case of the Fermi distribution case as shown below.

Distribution



Entropy



Major shortcoming of density functional schemes

- The theoretical foundations of the Kohn-Sham scheme are uncertain:
  Whereas the Hohenberg-Kohn theorem is rigorous, the Kohn-Sham theorem is based on an unproven assumption, namely the so-called non-interacting V-representability condition. This condition assumes that there exists for the density of any system of interacting electrons a potential such that the density of non-interacting electrons moving in this potential is identical to the density of the interacting electrons.

- Failure to describe spin multiplicity:
  This was discussed before

- Failure to describe strongly correlated systems:
  There exists at present no functional that could describe in a qualitatively correct way strongly correlated systems

- Failure to describe transition states:
  The total energies predicted for transition states are very unreliable. Both in this case and in the case of strongly correlated systems the true wave-function has not only one dominating determinant but several important determinants. This multi-configuration character of the wave-function becomes visible in the occupation

numbers. Since the occupation number of a certain natural orbital is equal to the sum of the squared coefficients of all the determinants that contain this natural orbital, there will be occupation numbers that are neither close to one or zero. If there is only one dominating determinant, the occupation number of the occupied natural orbitals contained in this determinant will be close to one whereas the occupation numbers of all the other virtual orbitals are very small. That the Kohn-Sham methods fail in this case is not so surprising. The Kohn-Sham kinetic energy term (Eq. 90) does not any more give high precision for the kinetic energy of the system of interacting electrons which, according to Eq. 82 will have important contributions from some virtual orbitals.

- Failure to describe van-der Waals forces:
  Even though van-der Waals forces are important for a wide range of problems such as the interaction of biomolecules, standard functionals do not include them. Functionals that include van-der Waals interaction are however now beginning to appear but they are numerically costly and their accuracy is not yet firmly established.

- No integer preference:
  When a molecule is torn apart into two fragments each fragment will have an integer number of electrons. If one pulls apart for instance the two atoms in a NaCl molecule then each atom will either be neutral or perhaps Cl has captured one of

the electrons of the Na. If one pulls apart in density functional theory the NaCl molecule the two atoms will have non-integer charges.

- Gap problem:
  The particle gap of a solid is defined as the difference between the electron affinity and the ionization energy. For a finite system this quantity can be calculated as

  $$(E(N+1)-E(N))-(E(N)-E(N-1))$$

  and DFT results are rather reasonable. It can be shown that for an infinite system this quantity is given by the difference $\varepsilon_{LUMO}-\varepsilon_{HOMO}$, where ε_{HOMO} is the Kohn-Sham eigenvalue of the Highest Occupied Molecular Orbital and ε_{HOMO} the eigenvalue of the Lowest Unoccupied Molecular Orbital. Both the HOMO and LUMO orbital satisfy Eq. 98. The HOMO is the orbital number N ($N/2$ for a closed shell system) and the LUMO is the orbital number $N+1$ ($N/2+1$). The HOMO enters into the expressions for the total energy, whereas the LUMO does not. The difference $\varepsilon_{LUMO}-\varepsilon_{HOMO}$ is much smaller than the experimental particle gap in most density functional schemes.

- Imperfect cancellation of the self-interaction:
  As we have seen in the discussion of the HF method one orbital does not interact with itself. This kind of unphysical self-interaction is however contained in all standard density functional schemes. The problem can most easily be seen in the case

of the hydrogen atom. In principle the exchange correlation potential should cancel the Hartree potentials, but it does not completely in practice. As a consequence density functional theory does not give the correct total energy of -1/2 for the hydrogen atom. The consequences of the incomplete cancellation of the self-interaction are most detrimental for single electron systems and for open shell systems that have a localized reactive unpaired electron. The imperfect selfinteraction cancellation also prevents the stable bonding of an additional electron to a neutral atom, since the Kohn-Sham potential decays exponentially to zero whereas it should decay to zero like $1/r$.

- Can not be improved in a systematic way:
  The wave-function methods that were previously discussed can be improved in a systematic way. In a CCSD calculation one can for instance include triple excitations as well to get a more accurate answer if the computing time allows for it. Such improvements are not possible in density functional theory.

In spite of all these shortcomings, it must be stressed again that density functional theory is at present the most widely used method for electronic structure calculations, because it gives surprisingly good energies with an acceptable numerical effort as can be seen from the two tables below.

The first table shows typical errors for various quantities. QCISD is another truncated CI method similar to the CCSD method. The error in the barrier is for the chemical reaction $H_2 + H \rightarrow H + H_2$. The experimental barrier is 9.7 and LDA predicts no barrier at all for this reaction. All other errors are averages over certain test sets containing various molecules.

Error in:	HF	MP2	QCISD	LDA	GGA
bond length (Angstroem)	0.020	0.014	0.013	0.021	0.020
bond angle (degrees)	2.0	1.8	1.8	1.9	2.3
frequencies (cm^{-1})	59	63	37		45
atomization energies (kcal/mol)	86	22	29	36	6
barrier (kcal/mol)	7.9	3.5	.2	-13	-5.6

The table below gives some indication of the size of the system that can be treated with different methods. Tight Binding (TB) is an approximative quantum mechanical method and Force Fields (FF) represent a parameterization of the Born Oppenheimer surface.

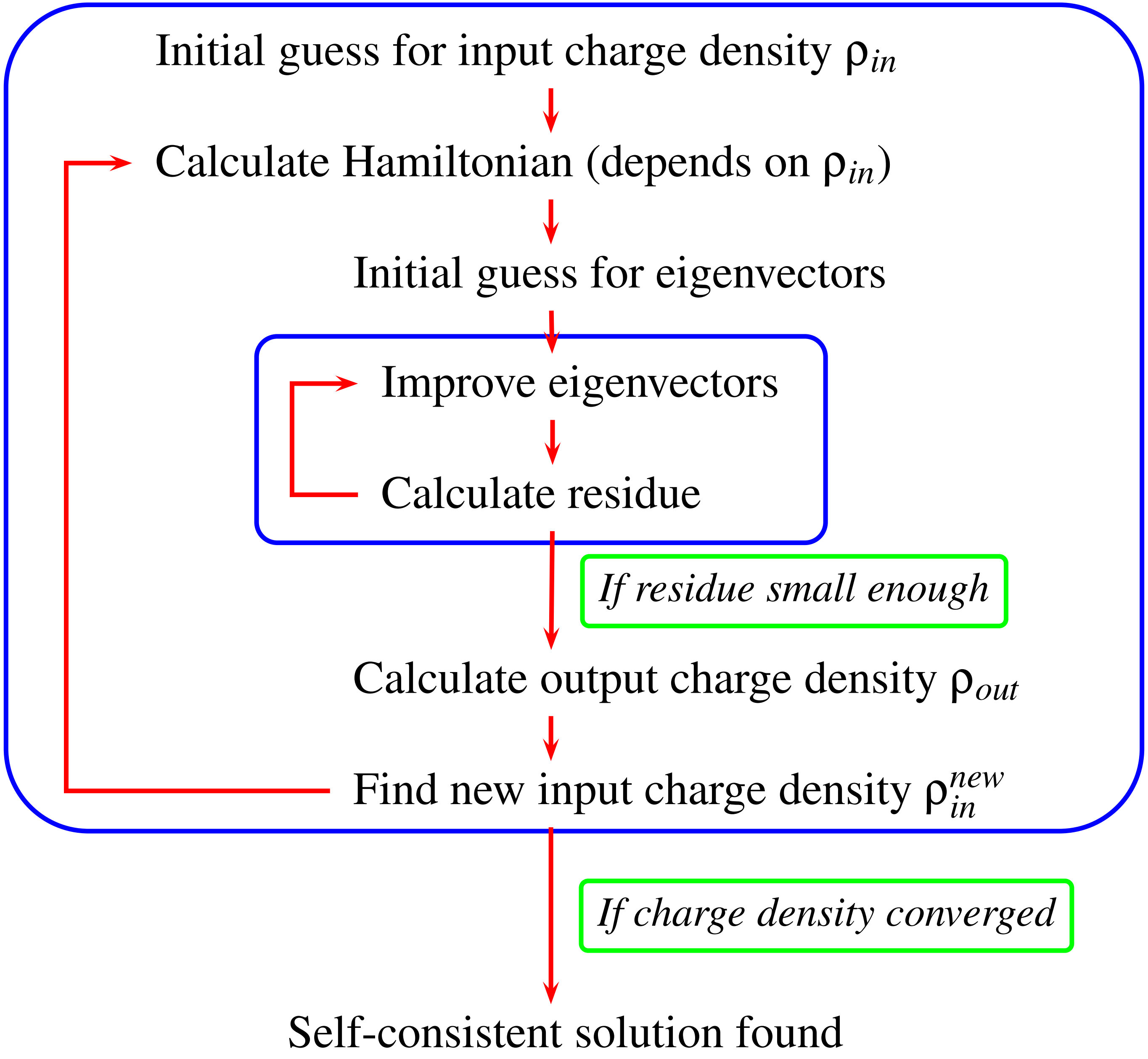
	HF	MP2	QCISD	LDA	GGA	TB	FF
Number of atoms	200	50	10	500	500	5000	100 000

Basic approaches to the numerical solution of the Kohn-Sham equations

There are two possibilities to solve the Kohn-Sham equations

- Direct numerical minimization of the total energy:
  One calculates the gradient (Eq. 96) and ”goes down” along this gradient until it vanishes

- Self-consistent eigenvalue problem:
  One solves the eigenvalue problem of Eq. 98. The equation is however more difficult than a simple eigenvalue problem. Note that the Hartree and exchange correlation potentials depend on the charge density. The correct charge density which is given in terms of the solution by Eq. 92 is of course not yet available at the start of the calculation. This problem can be circumvented by self-consistency iterations. One repeats the solution of the eigenvalue problem until the output charge density calculated form the eigenorbitals is equal to the input charge density that was used for the calculation of the potential. In the resulting flowchart shown below it is also assumed that the eigenvalue problem is solved by an iterative eigenvalue problem solver. The eigenvalue iterations are indicated by the small blue inner box, the self-consistency iterations by the large blue outer box.



Initial guess for input charge density ρ_{in}

Calculate Hamiltonian (depends on ρ_{in})

Initial guess for eigenvectors

Improve eigenvectors

Calculate residue

If residue small enough

Calculate output charge density ρ_{out}

Find new input charge density ρ_{in}^{new}

If charge density converged

Self-consistent solution found

Discretization of continuous differential equations

Up to now we have used in this course the language of traditional calculus, i.e. continuous functions and differential and integral operators acting on these functions. This language is not suited for numerical work, since one can not represent a continuous function on a computer. All one can represent are vectors of numbers. For numerical processing we consequently have to discretize our equations. This is done by expanding the wave-function in terms of a finite number M of basis functions $U_k(r)$

$$\psi_i(\mathbf{r}) = \sum_{k=1}^{M} u_i(k) U_k(\mathbf{r}) \quad ; \quad \psi_i^*(\mathbf{r}) = \sum_{k=1}^{M} u_i^*(k) U_k^*(\mathbf{r}) \tag{109}$$

This gives us an energy expression that depends on N finite vectors $u_i(k)$ of length M. Taking then the partial derivatives with respect to the degrees of freedom $u_i(k)$ of the numerical wave-function under the orthogonality constraints gives a set of discrete Euler-Lagrange equations, whose solution will provide us with the numerical wave-function that minimizes the energy within this basis set, i.e. with the numerical solution. The partial derivatives containing the orthogonality constraints can be calculated using the chain rule

$$\frac{\partial E}{\partial u_i^*(k)} = \int \frac{\delta E}{\delta \psi_i^*(\mathbf{r})} \frac{\partial \psi_i^*(\mathbf{r})}{\partial u_i^*(k)} d\mathbf{r} = \int \frac{\delta E}{\delta \psi_i^*(\mathbf{r})} U_k^*(\mathbf{r}) d\mathbf{r} \tag{110}$$

In analogy to Eq. 64 we defined the unconstrained gradients

$$
\begin{aligned}
d_j(k) &= \frac{1}{2}\int \frac{\delta E}{\delta\tilde{\psi}_j^*(\mathbf{r})}\frac{\partial\tilde{\psi}^*(\mathbf{r})}{\partial u_j^*(k)}d\mathbf{r} = \int d_j(\mathbf{r})U_k^*(\mathbf{r})d\mathbf{r} && (111)\\
d_j^*(k) &= \frac{1}{2}\int \frac{\delta E}{\delta\tilde{\psi}_j(\mathbf{r})}\frac{\partial\tilde{\psi}(\mathbf{r})}{\partial u_j(k)}d\mathbf{r} = \int d_j^*(\mathbf{r})U_k(\mathbf{r})d\mathbf{r} && (112)
\end{aligned}
$$

We have not introduced here coefficients $\tilde{u}_j(k)$ since, as explained previously, we immediately switch back to orthonormal orbitals after calculating derivatives. Using the chain rule of Eq. 110 in the expression for the constrained gradient Eq. 67 we obtain

$$
\begin{aligned}
\frac{1}{2}\frac{\partial E}{\partial u_i^*(k)} = \int d_i(\mathbf{r})U_k^*(\mathbf{r})d\mathbf{r} &- \frac{1}{2}\sum_j\left(\int d_j^*(\mathbf{r}')\psi_i(\mathbf{r}')d\mathbf{r}'\right)\int \psi_j(\mathbf{r})U_k^*(\mathbf{r})d\mathbf{r}\\
&- \frac{1}{2}\sum_j\left(\int d_i(\mathbf{r}')\psi_j^*(\mathbf{r}')d\mathbf{r}'\right)\int \psi_j(\mathbf{r})U_k^*(\mathbf{r})d\mathbf{r}
\end{aligned}
$$

Using the definition of the discrete unconstrained gradient of Eq. 111 the integrals can be replaced by sums in the Lagrange multiplier matrix

$$
\Lambda_{i,j} = \int d_j^*(\mathbf{r}')\psi_i(\mathbf{r}')d\mathbf{r}' = \sum_k u_i(k)\int d_j^*(\mathbf{r}')U_k(\mathbf{r}')d\mathbf{r}' = \sum_k u_i(k)d_j^*(k) \qquad (113)
$$

The remaining two integrals can be expressed in terms of the overlap matrix S

$$\int U_k^*(\mathbf{r})\psi_j(\mathbf{r})d\mathbf{r} = \sum_l u_j(l) \int U_k^*(\mathbf{r})U_l(\mathbf{r})d\mathbf{r} = \sum_l S(k,l)u_j(l) \tag{114}$$

Note that this overlap matrix $S(k,l) = \int U_k^*(\mathbf{r})U_l(\mathbf{r})d\mathbf{r}$ is the overlap matrix between the basis functions and it should not be confused with the overlap matrix between the orbitals $\mathcal{S}_{i,j} = \int \psi_i(\mathbf{r})\psi_j(\mathbf{r})d\mathbf{r}$. We thus obtain the final expression for the discrete gradient under the orthogonality constraints

$$\frac{1}{2}\frac{\partial E}{\partial u_i^*(k)} = d_i(k) - \frac{1}{2}\sum_j \left(\Lambda_{i,j} + \Lambda_{j,i}^*\right)\sum_l S(k,l)u_j(l) \tag{115}$$

where $\Lambda_{i,j}$ is given by Eq. 113. Eq. 115 is the central equation for numerical work. It is valid for any independent particle scheme such as HF or Density Functional theory. What changes from one independent particle scheme to the other is only the form of the unconstrained gradient d_i. In the discrete case the solution vectors u_i are again not uniquely defined. Any other set of vectors that is related to the original set of vectors by an unitary transformation is an equally valid solution. There exists therefore a set of solution vectors that diagonalizes Λ and satisfies

$$d_j(k) - \varepsilon_j \sum_l S(k,l)u_j(l) = 0 \tag{116}$$

Eq. 116 is the basic equation if a selfconsistent diagonalization approach is adopted whereas Eq. 115 is the basic gradient expression in a numerical minimization procedure. Once one has found the solution vectors of Eq. 115 one can transform the set of solution vectors u_i into the canonical set satisfying Eq. 116 if the canonical orbitals are desired.
It is instructive to consider the simplest case of a system of N non-interacting electrons in an external potential $V_{ext}(\mathbf{r})$. The total energy of such a system is given by

$$E = \sum_{i=1}^{N} \int \psi_i(\mathbf{r}) \left(-\frac{1}{2}\nabla^2 + V_{ext}(\mathbf{r}) \right) \psi_i(\mathbf{r})\, d\mathbf{r} = \sum_{i=1}^{N} \varepsilon_i \tag{117}$$

The unconstrained gradient of Eq. 117 is

$$d_i(\mathbf{r}) = \left(-\frac{1}{2}\nabla^2 + V_{ext}(\mathbf{r}) \right) \psi_i(\mathbf{r}) = \mathcal{H}\psi_i(\mathbf{r})$$

Hence the discrete unconstrained gradient of Eq. 111 is

$$d_i(k) = \sum_l u_i(l) \int U_k^*(\mathbf{r}) \mathcal{H} U_l(\mathbf{r}) = \sum_l H(k,l) u_i(l)$$

and the constrained gradient of Eq. 115 becomes

$$\frac{1}{2}\frac{\partial E}{\partial u_i^*(k)} = \sum_l H(k,l)u_i(l) - \frac{1}{2}\sum_j \left(\Lambda_{i,j} + \Lambda_{j,i}^* \right) \sum_l S(k,l) u_j(l) \tag{118}$$

The canonical equations Eq. 116 become

$$\sum_l H(k,l)u_i(l) - \varepsilon_i \sum_l S(k,l)u_i(l) = 0 \tag{119}$$

Eq. 119 is just the generalized eigenvalue problem known from elementary quantum mechanics. In matrix notation it reads

$$H\vec{u}_i - \varepsilon_i S\vec{u}_i = 0$$

If the basis set $U_k(\mathbf{r})$ is taken to be orthonormal, i.e. if $S(k,l) = \int U_k^*(\mathbf{r})U_l(\mathbf{r})\,d\mathbf{r} = \delta_{k,l}$ this general eigenvalue problem becomes an ordinary eigenvalue problem

$$\sum_l H(k,l)u_i(l) - \varepsilon_i u_i(k) = 0$$

or in matrix notation

$$H\vec{u}_i - \varepsilon_i \vec{u}_i = 0$$

In the case of a density functional or Hartree Fock calculation the equations are identical except that one has to replace the Hamiltonian matrix either by the Kohn-Sham matrix or the Fock matrix

$$H_{i,j}^{KS} = \int U_i(\mathbf{r})\mathcal{H}^{KS}U_j(\mathbf{r})\,d\mathbf{r} \quad ; \quad F_{i,j} = \int U_i(\mathbf{r})\mathcal{F}U_j(\mathbf{r})\,d\mathbf{r}$$

Gaussian basis sets

In the chemistry community the most popular basis set are atom centered Gaussian type orbitals (GTOs). On each atom i with position $\mathbf{R}_i$ several Gaussian like functions of form

$$G(\mathbf{r}-\mathbf{R}_i)$$

are centered. In the following discussion of the nomenclature of these orbitals G we will assume for simplicity that $\mathbf{R}_i = 0$

- Primitive Gaussian type orbitals:
  If the radial part is just a single Gaussian specified by the exponent α it is called a primitive Gaussian. The angular part can can either be given by spherical harmonics. i.e
  $$G(\mathbf{r}) = Y_{l,m}(\theta,\phi)\exp\left(-\alpha(x^2+y^2+z^2)\right)$$
  or by cartesian coordinates in which case
  $$G(\mathbf{r}) = x^{l_1}y^{l_2}z^{l_3}\exp\left(-\alpha(x^2+y^2+z^2)\right)$$
  The quality of the basis depends on how many Gaussians with different exponents are used for each atomic shell of the atom on which they are centered.

- Minimal basis sets:
  If only one exponent is used the basis set is called minimal basis set. For hydrogen one consequently has only one basis function per atom, for the first row atoms one has two s type functions describing the $1s$ and $2s$ electrons and one set of p_x, p_y, p_z functions describing the $2p$ electrons.

- Double Zeta basis sets:
  In a double Zeta basis set the number of exponents is doubled with respect to the minimal basis set. For first row atoms one has thus a $1s$, $1s'$, $2s$, $2s'$, $2px$, $2px'$, $2py$, $2py'$, $2pz$ and $2pz'$ functions

- Triple Zeta basis sets:
  The number of exponents is three times that of the minimal basis set.

- Polarization basis functions:
  Gaussians that have angular momenta higher than the occupied atomic shells (i.e. d type functions for first row atoms) are called polarization functions. In the context of density functional calculations polarization functions are important if there are complicated bonding geometries. Let's for instance consider first the CH_4 molecule. The bonds that the carbon atoms form with the 4 hydrogens are formed from the $sp3$ hybridized atomic orbitals. These 4 orbitals are just simple linear combinations

of s and p type functions

$$\frac{1}{2}(s+p_x+p_y+p_z)$$
$$\frac{1}{2}(s+p_x-p_y-p_z)$$
$$\frac{1}{2}(s-p_x+p_y-p_z)$$
$$\frac{1}{2}(s-p_x-p_y+p_z)$$

Let us now imagine that one hydrogen is replaced by a fluorine which has a higher electronegativity than hydrogen. As a consequence the entire electronic charge of the molecule will be pulled towards the fluorine and all the bond angles will be distorted. To describe such distorted bond angles higher angular momenta are necessary.

A large number of polarization functions is required for correlated wavefunction methods (CI, CC etc) in order to describe the angular correlation on each atom.

- Diffuse basis functions:
  Diffuse functions are Gaussians that have a small exponent and that are therefore fairly delocalized. Diffuse functions are necessary if one has a charged system with

an extra electron or if one wants to calculate properties that depend on the tail of the wavefunction such as the polarizability.

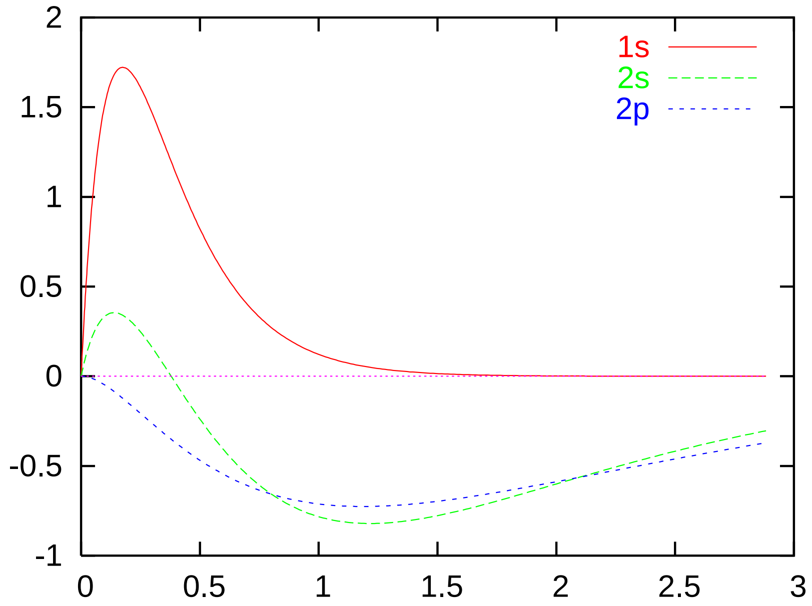
- Contracted Gaussian type orbitals:
  Gaussians are in principle quite terrible functions for describing atomic wavefunctions. They have the wrong behaviour both at the origin and far away from the origin. The true atomic wavefunction is similar to an exponential times a spherical harmonic. At the origin it therefore has a cusp whereas the Gaussian does not. Far away the Gaussian is decaying much faster than the exponential. The reason why Gaussian are nevertheless widely used is that all the integrals can easily and efficiently be evaluated. Using a linear combination of a few Gaussians, it is however possible to obtain a wavefunction that is rather similar to an atomic wavefunction and for which the integrals are still simple to evaluate. If such a linear combination is used as a basis function, it is called a contracted Gaussian. The contraction is always over radial Gaussians with several exponents that all share the same spherical harmonics.

  $$G(\mathbf{r}) = Y_{l,m}(\theta,\phi)\sum_k c_k \exp\left(-\alpha_k(x^2+y^2+z^2)\right)$$

  Depending on how many contracted Gaussian G are used to describe each shell one calls such contracted basis sets again single , double or triple zeta basis sets.

- Split valence basis sets:

Using a double or triple zeta contracted basis set is wasteful. The core electrons are chemically inert and do not change if one puts an atom into different chemical environments. Therefore one single contracted Gaussian is enough to describe it if the contraction contains a sufficient number of terms. The part of the valence function that is in the core region is also rigid since it has to be orthogonal to the core wave-function. It can therefore also be described by a single contracted Gaussian. What varies significantly in different chemical environments is the part of the valence function in the valence region. Split valence basis sets take this fact into account and give more flexibility to the valence wavefunction in the valence region than in the core region.



The Radial wavefunctions of carbon multiplied by r

Split valence basis sets are among the best Gaussian basis sets and the most widely used are denoted by

  - 3-21G: The core orbital is a contraction of 3 primitive Gaussians, the core part of the valence is a contraction of 2 primitive Gaussians and the valence part is a single primitive Gaussian
  - 6-31G: The core orbital is a contraction of 6 primitive Gaussians, the core part of the valence is a contraction of 3 primitive Gaussians and the valence part is a single primitive Gaussian
  - 6-311G: The core orbital is a contraction of 6 primitive Gaussians, the core part of the valence is a contraction of 3 primitive Gaussians and the valence part consists of two primitive Gaussian

Diffuse and polarization Gaussians can be added to these split valence basis sets.

- Correlation consistent basis sets:
  In correlated wavefunction methods the basis set has to describe both angular and radial correlation effects. It does not make sense to describe for instance radial correlation with high accuracy by including Gaussians of high angular momentum l while neglecting radial correlation by including only a few different radial Gaussians for the various angular momenta taken into account. Correlation consistent basis sets are balanced in their description of radial and angular correlation.

Numerical atomic orbitals (NAO's)

As in the case of Gaussian basis set, each atom carries a certain number of basis function, but in the case of NAO they are the numerical solution of a free atom. To form a basis set both the orbitals that are occupied and unoccupied in the free atom have to be used. Since in chemical bonding the atomic wavefunction are only relatively weakly modified, such a basis set allows us to obtain accurate results with basis sets of moderate size. Another advantage over the Gaussian basis sets is that the NAO's decribe correctly both the electron-nucleon cusp of the wavefunction as well as the exponential assymptotic decay for large radii. While the orbitals on a single atom are by construction orthogonal, orbitals centered on different atoms are not orthogonal to each other. Hence numerical instabilities due to the overcompleteness will also occur as one goes to very large basis set sizes. Like for Gaussians, basis set superposition errors also exist. However practice shows that both of these problems are less severe for NAO's than for Gaussians. The only minor disadvantage compared to Gaussians is that some integrals such as the kinetic energy can not be calculated analytically.

Problems of atomic basis sets

- With a systematic basis set one can approach the true (density functional) energy and wavefunction with arbitrarily small error if the basis set is sufficiently big. Atomic basis sets such as Gaussians are not systematic basis sets.

- With atomic basis sets one encounters so-called basis set superposition errors. They are due to the fact that the basis set for one specific atom becomes better if there are other atoms close by whose basis function cover also the region of the atom under consideration and act thus as additional basis functions. As a consequence binding energies are generally too large in such calculation.

- The Hellmann-Feynman theorem does not hold anymore for incomplete atom centered basis functions. In addition to the Hellmann-Feynman force so-called Pulay forces arise. Even though they are difficult to calculate they are implemented in most Gaussian type electronic structure codes.

Plane wave basis sets

Plane waves are the natural basis set for electronic structure calculations in periodic crystals. Bloch's theorem tells us that the wavefunction of the j-th band at the Brioullin wavevector $\mathbf{k}$ can be written as

$$\psi_{\mathbf{k},j}(\mathbf{r}) = \exp(I\mathbf{k}\cdot\mathbf{r})\left(\sum_{\mathbf{K}} C_{\mathbf{k},j}(\mathbf{K})\exp(I\mathbf{K}\cdot\mathbf{r})\right)$$

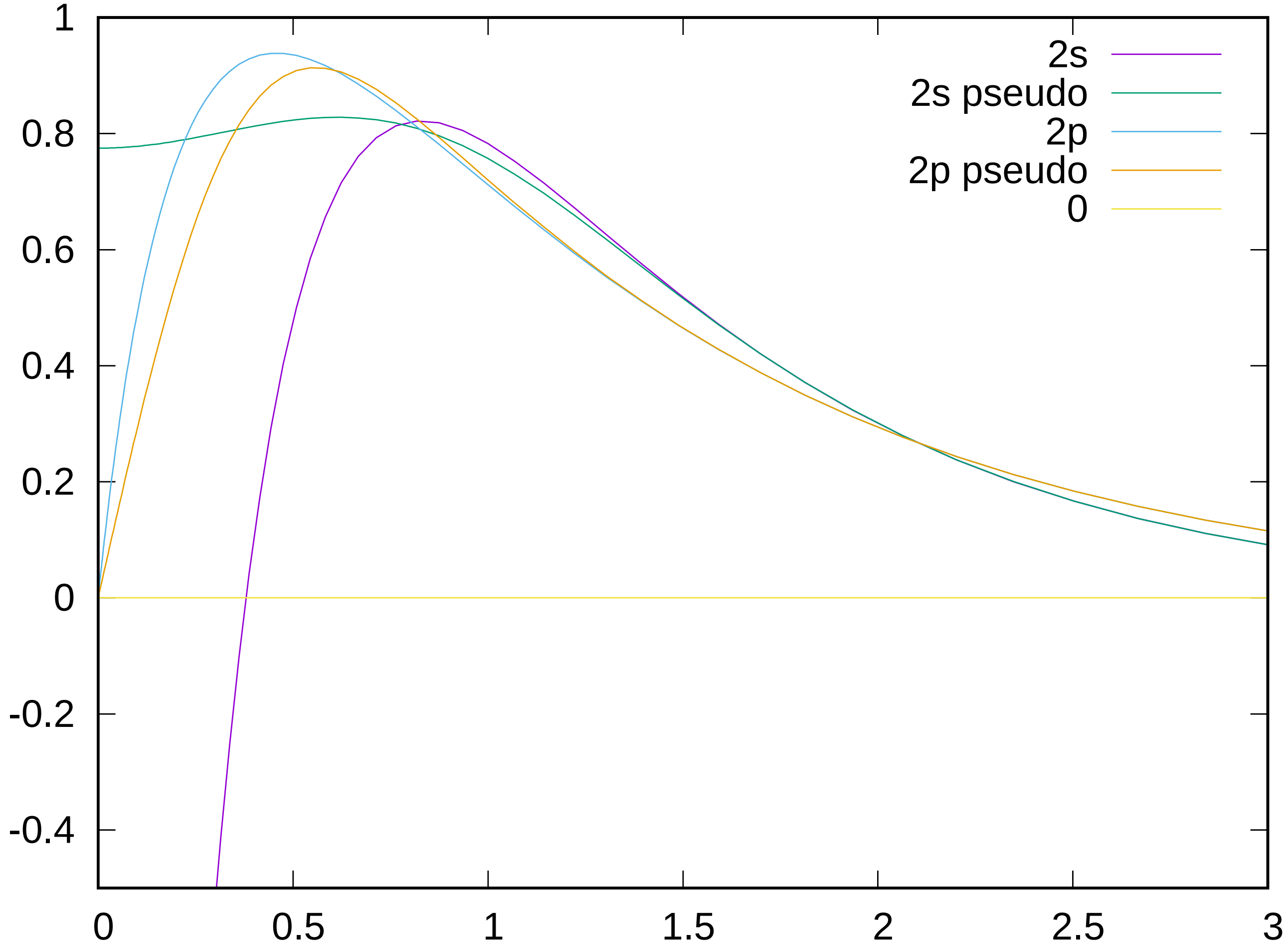
The expansion coefficients $C_{\mathbf{K},j}$ of the part that is periodic with respect to the Bravais lattice cell (part within the large brackets) satisfy the Schrödinger equation

$$\frac{1}{2}(\mathbf{k}+\mathbf{K})^2 C_{\mathbf{k},j}(\mathbf{K}) + \sum_{\mathbf{K}'} V_{\mathbf{K}-\mathbf{K}'} C_{\mathbf{k},j}(\mathbf{K}) = \varepsilon_j(\mathbf{k}) C_{\mathbf{k},j}(\mathbf{K}) \qquad (120)$$

where $V_{\mathbf{K}}$ is a Fourier component of the Kohn-Sham potential. In the above equation there is no overlap matrix since plane waves form an orthogonal basis set.
Since a plane wave extends over the whole computational volume, plane wave basis sets do not allow for an adaptive resolution. This would be necessary in an all electron calculation, i.e. in a calculation where one includes both the chemically relevant valence electrons and the background core electrons. The core regions would require an huge overall resolution that would render plane wave electronic structure calculations hopelessly slow.

For this reason pseudopotentials are used in all plane wave calculations. A pseudopotential describes some pseudoatom whose core electrons are eliminated but whose valence electrons have the same chemical properties as the valence atoms of the true atom. Even with pseudopotentials the number of plane waves is still significant (of the order of a few hundred per atom), but since plane waves can be handled numerically very efficiently, they give rise to quite fast electronic structure codes. Obviously the kinetic energy part is diagonal in plane waves and therefore trivial to handle. The potential energy part is a convolution that can easily be handled by Fast Fourier Transformation (FFT) techniques. Even though the use of pseudopotentials is mainly a numerical necessity, it also has a physical advantage. Relativistic effects can easily be incorporated into a non-relativistic electronic structure calculation through relativistic pseudopotentials. This is due to the fact that only the core electrons feel strong relativistic effects.

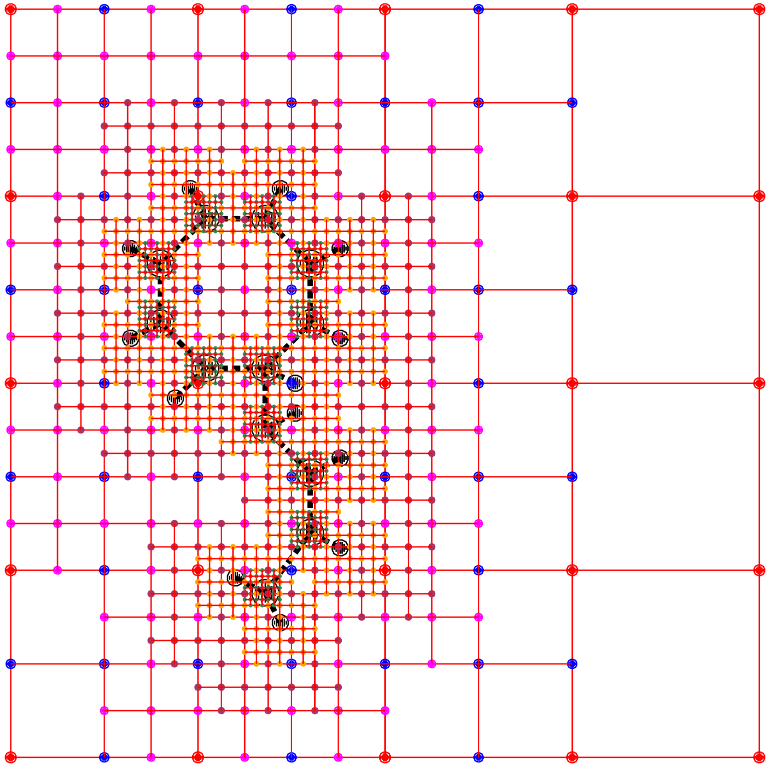
The figure below shows the all-electron and pseudo valence orbitals of carbon. Outside the covalent radius the all-electron and pseudo valence orbitals have to coincide in order to describe correctly the rearrangements of the the wavefunctions during chemical bonding. In the chemically inert core region they may however differ. In contrast to the all-electron 2s function the pseudo 2s function has no node and is therefore ’softer’, i.e requires a smaller number of plane waves to represent it.



Wavelet basis sets

Wavelets are a new type of mathematical functions that were introduced some 20 years ago. Because of their unique properties, they have found widespread use in many areas ranging from image compression to predictions of the solar cycle. Wavelets are also a promising basis set for Schrödingers equation. Wavelets

- are localized both in real and in Fourier space
- allow for adaptivity, i.e. close to the nuclei the resolution is higher



- are a systematic basis set

The zoo of wavelet families

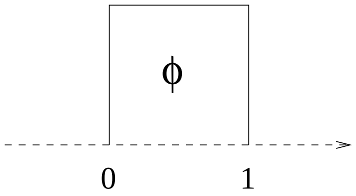
The notion of wavelet does not refer to a specific function but to a large class of various functions. One can classify wavelets into orthogonal, bi-orthogonal and multi-wavelet types. Within each class there exist various families. Each family is characterized by 2 fundamental functions, the scaling function and the wavelet. Within a family there are sub-families of different degrees. Any orthogonal (sub-)family is characterized by 2 fundamental functions, the scaling function ϕ and the wavelet ψ. We will first discuss the simplest wavelet family, the Haar wavelet, and then go on to more complicated and more powerful wavelet families such as Daubechies wavelets. A wavelet basis set is generated by translations and dilatations of the 2 fundamental functions.

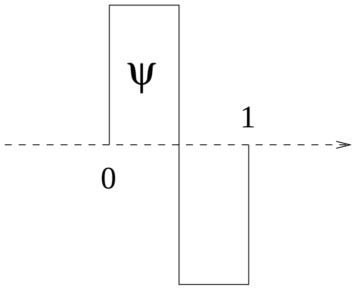
$$\begin{aligned} \phi_j^k(x) &\propto \phi(2^k x - j) \\ \psi_j^k(x) &\propto \psi(2^k x - j) \end{aligned}$$

The index j indicates the localization in time (space), whereas the index k the localization in frequency (Fourier space). Infinite resolution is obtained in the limit $k \to \infty$

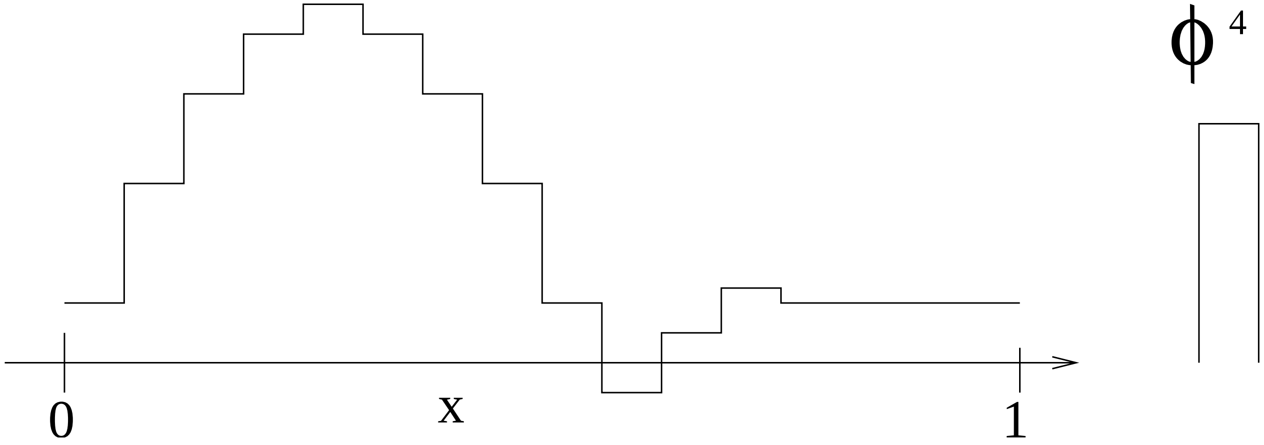
The Haar wavelet family

The Haar wavelet ψ and scaling function ϕ on resolution level 0 are shown below



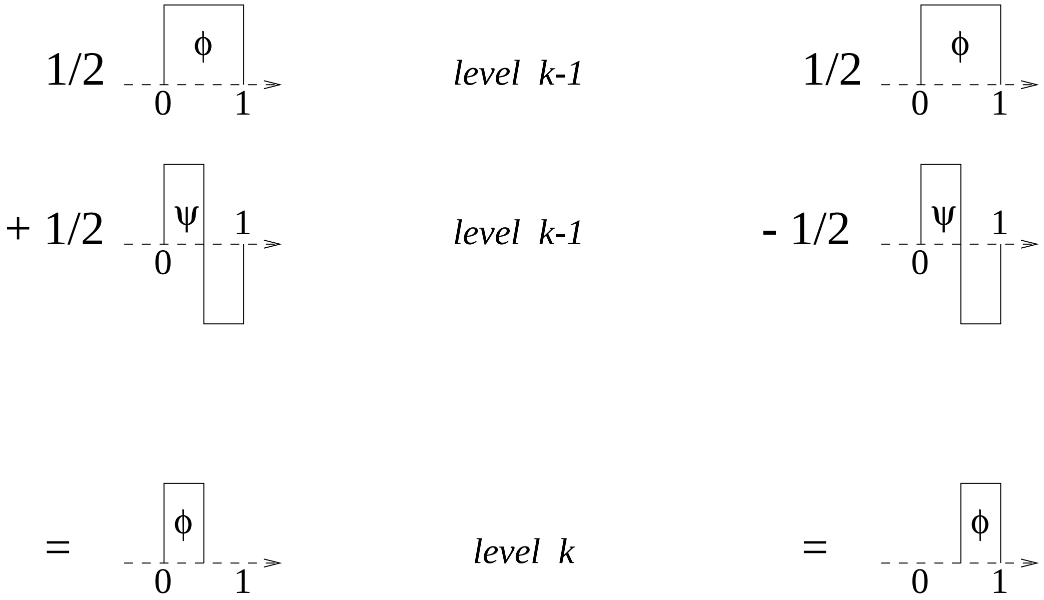


Let us now assume that we have a piecewise constant function in the interval between 0 and 1 which consists of 16 equidistant plateaus. Hence it can be expressed exactly by a linear combination of scaling functions at resolution level 4. The expansion coefficients s_i^4 are just the values of the plateaus.



$$f = \sum_{i=0}^{16} s_i^4 \, \phi_i^4(x) \qquad ; \qquad s_i^4 = f(i/16) \tag{121}$$

The figure below shows the fundamental transformation between scaling functions and wavelets on different resolution levels. A scaling function on resolution level k is a linear combination of a scaling function and a wavelet on a lower resolution level $k-1$,



$$\begin{aligned} \phi^k_{2i}(x) &= \frac{1}{2}\phi^{k-1}_i(x) + \frac{1}{2}\psi^{k-1}_i(x) && (122) \\ \phi^k_{2i+1}(x) &= \frac{1}{2}\phi^{k-1}_i(x) - \frac{1}{2}\psi^{k-1}_i(x) && (123) \end{aligned}$$

The two above relations 122,123 can be used to determine the matrix elements of the matrix T which transforms the scaling function and wavelet expansion coefficients from level $k-1$, s_i^{k-1} and d_i^{k-1}, to scaling function coefficients s_{2i}^k and s_{2i+1}^k at level k:

$$\left(\begin{array}{c} s_{2i}^k \\ s_{2i+1}^k \end{array} \right) = \left(\begin{array}{cc} T_{1,1} & T_{1,2} \\ T_{2,1} & T_{2,2} \end{array} \right) \left(\begin{array}{c} s_i^{k-1} \\ d_i^{k-1} \end{array} \right)$$

since they give rise to the two conditions

$$\left(\begin{array}{c} 1 \\ 0 \end{array} \right) = \left(\begin{array}{cc} T_{1,1} & T_{1,2} \\ T_{2,1} & T_{2,2} \end{array} \right) \left(\begin{array}{c} 1/2 \\ 1/2 \end{array} \right) \quad ; \quad \left(\begin{array}{c} 0 \\ 1 \end{array} \right) = \left(\begin{array}{cc} T_{1,1} & T_{1,2} \\ T_{2,1} & T_{2,2} \end{array} \right) \left(\begin{array}{c} 1/2 \\ -1/2 \end{array} \right)$$

The result is

$$T = \left(\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array} \right)$$

which gives rise to the following backward transformation rule

$$s_{2i}^{k+1} = s_i^k + d_i^k \qquad ; \qquad s_{2i+1}^{k+1} = s_i^k - d_i^k \tag{124}$$

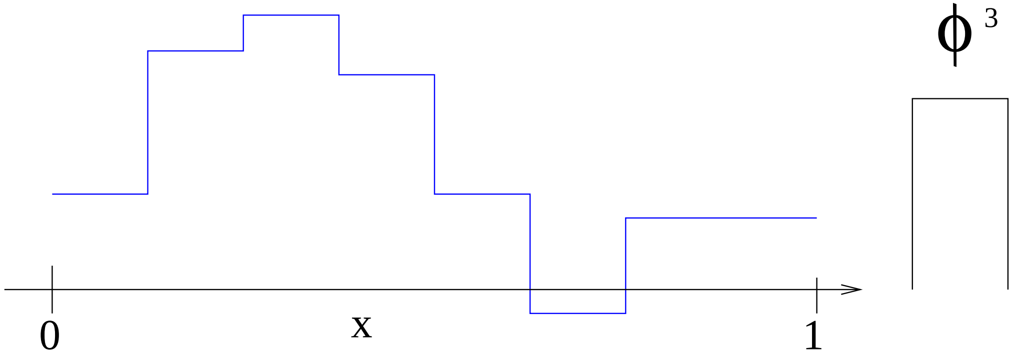
The transformation rule can be inverted to obtain the forward transformation rule

$$s_i^{k-1} = \frac{1}{2} s_{2i}^k + \frac{1}{2} s_{2i+1}^k \qquad ; \qquad d_i^{k-1} = \frac{1}{2} s_{2i}^k - \frac{1}{2} s_{2i+1}^k \tag{125}$$

After one forward transformation step (Eq. 125), the function from Eq. 121, is represented in an equivalent way by scaling functions and wavelets on resolution level 3.

$$f = \sum_{i=1}^{8} s_i^3 \phi_i^2(x) + \sum_{i=1}^{8} d_i^3 \psi_i^3(x) \tag{126}$$

The (blue) scaling function only part is a smoothed version of the original function and is shown below. The (red) wavelet part represents the rapidly varying difference between the orginal and smoothed function.

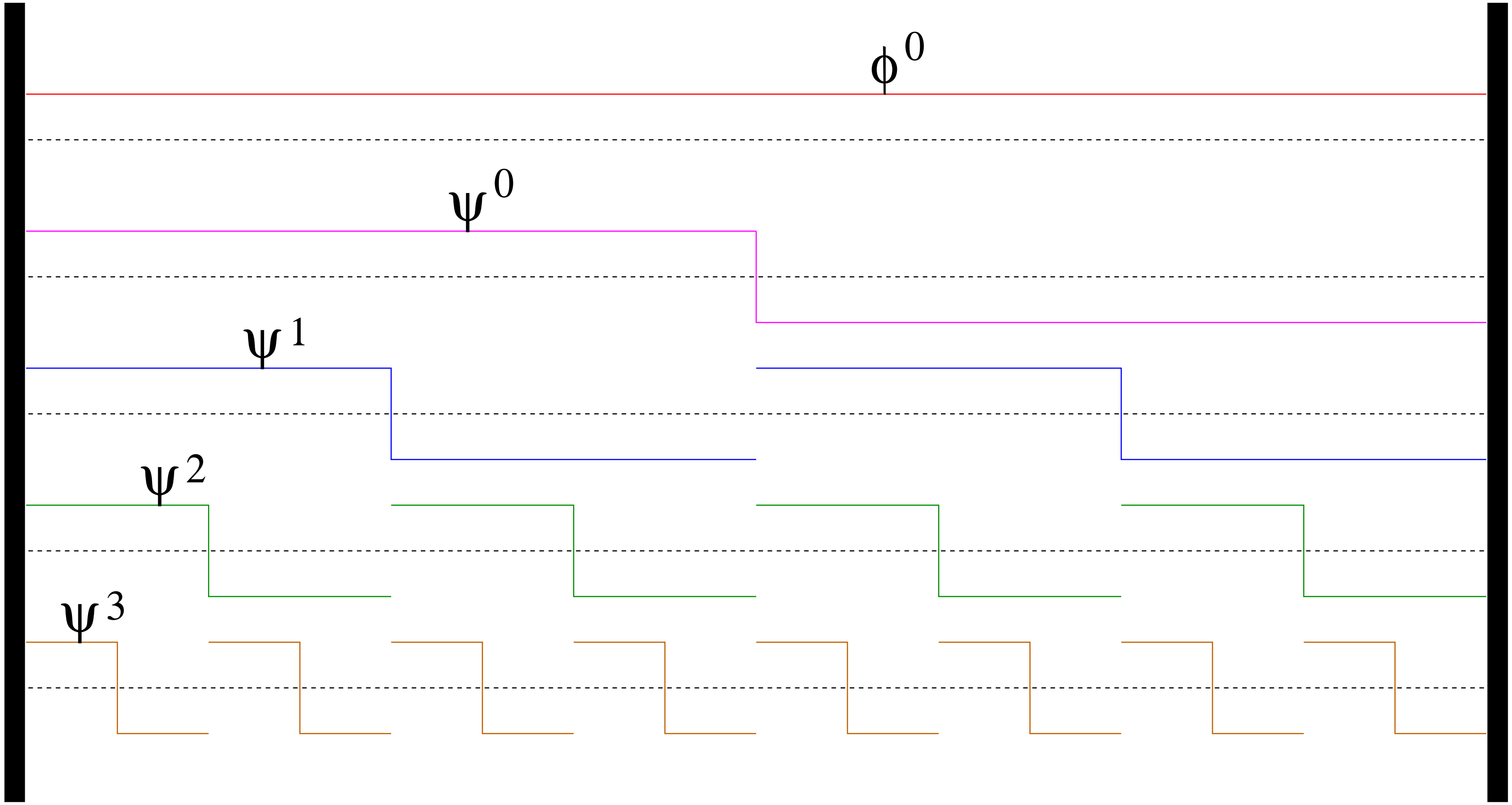


The scaling functions on resolution level 3 of Eq. 126 can now be decomposed recursively into scaling functions and wavelets on even higher resolution levels until only one scaling function is left. This single scaling function then represents only the average value of the function and is thus similar to the zeroth Fourier component in a plane wave representation.

The resulting representation is called the wavelet representation of the function.

$$f = s_1^0 \phi_1^0(x) + d_1^0 \psi_1^0(x) + \sum_{i=1}^{2} d_i^1 \psi_i^1(x) + \sum_{i=1}^{4} d_i^2 \psi_i^2(x) + \sum_{i=1}^{8} d_i^3 \psi_i^3(x) \qquad (127)$$

The basis functions from the above equation are sketched in the same colors below.

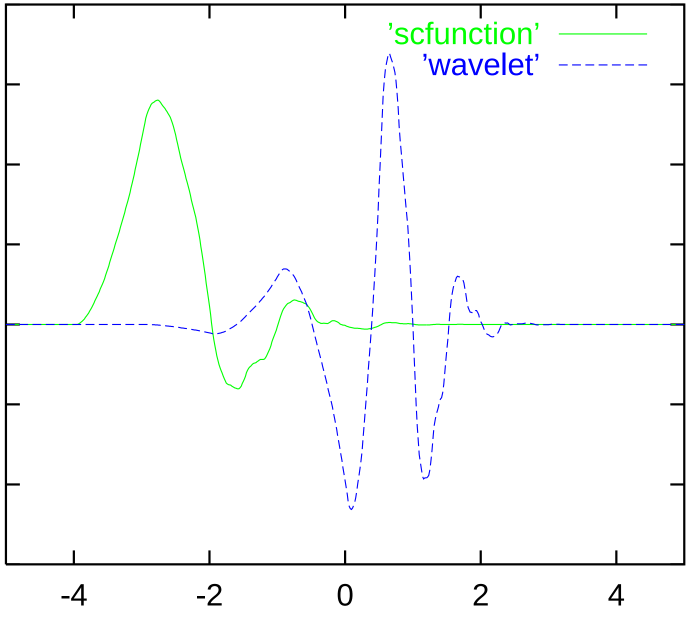


To go back from the wavelet representation (Eq. 127) to the original scaling function representation (Eq. 121) the backward transformation (Eq. 124) can be applied.

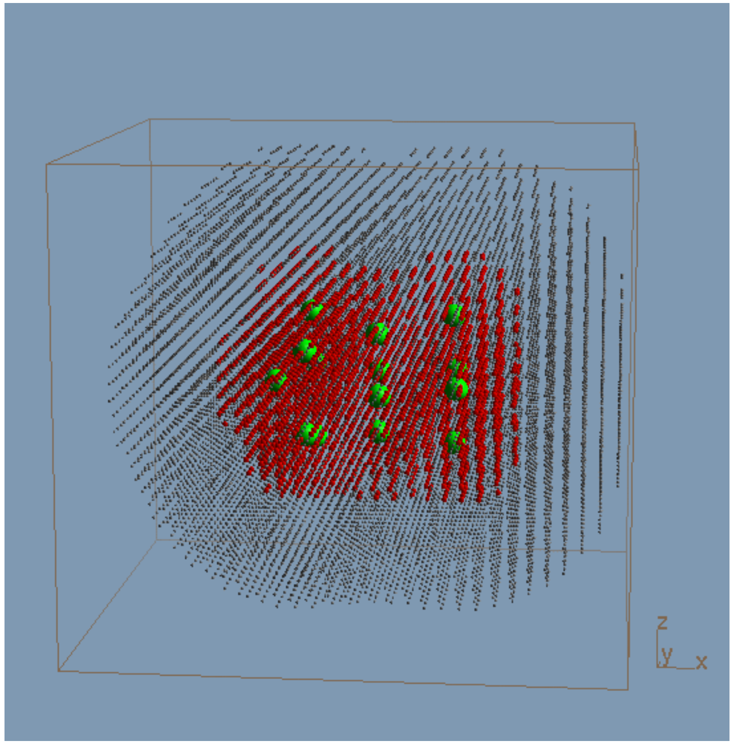
As described above the scaling function representation (Eq. 121) and the wavelet representation (Eq. 127) have the same number of expansion coefficients, namely 16 in our example. On a computer the storage requirements would thus be identical. However if the function varies more rapidly in one region than in another one the wavelet respresentation is advantageous. The function shown on top of equation 121 is for instance constant in the last 3 intervals. Hence the last 4 expansion coeffients are constant, $s^4_{16} = s^4_{15} = s^4_{14} = s^4_{13}$, but this does however not eliminate the need to store them. In a wavelet representation,on the other hand, the coefficients d^3_8 and d^3_8 are zero and consequently it is not necessary to store them. 14 coefficients are thus only required instead of 16 and the wavelet representation is more compact. This lead to the concept of adaptivity in a wavelet representation. High resolution wavelet coefficients are only needed in regions where the function varies rapidly. In slowly varying regions one knows from the beginning that these coefficient are zero (or at least very small) and hence the need not be stored.

Daubechies wavelets

Daubechies wavelets are orthogonal wavelets. They are higher order generalizations of the discontinuous Haar wavelet. It was for a long time deemed impossible to construct a set of continuous functions that have compact support and are orthogonal. As such the construction of the Daubechies wavelets was a major breakthrough in mathematics. The mathematical theory of Daubechies wavelets is fairly complicated and will not be explained in this course. Below is shown the 8-th order Daubechies family. Even though it looks kind of ugly, it can be shown that a linear combination of Daubechies scaling functions can exactly represent polynomials up to a certain order.



Daubechies wavelets are used as a basis set in the BigDFT electronic structure code. Since pseudopotentials are used, the very rapidly varying core electron wavefunctions are eliminated. Hence there is no need for a large degree of adaptivity and only 2 resolution levels are used. The high resolution region is chosen to coincide with the region where bonding takes place. The low resolution region describes the region where the tails of the wavefunctions decay slowly towards zero. The situation is shown below for a cluster of silicon atoms shown as green spheres. The red grid points carry both scaling functions and wavelets and cover therefore the high resolution region, the brown grid points carry only scaling functions and cover therefore the low resolution region.



Tight binding methods

In order to illustrate the concept of tight binding scheme let's focus onto a specific material for which many tight binding schemes exist, namely silicon. In tight binding schemes the cores electrons are eliminated as they are in a pseudopotential scheme. Hence we are left with 4 valence electrons that are distributed among the $3s$ and $3p_x$, $3p_y$ and $3p_z$ orbitals. In a pseudopotential calculation with a minimal basis set we would thus need 4 orbitals per atom. If the electronic structure calculation is done in the selfconsistent diagonalization scheme we have to diagonalize in each step of the selfconsistency iteration a Hamiltonian matrix of dimension $4N_{at}$ times $4N_{at}$, where N_{at} is the number of silicon atoms. Assuming that our silicon structure is a closed shell system, we have to calculate the lowest $2N_{at}$ eigenvalues and eigenvectors of this matrix. Once selfconsistency has been reached the total energy of the system can be obtained from the eigenvalues ε_i of the Kohn-Sham matrix and the charge density by the following formula which corresponds to Eq.75 of the HF method

$$\begin{aligned} E = 2\sum_{i=1}^{N/2} \varepsilon_i \quad &+ \quad \int E_{xc}^{LDA}(\rho(\mathbf{r})) - v_{xc}^{LDA}(\rho(\mathbf{r}))\rho(\mathbf{r})d\mathbf{r} \\ &- \quad \frac{1}{2}\int\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + \sum_{i<j\leq N_{at}} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} \end{aligned} \tag{128}$$

Empirical Tight binding schemes now give some simple rules that allow to approximately predict the matrix elements of the Hamiltonian without calculating them explicitly from some set of orbitals and the selfconsistent potential. From the eigenvalues ε_i^{TB} of this tight binding Hamiltonian H^{TB} we can then obtain the first term of the DFT total energy expression of Eq. 128. The additional terms of Eq. 128 are in most tight binding schemes replaced by a short-range pairwise potential P. This is justified since the exchange correlation terms are short range and since the long range Hartree terms in the second line of Eq. 128 cancel if the total charge density is a superposition of well localized atomic charge denities. Hence the tight binding energy expression is

$$E = 2\sum_{i=1}^{N/2} \varepsilon_i^{TB} + \sum_{i<j\leq N_{at}} P(|\mathbf{R}_i - \mathbf{R}_j|) \tag{129}$$

In empirical tight binding schemes one never sets up an explicit basis set. It is therefore not possible to calculate a wavefunction or charge density in real space. As a consequence the tight binding energy expression of Eq. 129 depends only the eigenvalues and not on the eigenvectors of the Hamiltonian.

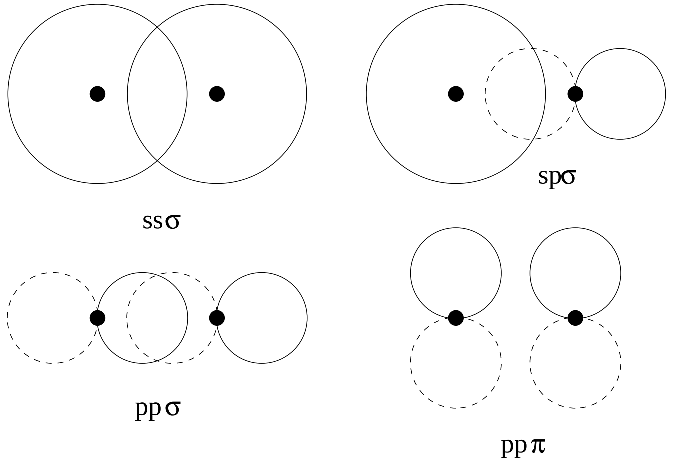
Let us now discuss in more detail the rules for setting up H^{TB}. In this discussion we will assume that there exists some basis set that consists of localized orbitals that are products of spherical harmonics and radial functions. We however never have to specify the explicit form of the radial functions. In addition we will assume that this basis set is orthogonal. Denoting a basis function centered on atom i with angular momenta l,m by $B_{i,l,m}$ we would have to solve the integral

$$H^{TB}_{i,l,m\,;\,i',l',m'} = \int B_{i,l,m}(\mathbf{r}) \left(\frac{-1}{2} \nabla^2 + V(\mathbf{r}) \right) B_{i',l',m'}(\mathbf{r}) d\mathbf{r} \tag{130}$$

To effectuate the integration one goes into a cylindrical coordinate system where the z axis connects the two atoms. For the kinetic energy part the integration over the known angular part (given by real spherical harmonics) can be done analytically. For the potential energy part this integration over the angular part can only be done analytically if the so-called two-center approximation is adopted. In this approximation one assumes that the total potential $V(\mathbf{r})$ can be decomposed into a sum of radially symmetric atom centered atomic potentials

$$V(\mathbf{r}) = \sum_{j=1}^{N_{at}} v_j(|\mathbf{r} - \mathbf{R}_j|)$$

In the calculation of $H^{TB}_{i,l,m;i',l',m'}$ one then neglects all the terms j that are not equal to either i or i'. The final result for angular integration part of Eq. 130 is given by the so-called Slater-Koster rules. For a system where the atom centered basis functions have only s and p character the Slater-Koster rules state that all the integrals can be expressed by 4 quantities $V_{s,s,\sigma}$, $V_{s,p,\sigma}$, $V_{p,p,\sigma}$ and $V_{p,p,\pi}$ illustrated below. In addition they give simple formulas for expressing arbitrary integrals, where the p orbitals are not necessarily parallel or perpendicular to the axis connecting the atomic centers, in terms of these 4 quantities.



Basic types of integrals in the Slater Koster formalism for s and p orbitals. The integrals are classified according to the two participating orbitals and according to the angular momentum with respect to the axis. σ *indicates* $m = 0$ *and* π *indicates* $m = 1$. *The positive lobe of the p orbitals is shown by a solid line and the negative lobe by a dashed line.*

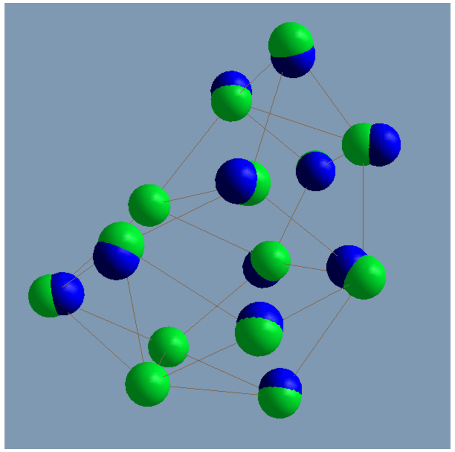
What is left unspecified by the Slater-Koster rules is the dependence of these 4 quantities on the distance between atom i and i'. In empirical tight binding schemes this dependence is found by fitting to known theoretical or experimental data. The form of the function P in Eq. 129 is found by fitting as well. The quantities $V_{s,s,\sigma}$, $V_{s,p,\sigma}$, $V_{p,p,\sigma}$ and $V_{p,p,\pi}$ generally tend rapidly to zero with increasing distance between two atoms. The resulting Hamiltonian matrix is therefore sparse.

Tight binding is the fastest and crudest quantum mechanical method. It is faster than a DFT calculation in a minimal basis set because

- The tight binding matrix H^{TB} can be calculated much faster than the Kohn-Sham matrix. This is due to the fact that the matrix is sparse and to the fact that calculating a matrix element by an integration over 3-dim space is much more expensive than deriving it from the Slater-Koster rules together with the fitted radial integral dependence.

- It is not necessary to do a selfconsistency iteration and the result is thus obtained after a single diagonalization.

The speed has of course its price. The accuracy of empirical tight binding schemes is less than the accuracy of DFT schemes. In addition tight binding schemes are only available for a few technically important materials. For these materials tight binding schemes are frequently the only way to do quantum mechanical calculations on large systems. The figure below shows the equilibrium geometry of a silicon cluster obtained with a tight binding scheme (blue spheres) and superimposed the DFT result (green spheres). The structures are qualitatively similar even though they are quantitively different.



In addition to the orthogonal empirical tight binding scheme outlined up to now there are still the following other tight binding schemes.

- Non-orthogonal tight binding schemes: The tight binding eigenvalues are obtained from a generalized eigenvalue problem. The overlap matrix is set up with rules similar to the rules for the Hamiltonian matrix in a orthogonal tight binding scheme. The fact that there exist two matrices allows for more flexibility in the fitting procedure.

- Selfconsistent tight binding schemes: These schemes give some simple rules for the charge transfer in a system and account for the resulting electrostatic repulsion or attraction in an approximative way. They are however numerically more expensive since a selfconsistency iteration has to be done.

Variational Quantum Monte Carlo

The variational Quantum Monte Carlo method solves in principle the many body Schrödinger equation exactly. The method is conceptually simple. One writes down a functional form for the wave-function that contains several variational parameters. Based on the variational principle one then finds the set of parameters that minimizes the energy. The difficult numerical problem is how to evaluate the energy expectation value. Since it is given by a high-dimensional integral it can only be calculated by Monte Carlo integration with importance sampling. We will illustrate the variational Quantum Monte Carlo method for the calculation of the singlet state of the H_2 molecule. A two electron singlet wave-function has the following form.

$$\Psi(\mathbf{x}_1,\mathbf{x}_2) = \Phi(\mathbf{r}_1,\mathbf{r}_2)\left(\alpha(s_1)\beta(s_2) - \alpha(s_2)\beta(s_1)\right)$$

Since the spin part is antisymmetric, the spatial part $\Phi(\mathbf{r}_1,\mathbf{r}_2)$ has to be symmetric, i.e. $\Phi(\mathbf{r}_1,\mathbf{r}_2) = \Phi(\mathbf{r}_2,\mathbf{r}_1)$. Φ satisfies the 2-particle Schrödinger equation

$$\mathcal{H}^{2el}\Psi(\mathbf{r}_1,\mathbf{r}_2) = E\Psi(\mathbf{r}_1,\mathbf{r}_2) \tag{131}$$

For 2 nuclei at $\mathbf{R}_A$ and $\mathbf{R}_B$ the Hamiltonian is

$$\mathcal{H}^{2el} = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} - \frac{1}{|\mathbf{r}_1 - \mathbf{R}_A|} - \frac{1}{|\mathbf{r}_1 - \mathbf{R}_B|} - \frac{1}{|\mathbf{r}_2 - \mathbf{R}_A|} - \frac{1}{|\mathbf{r}_2 - \mathbf{R}_B|} \tag{132}$$

An usual ansatz for a variational Quantum Monte Carlo wave-function is a product of a part that contains only single particle orbitals and a part that contains a so-called Jastrow factor. With a good ansatz one can get energies E that are very close to the exact eigenvalues of Eq. 131. The Jastrow factor depends on inter-electronic coordinates and represents in this way correlation effects among the electrons. In particular the Jastrow factor has to be chosen such that it satisfies the electron-electron cusp condition. When two electrons come very close the potential energy goes to infinity. Since the eigenvalue is constant and finite throughout space, the kinetic energy has to go to minus infinity. This leads to a discontinuity of the first derivative when $\mathbf{r}_1 = \mathbf{r}_2$. If an ansatz for a wave-function is chosen that does not satisfy the cusp condition the exploding of the potential energy leads to numerical difficulties. To derive this cusp condition let us introduce a new set of coordinates

$$\begin{aligned} \mathbf{R} &= (\mathbf{r}_1 + \mathbf{r}_2)/2 \\ \mathbf{d} &= \mathbf{r}_1 - \mathbf{r}_2 \end{aligned}$$

The kinetic energy of Eq. 131 then becomes

$$-\frac{1}{4}\nabla_R^2 - \nabla_d^2$$

Let us now consider the case where both electrons are very close to each other but not very close to the nucleus. In this regime the wave-function will essentially only depend on **d** and the Schrödinger equation (Eq. 131) locally takes on the form

$$\left(-\nabla_d^2 + \frac{1}{|\mathbf{d}|}\right)\Psi = E\Psi$$

This is the well known hydrogenic one particle Schrödinger equation for a repulsive Coulomb potential with $Z = 1/2$. It has therefore locally the solution

$$\exp(Z|\mathbf{d}|) = \exp(|\mathbf{d}|/2)$$

Such a wave-function satisfies the cusp condition

$$\frac{1}{\Psi}\frac{\partial\Psi}{\partial d}\bigg|_{d=0} = \frac{1}{2} \tag{133}$$

The above Eq. 133 is called the electron-electron cusp condition for anti-parallel electrons. It is satisfied by any exact wave-function and should be satisfied by any approximate Monte Carlo wave-function for reasons of numerical stability.
The energy expectation value $< E >$ (that is a function of the variational parameters in the wave-function Φ)

$$< E >= \frac{\int \Phi(\mathbf{r}_1,\mathbf{r}_2)\mathcal{H}^{2el}\Phi(\mathbf{r}_1,\mathbf{r}_2)d\mathbf{x}_1 d\mathbf{x}_2}{\int \Phi(\mathbf{x}_1,\mathbf{x}_2)\Phi(\mathbf{x}_1,\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2} \tag{134}$$

is calculated using Monte Carlo integration with importance sampling. Importance sampling is here essential since the wave-function is vanishingly small in most parts of space. In order to do the importance sampling we define a local energy E_L

$$E_L(\mathbf{r}_1, \mathbf{r}_2) = \frac{\mathcal{H}^{2el}\Phi(\mathbf{r}_1, \mathbf{r}_2)}{\Phi(\mathbf{r}_1, \mathbf{r}_2)} \tag{135}$$

This local energy would be a constant if $\Phi(\mathbf{r}_1, \mathbf{r}_2)$ was the exact wave-function. Since this is in general not the case $E_L(\mathbf{r}_1, \mathbf{r}_2)$ varies in space, but its variation is smaller the better $\Phi(\mathbf{r}_1, \mathbf{r}_2)$ is. In terms of the local energy the energy expectation value of Eq. 134 becomes

$$< E >= \frac{\int E_L(\mathbf{r}_1, \mathbf{r}_2)|\Phi(\mathbf{r}_1, \mathbf{r}_2)|^2 d\mathbf{r}_1 d\mathbf{r}_2}{\int |\Phi(\mathbf{r}_1, \mathbf{r}_2)|^2 d\mathbf{r}_1 d\mathbf{r}_2}$$

Using the Metropolis algorithm we can generate a sequence of configurations $\mathbf{r}_1, \mathbf{r}_2$ distributed according to

$$\frac{|\Phi(\mathbf{r}_1, \mathbf{r}_2)|^2}{\int |\Phi(\mathbf{r}_1, \mathbf{r}_2)|^2 d\mathbf{r}_1 d\mathbf{r}_2} \tag{136}$$

The energy expectation value then becomes a simple sum

$$< E >= \frac{1}{T} \sum_{t=1}^{T} E_L(\mathbf{r}_1(t), \mathbf{r}_2(t))$$

Even though Variational Quantum Monte Carlo is exact in principle, it is not possible in practice to come very close to the exact wavefunction. The reason for this is that the minimization problem becomes too difficult for a large number m of variational parameters because the wavefunction depends on these parameters in a complicated non-quadratic form. For this reason Variational Quantum Monte Carlo is typically the first part of a Quantum Monte Carlo calculation. The second part is a diffusion Monte Carlo simulation that is based on the results of the Variational Quantum Monte Carlo simulation.

Diffusion Quantum Monte Carlo

Diffusion Quantum Monte Carlo is based on the relation between Schrödingers equation and the diffusion equation. For the 1-dim case the diffusion equation for a quantity ρ (which is not an electronic charge density) is given by

$$\frac{\partial\rho}{\partial t} = \frac{1}{2}\frac{\partial^2\rho(x,t)}{\partial x^2} + (E_T - V(x))\rho(x) \tag{137}$$

The first term on the right hand side is the ordinary diffusion term arising from Fick's law, the second is a growth/decay term that does, in contrast to the first one, not conserve the norm of the quantity ρ. When a stationary state is reached, i.e. when $\frac{\partial\rho}{\partial t} = 0$, ρ satisfies the (in this case 1-dim) Schrödinger equation.

$$-\frac{1}{2}\frac{\partial^2\rho(x,t)}{\partial x^2} + V(x)\rho(x) = E_T\rho(x)$$

Thus one can obtain the solution of Schrödinger's equation by simulating a diffusion process until it has reached equilibrium. Note that the existence of a stationary state with a finite norm requires a particular value of E_T. If E_T is too large the quantity ρ will grow exponentially, if it is too small, it will decay to zero.

The short time step propagator for the diffusion equation (Eq. 137) is known analytically.

$$G(x,y;\Delta t)=\frac{1}{\sqrt{2\pi\Delta t}}e^{-(x-y)^2/(2\Delta t)}e^{-\Delta t(V(y)-E_T)} \tag{138}$$

This propagator gives a quadratic time propagation error

$$\rho(x,t+\Delta t)=\int dy\,\rho(y,t)G(x,y;\Delta t)+O(\Delta t^2)$$

This short time propagator can be simulated by a random walk involving several walkers, i.e configurations representing the distribution ρ as follows:

- Initialize a population of walkers.
- Move these walkers according to the Gaussian distribution

  $$\frac{1}{\sqrt{2\pi\Delta t}}e^{-(x-y)^2/(2\Delta t)}$$

  from the initial configuration y to the new configuration x. This move represents pure diffusion. The term $q=e^{-\Delta t(V(y)-E_T)}$ in Eq. 138 represents a so-called branching process that gives rise to the birth or death of walkers. It is simulated as follows: If q is less than 1 the walker dies with a probability of $1-q$. If q is greater than 1 the

> walker gives birth to either $[q-1]$ or $[q]$ new walkers, where $[\ ]$ represents integer truncation of the argument. The probability for having $[q]$ new walkers is given by $q-[q]$ and $[q-1]$ new walkers will come into existence with complementary probability $1+[q]-q$. Equivalently, one can say that the new number of walkers is given by $[q+r]$, where r is a random number. If $[q+r]$ is greater than one, $[q+r]-1$ walkers are born, otherwise the walker dies.

During this diffusion/branching process one monitors the population of the walkers. If it grows E_T is reduced, otherwise it is increased until one finds a value of E_T where the population remains stable. This value is the eigenvalue of the Schrödinger equation.
In the limit of an infinitesimal small time-step, Diffusion Quantum Monte Carlo is an exact method for bosonic ground states. Problems arise for fermionic systems. The diffusion equation (Eq. 137) assumes that ρ is a positive quantity. A fermionic many body wavefunction has however nodes where its sign changes. It is thus not everywhere positive. As a consequence one has to introduce the so-called fixed node approximation. The nodes of the wavefunction are assumed to be known before one starts the simulation. In most cases the nodes of a variational Quantum Monte Carlo wavefunction are used. Two different diffusion Quantum Monte Carlo simulations are then done in the sub-volume where the wavefunction is positive and where it is negative. Moves where walkers want to cross from one sub-volume to the other are always rejected. It is clear that the result of a diffusion Quantum Monte Carlo simulation can only be as good as the nodes that were

used as the input. Since the nodes obtained from variational Quantum Monte Carlo are usually quite good diffusion Quantum Monte Carlo gives very accurate answers.

Overview over some electronic structure codes

1 Gaussian

- The most widely used package in chemistry oriented fields
- Uses an Gaussian basis set (choosing the right basis set is a black art)
- Can do both DFT as well as wavefunction based methods as well as other things (semiempirical methods, ...)
- All-electron calculations for light elements, pseudopotential calculations for heavy elements.
- Does not perform very well on massively parallel computers
- Commercial

2 VASP

- The most widely used package in solid state type applications
- Uses a plane wave basis set
- Can perform calculations with various exchange correlation functionals, but heavily focused on the PBE functional
- Has high quality PBE pseudopotentials for most elements of the periodic table
- Performs well on parallel computers
- Commercial

3 FHI-aims

- Uses as basis set atomic orbitals
- Significantly higher accuracy for a given basis set size than for a Gaussian basis set
- Full all-electron calculations also for heavy atoms, approximate relativistic methods available
- Commercial

4 BigDFT

- Uses wavelets as basis functions
- Can perform density functional calculation with LDA and GGA (Generalized Gradient Approximation) functionals
- Is targeted towards structure prediction calculations
- Free software

5 Standard output quantities of electronic structure codes

- Energy
- Forces and higher derivates of the energy with respect to the atomic positions such as the Hessian matrix
- MD trajectories, geometry optimizations, vibrational frequencies and vibrational modes
- Electronic charge densities and related quantities such as dipoles and quadrupoles
- Ionization energies and electron affinities
- Information about the insulating/metallic character of the system
- Response properties such as electron-phonon coupling that allows to estimate critical temperatures for superconductivity