Chapter 2

Theory, Methods, . . .

Contents

2.1 Electronic structure methods . . . 8

2.2 Density Functional Theory – DFT . . . 9

2.2.1 Foundations of DFT – Thomas-Fermi theories and Hohenberg-Kohn theorems . 10

2.2.2 Kohn-Sham scheme . . . 11

2.2.3 Exchange-correlation functional . . . 13

2.2.4 Generalised gradient approximation – GGA . . . 16

2.2.5 Other approaches to the exchange-correlation term . . . 19

2.3 Methods . . . 21

2.3.1 Super-cell and Brillouin zone . . . 21

2.3.2 Brillouin zone summation . . . 23

2.3.3 Fractional occupation numbers . . . 24

2.3.4 Pseudo potentials . . . 26

2.4 Implementation . . . 31

2.4.1 Basis sets — plane waves . . . 31

2.4.2 The Car-Parrinello method – iterative schemes . . . 37

2.4.3 An implementation – Faust . . . 39

2.5 Output from the calculations . . . 40

2.5.1 Lattice constant, bulk modulus and the cohesive energy . . . 41

2.5.2 Adsorption energies . . . 41

2.5.3 Density of States – DOS . . . 42

2.5.4 Projected or Local Density of States – PDOS or LDOS . . . 44

2.5.5 Work function . . . 44

2.5.6 Blyholder model for adsorption of carbon monoxide . . . 44

2.5.7 Iso-electronic reactivity function . . . 46

2.5.8 Density differences / induced density . . . 47

2.5.9 Interaction density / energy . . . 47

2.5.10 Electro-static potential / electric field . . . 48

2.5.11 Pseudo scanning tunnelling microscope image . . . 48

2.1 Electronic structure methods

The interactions which govern a system consisting of electrons and atomic nuclei are of electrostatic nature. The Coulomb interaction, although simple in its form, becomes very complicated to handle already for a few particle ensembles and soon impossible to treat exactly with the current computational resources; thus approximations are called for. The ionic and electronic masses differ by a factor over three orders of magnitude so the ionic movement is much slower than the electronic and the nuclei are much more localised than the electronic wave function; therefore the nuclei move slowly in the instantaneous potential created by surrounding, more de-localised electrons. The electronic and ionic degrees of freedom can be de-coupled using the Born-Oppenheimer approximation which states that the electrons sense the instantaneous positions of the ions and adapt to these before the ions have moved significantly; thereby the total many-body wave function $\Psi\left[\{\mathbf{r}_i\},\{\mathbf{R}_I\}\right]$ can be factorised into an ionic $\Phi\left[\{\mathbf{R}_I\}\right]$ and electronic $\chi\left[\{\mathbf{r}_i\},\{\mathbf{R}_I\}\right]$ wave functions; $\mathbf{r}_i$ are the electronic and $\mathbf{R}_I$ the ionic coordinates:

$$\Psi\left[\{\mathbf{r}_i\},\{\mathbf{R}_I\}\right] = \Phi\left[\{\mathbf{R}_I\}\right]\chi\left[\{\mathbf{r}_i\},\{\mathbf{R}_I\}\right] \ . \tag{2.1}$$

Now the electronic degrees of freedom can be solved with the nuclei held fixed and the ions can be regarded as classical point charges:

$$\left[-\frac{1}{2}\nabla^2 + V\right]\Phi = E_{\mathrm{tot}}\Phi \tag{2.2}$$

$$\ddot{\mathbf{R}}_I = \mathbf{F}/M_I \ , \tag{2.3}$$

where the forces on the ions are obtained from the Hellmann-Feynman theorem

$$\mathbf{F}_I = -\frac{dE}{d\mathbf{R}_I} = -\frac{\partial E}{\partial \mathbf{R}_I} \ ; \tag{2.4}$$

the total energy E contains also the ion-ion repulsion in addition to the electronic total energy E_{tot}. The potential for the electrons, V, includes the electron-electron repulsion and the ion-electron attraction. The Born-Oppenheimer approximation usually works well but sometimes it fails for the nuclei of the light atoms, especially hydrogen.

Despite these simplification the electron structure is a many-body problem, and thus still not exactly solvable in a reasonable time. Computational procedures exist which approach the exact result but are too demanding for all except few-atom systems; in physics community the Quantum Monte Carlo – *QMC* – (Ceperley & Alder, 1986; Hammond, Lester & Reynolds, 1994; Ceperley & Mitas, 1996; Mitas, 1996) and in quantum chemistry community the Configuration Interaction – *CI* – (Fulde, 1991; Szabo & Ostlund, 1989) have existed already for decades but still have their limitations. In the quantum Monte Carlo method the many-body electron wave function is probed stochastically and the physical observables are obtained as statistical averages. Different algorithms exist, of different accuracy: the Variational Monte Carlo – *VMC* –, the Diffusion Monte Carlo – *DMC* – and the Green's Function Monte Carlo – *GFMC*. The so-called sign problem (Schmidt & Kalos, 1984; Anderson, 1995) arises from the fermionic exclusion principle and makes the methodology much more difficult than for bosons. The quantum Monte Carlo algorithms have progressed in the recent years in the speed for calculation (Umrigar, 1993; Umrigar, Nightingale & Runge, 1993; Harju *et al*, 1997), but still the quantum Monte Carlo calculations are not broadly used for studies of realistic materials due to the higher complexity in the practical implementation than *e.g.* in the density functional calculations.

The methods in quantum chemistry tackle the problem of interacting particles from a different perspective; in them instead of the many-body wave function $\Phi(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N)$ single-particle wave functions $\Phi(\mathbf{r})$ are employed; a single-particle wave function depends only on the coordinates of a single electron, and is thus much more straight-forward to treat than the many-body wave function which depends explicitly on all the particle coordinates. The total many-body wave function can be regarded to consist of combinations of the single-particle orbitals, usually of determinants or a determinant of them like in the Hartree-Fock method (Ashcroft & Mermin, 1976); the determinant is convenient because it automatically fulfils the Pauli exclusion principle. The Hartree-Fock method is variational because the Hamiltonian is the exact many-body operator and a trial function is assigned for the total wave function which minimises the total energy and thus the approximate wave functions can only yield an energy higher than the exact energy. The Hartree-Fock – or Hartree-Fock-Slater – determinant leads to the single-particle equations

$$-\frac{1}{2}\nabla^2\psi_i(\mathbf{r}) + \sum_I \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|}\,\psi_i(\mathbf{r}) + \int_{\mathbf{r}'} \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}'\,\psi_i(\mathbf{r}) + \sum_j \int_{\mathbf{r}'} \frac{\psi_i(\mathbf{r}')\,\psi_j(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}'\,\psi_j(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}) \ , \quad (2.5)$$

where $\mathbf{R}_I$ are the nuclear positions, $n(\mathbf{r})$ is the electron density $n(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2$ and ε_i the eigenvalue of orbital i. The fact which makes these equations time-consuming to solve is that the potential of orbital i contains explicitly a contribution from orbital j, and the integral operator in the last term – called the *exchange* – on the left hand sign. What is still missing from the Hartree-Fock equations is the *correlation* (please see *e.g.* Fulde, 1991) which describes the dependence in the movement of an electron on the movement of all the other electrons. The correlation is usually of a smaller magnitude than the other terms but significant for a reliable and qualitative analysis; unfortunately it's analytic form is not known and therefore cannot be exactly implemented numerically. In the configuration interaction method the variational Hartree-Fock wave functions – the ground and excited states – are mixed and the result is still guaranteed to be limited from below with the exact energy. The most accurate method of this approach is the Full Configuration Interaction – *FCI* – which include all the configurations in the given basis set.

The mostly used approach in physics and physical chemistry nowadays is the density functional theory which will be shortly introduced in the following Section.

2.2 Density Functional Theory – DFT

Density Functional Theory – *DFT* – has been the cornerstone of realistic, atomic-scale condensed matter investigations since mid-1960's. It is currently the only method which can be applied to problems of moderate size ($\approx 10 - 1000$ atoms) still explicitly accounting for the electronic degrees of freedom. The density functional theory originates as a continuation of the earlier work on approximate density-only theories – Thomas-Fermi, Thomas-Fermi-Dirac – and quantum-chemical approaches – Hartree, Hartree-Fock – yet unlike them it is not approximate *per se* but exact by construction. The theoretical foundations of the density functional theory were introduced by Hohenberg and Kohn (1964), and it was soon thereafter given a practically tractable form by Kohn and Sham – *KS* – (1965). Since then a lot of work has been done to improve on the still currently necessary approximations to the many-body terms in the density functional theory but the basic principles of the Hohenberg-Kohn-Sham formalism are unchanged since their introduction.

As there exist many good textbooks (Lundqvist & March, 1983; Parr & Yang, 1989; Dreizler & Gross, 1990) and reviews (Callaway & March, 1984; Jones & Gunnarsson, 1989) about the density functional theory in the literature we shall here discuss only the aspects which are necessary to understand the practical scheme applied in the present work.

2.2.1 Foundations of DFT – Thomas-Fermi theories and Hohenberg-Kohn theorems

The density functional theory treats the many-body problem by employing the electronic charge density as its basic variable. In this respect it closely resembles the Thomas-Fermi theories which were constructed approximating the kinetic energy contribution E_{kin} to the Coulomb-interacting many-body total energy expression,

$$E_{\text{tot}} = E_{\text{kin}} + E_{\text{ext}} + E_{\text{e-e}} \ , \tag{2.6}$$

by its correspondent for the local, non-interacting energy gas. Here $E_{\text{ext}} = \int_{\mathbf{r}} n(\mathbf{r}) V_{\text{ext}}(\mathbf{r})\, d\mathbf{r}$ is the energy arising from an external potential $V_{\text{ext}}(\mathbf{r})$, usually the attractive Coulomb interaction with ionic cores; $E_{\text{e-e}}$ is the electron-electron interaction, which for convenience is usually split into two parts, the classical Hartree energy

$$E_{\text{Hartree}} = \frac{1}{2} \int_{\mathbf{r}} \int_{\mathbf{r}'} \frac{n(\mathbf{r})\, n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}'\, d\mathbf{r} \tag{2.7}$$

and the so-called *exchange-correlation* functional $E_{\text{xc}}[n]$, which will be discussed below.

The density functional theory and the Kohn-Sham method as its application is theoretically well justified. The basis for the density functional theory is constituted by the two mathematical theorems of Hohenberg and Kohn (1964) for a system with non-degenerate electronic ground state:

1. The potential V_{ext} is – to within a constant shift – a unique functional of the density $n(\mathbf{r})$; especially the ground state density $n_0(\mathbf{r})$ corresponding to an external potential V_{ext} cannot be reproduced using any other potential V'_{ext} and thus, since the density is uniquely related to the external potential it also is sufficient to determine the ground state and all other electronic properties or observables of the system.

2. The correct ground state density $n_0(\mathbf{r})$ minimises the total energy functional

$$E_{\text{tot},V}[n] = E_{\text{kin}}[n] + E_{\text{e-e}}[n] + E_{\text{ext}}[n] = F_{\text{HK}}[n] + E_{\text{ext}}[n] \ , \tag{2.8}$$

now written using the universal – *i. e.* independent of any particular external potential – functional $F_{\text{HK}}[n]$ defined by Hohenberg and Kohn; the sub-script V reminds here that the energy is defined for a given external potential.

The essential content of the Hohenberg-Kohn theorems is that to find the electronic properties of the system, defined by V_{ext}, it is enough to minimise the total energy expression with respect to all the densities $n(\mathbf{r})$. The one giving the lowest energy is the ground state density and it can then be used to evaluate the wanted observables. However, this still remains a formidable task due to its many-body nature and a practical scheme is highly welcome. Such was described by Kohn and Sham (1965) and will be shortly introduced below.

We stress here that the density functional theory is thus a theory for the ground state. It can, however, in particular cases be used to extract properties of excited states; please see *e. g.* Dreizler and Gross (1990). For extensions of Hohenberg-Kohn theorems for systems with degenerate ground states, spin-polarised, magnetic, time-dependent systems *etc* please see the textbooks mentioned above and references therein.

2.2.2 Kohn-Sham scheme

Kohn and Sham (1965) conducted a scheme which employs single-particle wave functions, similarly to the Hartree and Hartree-Fock methods. However, the following remarks are in order:

- The single-particle Kohn-Sham orbitals serve only as an auxiliary basis set for the calculation, they are not guaranteed to bear any resemblance to the exact single-particle orbitals.
- The same comment holds for the single-particle eigenvalues, except the eigenvalue of the highest occupied orbital and ionisation energy (Perdew *et al*, 1982). This point cannot be over-emphasised, in particular in the semi-conductor community, where for example the "gap problem" has received lots of concern.
- The Kohn-Sham formalism is still exact by construction. In practice an approximation is necessary to account for the many-particle exchange-correlation effects but this approximation is well-controlled and many exact properties for them are known.

Nowadays the Kohn-Sham scheme constitutes the main-stream of the applications of the density functional theory to real-world systems, and even a limited list of system where Kohn-Sham formalism has been applied would be overwhelming. The explanation behind the success of Kohn-Sham method lies in its simplicity and relatively low computational cost needed to access moderately accurate results.

The Kohn-Sham scheme can be thought as a mapping of the many-body problem of fully-interacting particles in an external potential V_{ext} to a system of non-interacting particles in an effective potential V_{eff} giving the same ground state density; thus the Hamiltonian $\hat{H} = \hat{T} + \hat{V}_{\text{e-e}} + V_{\text{ext}}$ gets replaced by $\hat{H}_s = \hat{T}_s + V_{\text{eff}}$. $\hat{T}$ and $\hat{T}_s$ are the kinetic energy operators of interacting and non-interacting particles, respectively, and $V_{\text{e-e}}$ is the electron-electron interaction. The effective potential must be local and it is the same for all the orbitals, leading to a mean-field kind of description of the interactions.

In the framework of one-particle orbitals $\psi_i(\mathbf{r})$ the kinetic energy of non-interacting particles can easily be calculated exactly:

$$E_{\text{kin},s}[n] = \sum_i^N \int_{\mathbf{r}} \psi_i(\mathbf{r})^* \left(-\frac{1}{2}\nabla^2\right) \psi_i(\mathbf{r})\, d\mathbf{r} \ , \tag{2.9}$$

where the sum runs over the occupied states up to the N electrons. That a density functional formalism contains explicitly orbital-dependent terms is not a conceptual problem since if one first assumes the existence of the effective potential V_{eff} generating the ground state density n, then according to the first Hohenberg-Kohn theorem it is unique and thus the single-particle orbitals, being a property of the ground state, are functionals of the density: $\psi_i(\mathbf{r}) = \psi_i([n];\mathbf{r})$. The improved accuracy of the Kohn-Sham method over the Thomas-Fermi approaches lies just at the exact description of the kinetic energy; thus the large extra cost due to the single-particle orbital treatment compared to the density-only schemes is justified.

The single-particle equations of the Kohn-Sham method can be derived from the total energy expression. The functional $F_{\text{HK}}[n]$ is first divided into $E_{\text{kin},s}[n]$ and a new universal functional $G[n]$:

$$\begin{aligned} E_{\text{tot}}[n] &= F_{\text{HK}}[n] + E_{\text{ext}}[n] \\ &= E_{\text{kin},s} + G[n] + E_{\text{ext}}[n] \ . \end{aligned} \tag{2.10}$$

The functional $G[n]$ is further split into the classical Hartree electron-electron interaction term $E_{\text{Hartree}}[n]$ of the Equation 2.7 and a term containing the many-body interactions, $E_{\text{xc}}[n]$, the exchange-correlation energy which was mentioned above. Formally $E_{\text{xc}}[n]$ is thus defined as

$$E_{\text{xc}}[n] = F_{\text{HK}}[n] - E_{\text{Hartree}} - E_{\text{kin},s} \; . \tag{2.11}$$

The final total energy functional reads now

$$E_{\text{tot}}[n] = E_{\text{kin},s}[n] + E_{\text{Hartree}}[n] + E_{\text{xc}}[n] + E_{\text{ext}}[n] \; . \tag{2.12}$$

Note that once a practical scheme for obtaining the single-particle orbitals $\psi_i(\mathbf{r})$ (for $E_{\text{kin},s}$) exists and the form of $E_{\text{xc}}[n]$ is known the exact ground state energy can be easily obtained from the expression 2.12. However, as will be discussed below, although the first part – finding $\psi_i(\mathbf{r})$ – is straight-forward the exchange-correlation part remains a challenge and currently requires an approximation.

When the single-particle orbitals are employed they are in principle functionals of the density as just mentioned. The density is obtained using the orbitals as

$$n(\mathbf{r}) = \sum_i^N \psi_i(\mathbf{r})^* \psi_i(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2 \; . \tag{2.13}$$

The requirement of the minimisation of the energy with respect to a change in the wave function lead into a set of single-particle equations known as the Kohn-Sham equations:

$$\left(-\frac{1}{2}\nabla^2 + V_{\text{Hartree}} + V_{\text{xc}} + V_{\text{ext}}\right)\psi_i = \varepsilon_i \psi_i(\mathbf{r}) \; . \tag{2.14}$$

This corresponds to a partial differential equation with a Hamiltonian

$$H_s = -\frac{1}{2}\nabla^2 + V_{\text{Hartree}} + V_{\text{xc}} + V_{\text{ext}} = -\frac{1}{2}\nabla^2 + V_{\text{eff}}[n] \; , \tag{2.15}$$

The effective potential depends on the density, the Kohn-Sham orbitals ψ_i on the effective potential, and the density on the orbitals due to its definition 2.13; thus this set of equations form a loop which has to be solved self-consistently: One starts *e.g.* with an approximate density, constructs the effective potential and solves the Kohn-Sham equations to obtain the orbitals which are then used to calculate the new density *etc*. This procedure is continued until the density, orbitals, total energy and so on no longer change from the previous iteration.

As each of the Kohn-Sham equations is of the same form as the Schrödinger equation the same methods can be used to solve the Kohn-Sham orbitals. However, because the Kohn-Sham scheme consists of the self-consistent loop this may cause different algorithms which do not perform the full solution of the Kohn-Sham equations at each iteration to become more efficient than the traditional, straight-forward matrix diagonalisation. We shall shortly discuss some different algorithms when discussing the practical implementation of the Kohn-Sham method.

Studying the form of the Kohn-Sham equations 2.14 one can easily derive an alternative expression to the Equation 2.12 for the total energy,

$$E_{\text{tot}} = \sum_i^{\text{occ}} \varepsilon_i - E_{\text{Hartree}} + E_{\text{xc}} - \int n(\mathbf{r}) V_{\text{xc}}(\mathbf{r}) \, d\mathbf{r} \; . \tag{2.16}$$

The terms additional to the sum over the occupied eigenvalues are called the double-counting correction terms and are caused by the non-linearity of the total energy expression 2.12 with respect to the density – *e.g.* the Hartree energy contains the square of the density – so that the sum over eigenvalues

$$\sum_{i}^{N} \langle\psi_i| H |\psi_i\rangle \neq E_{\mathrm{tot}} \; . \tag{2.17}$$

Above we have discussed only the spin-polarised, integer-occupied formulation of the Kohn-Sham scheme; more general discussion will follow, but to introduce the theoretical background a simplified framework was chosen. For the detailed derivation of the more general Kohn-Sham formalism please see *e.g.* Dreizler and Gross (1990).

2.2.3 Exchange-correlation functional

The term of exchange and correlation functional was introduced into the Kohn-Sham theory when the kinetic energy of non-interacting particles $E_{\mathrm{kin,s}}$ was inserted to replace the true kinetic energy, and the classical Hartree interaction 2.7 was extracted from the full electron-electron interaction in order to enable an exact evaluation of the terms, hoping that the remaining terms are small in magnitude or can be approximated sufficiently accurately. Therefore it might first be surprising that the exchange and correlation constituents have clear physical origins: The exchange arises from the Pauli principle that two electrons with parallel spins cannot occupy the same region in space, and correlation because the motion of an electrons is correlated, *i.e.* the instantaneous position of an electron depends on all the other electronic positions as well. While the exchange energy is known exactly and used in the Hartree-Fock method, correlation is analytically only moderately known; it can, however, be obtained in some limiting cases and calculated to a very high accuracy using the quantum chemistry (Fulde, 1991) or quantum Monte Carlo (Ceperley & Alder, 1980) methods. Note that the definition of correlation associated with the Hartree-Fock scheme is somewhat different from the definition within density functional theory: in Hartree-Fock-correlated schemes the correlation energy is a correction to the Hartree-Fock energy, *i.e.* the wave functions are calculated to self-consistency without a correlation potential, but in density functional theory it is included in the self-consistent effective potential.

We shall briefly discuss some approximative exchange and correlation functional schemes. The main emphasis will, however, be on the generalised gradient approximation as the main bulk of the result of the present work was produced with it. Local density approximation is discussed in length due to its simplicity and wide use. Due to restrictions in the length we shall discuss only the spin-unpolarised case below unless otherwise mentioned; the generalisation to spin-polarised case is straight-forward – please see *e.g.* Parr and Yang (1989).

Local density approximation – LDA

The most straight-forward and frequently used approximation is the Local Density Approximation – *LDA* – (Kohn & Sham, 1965), or its spin-polarised version Local Spin Density Approximation – *LSDA* or *LSD* –, where the exchange-correlation energy and potential at each point of space are replaced by their correspondents for the homogeneous electron gas, *i.e.* as if at each point separately there was a constant density, independent of all the other points. The exact expression for exchange energy of homogeneous electron gas is easily derivable analytically (Dirac, 1930; Gáspár, 1954) and highly accurate quantum Monte Carlo calculations can be done for the correlation energy (Ceperley & Alder, 1980) in addition to some exact analytical limits (Wigner, 1934; Carr, 1961; Nozieres &

Pines 1966; Gell-Mann & Brueckner, 1957; Carr & Maradudin, 1964) – please see *e. g.* Mahan (1981). The local density approximation is expected to give good results only when the variations of homogeneity are small or slow; slow means the density changes are small within the Fermi wave momentum $k_F = \left[3\pi^2 n\left(\mathbf{r}\right)\right]^{1/3}$. However, neither of these requirements are ful-filled in real systems. Thus the high level of accuracy in calculations employing local density approximation – reviewed *e. g.* by Jones and Gunnarsson (1989) and Parr and Yang (1989) – is really surprising. Part of this success can be explained by the exact properties obeyed by the local density approximation, like the sum rule that the exchange-correlation hole around an electron contains exactly minus one electron – *i. e.* one hole –

$$\int_{\mathbf{r}'} n_{\mathrm{xc}}\left(\mathbf{r},\mathbf{r}'\right)\, d\mathbf{r}' = -1 \;\forall \mathbf{r} \tag{2.18}$$

and correct scaling behaviours, and partially by error cancellation; for example the errors in exchange have partially different sign than errors in correlation. Also helpful is that the region where typically most of the electrons reside, the core electrons close to the nuclei, are chemically inert and thus the error *e. g.* in total energy between two similar systems – like a molecule and its atoms separated infinitely far apart – is smaller than the errors in total energies. The amount of correlation, whose characteristics are least known, is also smallest; the accuracy of exchange energy and potential can at least be tested using exact functional like the Hartree-Fock method or the exact-exchange methods in the density functional theory (Talman & Shadwick, 1976; Aashamar, Luke & Talman, 1978).

The exchange-correlation energy in the local density approximation is given by

$$E_{\mathrm{xc}}^{\mathrm{LDA}}\left[n\right] = \int_{\mathbf{r}} n\left(\mathbf{r}\right) \varepsilon_{\mathrm{xc}}^{\mathrm{hom}}\left(n\left(\mathbf{r}\right)\right)\, d\mathbf{r} \tag{2.19}$$

and the potential by

$$V_{\mathrm{xc}}^{\mathrm{LDA}}\left(\mathbf{r}\right) = \frac{\partial\left(n\varepsilon_{\mathrm{xc}}^{\mathrm{hom}}\right)}{\partial n} = \varepsilon_{\mathrm{xc}}^{\mathrm{hom}} + \frac{n\partial\varepsilon_{\mathrm{xc}}^{\mathrm{hom}}}{\partial n}\ , \tag{2.20}$$

where $\varepsilon_{\mathrm{xc}}^{\mathrm{hom}}\left(n\left(\mathbf{r}\right)\right)$ is the energy density of homogeneous electron gas *per particle*. There exist several parametrisations of the correlation energy of the homogeneous electron gas. The most accurate and famous ones by Perdew and Zunger (1981), Vosko, Wilk and Nusair (1980) and Perdew and Wang (1992a) are fits to the quantum Monte Carlo results of Ceperley and Alder (1980). Recent accurate parametrisation of $\varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}\left(n\right)$is given in the Figure 2.1.

One clear case where the local density approximation fails qualitatively is the asymptotic decay of potential away from a density distribution; one can derive that the exact energy falls off like $\propto \frac{1}{r}$ outside the density distribution $n\left(\mathbf{r}\right)$ (Almbladh & von Barth, 1985) whereas in the local density approximation it decays exponentially away from n (Tong & Sham, 1966; Lang & Kohn, 1970; Lang & Kohn, 1971) due to the exponential decay of the density. Although the tail of the potential is essential in determining the ionisation energies of atoms and molecules or the work function in the case of a surface, the local density approximation often provides good results for these quantities. However, *e. g.* the negatively charged atoms are incorrectly not stable within the local density approximation due to this shortcoming in the asymptotic potential.

In overview general features of local density approximation are:

- binding energies, cohesion in case of solids, are usually overestimated, *i. e.* too much binding is given, often as much as 30 % – please see *e. g.* Parr and Yang (1989) or Dreizler and Gross (1990) – but in systems of weakly bound such as by hydrogen bonds – please see *e. g.* Laasonen, Csajka and Parrinello (1992) – or van der Waals forces even five times too large binding energies

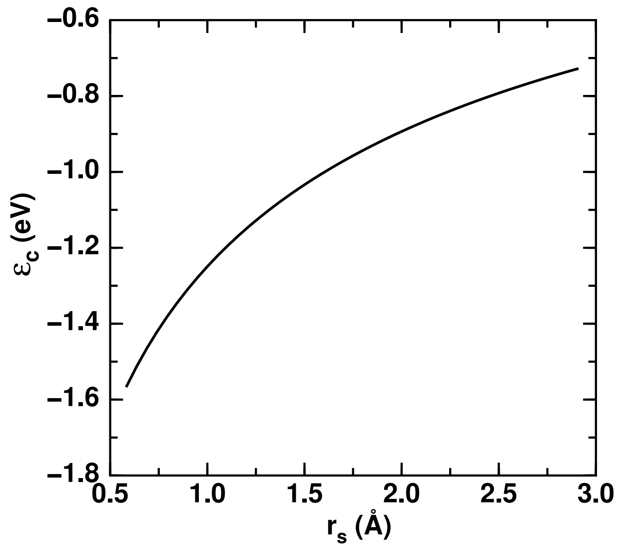


Figure 2.1: The parametrisation of the form $\varepsilon_c^{\mathrm{LDA}}(n) = -2A\,(1+\alpha_1 r_s)\log\left[1+\frac{1}{2a\left(\beta_1 r_s^{1/2}+\beta_2 r_s+\beta_3 r_s^{3/2}+\beta_4 r_s^2\right)}\right]$, with constants a, α_1, β_1, β_2, β_3, β_4, of the correlation energy for the homogeneous electron gas by Perdew and Wang (1992); $r_s = \left[\frac{3}{4\pi n(\mathbf{r})}\right]^{1/3}$ is the local Wigner-Seitz radius. Thus the correlation energy is obtained like in the local exchange: one inserts at a point the density and obtains from the functions the exchange and correlation energy density and potential, independently of the other points

are obtained (Pérez-Jordá & Becke, 1995; Montanari, Ballone & Jones, 1998; Zhang, Pan & Yang, 1997), for example in dimers of rare gas atoms

- lattice constants of solids are mostly too small due to under-binding of the core electrons and thus too diffuse cores seen by the valence electrons

- vibrational frequencies, phonons in solids, are well given; the bulk compressibility or bulk modulus is usually too large as a consequence of the too small a lattice constant

- the magnetic or spin-polarised properties are often non-satisfactorily predicted; for example the ground state of iron is erroneously not the ferromagnetic body-centred cubic phase but the non-magnetic hexagonal close-packed lattice

- reaction barriers are often underestimated *e.g.* for $H+H_2 \rightarrow H_2+H$ (Johnson *et al*, 1994) and H_2 dissociation on transition metal surfaces like Cu(111) (Hammer *et al*, 1994) and Al(110) (Hammer, Jacobsen & Nørskov, 1993); this was concluded to arise from incomplete Pauli repulsion between the states of the anti-bonding orbital of hydrogen, whose filling leads to the dissociation, and the states of the substrate

- *sd* transfer energies are too low (Gunnarsson & Jones, 1985)

- some negative ions are not bound although they are stable in experiments (Schwarz, 1978; Shore, Rose & Zaremba, 1977); this deficiency originates from the wrong asymptotic behaviour of the exchange and correlation as discussed above

The local density approximation has been the backbone of applications of density functional theory since the beginning. The straight-forward implementation of it and superior accuracy over the Hartree method with the same and comparable accuracy with Hartree-Fock method with a smaller computational cost have resulted in a vast amount of calculations using it. Much work has been devoted to go beyond the local density approximation but despite progress no method with overall, systematic improvement has been found yet. Moreover, the delicate error cancellation between exchange and correlation in the local density approximation aggravates the simultaneous testing and especially application of new exchange and correlation functionals difficult.

2.2.4 Generalised gradient approximation – GGA

The most obvious method to extend the local density approximation to inhomogeneous systems is to include also the changes or gradients of the density. The straight-forward attempt is to try a first-order Taylor expansion around the constant density and hope for the inhomogeneities to be either small or at least slowly varying, thus justifying the low-order Taylor expansion. The expansion has been derived for the kinetic, exchange and correlation energies, the former to be used in the density-only like Thomas-Fermi methods. For correlation, even though the analytical formula for the homogeneous density is missing, a gradient expansion is known. The expansion is expected to work only when (Hohenberg & Kohn, 1964)

$$\frac{|\nabla n(\mathbf{r})|}{2n(\mathbf{r})} << k_F(\mathbf{r}) \quad \frac{|\nabla^2 n(\mathbf{r})|}{2|\nabla n(\mathbf{r})|} << k_F(\mathbf{r}) \ . \tag{2.21}$$

In this approach, called the Gradient Expansion Approximation – *GEA* – (Hohenberg and Kohn, 1964), the terms with odd powers of the gradient of density, ∇n, must vanish due to rotational invariance. Thus the lowest-order gradient expansion approximation for the exchange reads as (Sham, 1971; Gross & Dreizler, 1981)

$$\begin{aligned} E_{\mathrm{x}}^{\mathrm{GEA}}(n, \nabla n) &= \underbrace{E_{\mathrm{x}}^{(0)}}_{=E_{\mathrm{x}}^{\mathrm{LDA}}} + E_{\mathrm{x}}^{(2)} = d_0 \int_{\mathbf{r}} n^{4/3}\, d\mathbf{r} + d_2 \int_{\mathbf{r}} \frac{|\nabla n|^2}{n^{4/3}}\, d\mathbf{r} \ , \\ & d_0 = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}, \ d_2 = -\frac{7}{432}\left(3\pi^5\right)^{-1/3} \ . \end{aligned} \tag{2.22}$$

Please note the orders of $|\nabla n|$ and n in this expression; it can be shown that the exact functional has the order $|\nabla n| / n^{4/3}$ in all the orders of the expansion. The zeroth order term is the local density approximation as expected.

Practical tests with the gradient expansion approximation applied have shown that many properties are less accurately produced than in the local density approximation. For example the correlation energy in helium and neon atoms has even a wrong sign (Burke, Perdew & Levy, 1995), the second-order exchange potential diverges in atoms both at large and small radii (Herman, van Dyke & Ortenburger, 1969) and the energy terms of the fourth or higher order also diverge for finite systems. One of the reasons for the shortcomings of the gradient expansion approximation is that the sum rule 2.18 is not obeyed, and the value of scaled gradient does not meet the expectations 2.21 set for the convergence of a Taylor series. Thus the gradient expansion does not converge after the first, lowest-order terms, invalidating the low-order Taylor expansion. Furthermore since the computation of higher gradients and gradients with increasing order of the expansion are finer and thus need higher a convergence with respect to the basis set this approach is not satisfactory and an alternative method has to be searched for.

A different way to include the gradient of the density is to construct a functional which still depends only on the local density $n(\mathbf{r})$ and the magnitude of the gradient $|\nabla n(\mathbf{r})|$ but whose functional form is no longer chosen to be the one of the exact gradient expansion. Therefore this approach is called the Generalised Gradient Approximation – *GGA* –, generalised because of the increased freedom in the choice of the form and approximation because the chosen function usually does not follow the exact lowest-order properties of the electron gas, whereas the gradient expansion approximation can be called a gradient correction – although the "correction" terms do not necessarily improve the results as discussed above. In the literature the term "gradient corrections" has often been misused to mean also generalised gradient approximations.

The choice of the functional form for the generalised gradient approximation is not unique and many different functions have been proposed. Here we shall follow the formalism of Perdew *et al* since their forms are the most *ab initio* in nature and they have also been used for the calculations of the present work. By the years they have proposed several parametrisations of which the so-called Perdew-Wang 1991 or GGA-II (Perdew, 1991; Perdew *et al*, 1992; Perdew, Burke & Wang, 1996) has been used most; it has lately been slightly modified in the newer Perdew-Burke-Ernzerhof (1996) parametrisation. This parametrisation is mostly just a simplification of the Perdew-Wang 1991 functional, sacrifying some exact properties for the sake of simplicity of the resulting functional. Although the parametrisation Perdew-Wang 1991 was used for the majority of the calculations in the present work we shall describe here the Perdew-Burke-Ernzerhof parametrisation because of the simpler form it possesses. They write the exchange and correlation energies somewhat differently due to the different spin dependence, as the correct exchange scales as (Oliver & Perdew, 1979)

$$E_{\mathrm{x}}\left[n_{\pm}\right]=\frac{1}{2}\left(E_{x}\left[2n_{+}\right]+E_{x}\left[2n_{-}\right]\right) \ , \tag{2.23}$$

where n_{+}, n_{-} are the spin up and down densities, respectively. In exchange an enhancement factor F^{PBE} over the local density approximation energy density is used, whereas in correlation the gradient-dependent part is an additive to the correlation of homogeneous density distribution:

$$\begin{aligned}
E_{\mathrm{x}}^{\mathrm{PBE}}\left(n\left(\mathbf{r}\right),\left|\nabla n\left(\mathbf{r}\right)\right|\right) &= \int_{\mathbf{r}} n\left(\mathbf{r}\right)\varepsilon_{\mathrm{x}}^{\mathrm{LDA}}\left(n\left(\mathbf{r}\right)\right)F_{\mathrm{x}}^{\mathrm{PBE}}\left(s\right)\,d\mathbf{r}\ ,\\
F_{\mathrm{x}}^{\mathrm{PBE}}\left(s\right) &= 1+\kappa-\frac{\kappa}{1+\frac{\mu s^{2}}{\kappa}}\\
\mu &= \beta\left(\frac{\pi^{2}}{3}\right)\\
E_{\mathrm{c}}^{\mathrm{PBE}}\left(n\left(\mathbf{r}\right),\left|\nabla n\left(\mathbf{r}\right)\right|\right) &= \int_{\mathbf{r}} n\left(\mathbf{r}\right)\left[\varepsilon_{\mathrm{xc}}^{\mathrm{LDA}}\left(n\left(\mathbf{r}\right)\right)+H_{\mathrm{c}}^{\mathrm{PBE}}\left(r_{s},\eta,t\right)\right]\,d\mathbf{r}\ ,\\
H_{\mathrm{c}}^{\mathrm{PBE}}\left(r_{s},\eta,t\right) &= \gamma\phi^{3}\ln\left[1+\frac{\beta}{\gamma}t^{2}\left(\frac{1+At^{2}}{1+At^{2}+A^{2}t^{4}}\right)\right]\ ,\\
A\left(r_{s},\eta\right) &= \frac{\beta}{\gamma}\frac{1}{e^{-\varepsilon_{\mathrm{c}}^{\mathrm{LDA}}/\gamma\phi^{3}}-1}\ ,\\
\phi\left(\eta\right) &= \frac{1}{2}\left[\left(1+\eta\right)^{2/3}+\left(1-\eta\right)^{2/3}\right]\ ,\\
\gamma &= \frac{1-\ln 2}{\pi^{2}}\ .
\end{aligned} \tag{2.24}$$

Here $s\left(\mathbf{r}\right)=\frac{|\nabla n|}{2k_{F}n}$ and $t\left(\mathbf{r}\right)=\frac{|\nabla n|}{2\phi k_{s}n}$, $k_{s}=\sqrt{4k_{F}/\pi}$, are dimensionless density gradients – of the correct form discussed above – and β comes from the generalised gradient expansion for the correlation (Perdew *et al*, 1992) and is $\approx 0.066\,725$ (Wang & Perdew, 1991). κ is formally set by the Lieb-Oxford bound (1981) for the exchange energy

$$E_{\mathrm{x}}\left[n_{\pm}\right]\geq E_{\mathrm{xc}}\left[n_{\pm}\right]\geq -1.679\int_{\mathbf{r}} n\left(\mathbf{r}\right)^{4/3}\,d\mathbf{r}\ . \tag{2.25}$$

However it has been lately suggested (Zhang & Yang, 1998; Perdew, Burke & Ernzerhof, 1996; Mortensen, 1998; Hammer, Hansen & Nørskov, 1999) that choosing a different, larger value of κ

than 0.804 derived by Perdew, Burke and Ernzerhof leads to improved energetics and thus has been used semi-empirically in calculations; however, the geometric properties like the lattice constant would be worsened.

Requirements in the construction of $F_{\mathrm{x}}^{\mathrm{PBE}}$ and $H_{\mathrm{c}}^{\mathrm{PBE}}$ are

- they exactly reduce to the homogeneous limit, the local density approximation, for constant density *i. e.* then $F \equiv 1$ and $H \equiv 0$; in the slowly varying limit the correlation is given by its second order gradient expansion whereas the exchange is set to cancel the correlation in this limit to recover the good linear response of the local spin density approximation

- the exchange-correlation hole integrates up exactly to minus one to satisfy the sum rule 2.18; this is achieved by a real-space cut-off construction (Perdew, Burke & Wang, 1996) where the exchange-correlation hole of the second-order gradient expansion is set to zero beyond a radius where the sum rule is obeyed. The Perdew-Wang 1991 and Perdew-Burke-Ernzerhof 1996 functionals are analytic fits to the obtained function

- in the rapidly varying limit $t \to \infty$ the correlation vanishes

- under uniform coordinate scaling $n(\mathbf{r}) \to \lambda^3 n(\lambda \mathbf{r})$ the exchange energy must scale like λ (Levy & Perdew, 1985) and the correlation scales to a constant when $\lambda \to \infty$ (Levy, 1989)

From the enhancement factor $F_{\mathrm{xc}}^{\mathrm{PBE}}$, defined by

$$E_{\mathrm{xc}}^{\mathrm{PBE}}\left(n(\mathbf{r}), |\nabla n(\mathbf{r})|\right) = \int_{\mathbf{r}} n(\mathbf{r})\, \varepsilon_{\mathrm{x}}^{\mathrm{LDA}}\left(n(\mathbf{r})\right) F_{\mathrm{xc}}^{\mathrm{PBE}}\left(r_s, \eta, s\right)\, d\mathbf{r} \tag{2.26}$$

or

$$F_{\mathrm{xc}}^{\mathrm{PBE}}\left(r_s, \eta, s\right) = \frac{\varepsilon_{\mathrm{xc}}\left(n(\mathbf{r})\right)}{\varepsilon_{\mathrm{x}}^{\mathrm{LDA}}\left(n(\mathbf{r})\right)}\ , \tag{2.27}$$

it can be seen that the gradient terms deepen the exchange-correlation energy because the exchange energy of the homogeneous electron gas is negative and the enhancement factor is larger than one; thus inhomogeneities are preferred.

Other popular parametrisations within the generalised gradient approximation are the exchange functional of Becke (1988) and the correlation functional functional of Lee, Yang and Parr (1988) which start from the Colle-Salvetti (1975; 1979) formula to be discussed below. The combination of these two – commonly denoted as *BLYP* – has been used widely in the computational chemistry although they both contain a fit parameter and the correlation energy does not reduce to the local density approximation for the homogeneous electron gas; yet the results are often pretty good (Johnson, Gill & Pople, 1993) for the light elements.

The generalised gradient approximations and especially the parametrisations of Perdew-Wang 1991 and Perdew, Burke and Ernzerhof 1996 have the following general properties:

- the core electrons are bound more strongly, thus the valence electrons see a better screened nucleus and are less bound than in local density approximation

- as a consequence the bond lengths in molecules, lattice constants of solids, tend to come out too large, with some exceptions like aluminium

- the bulk moduli of solids are usually too low, most probably due to the too large a lattice constant

- the accuracy of vibrational frequencies is similar to local density approximation

- reaction barriers are higher than in local density approximation but still often too low, *e. g.* for dissociation of H_2 on Cu (111) (Hammer *et al*, 1994) and $H+H_2 \Leftrightarrow H_2+H$ (Johnson *et al*, 1994)
- the work function of surfaces is usually too low whereas they are much closer to the experimental values within the local density approximation; the reason for this is still unknown (Martin Fuchs, personal communication)
- the surface energies are often lower than those from local density approximation or experiments; this has been suggested to be due to the too weak bonding in generalised gradient approximation so when the coordination is reduced by cutting the surface it does not cost as much energy, thus giving too low a surface energy
- the ground state of iron is correctly reproduced by the generalised gradient approximations
- the binding energy of rare gas dimers are in good agreement with experiments within the Perdew-Burke-Ernzerhof 1996 functional (Pérez-Jordá & Becke, 1995; Zhang, Pan & Yang, 1997) whereas they were even unbound with the older parametrisations of the exchange energy by Perdew and Wang (1986) and Becke (1988)

Thus generalised gradient approximation often does better than the local density approximation but has its pitfalls too. The interest is now shifting towards a generalised gradient approximation with second order gradients, and suggestions for such have been given by Engel and Vosko (1994), Proynov, Vela and Salahub (1994), Neumann and Handy (1997). Becke (1993; 1996) has introduced a functional where exact exchange from a Hartree-Fock-operator is mixed with the generalised gradient approximation for the exchange, and with the correlation functional of Lee, Yang and Parr this *B3LYP* approximation has become widely used in computational chemistry.

For the systems studied in the present work the generalised gradient approximation is more advantageous than harmful, and since the computational cost is only minimally higher than in local density approximation we preferred the generalised gradient approximation and used it for most of our calculations. However, the calculation of gradients and more complicated form of the exchange-correlation functional set a slightly higher convergence requirement for the basis set size.

2.2.5 Other approaches to the exchange-correlation term

There has naturally been a long-standing effort to find exchange-correlation functionals which would perform better than the local density or generalised gradient approximations. All of them lead to improved properties in some quantity, but no general improvement in the comparison to experiments has been achieved; only the method of exact exchange provides, as the name suggests, an over-all up-grade in the accuracy. However, it is computationally demanding and a correlation functional accurate enough has not been found yet.

- The correct potential acting on an electron does not explicitly depend on itself but only indirectly via the other electrons. However, the exchange-correlation potentials from *e. g.* local density and generalised gradient approximations do not cancel the contribution of the electron in the Hartree term. The method of the self-interaction correction – *SIC* – (Perdew & Zunger, 1981) corrects for this shortcoming explicitly, but leads to a more complicated set of single-particle equations, with orbital-dependent potentials so the eigenfunctions are no longer invariant with respect to an orthonormal rotation of the orbitals. Furthermore the correction from this scheme vanishes for Bloch waves so that alternative basis sets, *e.g.* Wannier functions, should be employed. These complications explain why the method has been seldomly used

- A truly non-local exchange-correlation scheme is the weighted density approximation – *WDA* – (Gunnarsson, Jonson & Lundqvist, 1977; 1979; Alonso & Girifalco, 1977; 1978) which is obtained from the exact expression for the exchange-correlation energy by replacing the pair correlation functional of the non-homogeneous electron distribution by its correspondent for the homogeneous electron gas, but the correct, non-local density pre-factor is remained unlike in the local density approximation. This leads to the improved asymptotic behaviour of the exchange-correlation term, although with the wrong proportionality factor. The scheme is computationally somewhat more involved and not as straight-forward to implement as the local density and generalised gradient approximations, and the original results were not too encouraging but recently Singh (1993) has reported newer calculations with good accuracy

- The Hartree-Fock contains the correct exchange energy, the Fock term. It can be used also in the Kohn-Sham method but the potential will be derived in a different manner, *e.g.* using the method of optimised effective potential – *OEP* –, also called optimised potential method – *OPM*. This leads to a local effective potential which is the same for all the orbitals, as required to remain in the framework of the Kohn-Sham scheme. This calculation is non-trivial and an approximation, which however still requires a large amount of computation, was derived by Krieger, Li and Iafrate – *KLI* – (1992). The exact exchange leads to energies very close to the Hartree-Fock method when used without the correlation functional, but the poor knowledge of the exact correlation functional makes the very accurate calculations using the Kohn-Sham scheme difficult. However, now that the accuracy provided by the first orbital-dependent correlation functional of Colle and Salvetti (1975; 1979) has been found to be inadequate the rather new functional (Görling & Levy, 1994; Engel *et al*, 1998) based on the perturbation theory similar to the MP2 correlation in the case of the Hartree-Fock method might provide improved results. However, the computational effort of this functional is even larger than that of the exact exchange

- Two methods, the screened exchange (Bylander & Kleinman, 1990; Seidl, 1995; Seidl *et al*, 1996) and local mass approximations (Engel & Pickett, 1996), which were recently proposed try to approximate the correct exchange-correlation term by splitting it into two parts which are known at the limiting cases of a non-homogeneous and a homogeneous electron system, and the sum of the two exchange-correlation contributions is known more accurately than the total term. The screened exchange method is computationally very demanding, whereas the local mass approximation needs less than twice the computational time of *e.g.* the corresponding calculation employing the local density approximation. The number of applications of these approximations is still scarse and has mainly concentrated on the gap problem in the semiconductors

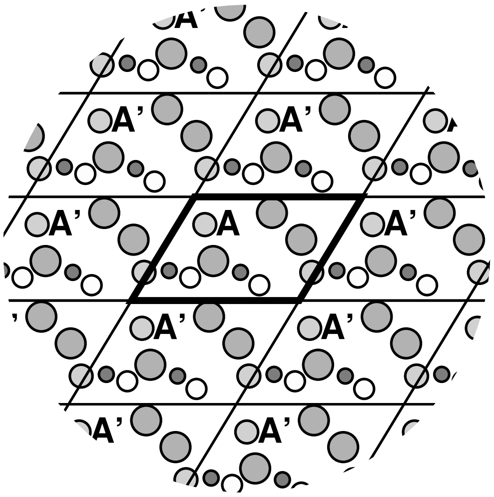


Figure 2.2: An illustration of two-dimensional periodic boundary conditions for atoms. The atom **A** and its periodic replicas **A'** are shown for one atom, and the super-cell has been emphasised

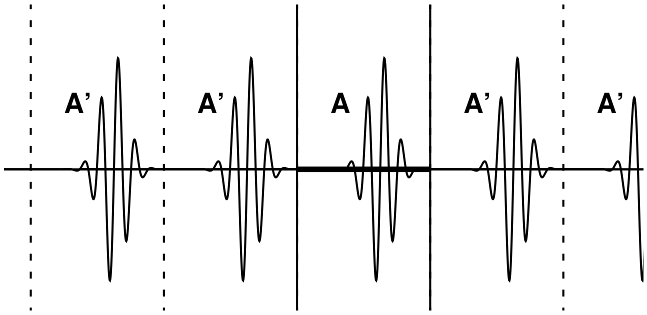


Figure 2.3: An illustration of one-dimensional periodic boundary conditions for a wave function. The super-cell is shown with the vertical solid lines, and the replicas of the wave function **A** with **A'**

2.3 Methods

2.3.1 Super-cell and Brillouin zone

When calculating the electronic structure atoms, molecules or clusters the object is isolated and does not interact with other particles. However, when studying crystals in one – *e. g.* wires, chains – two – surfaces, planes – or three dimensions the number of atoms is to be considered macroscopic along the continuous direction, leading to atom densities of 10^{23} 1/cm^3. Simulating this amount of particles is, a formidable task and a simplification needs to be applied. If the system is fully periodic an obvious choice is to use the periodic boundary conditions, *i. e.* the atomic and electronic structure is repeated along the three basis vectors $\mathbf{a}_{1,2,3}$ so that if there is an atom at position $\mathbf{R}_\mathrm{I}$ there are atoms also at $\mathbf{R}_\mathrm{I} + i\mathbf{a}_1 + j\mathbf{a}_2 + k\mathbf{a}_3$ where i, j and k are integers like shown in the Figure 2.2; likewise for the electronic wave functions, in the Figure 2.3 in a one-dimensional case.

The calculational saving comes if only the atoms and electrons inside the unit cell of the calculation – called the super-cell and shown in the Figure 2.2 with the emphasised line – needs to be explicitly considered and the replicas of these objects are incorporated implicitly. In the current calculations up to hundreds of atoms in the super-cell can be treated in the density functional theory and up to millions of atoms in classical molecular dynamics calculations where the electronic degrees of freedom are not explicitly included.

If non-homogeneities are introduced to the system the periodicity might be lost; this is the case *e.g.* for point defects in crystals, perpendicular to a surface, line defects *etc*. Four approaches have been widely used to deal with such systems:

- Green's function: This method provides the change in the electronic structure due to a perturbation; however, it is rather complicated and still contains approximations

- Embedding: Only the region with the perturbation is explicitly regarded and is smoothly joined to the unperturbed regions at the boundaries; this approach is also complicated and therefore has not been widely used

- Clusters: The region of interest and its immediate neighbourhood are modelled with a cluster; the quality of this approximation depends critically on the size of the cluster, but the size necessary for convergence is often believed to be too large making this approach un-practical

- Super-cell: The periodic boundary conditions are still preserved but the super-cell is chosen to be so large that the defects do not interact

An example of the last scheme is the slab geometry – sketched in the left-most panel of the Figure 2.4 – which is used in most of the calculations of surfaces nowadays: the system is treated periodically in all the three directions but perpendicular to the surface the replicas of the slab are isolated with a large enough region of vacuum in between them; also the thickness of the slab has to be carefully chosen to prevent artificial interactions between the two surfaces. The quality of the approximation depends on the system under study: if it is *e.g.* a periodic structure of adsorbates like in the middle panel of the Figure 2.4 the lateral ordering is correctly taken into account but if an isolated molecule or a reaction as in the right-most panel of the Figure 2.4 needs to be studied a large unit cell is required.

The reciprocal lattice of a periodic real space lattice with the three basis vectors $\mathbf{a}_{1,2,3}$ consists of the points $\mathbf{K}$ which yield

$$e^{\mathbf{i}\mathbf{K}\cdot\mathbf{R}} = 1 \; ; \tag{2.28}$$

$\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$ is a point in the real space lattice with integer coefficients $n_{1,2,3}$. Bloch's Theorem – please see *e.g.* Tinkham (1964) – states that the solutions of the one-electron Hamiltonian H can be chosen as

$$\psi_{i\mathbf{k}}(\mathbf{r}) = e^{\mathbf{i}\mathbf{k}\cdot\mathbf{r}} u_{i\mathbf{k}}(\mathbf{r}) \;\; , \tag{2.29}$$

where i is the band index, $\mathbf{k}$ is a point in the first Brillouin zone *i. e.* in the region of the reciprocal lattice closer to $\mathbf{k}= \mathbf{0}$ than any integer multiple $\mathbf{K} = m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3$ of the reciprocal lattice vectors and $u_{i\mathbf{k}}(\mathbf{r})$ is a function with the periodicity of the real space lattice so that

$$u_{i\mathbf{k}}(\mathbf{r}+\mathbf{R}) = u_{i\mathbf{k}}(\mathbf{r}) \;\; , \;\; \mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 \;\; . \tag{2.30}$$

From this form it follows that

$$\psi_{i\mathbf{k}}(\mathbf{r}+\mathbf{R}) = e^{\mathbf{i}\mathbf{k}\cdot\mathbf{R}} \psi_{i\mathbf{k}}(\mathbf{r}) \;\; , \tag{2.31}$$

or that the translation by a multiple of a lattice vector introduces a modulation of the phase of the wave function with the periodicity of the real space lattice, with equivalent densities:

$$\left|\psi_{i\mathbf{k}}(\mathbf{r}+\mathbf{R})\right|^2 = \left|\psi_{i\mathbf{k}}(\mathbf{r})\right|^2 \;\; . \tag{2.32}$$

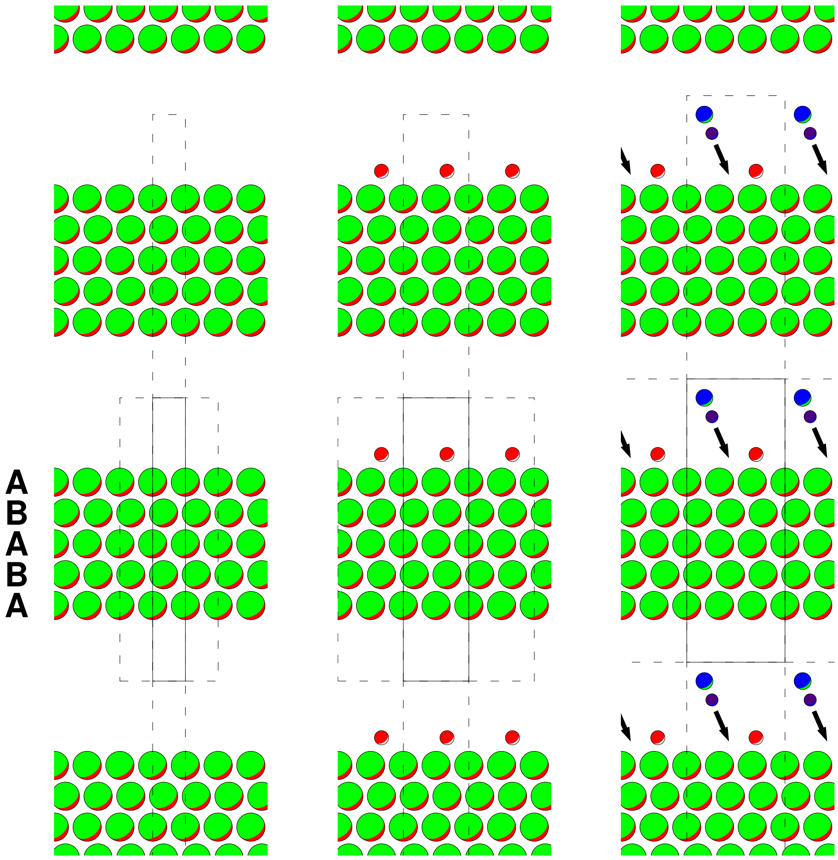


Figure 2.4: *Left* Side view of a slab geometry for a five-layer hcp slab; *Middle* The slab with adsorbates; *Right* Modelling surface reactions in the slab geometry

The attractivity of Bloch's Theorem can be understood as if the sum of all wave functions extending in the whole crystal is transformed into a – still infinite – set of wave functions of a plane wave times a periodic function; the quantities containing the full set of wave functions are transformed into integrals over the $\mathbf{k}$ points in the first Brillouin zone and for example the kinetic energy and density can thus be regarded as integrals

$$E_{\mathrm{kin,s}}\left[\{\psi\}\right] = \sum_i \int_{\mathbf{k}\in 1^{\mathrm{st}}\ \mathrm{BZ}} \int_{\mathbf{r}} \psi^*_{i\mathbf{k}}\left(\mathbf{r}\right)\left(-\frac{1}{2}\nabla^2\right)\psi_{i\mathbf{k}}\left(\mathbf{r}\right)\, d\mathbf{r}\, d\mathbf{k} \tag{2.33}$$

$$n\left[\{\psi\}\right](\mathbf{r}) = \sum_i \int_{\mathbf{k}\in 1^{\mathrm{st}}\ \mathrm{BZ}} \psi^*_{i\mathbf{k}}\left(\mathbf{r}\right)\psi_{i\mathbf{k}}\left(\mathbf{r}\right)\, d\mathbf{k}\ . \tag{2.34}$$

The advantage comes from the smooth behaviour of the integrands in these equations and thus the integrals can be transformed into sums over a finite set of $\mathbf{k}$ points like discussed in the next Section.

2.3.2 Brillouin zone summation

The sums in the Equations 2.33 and 2.34 contain an integral over the first Brillouin zone. In the general case this cannot be performed exactly but has to be approximated. Two approaches have been widely used:

1. The Brillouin zone is divided into sub-volumes and the energy dependence of the bands is approximated by Taylor expansion at the corner points. The mostly used approach is the linear tetrahedron method (Jepsen & Andersen, 1971; Lehmann & Taut, 1972), where the sub-

volumes are tetrahedra with one **k** point at each corner of them and the energy is expanded to linear order around these points. It turns out that the weights on the points depend only on the volume of the tetrahedra, not their shape. Recently Blöchl (1989; Blöchl, Jepsen & Andersen, 1994) introduced the second-order tetrahedron method

2. Special **k** point scheme; the integral is replaced with a finite, weighted sum over a set of discrete, sometimes called "representative" points. Nowadays this method is used more often than the tetrahedron method due to its simplicity, and the latter is now employed mainly for the calculation of the density of states. Earlier metals constituted a problem within the special point scheme due to the abrupt step in occupation numbers between occupied and unoccupied states and thus the approach was mainly used for semi-conductors and insulators only, but the introduction of the broadening of the occupation numbers – please see the Section 2.3.3 – has enabled also the treatment of metals within the special **k** point scheme. It has been used throughout the current work and will be described below

In the special **k** point scheme a finite, weighted sum substitutes the original, continuous integral, *i. e.*

$$\frac{1}{(2\pi)^3} \int_{\mathbf{k}} f(\mathbf{k})\ d\mathbf{k} \approx \frac{1}{\Omega} \sum_{\mathbf{k}_i} w_{\mathbf{k}_i} f(\mathbf{k}_i) \ , \tag{2.35}$$

where Ω the size of the super-cell, $\mathbf{k}_i$ form the set of **k** points and $w_{\mathbf{k}_i}$ are the corresponding weights chosen to weighten the contribution from the different **k** points correctly; the **k** points are chosen to span only the irreducible part of the Brillouin zone, defined as the smallest wedge of the first Brillouin zone which can be reduced with the symmetry operations, and the sum is normalised

$$\sum_{\mathbf{k}_i} w_{\mathbf{k}_i} = 1 \ . \tag{2.36}$$

Please note that for isolated systems like atoms in three, lines or chains in two and surfaces in one directions the electron system does not interact over the periodicity – provided the vacuum separating the systems is large enough – and therefore in these directions there cannot be any hybridisation of the wave functions and thus no dispersion. Thereby the integration 2.35 is exact already with a sampling of a single point, and *e.g.* the Γ point $(0, 0, 0)$ alone is enough for the summation over the Brillouin zone for an atom or molecule; in the case of surfaces the so-called surface Brillouin zone is two-dimensional with the periodicity corresponding the directions along the surface plane in the real space.

We shall explore the special **k** point scheme more thoroughly in the Appendix, where especially the case of Brillouin zone summation in the two-dimensional hexagonal cell, for which we have generalised a previous scheme, will be discussed.

2.3.3 Fractional occupation numbers

In defect-free insulators and semi-conductors there is an energy gap separating the occupied and unoccupied single-electron states; however, in metals the density of states is continuous across the Fermi energy and there are plenty of states above and below the Fermi energy. At zero temperature, where most of the calculations are performed, the occupation of the states would be a step function, and an excitation of infinitesimal energy would change the occupation numbers. Thus during the self-consistent iterations, when the wave functions and subsequently the density are not yet converged, the occupation numbers would change abruptly leading to instability, which is not present in the calculations for insulators and semi-conductors; also more **k** points are usually needed when calculating metals in order to approximate the integral over the Brillouin zone because of the accurate mapping of the Fermi surface, which is the iso-surface in the reciprocal space where the bands cross the Fermi energy. Traditionally two approaches have been used to prevent the oscillations:

- Interpolation of the density of states between the **k** points with the linear or quadratic tetrahedron method; please see also the Section 2.3.2
- Using fractional occupation numbers; the occupation numbers vary smoothly near the Fermi energy, thus smoothing the changes in the density and leading to a stabilised convergence to self-consistency

The tetrahedron method is still used sometimes nowadays for the densities of states, but in the self-consistent procedure the broadening schemes are employed more often. In the following we shall explain the broadening schemes used in the present work.

The use of fractional occupation numbers is an artificial method to improve the convergence of the self-consistent calculation; only the Fermi broadening scheme has any real counterpart, being the physical occupation at finite temperatures. This property has been used to generalise the density functional theory to non-zero electronic temperature by Mermin (1965). Usually the broadening is, however, much larger than the physical temperature and the ions are at zero temperature anyway; thereby the broadening introduces an approximations which has to be subsequently corrected if possible. The corrections are called *extrapolation to zero temperature* and exist explicitly for the energy (Gillan, 1989; Kresse & Furthmüller, 1996a) in the first order but no straight-forward method exist for the forces. In the case of the Fermi broadening the occupation number of the state i at the **k** point **k** is given by

$$f_{i\mathbf{k}}(x) = \frac{1}{e^x + 1} , \tag{2.37}$$

where $x = (\varepsilon_{i\mathbf{k}} - \varepsilon_{\mathrm{F}})/\sigma$; ε_{F} is the Fermi energy, determined by requiring the correct amount of electrons, and σ is the width or temperature in the broadening. The free energy, which includes the artificial entropy from the electronic degrees of freedom

$$S(f_{i\mathbf{k}}) = -\left[f_{i\mathbf{k}} \ln f_{i\mathbf{k}} + (1 - f_{i\mathbf{k}}) \ln (1 - f_{i\mathbf{k}})\right] , \tag{2.38}$$

is

$$F = E + \sum_{i\mathbf{k}} \sigma S(x) . \tag{2.39}$$

The energy extrapolated to zero temperature is in the case of the Fermi broadening given by

$$E_{\sigma=0} \approx \frac{1}{2}\left[F(\sigma) + E(\sigma)\right] . \tag{2.40}$$

Other methods for the broadening are the Gaussian and Methfessel-Paxton (1989) broadenings. The scheme of Gaussian broadening is a special case – the zeroth order function – of the Methfessel-Paxton scheme; these two usually need no correction to the forces. The latter leads, however, to occupation number slightly larger than one and less than zero.

The stability in the initial phase of the calculation or after an ionic movement is often rather poor since the electronic structure changes still remarkably and thus too large variations in the density need to be avoided. If the eigenvalues near the Fermi energy vary a lot the changes in the occupation numbers can be limited using the damping of the eigenvalues introduced by Stumpf (1993; Stumpf & Scheffler, 1994); as the self-consistent solution is approached the occupation numbers approach their non-damped values and this scheme leads to no additional error.

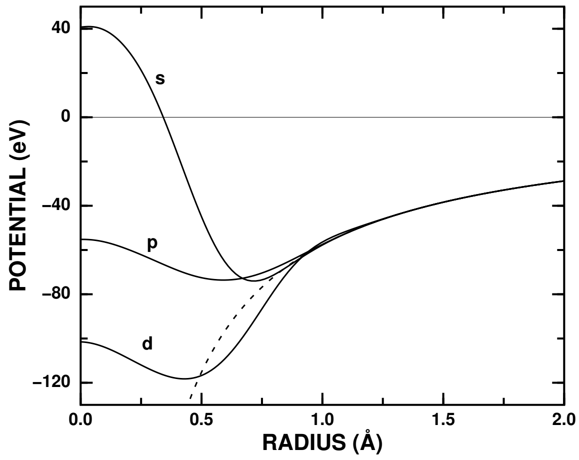


Figure 2.5: An example of a pseudo potential: A pseudo potential for silicon; the dashed line is the hard-core potential -Z_v/r

2.3.4 Pseudo potentials

Because deep in energy, highly localised and thus not available for bonding, the core electrons do not usually contribute significantly to the changes in the total energy or density in different chemical environments. Therefore their inclusion explicitly in a calculation is not necessary if the potential screening the nuclear contribution $-Z/r$ arising from the core electrons and the orthogonalisation of valence electrons against the core orbitals can be arranged implicitly. This is the goal of pseudo potentials; an excellent review about the pseudo potential concept and its development has been given by Pickett (1989). The modern pseudo potentials, created usually using the neutral, free atom as the reference state, are able to reproduce accurately the valence orbitals in the chemically significant region close to the states in an all-electron calculation, *i. e.* they are *transferable* into different chemical environments.

An example of a pseudo potential is shown in the Figure 2.5: it shows a pseudo potential for a silicon atom, whose valence configuration is $3s^2 4p^2$. The potentials for different l components are different, but all approach the asymptotic limit $-Z_v/r = -4/r$ beyond the core radii inside which the potentials and pseudo wave functions are allowed to differ from their all-electron counterparts. The pseudo wave functions of the valence orbitals, *i. e.* the ground state of the Kohn-Sham equation of that angular momentum channel l with the pseudo potential as the external potential, are node-less by construction in order to make them as smooth as possible; therefore a smaller basis set is sufficient to represent the valence orbitals, which will be close to the pseudo wave function at the core region.

The break-through in transferable pseudo potentials for the density functional calculations came with the scheme of Hamann, Schlüter and Chiang (1979): they devised a straight-forward procedure to generate transferable pseudo potentials which was followed by Bachelet, Hamann and Schlüter (1982) to generate a list of pseudo potentials going through most of the Periodic Table. In their scheme Hamann, Schlüter and Chiang required the following properties:

- The ground state eigenvalues of the pseudo and all-electron potentials shall be equivalent in the reference configuration

- The pseudo and all-electron wave functions match out-side the core radius

- The norm of the corresponding wave functions

$$4\pi \int_{r=0}^{r_c} |\phi(r)|^2 r^2 \, dr \tag{2.41}$$

is the same for the pseudo and all-electron orbital inside the core radius r_c; this norm-conservation guarantees that the electro-static potential felt by the pseudo valence function out-side the core radius is identical to the all-electron one. The norm-conservation further ensures that the logarithmic derivative of the pseudo and all-electron wave function and their first energy derivatives agree out-side the core radius; this is important for the scattering properties of the electrons and thus for the over-all transferability of the pseudo potential.

Additionally the pseudo wave function and potential are constructed to be continuous up to some given derivative over the core radius. Hamann, Schlüter and Chiang used functions inside the core radii which decayed smoothly towards the all-electron counterparts and thus are not yet identical at the core radius; this might lead to errors in some cases (Martin Fuchs, personal communication).

The scheme of Hamann, Schlüter and Chiang was generalised by Hamann (1989) to allow the construction of a pseudo potential for the un-occupied or un-bound channels; Bachelet, Hamann and Schlüter used in their list of pseudo potentials a positively ionised atom as the configuration state for the potentials corresponding to those channels and neutral for the occupied ones, but Hamann's scheme allows one to generate a pseudo potential for all the channels in the neutral ground state.

The pseudo potentials generated using the concepts of Hamann, Schlüter and Chiang have the unpleasant feature that they are often very *hard*, *i.e.* they have rather steep maxima like in the Figure 2.6 which shows the pseudo potential for ruthenium. This sets higher requirements for the basis set used in the calculation for the solid; especially for the plane wave basis which will be introduced later this means a large increase in basis set size. The hard pseudo potentials occur mostly for the right-most second row and transition metal elements, where the 2*p* and 3*d*, 4*d* or 5*d* orbitals are very localised, especially the 2*p* and 3*d* where there is no orbital lower in energy with the same angular momentum to against which the orbital should be orthogonal and thus the orbital can remain more localised. To overcome the problem with too hard pseudo potentials new schemes were devised of which the ones by Kerker (1980), Vanderbilt (1985), Rappe *et al* (1990) and Troullier and Martins (1991) are examples; the Figure 2.6 shows also the 4*d* pseudo orbital of ruthenium generated with the scheme of Troullier and Martins, and the dip in the Hamann-Schlüter-Chiang is smoothed away. The smoother, *soft* pseudo potential are achieved by increasing the core radii, thus allowing for a larger freedom in the generation – however, the pseudo potential and wave function match now exactly the all-electron counterparts at the core radius without a decaying difference function –, by choosing a form which leads to a softer potential and in some cases setting a *penalty function* which affects the generation of the pseudo potential, trying to maximise the smoothness. The result is a large reduction in the plane wave basis set as shown in the Figure 2.7, but usually the transferability of the pseudo potentials generated according to Hamann, Schlüter and Chiang is better due to the smaller core radii used.

A more radical approach to increase the smoothness was invented by Vanderbilt (1990) and first implemented by Kari Laasonen (Laasonen *et al*, 1993). The Vanderbilt pseudo potentials are not norm-conserving in the sense that the pseudo wave function does not have to have the same norm inside the core radius as the all-electron wave function as seen in the Figure 2.8; the pseudo wave function may thus avoid the peak in the orbital – in this case the 4*d* – and thus a much

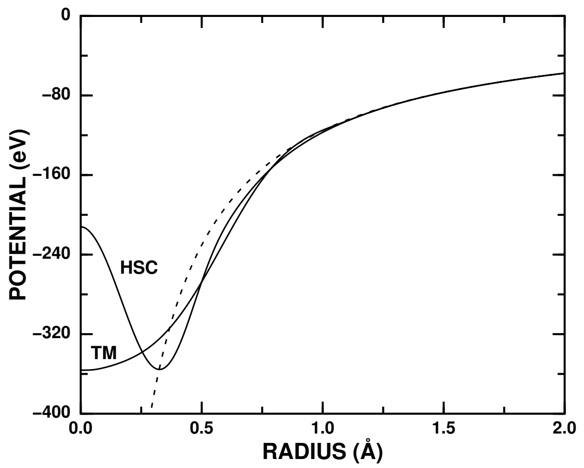


Figure 2.6: An example of a hard Hamann-Schlüter-Chiang and a soft Troullier-Martins pseudo potential for the 4*d* state of ruthenium; the dashed line is the hard-core potential -Z_v/r

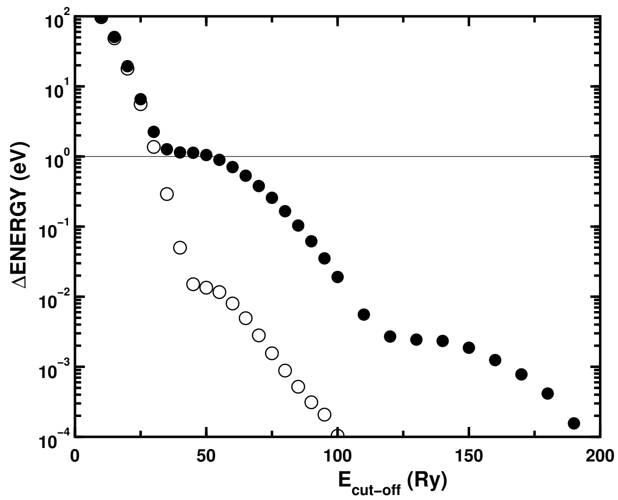


Figure 2.7: The convergence of the total energy with the plane wave basis size for different pseudo potentials in bulk platinium; a Hamann – filled circles – and a Troullier-Martins – empty circles – pseudo potential was used

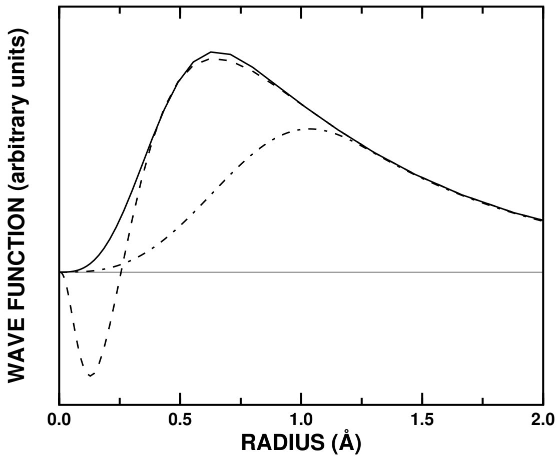


Figure 2.8: An example of the pseudo wave function – times radius r – of the Troullier-Martins – solid line – and Vanderbilt ultra-soft – dot-dashed line – pseudo potential for the 4*d* orbital of ruthenium. The all-electron wave function is shown with a dashed line

softer pseudo wave function results. The charge missing from the square of the wave functions is later augmented with the missing charge. The complications arising from the charge augmentation partially diminishes the advantage gained from the lower basis set size needed as evidenced in the Figure 2.7 but still in most cases with very localised orbitals the Vanderbilt construction would save computational cost; the main obstacle for using it is the additional programming required.

The pseudo potential concept relies on the frozen-core approximation, *i. e.* that the core orbitals do not change significantly in different chemical environments. On the other hand the pseudo potential must ensure that the valence orbitals and in the reference state unoccupied orbitals are at least reasonably well reproduced since in a bonding situation these might become partially occupied or the orbitals gain some character with their angular momentum. If there are, however, energetically not much more bound states than the valence orbitals – called *semi-core* states – they may easily at least polarise and should thus be regarded as valence states; thus the same pseudo potential angular momentum channel should describe both the semi-core and unoccupied state although they are energetically rather far separated. A problematic example is cesium where the 4*f* states affect the screening of 5*s* and 5*p* electrons, and the originally unoccupied 6*p* is not much higher in energy and becomes thus partially occupied in the calculation for the solid.

A convenient form to calculate the action of the pseudo potential to the orbitals $\psi_{i\mathbf{k}}$ would be fully *non-local* around the atom I:

$$V^I_{\mathrm{ps}}(\mathbf{r}) = V_{\mathrm{local}}(\mathbf{r}-\mathbf{R}_I) + \sum_{lm} |\eta_l(|\mathbf{r}-\mathbf{R}_I|)\rangle \left|lm\left(\widehat{\mathbf{r}-\mathbf{R}_I}\right)\right\rangle \left\langle lm\left(\widehat{\mathbf{r}'-\mathbf{R}_I}\right)\right| \langle \eta_l(|\mathbf{r}'-\mathbf{R}_I|)| \ , \tag{2.42}$$

since then the matrix elements become – the ionic coordinate $\mathbf{R}_I$ has been shifted to the origin for clarity –

$$\begin{aligned} \langle\chi_1(\mathbf{r})|\, V_{\mathrm{ps,nl}}\, |\chi_2(\mathbf{r}')\rangle &= \langle\chi_1(\mathbf{r})|V_{\mathrm{local}}(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')|\chi_2(\mathbf{r}')\rangle \\ &+ \sum_{lm} \left[\langle\chi_1(\mathbf{r})|\eta_l(|\mathbf{r}|)\rangle|lm(\widehat{\mathbf{r}})\rangle\right] \left[\langle lm(\widehat{\mathbf{r}'})|\langle\eta_l(|\mathbf{r}'|)|\chi_2(\mathbf{r}')\rangle\right] \end{aligned} \tag{2.43}$$

now the inner products $\langle\chi_1(\mathbf{r})|\eta_l(|\mathbf{r}|)\rangle|lm(\widehat{\mathbf{r}})\rangle$ can be calculated independently from $\langle lm(\widehat{\mathbf{r}'})|\langle\eta_l(|\mathbf{r}'|)|\chi_2(\mathbf{r}')\rangle$ – they are naturally the complex conjugates of each other – and thus the calculation scales only linearly with the number of basis functions. This is in contrast to the so-called semi-local form, where a matrix element involve integrals over both χ_1 and χ_2. This kind of construction was achieved by Kleinman and Bylander (1982) and is in use in most of the pseudo potential calculations. The pseudo potential term of Kleinman and Bylander reads as

$$\begin{aligned} V_{\mathrm{ps}}(\mathbf{r}) &= V_{\mathrm{local}}(\mathbf{r}) + \sum_{lm} \left| \Delta V_{\mathrm{ps}}^{l}(r)\, \phi_l(r) \right\rangle |lm\rangle\, E_{\mathrm{KB}}^{l}\, \langle lm| \left\langle \Delta V_{\mathrm{ps}}^{l}(r')\, \phi_l(r') \right| \\ E_{\mathrm{KB}} &= \frac{1}{\left\langle \phi_l(r)\, \Delta V_{\mathrm{ps}}^{l}(r)\, \phi_l(r) \right\rangle}\,. \end{aligned} \tag{2.44}$$

In addition to the calculational saving the Kleinman-Bylander construction unfortunately brings a disadvantage: it can lead to wrong ordering of the eigensolutions of the pseudo potential, *i.e.* there is a state – called a *ghost* state – lower in energy than the node-less wave function of reference state; this has to have nodes in order to be orthogonal to the wave function of the reference state. The origin of and ways to avoid the ghost states has been studied by Gonze and co-workers (Gonze, Käckell & Scheffler, 1990; Gonze, Stumpf & Scheffler, 1991), and generalised to the pseudo potentials with multiple projectors with the same angular quantum number by Khein (1995). In the semi-local form a corollary of the Wronskian theorem (Messiah, 1974) prevents the ghost states but this theorem is not valid in the fully separable Kleinman-Bylander form. The existence of a ghost state is disastrous in the calculation for a solid and therefore they must not exist in the pseudo potentials; a careful construction of the pseudo potential and choice of the local component enable the use of the Kleinman-Bylander construction free of ghost states for about all the elements of the Periodic Table (Stumpf, Gonze & Scheffler, 1990). Furthermore, there are now straight-forward tests which indicate the existence of ghost states already before the pseudo potential has been used in a calculation for a solid. However, the fully non-local form has always a worse quality than the semi-local form and although there would not exist any real ghost states sometimes a state with node is in energy close above the reference state, called a *pseudo* ghost state, so although it is not classified as a real ghost it affects the properties calculated using the pseudo potential.

A good transferability for a pseudo potential is achieved if the pseudo potential reproduces well the results which would be obtained from an all-electron calculation. A failure in the logarithmic derivatives, which were earlier used to justify the quality of a pseudo potential, is indicative of serious problems but not a real test for the transferability. As another indicator of the transferability the change of the properties of the pseudo potential and an all-electron atom can be compared when the system is perturbed from its reference state. Such a test – called the *hardness* test due to its similarity to the hardness/softness concept in chemistry – is to follow the change in total energy and Kohn-Sham eigenvalues when the occupation numbers are changed. Teter (1993; Filippetti *et al*, 1997) suggested using the hardness as an ingredient in the process of creating the pseudo potential in order to guarantee a high transferability. The traditional definition of hardness does consider only spherical perturbations and therefore does not fully probe the transferability of a pseudo potential; an extension to non-spherical perturbations was discussed by Filippetti *et al* (1995).

In summary, the pseudo potential concept consists of the following approximations:

- Frozen-core approximation: the core electrons are not allowed to change in different environments
- Pseudisation: the potential seen by the valence electrons has been modified, and they no longer contain the correct radial nodes and are not explicitly orthogonalised against the core electrons

- Linearisation: the exchange-correlation functional is not linear in the un-screening process; the error arising from this can, however, be partially or fully avoided by using the pseudo core or true core electron density (Louie, Froyen & Cohen, 1982)
- Separation to non-local form: this is not always done but in most calculations the Kleinman-Bylander construction is used

In the current work we have used the Troullier-Martins pseudo potentials since we are mainly investigating solids with 4*d* and 2*p* elements. The potentials were generated using the fully relativistic generator written by José-Luis Martins (Troullier & Martins, 1991) and the program `psgen` in the package `gncpp` – for *Generalised Norm-Conserving Pseudo Potentials* – maintained by Martin Fuchs (Martin Fuchs, personal communication; Fuchs & Scheffler, 1998) based on the code by Hamann (1989). The potentials were tested against ghost states with the program `pswatch` in `gncpp`, and the transferability tests were performed in different atomic electronic states with the scripts provided in `gncpp`. Kleinman-Bylander form was used for all the atoms except the few Vanderbilt pseudo potentials employed during this work.

2.4 Implementation

Before the Kohn-Sham equations can be solved numerically a few approximations must be introduced. We will in the following discuss some of the major issues used to achieve the computationally attainable approach of the present work.

2.4.1 Basis sets — plane waves

Whereas the Kohn-Sham equations are continuous the computers work with a discrete and finite set of numbers, a task which is achieved employing a limited set of basis functions, called the basis set. Thus the *i*th state at the point $\mathbf{k}$ in the Brillouin zone is represented as

$$\psi_{i\mathbf{k}}(\mathbf{r}) = \sum_{n=1}^{N_{\text{basis}}} c_{i\mathbf{k},n}\,\chi_n(\mathbf{r}) \ , \tag{2.45}$$

where n goes through the basis functions $\chi_n(\mathbf{r})$ up to a finite count N_{basis}, and $c_{i\mathbf{k},n}$ are the expansion coefficient, *i. e.* plain numbers. The idea is that the operations acting on χ_n needed to solve the Kohn-Sham equations can be calculated, either analytically or numerically, and thus a straightforward implementation leads to the equations

$$\sum_n c_{i\mathbf{k},n} H \chi_n(\mathbf{r}) = \varepsilon_i \sum_n c_{i\mathbf{k},n} \chi_n(\mathbf{r}) \ . \tag{2.46}$$

Multiplying with χ_m^* from the left and integrating over the space these equations become

$$\sum_n c_{i\mathbf{k},n} H_{mn} = \varepsilon_i \sum_n c_{i\mathbf{k},n} \mathcal{O}_{mn} \ , \qquad H_{mn} = \langle \chi_m | H | \chi_n \rangle \ , \qquad \mathcal{O}_{mn} = \langle \chi_m | \chi_n \rangle \ . \tag{2.47}$$

These equations can be expressed in a matrix form, noting that the $\mathbf{k}$ points are independent, as

$$\underline{\mathbf{Hc}} = \underline{\varepsilon \mathcal{O} \mathbf{c}} \ , \qquad \underline{\mathbf{H}}_{mn} = H_{mn} \ , \qquad \underline{\mathbf{c}}_{ni} = c_{i\mathbf{k},n} \ , \qquad \underline{\mathbf{O}}_{mn} = \mathcal{O}_{mn} \ . \tag{2.48}$$

This is a generalised – due to the over-lap matrix $\mathcal{O}$ – matrix eigenvalue problem which can be solved by a direct diagonalisation of the matrix $\underline{\mathbf{H}}$. However, the diagonalisation scales in the third power and the storage needed in the square of the matrix size – in this case the matrix size is given by the number of basis functions N_{basis} – and an iterative method avoiding the direct diagonalisation will become advantageous in large systems; we will return to this later when discussing the Car-Parrinello method.

The following list governs some of the different kind of basis set functions in use nowadays.

- "Basis-set free" methods: logarithmic meshes, an equidistance grid; these can, however, also be regarded as basis sets with δ-functions as the basis functions
- Atom-centred functions: Gaussians, atomic orbitals, mixed basis, ...
- Equidistance mesh and adaptions: wavelets, plane waves, ...

The choice of the optimal basis set for the system at hand depends on the system and problem at hand.

One of the most widely used basis set is the plane wave basis. This is because of its simplicity and high efficiency, although the latter advantage is diminished by the large number of plane waves needed. The plane wave is a periodic function and is as such suitable as $u(\mathbf{r})$ in the Bloch theorem; thereby the Kohn-Sham state can be expressed as

$$\psi_{i\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}}^{\mathbf{G}_{\max}} c_{i\mathbf{k}}(\mathbf{G})\, e^{\mathbf{i}(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \ , \tag{2.49}$$

where the set of $\mathbf{G}$ vectors is limited to those which have the periodicity of the real space unit cell; thus

$$\mathbf{G} = i\mathbf{b}_1 + j\mathbf{b}_2 + k\mathbf{b}_3 \tag{2.50}$$

with $\mathbf{b}$'s being the reciprocal lattice vectors, *i. e.* the $\mathbf{G}$ vectors are simply integer multiples of the $\mathbf{b}$'s. The plane waves are also often written in the bra-ket notation as

$$e^{\mathbf{i}(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} = |\mathbf{k}+\mathbf{G}\rangle \quad , \qquad \left[e^{\mathbf{i}(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}\right]^* = e^{-\mathbf{i}(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} = \langle\mathbf{k}+\mathbf{G}| \ . \tag{2.51}$$

The plane waves have the following useful properties:

- They are orthogonal,[1] *i. e.*

  $$\langle\mathbf{g}|\,\mathbf{g}'\rangle = \int_{\mathbf{r}} e^{-\mathbf{i}\mathbf{g}\cdot\mathbf{r}} e^{\mathbf{i}\mathbf{g}'\cdot\mathbf{r}}\, d\mathbf{r} = \delta_{\mathbf{g}\mathbf{g}'} \ ; \tag{2.52}$$

  thereby the overlap matrix in the Equation 2.47 becomes diagonal unit matrix and therefore vanishes from the eigenvalue problem.
- The plane wave set is complete, *i. e.* any function can be unambiguously represented using them
- The discrete Fourier transform can be used to transfer functions from real to reciprocal space; especially beneficial is the Fast Fourier Transform – FFT – variant of the discrete Fourier transform; this will be discussed below

Some terms of the Kohn-Sham energy and potential are easily evaluated using the plane wave basis in the reciprocal space – one example is the kinetic energy

$$E_{\text{kin}} = \sum_{i\mathbf{k}} f_{i\mathbf{k}} w_{\mathbf{k}} \int_{\mathbf{r}} \psi^*_{i\mathbf{k}}(\mathbf{r}) \left(-\frac{1}{2}\nabla^2\right) \psi_{i\mathbf{k}}(\mathbf{r})\, d\mathbf{r} = \sum_{i\mathbf{k}} f_{i\mathbf{k}} w_{\mathbf{k}} \sum_{\mathbf{G}}^{N_{\mathbf{G}}} |\mathbf{k}+\mathbf{G}|^2 \, |c_{i\mathbf{k}}(\mathbf{G})|^2 \tag{2.53}$$

[1] It is common to leave out the index $\mathbf{k}$ from $\mathbf{k}+\mathbf{G}$ and think of $\mathbf{G}$ as a generalised wave vector index no longer as discrete multiples of the reciprocal lattice vectors; in such context we shall use lowercase $\mathbf{g}$ to denote them

– since the expression is *diagonal* in the reciprocal space like the matrix element of the Hamiltonian,

$$\langle \mathbf{k}+\mathbf{G}|\,\hat{T}\,|\mathbf{k}+\mathbf{G}'\rangle = \underbrace{\left.\begin{pmatrix} |\mathbf{k}+\mathbf{G}_0|^2 & 0 & 0 & \dots & 0 \\ 0 & |\mathbf{k}+\mathbf{G}_1|^2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & 0 & \dots & 0 & |\mathbf{k}+\mathbf{G}_{N_\mathbf{G}}|^2 \end{pmatrix}\right\}\mathbf{G}}_{\mathbf{G}'} \,, \tag{2.54}$$

i. e. it contains only a single summation over $\mathbf{G}$ and not over both $\mathbf{G}$ and $\mathbf{G}'$; it could be also written as include the sum

$$\sum_{\mathbf{G}\mathbf{G}'} \delta_{\mathbf{G}\mathbf{G}'} \,. \tag{2.55}$$

The early approach until the early 1980's was to calculate all the terms in the reciprocal space but *e. g.* the energy from the local part of the pseudo potential becomes a convolution

$$E_{\text{ps,local}} = \int_\mathbf{r} n(\mathbf{r})\, V_{\text{ps,local}}(\mathbf{r})\, d\mathbf{r} = \sum_{\mathbf{G}\mathbf{G}'} n^*(\mathbf{G})\, V_{\text{ps,local}}(\mathbf{G}-\mathbf{G}') \,, \tag{2.56}$$

where $n(\mathbf{G})$ and $V_{\text{ps,local}}(\mathbf{G})$ are the Fourier transforms of the density and local part of the pseudo potential into the reciprocal space, respectively. Now this term contains a double sum and is very costly to calculate, and furthermore scales in the square power of the number of the plane waves whereas in the real space the operation scales linearly. It was thanks to Roberto Car and Michele Parrinello (1985) to take advantage of the FFT in this context; namely, they used the fact that the convolutions in the reciprocal space are diagonal in the real space and vice versa. Thus one calculates the terms diagonal in the space where the term is diagonal and uses FFT to transform between the two spaces when needed. The FFT scales only as $N_\mathbf{G} \ln(N_\mathbf{G})$ with the number of plane waves and thus quickly leads to large savings in the computational time.

It is interesting to further notice that an inner product can be calculated as easily in the real as in the reciprocal space since

$$\begin{aligned} \langle f(\mathbf{r})\,|g(\mathbf{r})\rangle_\mathbf{r} &= \int_\mathbf{r} f^*(\mathbf{r})\, g(\mathbf{r})\, d\mathbf{r} = \int_{\mathbf{r}'}\int_\mathbf{r} f^*(\mathbf{r}) \underbrace{\int_\mathbf{g} e^{-i\mathbf{g}\cdot\mathbf{r}} e^{i\mathbf{g}\cdot\mathbf{r}'}\, d\mathbf{g}}_{\delta(\mathbf{r}-\mathbf{r}')}\, g(\mathbf{r}')\, d\mathbf{r}'\, d\mathbf{r} \\ &= \int_\mathbf{g} f^*(\mathbf{g})\, g(\mathbf{g})\, d\mathbf{g} = \langle f(\mathbf{g})\,|g(\mathbf{g})\rangle_\mathbf{g} \end{aligned} \tag{2.57}$$

The number of plane waves is typically limited in the calculations using a so-called cut-off energy E_{cut}, which is defined as the kinetic energy of the largest frequency or longest $\mathbf{G}$ vector in the expansion:

$$\frac{1}{2}|\mathbf{k}+\mathbf{G}|^2 \leq E_{\text{cut}} \tag{2.58}$$

at each $\mathbf{k}$. Thus the cut-off energy defines a sphere – illustrated in the Figure 2.9 in two dimensions – in the reciprocal space which contains all the $\mathbf{G}$ vectors. The cut-off energy is mainly set by the density close to the nuclei even when employing pseudo potentials, and thus depends on the material in use: the more localised states it contains, the higher the cut-off required. Especially demanding for a high cut-off are hydrogen, the 2*p*, 3*d* materials, and always the materials where semi-core states have to be included in the plane wave expansion. Thus especially the materials where a new *l* channel is only partially occupied are difficult since it is more localised than in the elements below them in the Periodic Table where the orthogonalisation against the lower-lying orbital with the same *l* quantum number forces the orbital to be more de-localised. The cut-off energy

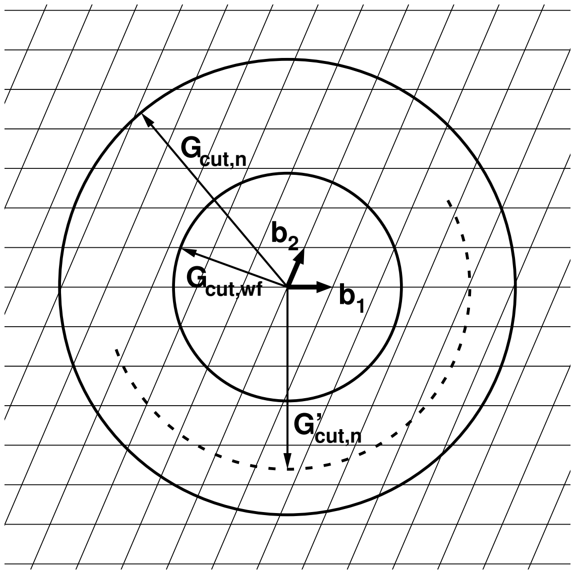


Figure 2.9: The construction of **G** vectors in two-dimensions. $\mathbf{b}_1$ and $\mathbf{b}_2$ are the reciprocal lattice vectors and all the crossings of the line give a **G** vector. The cut-off spheres for the wave functions $G_{\text{cut,wf}}$ and density $G_{\text{cut,n}} = 2G_{\text{cut,wf}}$ as well as for density with duelling $G_{\text{cut,n}} < 2G_{\text{cut,wf}}$ are shown, the latter partially, with a dashed line. Please note that $\mathbf{k} = 0$ was used for the cut-off sphere of the wave functions

is traditionally expressed in Rydberg units; this is convenient since the square of the maximum length of the **G** vectors is the same number in Hartree's.

As can be deduced from the definition of the cut-off energy the number of plane waves inside the cut-off sphere scales linearly with the volume of the system as for example doubling the volume would correspond to halving one of the reciprocal lattice vectors **b** and thus doubling the density of **G** vectors within the cut-off sphere. This holds also for increasing the empty vacuum in *e. g.* cluster or surface systems; this is not the case for the basis sets whose basis functions are localised at the atoms. This is one of the main reasons which lead Francois Gygi (1992) to apply the method of adaptive coordinates to the electronic structure calculations; there the problem is largely intervened by having grid points only in the physically relevant area. Another often appearing disadvantage of plane waves is that they always have the periodicity of the super cell used, even for clusters and surfaces parallel to the surface; then the only chance to avoid artificial interactions is to separate the objects with enough space to decouple the electron densities.

The density from the Kohn-Sham orbitals is easily calculated real space, and it is diagonal in the **k** space:

$$n(\mathbf{r}) = \sum_{i\mathbf{k}} f_{i\mathbf{k}} w_{\mathbf{k}} \left|\psi_{i\mathbf{k}}(\mathbf{r})\right|^2 \quad . \tag{2.59}$$

Because the density no longer contains the **k** points the Fourier transform of the density contains only reciprocal vectors **G** but now to two times larger wave vectors as easily demonstrated by investigating the square of function $\cos x + \sin x$:

$$(\cos x + \sin x)^2 = 1 + \sin(2x) \quad , \tag{2.60}$$

i. e. the oscillations of a square of a function can contain frequencies twice higher than the function itself. Similarly to the density all the potential terms except the non-local part of the pseudo potential contain wave vectors on up to $|\mathbf{G}_{\text{cut,n}}| = 2|\mathbf{G}_{\text{cut,wf}}|$, and thus $E_{\text{cut,n}} = 4E_{\text{cut,wf}}$. The factor 4 can usually be reduced from its ideal value to about 2.5. . . 3 – a trick to save some computer time called dualling (Remler & Madden, 1990) and depicted in the Figure 2.9 – when employing the local density approximation but should be not reduced at least as much when using more complicated exchange-correlation functionals like *e. g.* the generalised gradient approximation because these require a higher precision due to their higher non-linearity. Since all the operations on wave functions involve also the density or local potential all the FFT's have to performed in the grid which comprises all the **G** vectors of the density, *i. e.* up to $G_{\mathrm{cut,n}}$. Therefore the duelling is very efficient way to save computational time with small loss of accuracy: the wave functions are stored accurately, only the length of the FFT's is reduced; the errors are in the density and potentials.

An interesting property of the FFT is that it is exact on the grid points, *i. e.* a function in the reciprocal space is correctly transformed to the real space grid points and vice versa – the finite cut-off causes errors only in between the grid points in real space and in wave vectors beyond the cut-off energy. In principle one can use a finer FFT grid than required by the cut-off energy but as soon as the back-transformed functions are used in the reciprocal space the advantage is lost. A similar trick of "double grid", when executed consistently – like often done when employing the Vanderbilt pseudo potentials (Laasonen *et al*, 1993) – , leads to an improved accuracy, though.

A three-dimensional FFT consists of three subsequent one-dimensional FFT's along each of the lattice vectors; this is visualised in the Figure 2.10. It is easy to understand this if one looks at the definition of the discrete Fourier transform:

$$\begin{aligned}\mathcal{F}\left[f\left(\mathbf{r}\right)\right]\left(\mathbf{g}\right) &= f\left(\mathbf{g}\right)=\sum_{\mathbf{r}} f\left(\mathbf{r}\right)e^{\mathbf{i}\mathbf{g}\cdot\mathbf{r}} \\ &= \sum_{x,y,z} f\left(x,y,z\right)e^{\mathrm{i}\left(g_x x+g_y y+g_z z\right)}=\sum_z\left\{\sum_y\left[\sum_x f\left(x,y,z\right)e^{\mathrm{i}g_x x}\right]e^{\mathrm{i}g_y y}\right\}e^{\mathrm{i}g_z z} \qquad (2.61)\end{aligned}$$

and similarly for a non-rectangular case. This can be made use of in the parallelisation of the FFT as will be discussed below.

Because the radius of the cut-off sphere for the wave functions is only half of the radius for the density and the FFT's are three-dimensional the computational time can be slightly reduced by in the first two dimensions neglecting the transforms which contain only regions with zeroes; this is illustrated in the Figure 2.11.

One very useful benefit from working in the reciprocal space is that the calculation of atom-centred properties becomes very efficient and continuous with respect to a move of the atoms. This is because if a function f depends on the position $\mathbf{R}_{\mathrm{I}}$ of the ion I then

$$\begin{aligned}\mathcal{F}\left[f\left(\mathbf{r}-\mathbf{R}_{\mathrm{I}}\right)\right]\left(\mathbf{g}\right) &= \int_{\mathbf{r}} f\left(\mathbf{r}-\mathbf{R}_{\mathrm{I}}\right)e^{\mathbf{i}\mathbf{g}\cdot\mathbf{r}}\,d\mathbf{r}=\int_{\mathbf{r}'} f\left(\mathbf{r}'\right)e^{\mathbf{i}\mathbf{g}\cdot\left(\mathbf{r}'+\mathbf{R}_{\mathrm{I}}\right)}\,d\mathbf{r}' \\ &= e^{\mathbf{i}\mathbf{g}\cdot\mathbf{R}_{\mathrm{I}}}\int_{\mathbf{r}'} f\left(\mathbf{r}'\right)e^{\mathbf{i}\mathbf{g}\cdot\mathbf{r}'}\,d\mathbf{r}'=e^{\mathbf{i}\mathbf{g}\cdot\mathbf{R}_{\mathrm{I}}}\mathcal{F}\left[f\left(\mathbf{r}\right)\right]\ , \qquad (2.62)\end{aligned}$$

i. e. the atomic position enters only as a phase factor; in the real space an interpolation to the grid points would be necessary. Also the gradients with respect to **r** and ionic positions $\mathbf{R}_{\mathrm{I}}$ are trivial:

$$\begin{aligned}\frac{\partial}{\partial\mathbf{r}}f\left(\mathbf{r}\right) &= \sum_{\mathbf{g}}\mathbf{i}\mathbf{g}\,f\left(\mathbf{g}\right)e^{\mathbf{i}\mathbf{g}\cdot\mathbf{r}}=\mathcal{F}^{-1}\left[\mathbf{i}\mathbf{g}\,f\left(\mathbf{g}\right)\right] \\ \frac{\partial}{\partial\mathbf{R}_{\mathrm{I}}}f\left(\mathbf{r}-\mathbf{R}_{\mathrm{I}}\right) &= \sum_{\mathbf{g}}\mathbf{i}\mathbf{g}\,e^{\mathbf{i}\mathbf{g}\cdot\mathbf{R}_{\mathrm{I}}}f\left(\mathbf{g}\right)e^{\mathbf{i}\mathbf{g}\cdot\mathbf{r}}=\mathcal{F}^{-1}\left[\mathbf{i}\mathbf{g}\,e^{\mathbf{i}\mathbf{g}\cdot\mathbf{R}_{\mathrm{I}}}f\left(\mathbf{g}\right)\right]\ . \qquad (2.63)\end{aligned}$$

The latter equation guarantees also a very convenient way of calculating the forces acting on the ions. The evaluation of the forces is extremely trivial because the basis set does not depend on the

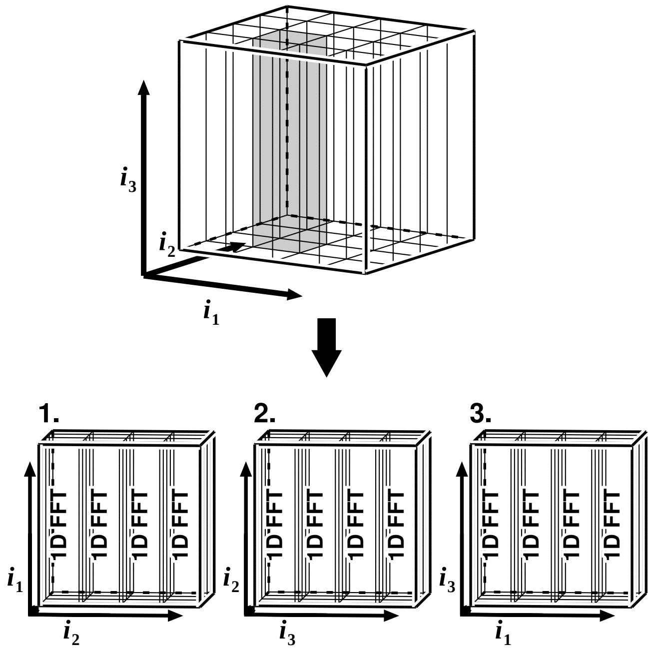


Figure 2.10: The three-dimensional fast Fourier transform as three subsequent one-dimensional transforms. The directions i are aligned along the three directions of the fast Fourier transform, and they are always orthogonal

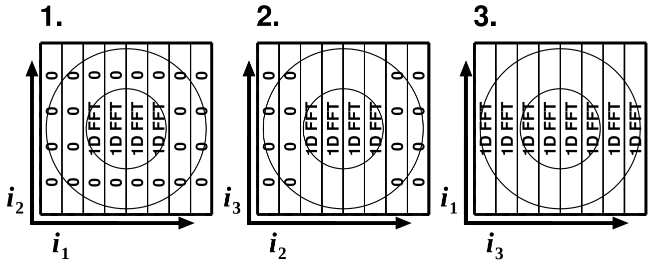


Figure 2.11: The savings in the 3D FFT for a wave function which are possible to achieve in passing from the reciprocal space to real space thanks to the zeroes outside the cut-off sphere for the wave function; the transfer from real to reciprocal space would proceed in the same manner from right to the left

atomic positions and is complete; therefore the Hellmann-Feynman theorem states that

$$\mathbf{F}_{\mathrm{I}} = -\frac{\partial E_{\mathrm{KS}}}{\partial \mathbf{R}_{\mathrm{I}}} \tag{2.64}$$

does not contain any Pulay corrections (Pulay, 1969), which are often difficult to evaluate, to the forces.

2.4.2 The Car-Parrinello method – iterative schemes

A major breakthrough for the realistic simulation of materials using electronic structure calculations was achieved in 1985 when Roberto Car and Michele Parrinello introduced a scheme which iteratively solved the electron structure while the ions were moved according to molecular dynamics; iteratively means that the Hamiltonian matrix is not diagonalised exactly at each electronic iteration. The iterative approach is natural also because the problem is self-consistent, therefore many exact diagonalisations are required, and the scheme of Car and Parrinello was able to make very efficient use of the Fast Fourier Transform – *FFT*. In the framework of the iterative procedure the wave functions are up-dated using information of the current wave function and possibly the previous iterations. Most of the algorithms proposed after the original work of Car and Parrinello are designed to optimise the electron structure – *i. e.* minimise the total energy with respect to the electronic degrees of freedom; they include the Williams-Soler algorithm (1987), second-order equations of motion (Payne *et al*, 1992; Tassone, Mauri & Car, 1994), conjugate gradient method (Gillan, 1989; Arias, Payne & Joannopoulos, 1992; Kresse, 1993; Kresse & Furthmüller, 1996a,b) and direct inversion in iterative sub-space or the residual minimisation (Pulay, 1980; Wood & Zunger, 1985; Hutter, Lüthi & Parrinello, 1994; Kresse & Furthmüller, 1996a). The original Car-Parrinello dynamics has also been enhanced (Tassone, Mauri & Car, 1994; Tuckerman & Parrinello, 1994a,b; Hutter, Tuckerman & Parrinello, 1995) and it has been extended into path integrals (Marx & Parrinello, 1994; 1996; Tuckerman *et al*, 1996). Here we shall describe mainly only the approach based on equations of motion to up-date the wave functions but many features are transferable to the case of the other schemes also; further details of them can be found *e. g.* in the work of Kresse *et al* (1993; Kresse & Furthmüller, 1996a,b).

In the iterative methods using "equations of motion" – this should not be confused with the real dynamics of the electrons – the wave function of the next iteration is a linear combination of the present wave function, possibly the wave functions of the previous iterations and the residual $R(\mathbf{r})$ acting on the current state, *e. g.*

$$\begin{aligned} \psi^{n+1}(\mathbf{r}) &= \psi^{n}(\mathbf{r}) + f(\mathbf{r})\, R(\mathbf{r}) \\ \psi^{n+1}(\mathbf{r}) &= -\psi^{n-1}(\mathbf{r}) + \psi^{n}(\mathbf{r}) + \alpha\dot{\psi}(\mathbf{r}) + f(\mathbf{r})\, R(\mathbf{r}) \ , \qquad \dot{\psi}(\mathbf{r}) = \psi^{n}(\mathbf{r}) - \psi^{n-1}(\mathbf{r}) \end{aligned} \tag{2.65}$$

for a first- and second-order equation of motion, respectively; n is the index for the iteration, α a coefficient equivalent to friction and $f(\mathbf{r})$ is a function whose magnitude controls the amount of change and shape the weighting in different parts of the cell; its form is in principle arbitrary and can be chosen to deliver fastest rate of convergence. Ideally it would resemble the inverse of the residual since then the wave functions would change approximately with the same amount everywhere and the magnitude of f is not limited by fast changing regions; this is called *pre-conditioning* and will be discussed below.

The residual is the difference between the current state $\psi_{i\mathbf{k}}(\mathbf{r})$ and the true eigen-state of the present Hamiltonian; however, as the latter is not known it is approximated by the product

$$R_{i\mathbf{k}}(\mathbf{r}) = -\left[H(\mathbf{r}) - \varepsilon_{i\mathbf{k}}\mathcal{S}\right]\psi_{i\mathbf{k}}(\mathbf{r}) \ . \tag{2.66}$$

$\mathcal{S}$ comes from the Vanderbilt construction and is a consequence of the generalised ortho-normalisation constraint (Laasonen *et al*, 1993). The residual can also be viewed as the force, or the partial derivative of the total energy with respect to the present orbital, acting on this orbital

$$R_{i\mathbf{k}}(\mathbf{r}) = \frac{\delta E_{\mathrm{tot}}}{\delta\psi^{*}_{i\mathbf{k}}(\mathbf{r})} \ ; \tag{2.67}$$

the eigenvalue $\varepsilon_{i\mathbf{k}} = \langle\psi_{i\mathbf{k}}|H|\psi_{i\mathbf{k}}\rangle$ in the Equation 2.66 comes from the constraint of ortho-normalisation. Clearly in the ground state the inner product of the residual $R_{i\mathbf{k}}$ with the all eigenstates ψ_j vanishes since $\langle\psi_{j\mathbf{k}}|H|\psi_{i\mathbf{k}}\rangle$ is diagonal and equal to $\varepsilon_{i\mathbf{k}}$.

The operation of the Hamiltonian on the state can be performed easiest employing the fast Fourier transform since some of the terms in the product $H\psi_{i\mathbf{k}}$ are diagonal, *i. e.* depend only either on $\mathbf{G}$ or $\mathbf{r}$ and not on pairs $\mathbf{G},\mathbf{G}'$ or $\mathbf{r},\mathbf{r}'$, in real and some in reciprocal space; the product in case of the fully non-local form of the pseudo potential is

$$H(\mathbf{r})\psi_{i\mathbf{k}}(\mathbf{r}) = -\frac{1}{2}\nabla^2\psi_{i\mathbf{k}}(\mathbf{r}) + \sum_{I}\sum_{l'm',lm} D^{\mathcal{I}}_{l'm',lm}\left|\beta^{I}_{l'm'}\right\rangle B^{I}_{lm,i} + V_{\mathrm{local}}(\mathbf{r})\psi_{i\mathbf{k}}(\mathbf{r}) \ , \tag{2.68}$$

where $D^{\mathcal{I}}_{l'm',lm}$ are scalars arising from the Vanderbilt construction, $\beta^{I}_{l'm'}(\mathbf{r}-\mathbf{R}_I) = \mathrm{i}^{l}\beta^{\mathcal{I}}_{l'}(|\mathbf{r}-\mathbf{R}_I|)\,Y_{l'm'}\left(\widehat{\mathbf{r}-\mathbf{R}_I}\right)$, $B^{I}_{lm,i} = \left\langle\beta^{I}_{lm}|\psi_{i\mathbf{k}}\right\rangle$ and $V_{\mathrm{local}}(\mathbf{r}) = V^{\mathrm{loc}}_{\mathrm{pp}}(\mathbf{r}) + V_{\mathrm{Hartree}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r})$; for norm-conserving Hamann or Troullier-Martins type of pseudo potentials $D^{\mathcal{I}}_{l'm',lm} = \delta_{ll'}\delta_{mm'}E_{\mathrm{KB},l}$. As described in Section 2.4.1 the kinetic energy term is calculated most conveniently in the reciprocal space; the wave functions are also most efficiently stored in the reciprocal space since less coefficients need to be stored there – in the reciprocal space the number of points inside the cut-off sphere is smaller than the number of points on the corresponding real-space grid – and for the operations with the local potential it should be transformed to the larger fast Fourier transform grid anyway. The inner products $B^{I}_{lm,i}$ have been calculated beforehand, either in the real or reciprocal space: in the real space the operations are localised and this can be used to reduce the calculational burden, on the other hand in the reciprocal space the implementation is straight-forward. Thus the operation of the Hamiltonian on the current state is eventually

$$\begin{aligned}\left[H\psi_{i\mathbf{k}}\right](\mathbf{G}) &= |\mathbf{k}+\mathbf{G}|^2\psi_{i\mathbf{k}}(\mathbf{G}) + \sum_{I}\sum_{l'm',lm} D^{\mathcal{I}}_{l'm',lm}\beta^{I}_{l'm'}(\mathbf{G})\,B^{I}_{lm,i} \\ &\quad +\mathrm{FFT}\left\{V_{\mathrm{local}}(\mathbf{r})\,\mathrm{FFT}^{-1}\left[\psi_{i\mathbf{k}}(\mathbf{G})\right](\mathbf{r})\right\}(\mathbf{G})\end{aligned} \tag{2.69}$$

or

$$\begin{aligned}\left[H\psi_{i\mathbf{k}}\right](\mathbf{G}) &= |\mathbf{k}+\mathbf{G}|^2\psi_{i\mathbf{k}}(\mathbf{G}) \\ &\quad +\mathrm{FFT}\left\{\sum_{I}\sum_{l'm',lm} D^{\mathcal{I}}_{l'm',lm}\beta^{I}_{l'm'}(\mathbf{r})\,B^{I}_{lm,i} + V_{\mathrm{local}}(\mathbf{r})\,\mathrm{FFT}^{-1}\left[\psi_{i\mathbf{k}}(\mathbf{G}')\right](\mathbf{r})\right\}(\mathbf{G})\end{aligned} \tag{2.70}$$

with the fast Fourier transforms the calculation is arranged as shown in the Figure 2.12.

The density is the absolute value of the wave function squared in the real space, thus

$$n(\mathbf{r}) = \sum_{\mathbf{k}} w_{\mathbf{k}}\sum_{i} f_{i\mathbf{k}}\left|\mathrm{FFT}^{-1}\left[\psi_{i\mathbf{k}}(\mathbf{G})\right](\mathbf{r})\right|^2 \ . \tag{2.71}$$

The other fast Fourier transforms during one self-consistent cycle – with exchange-correlation functionals not explicitly dependent on single states – do not involve single states and therefore one

Input: $\psi_{i\mathbf{k}}(\mathbf{G}) = c_{i\mathbf{k}}(\mathbf{G}),\ V_{\text{loc}}(\mathbf{r}),\ B^I_{lm,i},\ \beta^I_{l'm'}(\mathbf{G})$

Reciprocal space		Real space
$c_{i\mathbf{k}}(\mathbf{G})$	$\text{FFT}^{-1} \Longrightarrow$	$\psi_{i\mathbf{k}}(\mathbf{r})$
		$\Downarrow$
$H_{\text{kin+loc}}\psi_{i\mathbf{k}} = \lvert\mathbf{k}+\mathbf{G}\rvert^2 c_{i\mathbf{k}}(\mathbf{G}) + [V_{\text{loc}}\psi_{i\mathbf{k}}](\mathbf{G})$	$\Longleftarrow \text{FFT}$	$V_{\text{loc}}(\mathbf{r})\,\psi_{i\mathbf{k}}(\mathbf{r})$

$$H_{\text{pp}_{\text{non-loc}}}\psi_{i\mathbf{k}}(\mathbf{G}) = \sum_I \sum_{l'm',lm} D^{\mathcal{I}}_{l'm',lm}\beta^I_{l'm'}(\mathbf{G})\,B^I_{lm,i}$$

$$H\psi_{i\mathbf{k}} = H_{\text{kin+loc}}\psi_{i\mathbf{k}} + H_{\text{pp}_{\text{non-loc}}}\psi_{i\mathbf{k}}$$

Figure 2.12: The schematic illustration of the Hamiltonian operating on the wave function in the Car-Parrinello method implemented in the Kohn-Sham framework

iteration takes three fast Fourier transforms per electronic state, or two if the wave function Fourier transformed into the real space can be stored in the memory when calculating the density because the wave function is not changed before the first Fourier transform in $H\psi_{i\mathbf{k}}$. The density needs to be transformed into the reciprocal space for the Hartree and the local part of the pseudo potential, and its gradients when employing the generalised gradient approximations; the local potentials calculated in the reciprocal space are back-transformed and added to the exchange-correlation potential. Thereby the self-consistent cycle looks like shown in the Figure 2.13.

In the schemes not based on an equation of motion – *e. g.* the conjugate gradient or the residual minimisation – the basic iteration loop is very similar, only the details of the up-date of the wave functions have changed but the same gradient or residual of the Equation 2.66 is used. These methods are more similar in spirit to the original exact diagonalisation as they up-date the wave functions closer to a diagonal Hamiltonian and the density and potential are calculated less often. In the calculational effort both approaches are probably rather equal.

2.4.3 An implementation – Faust

The calculations of the present work were originally started using the `fhi93cp` code (Stumpf & Scheffler, 1994) and shortly the `fhi94md` version[2] until the author shifted to a program code obtained from Kari Laasonen (personal communication; please see also Laasonen *et al*, 1993); the code includes the possibility to use the Vanderbilt ultra-soft pseudo potentials. It has been actively maintained and developed by the author, and the working name for the code is "Faust" (Seitsonen, Boisvert & Laasonen, unpublished). The main goal was to achieve a code which is easily readable, portable and editable without sacrifying too much of the performance, and personally to obtain further experience in programming and electronic structure methods. Some features from other codes CASTEP/CETEP, da Capo (Bjørk Hammer/Denmark) `fhi94md`and preliminary versions of FINGER (Laboratory of Physics, Helsinki University of Technology, Finland) have been adapted to Faust, and $\mathbf{k}$ and $\mathbf{G}$ vector parallel versions have been written, the latter mainly with the help from Sami Pöykkö (personal communication). Recently the code was adapted from FORTRAN77 to FORTRAN90, with some modifications still to be done.

Some of the features in Faust include:

- Alternative compilation in fully double or partially single precision

[2]The newest version of the code, `fhi96md`, has been documented by Bockstedte *et al* (1997)

$$\text{Input: Initial } \psi_{i\mathbf{k}}(\mathbf{G})$$

$$\psi_{i\mathbf{k}}(\mathbf{G}) \overset{\mathrm{FFT}^{-1}}{\Longrightarrow} \psi_{i\mathbf{k}}(\mathbf{r})$$

$$n(\mathbf{r}) = \sum_{\mathbf{k}} w_{\mathbf{k}} \sum_i |\psi_{i\mathbf{k}}(\mathbf{r})|^2$$

$$V_{\mathrm{xc}}(\mathbf{r}) = V_{\mathrm{xc}}[n(\mathbf{r})]$$

$$n(\mathbf{r}) \overset{\mathrm{FFT}}{\Longrightarrow} n(\mathbf{G})$$

$$V_{\mathrm{H}}(\mathbf{G}) = V_{\mathrm{H}}[n(\mathbf{G})]$$

$$V_{\mathrm{ps_{local}}}(\mathbf{G}) = V_{\mathrm{ps_{local}}}[n(\mathbf{G})]$$

$$V_{\mathrm{H}}(\mathbf{G}) + V_{\mathrm{ps_{local}}}(\mathbf{G}) \overset{\mathrm{FFT}^{-1}}{\Longrightarrow} [V_{\mathrm{H}} + V_{\mathrm{ps_{local}}}](\mathbf{r})$$

$$V_{\mathrm{local}}(\mathbf{r}) = V_{\mathrm{xc}}(\mathbf{r}) + [V_{\mathrm{H}} + V_{\mathrm{ps_{local}}}](\mathbf{r})$$

$$\underline{\texttt{foreach } \mathbf{k}}$$

$$\underline{\texttt{foreach i}}$$

$$\text{Calculate } R_{i\mathbf{k}} = -[H - \varepsilon_{i\mathbf{k}} S]\,\psi_{i\mathbf{k}}$$

$$\mathrm{Up-date}\ \psi_{i\mathbf{k}}(t+\Delta t) = a\psi_{i\mathbf{k}}(t-\Delta t) + b\psi_{i\mathbf{k}}(t) + \mathtt{f}[R_{i\mathbf{k}}]$$

$$\underline{\texttt{end loop i}}$$

$$\mathrm{Ortho-normalise}\ \psi_{i\mathbf{k}}$$

$$\text{Calculate new } B^{I}_{lm,i}$$

$$\underline{\texttt{end loop } \mathbf{k}}$$

$$\text{Output: New } \psi_{i\mathbf{k}}(\mathbf{G})$$

Figure 2.13: The schematic illustration of the self-consistent cycle in the equation-of-motion-based, iterative Car-Parrinello method implemented in the Kohn-Sham framework

- Dynamic memory allocation
- Compiles and runs at IBM, SGI, DEC, Cray T3E, HP, linux and other architectures
- Conjugate gradient algorithm/sub-space rotation
- Run-time control of many of the calculational parameters
- The numerical and White-Bird evaluation of gradient approximations
- Weighted-density approximation, local mass approximation, “Meta-GGA” (Perdew *et al*, 1999) as more “exotic” exchange-correlation functionals
- Fermi/Gaussian/Methfessel-Paxton/“cold” smearing of occupation numbers
- Molecular dynamics for the ions
- Surface dipole correction (Neugebauer & Scheffler, 1992; Bengtsson, 1999)
- Semi-local form of the pseudo potential

2.5 Output from the calculations

Here we shall shortly explain the interpretation of the calculations, that is, what information can be won from the resulting output.

2.5.1 Lattice constant, bulk modulus and the cohesive energy

The lattice constant is in principle the value which minimises the Gibbs's free energy but due to the approximation to the exchange-correlation term the theoretical lattice constant is not the same as from the experiments. Usually the temperature and zero-point effects are also not included; these would increase the lattice constant somewhat, typically less than one per cent.

The lattice constant is usually determined by calculating the total energy of the bulk material at several values of the lattice constant and performing a fit to the resulting points. We have used the Murnagham (1944) fit, which also yields a value for the bulk modulus. The minimum of the total energy can be used to determine the cohesive energy, but this has seldom been calculated since for it the energy of an isolated atom is needed with an equivalent accuracy, and in the methods employing the atom has to be located in a large unit cell to avoid the interaction with its periodic replicas, and the calculation has to be performed with spin-polarisation, the occupation numbers are not necessarily integers and the charge density even not spherical (Kutzler & Painter, 1987).

We have checked the effect of the zero-point vibrations using the formulas of Moruzzi, Janak and Williams (1978). In metals like aluminium, rhodium and silver the change in the lattice constant due to the zero-point vibrations was found to be less than one per cent and we thus neglected this effect from the determination of the lattice properties for ruthenium.

2.5.2 Adsorption energies

The binding energy of an adsorbate to a surface is not always uniquely defined, as its value depends on the energy reference: For example atoms which form strongly bound molecules can overcome the repulsion which reigns between the adsorbed atoms *e.g.* when the surface is progressively heated but before the atoms would desorb they recombine. Also part of the energy gained in the recombination process might be partially adsorbed to the degrees of freedom of the surface. Thus the value of binding energy derived from the thermal desorption spectroscopy needs an accurate analysis of the spectra, and the value obtained from calculations should take into account a larger portion of the potential energy surface than only at the immediate vicinity of the adsorption sites. We usually refer the values to dimers in the case of atomic adsorbates of the $2p$ row, which are in gaseous form at normal conditions, but in the case of cesium the reference is a single atom. The binding energy in the case of with respect to an atom and to a dimer is respectively

$$\begin{aligned} E_{\text{bind}}^{A/S} &= E\left(\text{A/S}\right) - \left[E\left(\text{S}\right) + E\left(A_{\text{atom}}\right)\right] \\ E_{\text{bind}}^{A/S} &= E\left(\text{A/S}\right) - \left[E\left(\text{S}\right) + \frac{1}{2}E\left(A_2\right)\right] \ , \end{aligned} \tag{2.72}$$

where A stands for the adsorbate and S the substrate and E is the total energy of the system; please note that then negative values indicate an energy gain upon adsorption and positive repulsion.

Furthermore if the adsorbate-adsorbate interaction is repulsive then desorbing one adsorbate reduces the repulsion and this energy will increase the binding energy of the remaining adsorbates. As an example let us consider the adsorption of cesium, which is repulsive at all coverages: if the ordered structures formed are $(\sqrt{3}\times\sqrt{3})\text{R}30^\circ$, (2×2), $(\sqrt{7}\times\sqrt{7})\text{R}19.1^\circ$ and (3×3) at the coverages 1/3, 1/4, 1/7 and 1/9, the binding energy at the different coverages is obtained as

$$\begin{aligned} E_{\text{bind},1/3} &= \left[E_{1/3} - 3E_{\text{clean}}\right] - \left[E_{1/4} - 4E_{\text{clean}}\right] = E_{1/3} - E_{1/4} + E_{\text{clean}} \\ E_{\text{bind},1/4} &= \left[E_{1/4} - 4E_{\text{clean}}\right] - \left[E_{1/7} - 7E_{\text{clean}}\right] = E_{1/4} - E_{1/7} + 3E_{\text{clean}} \end{aligned} \tag{2.73}$$

$$E_{\text{bind},1/7} = \left[E_{1/7} - 7E_{\text{clean}}\right] - \left[E_{1/9} - 9E_{\text{clean}}\right] = E_{1/7} - E_{1/9} + 2E_{\text{clean}} \ ;$$

E_{clean} is the total energy of the clean surface in a 1×1 unit cell. Please note that if the binding energy for adding one further adsorbate to the structure is defined in the same manner the value can be different, as the repulsion in this case might increase by a different amount than it decreases when an atom is desorbed.

The coverage dependence of the interaction between the adsorbates can easily be characterised using the lattice gas model where the interaction energy E_{int} is decomposed into a sum of parameters which are independent of the coverage. Kreuzer and Payne (1997) write the energy as a sum of the binding energy of an isolated adsorbate E_0, nearest neighbour pairs of the adsorbates V_2, second nearest neighbour pairs V_2', third nearest pairs V_2'' *etc*, and triplets of atoms at the nearest neighbour sites V_t, V_t' and V_t'' depending on the angle between them, and so on:

$$\begin{aligned} E_{\text{int}} &= \sum_i n_i E_0 \\ &+ \sum_{i,j=\text{NN}} n_i n_j V_2 + \sum_{i,j=\text{NNN}} n_i n_j V_2' + \sum_{i,j=\text{3rdNN}} n_i n_j V_2'' + \dots \\ &+ \sum_{i,j=\text{NN},\angle=180^\circ} n_i n_j V_t + \sum_{i,j=\text{NN},\angle=120^\circ} n_i n_j V_t' + \sum_{i,j=\text{NN},\angle=60^\circ} n_i n_j V_t'' + \dots \ , \end{aligned} \tag{2.74}$$

where $n_i = 0, 1$ is the occupation of the site i. Only one adsorption site is assumed at all configurations. The sums have to be truncated after some terms, and care must be taken that the same interaction is not counted twice; the Figure 2.14 shows as an example the number of different interactions on a hexagonal lattice, like has been studied in the present work. We obtain values for the parameters by calculating the binding energy of the adsorbate at different coverage and then inverting the Equation 2.74 to obtain the parameters.

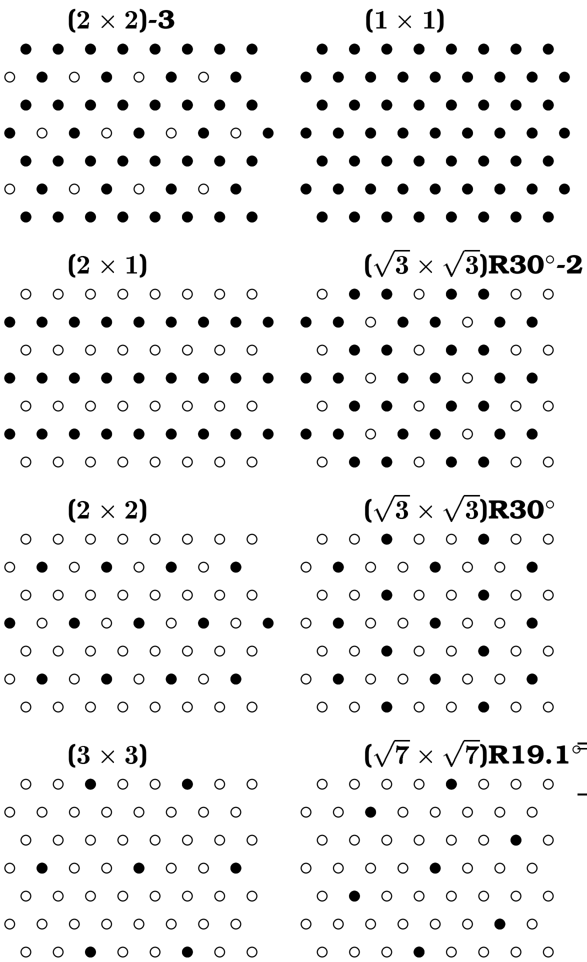
2.5.3 Density of States – DOS

The density of states, or the total density of states, describes how many electronic states are located at inside an infinitesimally small region $d\varepsilon$ around a given energy ε. In practice this is usually calculated from the Kohn-Sham eigenvalues $\varepsilon_{i\mathbf{k}}$ from the weighted sum

$$g(\varepsilon) = \sum_{\mathbf{k}} w_{\mathbf{k}} \sum_i \mathcal{B}\left(\varepsilon - \varepsilon_{i\mathbf{k}}\right) \ . \tag{2.75}$$

$\mathcal{B}$ is a broadening function necessary due to the discreteness of the eigenvalues; sometimes the tetrahedron method is used to interpolate the energies. We have used the Gaussian broadening in the present work; $w_{\mathbf{k}}$ is the weight associated with the $\mathbf{k}$ point as explained in the Section 2.3.2. The density of states is usually normalised to the Fermi energy in the case of metals and maximum of the valence band in semi-conductors; furthermore if the vacuum level of the potential or energy is known the energy can be normalised with respect to it. In the present work the energies are always referred to the Fermi energy if not explicitly mentioned.

The total density of states provides no direct information on the spatial location of the electronic states. Indirectly some of the states can often be associated to specific atoms if the energy range is well isolated from other features, *e.g.* the semi-core or deeply bound valence states and well-resolved peaks. A convenient quantity is obtained by subtracting from the density of states of the substrate with the adsorbate the density of states of the clean substrate; thereby the difference is only due to the adsorption. Subtracting also the density of states of the adsorbate only is usually not useful since it has only localised peaks which shift to different energies upon the adsorption.



Structure	E_0	V_2	V_2'	V_2''	V_2'''	V_t	V_t'	V_t''
(1 × 1)	1	3	3	3	6	3	6	6
(2 × 2)-3	1	2	2	3	4	2	2	2
($\sqrt{3} \times \sqrt{3}$)R30°-2	1	3/2	3	3/2	3	0	3	0
(2 × 1)	1	1	1	3	2	1	0	0
($\sqrt{3} \times \sqrt{3}$)R30°	1	0	3	0	0	0	0	0
(2 × 2)	1	0	0	3	0	0	0	0
($\sqrt{7} \times \sqrt{7}$)R19.1°	1	0	0	0	3	0	0	0
(3 × 3)	1	0	0	0	0	0	0	0

Figure 2.14: The number of different interactions per atom in the lattice gas model on a hexagonal lattice at different coverages. All the different configurations up to the fourth nearest neighbour and the trio interactions have been treated as separate

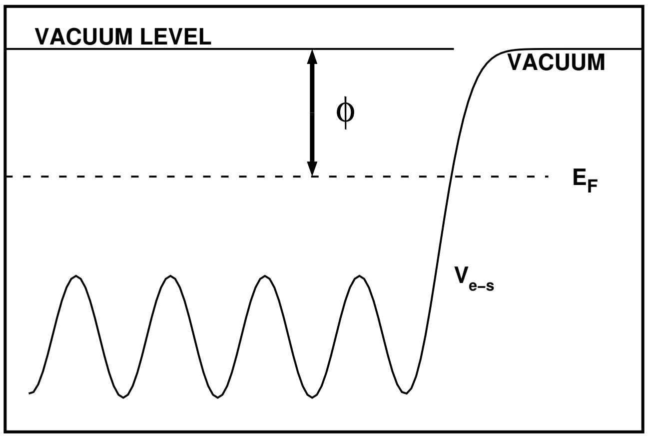


Figure 2.15: The definition of the work function $\phi = V_{\mathrm{e-s,vac}} - \varepsilon_F$; ε_F is the Fermi energy and $V_{\mathrm{e-s}}$ the electro-static potential

2.5.4 Projected or Local Density of States – PDOS or LDOS

The local density of states is similar to the total density of states but it contains also spatial information; this is achieved by projecting or calculating the overlap of localised orbitals centred on the atoms with the Kohn-Sham states of the under study:

$$g_{lm,I}(\varepsilon) = \sum_{\mathbf{k}} w_{\mathbf{k}} \sum_{i} f_{i\mathbf{k}} \mathcal{B}(\varepsilon - \varepsilon_{i\mathbf{k}}) \left| \langle \phi(|\mathbf{r}-\mathbf{R}|) | \left\langle Y_{lm} \left(\widehat{\mathbf{r}-\mathbf{R}} \right) \right| \psi_{i\mathbf{k}}(\mathbf{r}) \rangle \right|^2 \qquad (2.76)$$

Here $\phi(r)$ is a radial projection function, usually the atomic pseudo wave function; its use is unpleasant but unavoidable because the logical definition where the Kohn-Sham orbitals are projected only onto the angular functions would lead to semi-local expressions, which are computationally very demanding in large systems. Furthermore the radial function need to be cut off at a distance since its extent is very large and would otherwise over-lap the neighbouring atoms, leading to artificial contribution to the density of states. We have used a radial cut-off which is slightly smaller than half of the distance to the nearest atom.

The analysis of the local density of states is similar to the total density of states but it is now decomposed at each atom and to the angular momentum channels *l,m*.

2.5.5 Work function

The work function of a surface is defined as the minimum energy required to move an electron from the bulk to the vacuum outside the surface. In practice the value for the work function is obtained as illustrated in the Figure 2.15:

$$\phi = V_{\mathrm{e-s,vac}} - \varepsilon_F \ , \qquad (2.77)$$

where $V_{\mathrm{e-s,vac}}$ is the value of the electro-static potential in the vacuum, *i. e.* far away from the slab.

2.5.6 Blyholder model for adsorption of carbon monoxide

The frontier orbitals – the highest occupied and the lowest unoccupied molecular orbitals – of carbon monoxide are the 5σ and $2\pi^*$ in the Figure 2.16; the energy levels of the molecule are listed as well. In 1964 George Blyholder introduced a model according to which the bonding of CO onto surfaces happens via the emptying of the 5σ orbital and filling of the $2\pi^*$ orbitals as shown in the Figure 2.17.

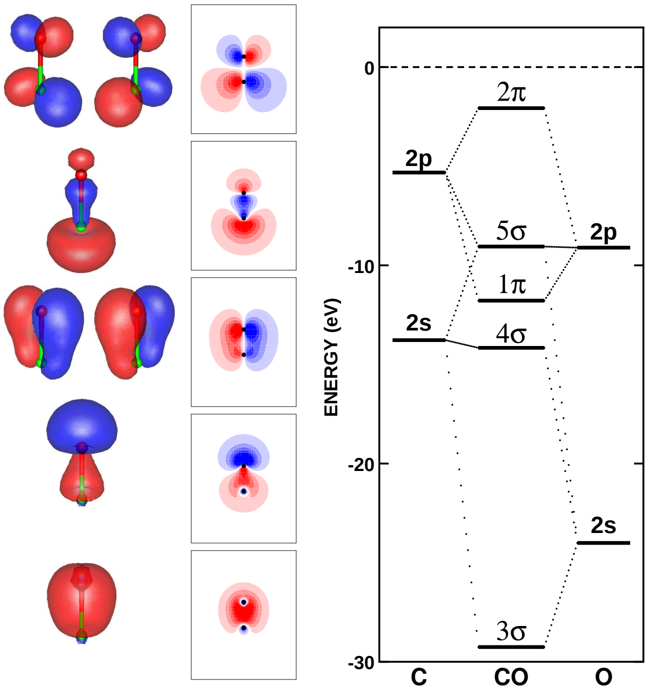


Figure 2.16: The significant Kohn-Sham single particle states of the carbon monoxide; right are the Kohn-Sham energy levels of the orbitals from our density functional calculation employing the generalised gradient approximation

The electrons are redistributed to the hybrids with the states of the substrate; the interaction with the electrons of s and p character is weak – CO is only very weakly bound on the surfaces of aluminium – and thus the main ingredient is the hybridisation with the d orbitals, and the bond is strongest on the surfaces of the transition metals which have an open d shell. On the contrary the interaction with the noble metals gold, silver and copper is weaker as there are no free – mainly only polarisable – d electrons, and thus the back-bonding channel to the $2\pi^*$ orbitals is not efficient. The adsorption site for CO depends on the nature of the d hybrids: On the on-top site the 5σ orbital overlaps with the d_{z^2} like states and the back-bonding occurs via the $d_{xz,yz}$ orbitals, whereas on the three-fold – and two-follow bridge site – sites this is vice versa.

Recently the Blyholder model has been criticised as over-simplified and thus improper. However, although we have also found a strong mixing of the 4σ and 5σ orbitals and the Kohn-Sham energy

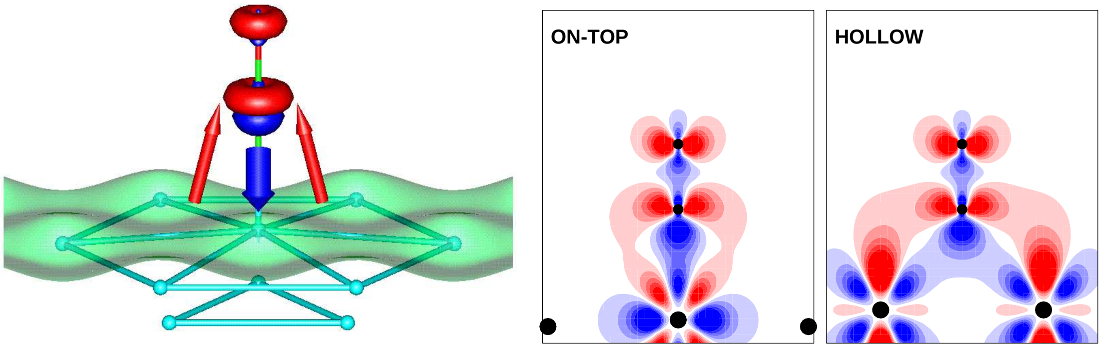


Figure 2.17: The Blyholder mechanism: The 5σ orbital is emptied and the $2\pi^*$ filled in the formation of hybrids with the substrate; right is the interaction mechanism with the different d like states on the on-top and hollow adsorption sites

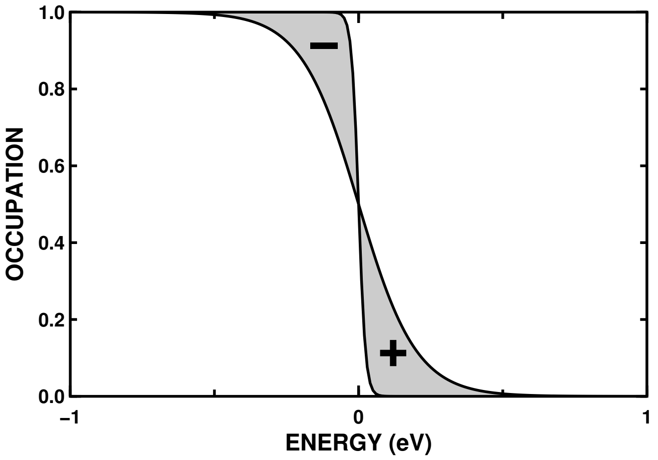


Figure 2.18: The occupation functions of broadenings $T_{\mathrm{el},1}$ = 0.1 eV and $T_{\mathrm{el},1}$ = 0.01 eV used in the reactivity function of Wilke; the positive and negative regions correspond mainly to occupied and unoccupied states, respectively

level of the orbital with mainly 5σ character moves below the level of the 1π orbital, thus making the 5σ no longer the highest occupied "molecular" orbital – however the 1π does not take part into the bonding as its overlap is small due to little overlap with the substrate states –, the main features of the CO adsorption are still described well by the Blyholder model: a model is only a model, and the details of the bonding go beyond the Blyholder model.

2.5.7 Iso-electronic reactivity function

The straight-forward study of surface reactions by direct mapping of the related multi-dimensional potential energy surface is not feasible with the current methods and computer capacity. Therefore it would be advantageous to find a measure which could predict the response of a substrate to adsorbates without the actual, complicated calculation. These reactivity theories have been put forward especially in the community of quantum chemistry, where the concept of frontier orbitals – using the concepts like the highest occupied and lowest unoccupied molecular orbital, HOMO and LUMO – has emerged. However, this does not directly generalise to the metallic case where there is no gap separating the occupied and unoccupied states and the density of states around the Fermi energy, *i.e.* where the states easily deform and change their occupation under different chemical environments, is large. An alternative scheme for this purpose was developed by Wilke, Cohen and Scheffler (1996): it explores the character of the states around the Fermi energy by changing the broadening of the occupation numbers in a scheme shown in the Figure 2.18 and defining the reactivity function as

$$\mathcal{W}^{\mathrm{N}} = \frac{1}{k_B^2 T_{\mathrm{el}}} \left[\frac{\partial n(\mathbf{r}, T_{\mathrm{el}})}{\partial T_{\mathrm{el}}} \right]_{V_{\mathrm{ext}}(\mathbf{r}),\mathrm{N}} \approx \frac{n(\mathbf{r}, T_{\mathrm{el},1}) - n(\mathbf{r}, T_{\mathrm{el},2})}{k_B^2 T_{\mathrm{el},1} \left[T_{\mathrm{el},1} - T_{\mathrm{el},2} \right]} , \tag{2.78}$$

where the super-script N denotes that the number of particles is conserved and T_{el} is the temperature used for the broadening of the occupation numbers in the Fermi function. Please note that we use two different finite temperatures whereas Wilke *et al* used $T_{\mathrm{el},2} = 0$. Both densities are calculated self-consistently.

The function yields a spatially resolved function, and the regions where it is positive the states are mostly unoccupied, becoming easily occupied and vise versa. It should be kept borne in mind, however, that the reactivity function is only an approximative estimate for the reactivity, and in chemisorbed systems the actual binding is likely to happen in an energy range not in the vicinity of the Fermi level.

2.5.8 Density differences / induced density

The density differences describe the charge transfer in the system upon a change in the atomic configuration. The change here comprises extraction or inclusion of atoms or molecules: the positions of the corresponding atoms in different calculations have to be the same, *i. e.* they must not be relaxed except in one case and the coordinates to the other calculations are taken from the relaxed one. Generally the system we are interested in is an adsorbate on a surface, and although we refer to the density difference as arising from the adsorption, the geometry is taken from the final state, *i. e.* the adsorbed system. Thus the density difference in this characteristic case would be calculated as

$$\Delta n = n^{\mathrm{A/S}}\left(\mathrm{A/S}\right) - n^{\mathrm{A/S}}\left(\mathrm{S}\right) - n^{\mathrm{A/S}}\left(\mathrm{A}\right) \ , \tag{2.79}$$

where

- $n^{\mathrm{A/S}}\left(\mathrm{A/S}\right)$ is the system with the adsorbate A on the substrate S, and the atoms have been relaxed
- $n^{\mathrm{A/S}}\left(\mathrm{S}\right)$ is the system with only the substrate S, but the coordinates of the substrate atoms have been taken from the calculation with the adsorbate; thus the super-script A/S
- $n^{\mathrm{A/S}}\left(\mathrm{A}\right)$ is the system with only the adsorbate A, but the coordinates of the adsorbate atoms have been taken from the calculation with the adsorbate on the surface; thus the super-script A/S

Usually the atoms match in the sense that from the adsorbed system is subtracted both the substrate and the adsorbate. However, in the case of an alkali metal we also want to see the so-called induced charge because the density of the outer-most, *s* electron differs notably from the atomic density. The induced density is defined similarly to the normal density difference above as

$$\Delta n = n^{\mathrm{A/S}}\left(\mathrm{A/S}\right) - n^{\mathrm{A/S}}\left(\mathrm{S}\right) - n^{\mathrm{A/S}}\left(\mathrm{A}_{\mathrm{semi-core}}\right) \ , \tag{2.80}$$

where, however, $n^{\mathrm{A/S}}\left(\mathrm{A}_{\mathrm{semi-core}}\right)$ contains only the semi-core states of the atomic adsorbate, if present; in our case of cesium the semi-core states are the 5*s* and 5*p* orbitals.

One further consideration is whether for the adsorbate the density is calculated self-consistently in the slab geometry – without the substrate – or as a super-position of the density of the atomic constituents. For laterally large unit cells the difference is negligible, and also in the reasonably small cells this choice the qualitative picture remains the same irrespective of the choice.

2.5.9 Interaction density / energy

The interaction density bears a close relationship to the density difference just explained but it is used to monitor the density changes upon co-adsorption on the substrate. It is defined as

$$\Delta n^{i} = n^{(\mathrm{A}_1+\mathrm{A}_2)/\mathrm{S}}\left((\mathrm{A}_1+\mathrm{A}_2)/\mathrm{S}\right) - n^{(\mathrm{A}_1+\mathrm{A}_2)/\mathrm{S}}\left(\mathrm{A}_1\right) - n^{(\mathrm{A}_1+\mathrm{A}_2)/\mathrm{S}}\left(\mathrm{A}_2\right) + n^{(\mathrm{A}_1+\mathrm{A}_2)/\mathrm{S}}\left(\mathrm{S}\right). \tag{2.81}$$

Here A_1 and A_2 refer to the two different co-adsorbates. The interaction density arises from the mutual interaction between the co-adsorbates and would vanish for distant or non-interacting co-adsorbates.

A related quantity is the interaction energy

$$-E\left[(\mathrm{A}+\mathrm{B})/\mathrm{S}\right]+E\left[\mathrm{A}/\mathrm{S}\right]+E\left[\mathrm{B}/\mathrm{S}\right]-E\left[\mathrm{S}\right] \ , \tag{2.82}$$

which is defined as the interaction density as a sum and difference of four energies, which in the case of energy can however be calculated all from relaxed systems and even the adsorption sites and coverages can be different and are generally chosen to be the minimum-energy structures. We have chosen the interaction energy to be positive in case of an attractive

2.5.10 Electro-static potential / electric field

The electro-static interaction is an important contribution in the chemical processes, especially when there is a large difference in the electro-negativity of the materials. One good example is the adsorption of an alkali metal to the surface, a system we have studied in the present work. The electro-static potential outside the core regions consists of the electron-electron repulsion and electron-ion attraction; in the calculations employing pseudo potentials the electro-static potential also includes the local and non-local contribution from the pseudo potentials inside the core regions, but since we are interested only in the regions outside the surface this is irrelevant.

Another interesting, related quantity is the electric field arising from the electro-static potential; this is simply the negative of the gradient of the electro-static potential:

$$\mathbf{E}\left(\mathbf{r}\right)=-\nabla V_{\mathrm{e-s}}\left(\mathbf{r}\right) \ . \tag{2.83}$$

This is a three-dimensional, *i. e.* vector quantity and thus difficult to visualise; we show either the magnitude of the field or the components gradient calculated in the cylindrical coordinates since our interest is on the field close to an alkali atom at the surface and this is largely a radial field.

2.5.11 Pseudo scanning tunnelling microscope image

The scanning tunnelling microscope is a powerful tool for studying the surfaces, structures and even reactions on them. However the atomic-scale features in the figures delivered from it are not as straight-forward to interpret as originally was supposed since the tunnelling current does not always simply correlate with the atomic geometry or better, the height of the atoms. A simple example is the oxygen atom on ruthenium, which is imagined as a depression although it resides above the top-most substrate layer. This apparent confusion derives from two effects: First of all, the tunnelling occurs only in the vicinity of the Fermi level and secondly, the tunnelling matrix elements contain many-body contributions and involve also the electronic – and this way, the geometrical – structure of the tip as well as the surface (Lang, 1989).

An easy model for obtaining qualitatively and often also quantitatively useful, but sometimes even qualitatively wrong, theoretical picture which then can be used to compare with the measured image comes from Tersoff and Hamann (1985). In this model the tunnelling current is proportional to the density of states of the surface only for a tip with an *s* like electronic state dominating the tunnelling; Chen (1990; 1994) has derived the corresponding expression for tips with different kinds of electronic structure. Since in our calculations the **k** sampling is finite we include the occupied

Kohn-Sham states from a region around the Fermi energy in the density of states and thus the pseudo image. We simulate the operation of the microscope in the constant current mode, *i. e.* we show the interpolated height distribution above the surface where the energy-restricted density of states remains constant. We use a value of the constant density of states which either reproduces an experimentally determined corrugation of an adsorption system, or a value which leads to a realistic distance of 6-7 Å above the surface.