

A 7 Car-Parrinello Molecular Dynamics and Reaction Kinetics

Kurt Schroeder

Institut für Festkörperforschung

Forschungszentrum Jülich GmbH

Contents

1	Introduction	2
2	Total energy calculations using density functional theory	3
2.1	Approximations for the exchange-correlation energy	4
2.2	Many degrees of freedom	6
3	The combined MD/DFT-method	6
4	Plane wave pseudopotential calculations	8
4.1	Pseudopotentials	9
4.2	Iterative methods	11
5	Dynamical properties of amorphous Si	12
6	Simulated annealing studies of Se clusters	14
7	Surfactant mediated growth: Kinetics of adatoms	15
A	Fast Fourier Transform	20
B	Davidson Kosugi method	21
C	Quasi-Newton methods	22

1 Introduction

The Density-Functional(DFT)-Method (see contribution of R. Zeller) has opened a whole new world for calculating ground state properties of condensed matter systems. Equilibrium structures of crystals, molecules, surfaces, adsorption layers, structural defects and dissolved impurities are just a few examples [1, 2].

But many materials problems are connected with dynamical properties. E.g., phase transitions are often governed by a redistribution of atoms over long distances. The same is true for crystal growth phenomena, where the atoms deposited to the surface find the appropriate sites for incorporation into the growing crystal only after visiting many surface sites. In both cases the diffusion of atoms is of interest, which is a thermally activated process. In addition, vibrational spectra are often used to identify symmetries of defects in bulk systems or at surfaces. These time and/or temperature dependent quantities are the realm of classical molecular dynamics simulation which are performed with thousands of atoms using analytical model interactions or empirical interactions between the atoms whose parameters are fitted to a set experimental results (see contribution of R. Winkler).

In many instances, the coordination and symmetry of the environment changes when an atom undergoes large displacements from ideal lattice positions, e.g. during a diffusion jump or in the neighborhood of defects. It is difficult to generate empirical potentials which are valid for all transition configurations encountered during a diffusion jump and estimate the validity of the calculated energy differences. Particularly, in covalent systems with directed electronic bonds the formation and character of bonds depends sensitively on the local atomic arrangement. The first challenge thus is to calculate the forces on the atoms from a realistic local electronic distribution which may drastically deviate from the ideal bulk electron density. An illustrative

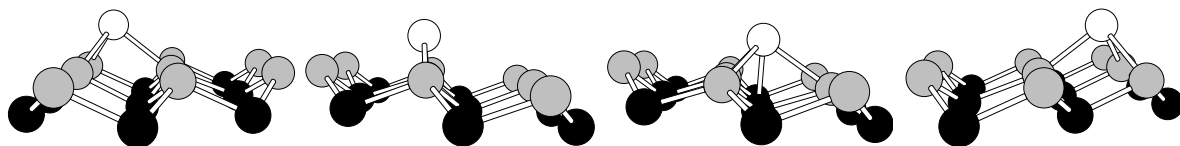


Fig. 1: Diffusion path of a Ge or Si adatom on the As-covered Si(111) surface. Configurations with different bond coordination are shown: (a) the stable position at the H3 site, (b) an intermediate position, (c) the saddle point at the T4 site, (d) H3 site in the neighboring cell. 2nd-layer Si: black, 1st-layer As: gray, adatom: white

example is shown in Fig. 1. Calculations using DFT [3] show that on an As-covered Si(111) surface the equilibrium configuration of a Ge (or Si) adatom is a so-called H3-site (on top of the hexagonal hole in the surface double layer), where the adatom binds to 3 As atoms of the terminating upper layer. In this configuration the bond angle to the As atoms are nearly ideal tetrahedron angles just as in bulk Si, and on the adatom an sp^3 -hybridization is favorable (where one of the orbitals is a singly occupied dangling bond). It is clear that some of these bonds have to be broken when the adatom performs a diffusion jump to an equivalent H3-site in a neighboring surface cell. In one intermediate configuration the Ge adatom has only two bonds to As atoms, which would suggest a planar sp^2 -hybridization on the Ge atom and very different bond angles. The saddle point configuration is the T4 site (on top of the second-layer Si), where the adatom again has three bonds to As atoms, but a repulsive interaction to the second-layer Si atom.

As we will see the DFT-approach provides the necessary information, since the ground state electron density can be calculated to a very good approximation for any atomic arrangement, and realistic forces on the atoms can be calculated for arbitrary configurations. For the characterization of the numerical procedures we closely follow R. Jones' and D. Hohl's contribution to an earlier IFF Spring School [4]. Before we turn to the treatment of dynamics using DFT we shortly summarize the basics of DFT for ground state properties.

2 Total energy calculations using density functional theory

The density functional theory is based on two theorems of Hohenberg and Kohn [5] which are discussed in detail in R. Zeller's contribution.

1. All ground state properties of a system of many electrons in an external field, V_{ext} (e.g. due the Coulomb potential of ions) are fully determined if the electron density, $n(\vec{r})$, is known, they can be written as "functionals" of the density. One example is the total electronic energy $E = E[n]$.
2. A variational principle exists for the energy: $E[n] \geq E_{GS}$, where E_{GS} is the exact ground state energy of the system. The ground state energy is obtained with the exact ground state density, i.e. $E[n_{GS}(\vec{r})] = E_{GS}$.

This variational principle can be used constructively if one decomposes the functional $E[n]$ as suggested by Kohn and Sham [6]:

$$E[n] = T_0[n] + \int d\vec{r} n(\vec{r}) \left(V_{ext}(\vec{r}) + \frac{1}{2}\phi(\vec{r}) \right) + E_{xc}[n] \quad (1)$$

where $T_0[n]$ is the kinetic energy of *non-interacting* electrons with the density $n(\vec{r})$, $\phi(\vec{r})$ is the classical Coulomb potential of the electrons, i.e. the Hartree potential

$$\phi(\vec{r}) = e^2 \int d\vec{r}' \frac{n(\vec{r}')}{|\vec{r}' - \vec{r}|}, \quad (2)$$

and $E_{xc}[n]$ is the *exchange-correlation* energy. The variational principle then yields

$$\frac{\delta E[n]}{\delta n(\vec{r})} = \frac{\delta T_0}{\delta n(\vec{r})} + V_{ext}(\vec{r}) + \phi(\vec{r}) + \frac{\delta E_{xc}[n]}{\delta n(\vec{r})} = \mu, \quad (3)$$

where the Lagrange parameter (chemical potential) μ is used to guarantee the condition of constant number of electrons: $\int d\vec{r} n(\vec{r}) = 2M$ (M is the number of occupied states for non-spinpolarized problems).

Equation (3) is valid for the system of electrons, including all Coulomb interactions. Suppose, we want to treat a (fictitious) system with non-interacting electrons which move in an effective potential $V_{eff}(\vec{r})$. Then the total energy variation yields

$$\frac{\delta E[n]}{\delta n(\vec{r})} = \frac{\delta T_0}{\delta n(\vec{r})} + V_{eff}(\vec{r}) = \mu, \quad (4)$$

A solution for this system can easily be obtained by solving the single-particle Schrödinger equation

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + V_{eff}(\vec{r}) \right) \psi_i(\vec{r}) = \epsilon_i \psi_i(\vec{r}) \quad (5)$$

and constructing the electron density from the M wavefunctions with the lowest eigenvalues (doubly occupied because of spin degeneracy):

$$n(\vec{r}) = 2 \sum_{i=1}^M |\psi_i(\vec{r})|^2. \quad (6)$$

The problems (3) and (4) are identical if one identifies

$$V_{eff}(\vec{r}) = V_{ext}(\vec{r}) + \phi(\vec{r}) + \frac{\delta E_{xc}[n]}{\delta n(\vec{r})}. \quad (7)$$

If this condition is fulfilled, e.g., by a selfconsistency cycle, we have mapped the complicated many-body electron problem *exactly* to the much easier problem for non-interacting electrons. All terms of the total energy, eq. (1), can be calculated rather easily except the exchange-correlation term, $E_{xc}[n]$. One can show that it is relatively small (at least for atoms and molecules) [2], and an exact expression can be written down [7, 8]

$$E_{xc}[n] = \frac{e^2}{2} \int d\vec{r} n(\vec{r}) \int d\vec{r}' \frac{n_{xc}(\vec{r}, \vec{r} - \vec{r}')}{|\vec{r}' - \vec{r}|}, \quad (8)$$

where

$$n_{xc}(\vec{r}, \vec{r} - \vec{r}') = n(\vec{r}') \int_0^1 d\lambda (g(\vec{r}, \vec{r}'; \lambda) - 1) \quad (9)$$

can be interpreted as the “*exchange-correlation-hole*”. $E_{xc}[n]$ may then be viewed as the Coulomb interaction of the electrons of density $n(\vec{r})$ with the corresponding holes. The function $g(\vec{r}, \vec{r}'; \lambda)$ is the pair-correlation-function of an electron system with density $n(\vec{r})$ and Coulomb interaction $\lambda\phi(\vec{r})$. The integral over λ in eq. (9) connects the electron correlations in the physical system with full Coulomb interaction, whose energy one would like to calculate, and the correlations in the fictitious system without interactions, for which the single-particle Schrödinger equation can be solved.

2.1 Approximations for the exchange-correlation energy

For practical purposes one has to use suitable approximations for the exchange-correlation term, $E_{xc}[n]$. The most commonly applied approximation, the so-called local-density-approximation (LDA), relates the exchange-correlation energy of a non-uniform system locally to that of a uniform electron gas with the same density:

$$E_{xc}^{LDA}[n] = \int d\vec{r} n(\vec{r}) \epsilon_{xc}^{hom}[n(\vec{r})] \quad (10)$$

Here $\epsilon_{xc}^{hom}[n(\vec{r})]$ is the exchange-correlation energy per particle of a homogeneous electron gas with the density $n(\vec{r})$. Quantum-Monte-Carlo results of Ceperley and Alder [9] for the function $\epsilon_{xc}^{hom}[n(\vec{r})]$ have been parameterized by Vosko, Wilk, and Nusair [10]. LDA works very well for systems with extended wavefunctions like bulk metals and semiconductors, but also for small molecules very reliable results can be obtained as is demonstrated in Fig. 2 for the energy surfaces of the low-lying electronic states of O_3 as a function of binding angle [11]. For O_3 the wavefunctions cannot be approximated by a single Slater-determinant. Thus, Hartree-Fock-calculations yield ground states with the wrong symmetry. On the other hand, LDA-calculations

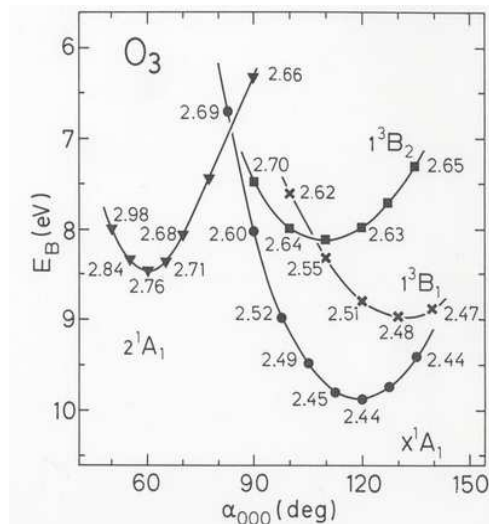


Fig. 2: Energy surfaces of low-lying electronic states with different symmetry for ozone, O_3 , as a function of angle α_{OOO} [11]. The numbers at the calculated points denote the O-O-binding length for the respective configurations.

yield a good description of the electronic ground state with a binding angle $\alpha = 120^\circ$ as well as of the energy difference to the low lying excited state with $\alpha = 60^\circ$.

Recently, more complicated approximations for the exchange-correlation energy have been developed, the so-called general gradient approximations (GGA) (see e.g. [12, 13]) which take into account the spatial variation of the electron density by terms depending on the gradient of the density, $\nabla n(\vec{r})$. With this approximation, reliable results for the conformation of organic molecules, both for free molecules in the gas phase as well as for molecules adsorbed at surfaces can be obtained. As an example we present in Fig. 3 a comparison of LDA and GGA calculations for the structure of a free glycinate radical [14]. While the bond lengths are of similar quality for LDA and GGA, the bond angles improve considerably with the GGA.

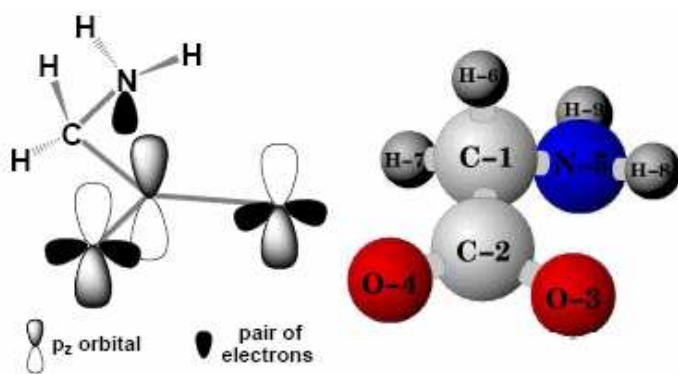


Fig. 3: Conformation of the free glycinate radical $COOCH_2NH_2$ [14]. The sketch on the left shows the character of the occupied electronic levels, especially for the carboxyl group, COO . In the table the calculated bond lengths for glycinate using LDA and GGA are compared with experimental values (from Ref. [15] for glycine $COOHCH_2NH_2$).

	Bondlengths Å		
	LDA	GGA	Exp
$C^1 - C^2$	1.487	1.456	1.520
$C^1 - N^5$	1.428	1.495	1.463
$C^2 - O^3$	1.315	1.256	1.223
$C^2 - O^4$	1.310	1.252	-
$C^1 - H^6$	1.148	1.139	1.098
$C^1 - H^7$	1.138	1.129	-
$N^5 - H^8$	1.067	1.067	1.000
$N^5 - H^9$	1.065	1.062	1.000

	Bond angles (degree)		
	LDA	GGA	Exp
$H^8 - N^5 - H^9$	109.2	111.6	112.5
$H^8 - N^5 - C^1$	111.8	114.9	113.8
$N^5 - C^1 - H^7$	110.9	111.5	111.7
$N^5 - C^1 - C^2$	109.7	111.1	110.6
$H^6 - C^1 - H^7$	94.5	104.8	107.4
$O^3 - C^2 - O^4$	117.5	118.3	-

2.2 Many degrees of freedom

The examples above show that one can indeed obtain reliable results for the conformation and energy surfaces of complex molecules. But how can we guarantee that we find the most stable conformation for the system, if the number of degrees of freedom becomes larger and no symmetry arguments can be used? Hoare and McInnes [16] have correlated the number of (relative) minima with the number of atoms in a cluster, when the interaction of the atoms can be described by a pair-potential of the Lennard-Jones form: $V_{ij}(r) = 4 \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]$ ($r = |\vec{r}_i - \vec{r}_j|$ for each pair of atoms (i, j)). Their result is displayed in Table 1. One can easily imagine that the number of (relative) minima increases exponentially with the number of degrees of freedom. Indeed, Wille and Vennik [17] have shown that for N -atom-clusters which interact via pair po-

Table 1: Number of minima in N -atom clusters with Lennard-Jones interaction [16]

N	6	7	8	9	10	11	12	13
LJ min	2	4	8	18	57	145	366	988

tentials there is no algorithm to determine the structure of the ground state, whose CPU-time demand only grows with a power of N . Such problems are called “NP-complete” by mathematicians, and they are characterized as “intractable” [18]. It is thus impossible to explore all minima, and it is a challenge to find a method which avoids most minima and yields a path to the lowest minimum, the ground state, directly. Kirkpatrick *et al.* [19] have suggested a “simulated annealing” method based on the principles of the Monte-Carlo method. Car and Parrinello [20] have pointed out that a combination of Molecular dynamics with DFT at elevated temperatures can sample a large fraction of the configuration space and filter out the low lying energy minima by following the dynamic trajectories.

3 The combined MD/DFT-method

In order to treat dynamical processes like atomic vibrations or diffusion of atoms one has to formulate and solve the equations of motion for the relevant degrees of freedom. Due to the large atomic mass the motion of the ions can be handled by classical Newtonian dynamics. Within the Born-Oppenheimer approximation the electronic dynamics can be decoupled from the atomic motion. However, the ground state energy of the electrons depends on the atomic configuration, and the forces on the ions are determined by the electron density. In principle, we thus have to solve the electronic selfconsistency problem for every atomic configuration on the atomic trajectory to obtain reliable forces [5]. If one uses the electron density determined selfconsistently for a neighboring configuration to calculate the forces on the ions, one makes an error which depends on the “distance” of the two configurations. The closer the configurations the smaller the error. This is indicated in Fig. 4, where the paths to selfconsistency and the minimal energy surface for two neighboring configurations are plotted. The envelope of all minimal energy surfaces is the “Born-Oppenheimer surface”, which yields the correct forces on the ions for all configurations.

Car and Parrinello [20] have suggested to treat electronic and ionic degrees of freedom on the same footing. They formulated a (fictitious) dynamics for the Kohn-Sham wavefunctions

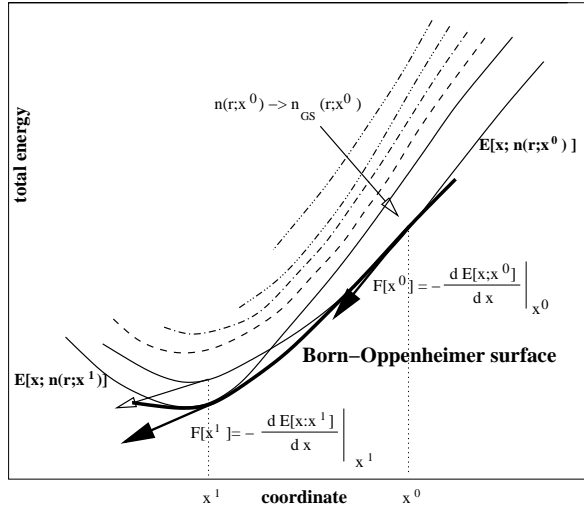


Fig. 4: Schematic one-dimensional sketch of the Born-Oppenheimer surface and paths to selfconsistency. Fixing the atomic configuration to x^0 and calculating the electron density, $n_{GS}(\vec{r}; x^0)$, selfconsistently one reaches the correct ground state (minimal) energy, $E[\vec{x}, n(\vec{r}; x^0)]$, which contains a point on the “Born-Oppenheimer surface” for this configuration. Then the correct forces on the atoms can be calculated ($F(x^0)$, bold arrow at x^0). If one uses the same selfconsistent electron density at the point x^1 one is not on the “Born-Oppenheimer surface” anymore. The forces (open arrow at point x^1) contain a small error. Only if the selfconsistent electron density at the point x^1 is calculated, the “Born-Oppenheimer surface” is reached again and the correct forces, $F(x^1)$, are calculated.

$\{\psi(\vec{r}_i)\}$, which describe the electronic degrees of freedom, and a classical Newtonian dynamics for the ionic degrees of freedom $\{\vec{R}_n\}$. As discussed above, the total energy of the system $E[\{\psi(\vec{r}_i)\}, \{\vec{R}_n\}]$ is a function(al) of both sets of variables. Using the following Lagrangian

$$\begin{aligned} \mathcal{L} = & \sum_i \mu \int_{\Omega} d\vec{r} |\dot{\psi}_i^* \dot{\psi}_i| + \sum_n \frac{1}{2} M_n \dot{\vec{R}}_n^2 \\ & - E[\{\psi(\vec{r}_i)\}, \{\vec{R}_n\}] - E_{ion}(\{\vec{R}_n\}) + \sum_{ij} \Lambda_{ij} \left(\int_{\Omega} d\vec{r} \psi_i^* \psi_j - \delta_{ij} \right) \end{aligned} \quad (11)$$

one obtains the equations of motion

$$\begin{aligned} \mu \ddot{\psi}_i(\vec{r}, t) &= - \frac{\delta E[\{\psi(\vec{r}_i)\}, \{\vec{R}_n\}]}{\delta \psi_i^*(\vec{r}, t)} + \sum_j \Lambda_{ij} \psi_j(\vec{r}, t), \\ M_n \ddot{\vec{R}}_n &= - \nabla_{\vec{R}_n} \left(E[\{\psi(\vec{r}_i)\}, \{\vec{R}_n\}] + E_{ion}(\{\vec{R}_n\}) \right). \end{aligned} \quad (12)$$

Here we have included the ion-ion Coulomb interaction $E_{ion}(\{\vec{R}_n\})$, which in a periodic crystal is calculated by the Ewald-method.

M_n are the ionic masses and μ is a fictitious mass for the electronic degrees of freedom. The Lagrange parameters Λ_{ij} are introduced to guarantee the orthonormality of the wavefunctions $\psi_i(\vec{r}, t)$. With these orbitals and the electron density $n(\vec{r}) = \sum_{i=1}^M 2|\psi_i(\vec{r})|^2$ we determine the energy $E[\{\psi(\vec{r}_i)\}, \{\vec{R}_n\}]$, which enters into the Lagrangian \mathcal{L} , eq. (11), as a classical potential. The fictitious Newtonian dynamics of second order for the wavefunctions $\{\psi(\vec{r}_i)\}$ with masses $\mu \ll M_n$ has two consequences: (i) The time step for the numerical solution of the equations of motion, eq. (12), has to be rather small compared to the one used for regular molecular dynamics. (ii) The transfer of (kinetic) energy between the classical ionic degrees of freedom,

$\{\vec{R}_n\}$, and the quantum mechanical electronic degrees of freedom, $\{\psi(\vec{r}_i)\}$, is very slow. This means that one stays close to the Born-Oppenheimer surface once it has been reached, and since the atomic configuration does not change a lot during one time step, one can proceed for quite a few time steps without necessity to fully relax the electronic system again.

One can solve the combined systems, eqs. (12), for the coupled dynamics of electrons and ions in three ways which we will discuss in the following:

1. Molecular dynamics calculations with the “correct” (LDA) interatomic potential: Here one really solves the equations of motion (12) and analyzes the dynamical trajectories. As an example we present the early calculations of Car and Parrinello [21] for amorphous Si (see section 5).
2. “Simulated annealing” calculations of (LDA) ground state geometries for molecules and small atomic clusters: In this case the trajectories of atoms and wavefunctions are followed at higher temperatures where the phase space can be explored (almost) completely, since (small) barriers between relative minima can be overcome by thermal motion. By slowly cooling the ensemble to low temperatures the low lying minima of the total energy, i.e., the ground state configurations of the clusters can be calculated. As an example we present calculations [22] for selenium-clusters (Se_n , for $n = 3, \dots, 8$) (see section 6).
3. “Molecular relaxation”: One can solve the electronic problem for a given atomic configuration. Using eqs. (3) and (6) for the total electronic energy, $E[\{\psi(\vec{r}_i)\}, \{\vec{R}_n\}]$ we can transform the equation of motion for the wavefunction $\psi(\vec{r}_i)$:

$$\mu \ddot{\psi}_i(\vec{r}, t) = F_i + F_{i,ortho} = - \frac{\delta E}{\delta \psi_i^*(\vec{r}, t)} + \sum_j \Lambda_{ij} \psi_j(\vec{r}, t) = - \mathcal{H} \psi_i(\vec{r}, t) + \sum_j \Lambda_{ij} \psi_j(\vec{r}, t) \quad (13)$$

with the single particle Hamiltonian \mathcal{H} from eq. (5), which contains the effective potential V_{eff} , eq. (7). Thus for $\ddot{\psi}_i(\vec{r}, t) = 0$, eq. (13) is identical (up to a unitary transformation) with eq. (5), and the eigenvalues of the Lagrangian multiplier matrix Λ_{ij} are the eigenvalues of the Kohn-Sham equations (5). The condition $\ddot{\psi}_i(\vec{r}, t) = 0$ is fulfilled, if selfconsistency is obtained, and this defines the Born-Oppenheimer surface for the chosen atomic configuration. Instead of really solving a time dependent equation, one can use any relaxation mechanism to obtain selfconsistency. With the final result for the total energy we can calculate the forces on the atoms according to the right hand side of eq. (12b). We then move the atoms to a new configuration by “Quasi-Newton” methods (see below) and calculate the electronic selfconsistency for each visited configuration. An energy minimum is reached (approximately), if all forces are smaller than a chosen tolerance. We have used this “molecular relaxation” method to determine the reaction paths of adatoms on the As-terminated Si(111) surface [3] (see section 7).

In all cases the “forces” on the wavefunctions have to be calculated many times until selfconsistency is reached. We now turn to the discussion of fast numerical algorithms to achieve this goal.

4 Plane wave pseudopotential calculations

If we are interested in minimal energy atomic configurations, the binding forces are of utmost importance. Only the electronic states of the outer shells, e.g. the overlapping valence states,

substantially contribute to the chemical binding. Thus, it is sufficient to represent the distribution of these valence electrons in the solid with high accuracy.

4.1 Pseudopotentials

The pseudopotential method [23] provides the framework for this purpose. Instead of describing all oscillations of the true valence wavefunctions due to the orthogonalization to the core states, one uses smoothed node-less “pseudo”-wavefunctions inside of a sphere around each ion with a chosen cut-off distance, r_c . By adjusting the potential due to the nuclear charges and the inner shell electrons in the sphere, i.e. replacing the true potential by a “pseudopotential”, one can guarantee that the valence eigenvalues are reproduced by the pseudo-wavefunctions. This idea is schematically represented in Fig. 5a. For the use of pseudopotentials in the framework

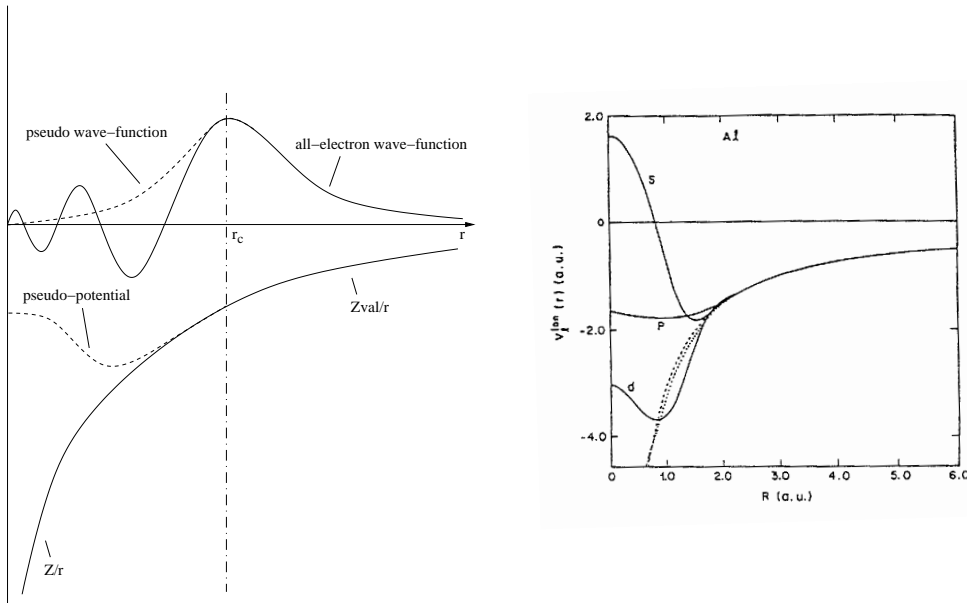


Fig. 5: (a) The concept of pseudopotentials for outer valence electrons. (b) calculated norm-conserving pseudopotentials for Al for $\ell = 0, 1, 2$, constructed according to the Bachelet-Hamann-Schlüter-scheme [25].

of DFT Bachelet, Chiang, Hamann and Schlüter [24, 25] have constructed norm-conserving pseudopotentials for the first time for all elements of the periodic table, and their construction principle is still in use. One has to generate a pseudopotential for each angular momentum ℓ by solving the Schrödinger equation for an isolated atom with DFT. By fixing the logarithmic derivative of the pseudo wavefunction for angular momentum ℓ for a given reference energy (e.g. a valence eigenvalue of an atom) to the correct value, calculated with the true (all electron, ae) wavefunction,

$$L_\ell^{ae}(r_c, \epsilon) = \frac{\partial \ln \phi_\ell^{ae}(r, \epsilon)}{\partial r} \Big|_{r=r_c} = \frac{\partial \ln \phi_\ell^{ps}(r, \epsilon)}{\partial r} \Big|_{r=r_c} = L_\ell^{ps}(r_c, \epsilon), \quad (14)$$

one can ensure that the scattering properties of the atom are correctly described for this ℓ -channel at the reference energy and in an interval around this energy. This criterion is directly connected to the norm-conservation of the wavefunction which is included in the construction principle of the pseudopotential. In Fig. 5b such “norm-conserving” pseudopotentials for

Al for $\ell = 0, 1, 2$ are shown. For each atom the full pseudopotential is a sum over angular momentum components. Following Kleinman and Bylander [26] one separates a common (ℓ -independent) local potential $V_{loc}^{ps}(r)$ and transforms the ℓ -dependent difference potentials $\Delta V_{\ell}^{ps}(r)\hat{P}_{\ell} = (V_{\ell}^{ps}(r) - V_{loc}^{ps}(r))\hat{P}_{\ell}$ (with \hat{P}_{ℓ} the projector to the ℓ -channel) into a separable projector form. The total external potential due to the atoms at sites $\{\vec{R}_n\}$ can then be written as a sum over all atoms and for each atom over angular momentum components

$$V_{ext}(\vec{r}) = \sum_n V^{ps}(\vec{r} - \vec{R}_n) = \sum_n \left(V_{loc}^{ps}(|\vec{r} - \vec{R}_n|) + \sum_{\ell} \left| \hat{T}_{\ell}^{ps} \right\rangle \langle \hat{T}_{\ell}^{ps} \right|_{(\vec{r} - \vec{R}_n)} \right) \quad (15)$$

The smooth pseudo wavefunctions can be expanded in a planewave basis quite efficiently. For periodic solids the pseudo wavefunctions obey Bloch's theorem and can be written as a sum over reciprocal lattice vectors \vec{G}^n :

$$\psi_i^{ps}(\vec{r}) = \psi_{\vec{k},\nu}(\vec{r}) = \sum_n c_{\vec{k},\nu}(\vec{G}^n) \exp \left(i(\vec{k} + \vec{G}^n) \vec{r} \right) \quad (16)$$

If one deals with non-periodic systems (like defects in crystals or surfaces), a periodically repeated “supercell” is used, which has to be large enough to avoid interactions between the defect images. Supercells containing up to a few hundred atoms can nowadays be used with parallel computers. The \vec{k} -vectors are restricted to the first Brillouin zone and the \vec{G}^n -vectors are reciprocal lattice vectors of the periodic system. The basis set used in the expansion, eq.(16), is limited by a cut-off, expressed as the kinetic energy of the plane wave with the largest \vec{G} -vector: $E_{cut} = \frac{\hbar^2}{2m} \vec{G}_{max}^2$.

The simple pseudopotential concept cannot be used for electrons of inner shells, since one needs a very large number of planewaves (E_{cut} large) to represent the rather localized wavefunctions of this type. The same is true for atoms which contain valence electrons occupying an ℓ -shell for the first time. Since these functions are not shielded by inner shells, they penetrate deep into the region of the attractive Coulomb potential of the nucleus. This applies to the atoms of the second period (B, C, N, O, F) with open $2p$ -shells, and transition metals (Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn) with open $3d$ -shells. A generalization of the pseudopotential method has been developed for these cases, the so-called projector-augmented-wave (PAW) method (see e.g. Ref. [27]). We will not deal with these complications here, very recent review articles can be found in the *Handbook of Materials Modeling* [1], in particular the articles by Car *et al.* [28] and Blöchl *et al.* [29].

With the planewave expansion, eq. (16), the single particle Schrödinger equation (5) for the pseudo wavefunctions is transformed into a diagonalization problem of a hermitian Hamiltonian matrix for the determination of the expansion coefficients $c_{\vec{k},\nu}(\vec{G}^n)$:

$$\sum_m H_{\vec{k},\nu}(\vec{G}^n, \vec{G}^m) c_{\vec{k},\nu}(\vec{G}^m) = \epsilon_{\vec{k},\nu} c_{\vec{k},\nu}(\vec{G}^n), \quad (17)$$

where $H_{\vec{k},\nu}(\vec{G}^n, \vec{G}^m)$ are the matrix elements of the Hamiltonian for plane wave states. The dimension of the Hamiltonian matrix for realistic systems (several hundred planewaves per atom and supercells containing up to several hundred atoms) is immense: $N \leq 100000$, and it is clear that we cannot solve such a problem by conventional methods. We can estimate the requirements for a direct diagonalization as follows: The memory necessary to keep a 100000×100000 -matrix is 10000 Mwords, and the number of multiplications for the direct

diagonalization of such a matrix, e.g. by the Householder method, scales as N^3 , which is prohibitively large. But, we only need the lowest M (= number of electrons/2) eigenvalues and eigenvectors, which are occupied in the electronic ground state and establish the ground state density according to eq. (6). They can be calculated by **iterative** methods.

4.2 Iterative methods

Iterative diagonalization methods are based on the following property of hermitian $N \times N$ matrices \mathbf{A} with a dominant eigenvalue λ_1 and associated eigenvector u_1 (see e.g. Ref. [30]): For each starting vector x_0 which has a non-zero projection to u_1 , the series $\{x_k\}$, $k \geq 1$, of the iterated, normalized vectors converges to u_1 :

$$\begin{aligned} k = 0 & \quad x_0; \langle x_0 | u_1 \rangle \neq 0 \\ k \geq 1 & \quad \tilde{x}_k = \mathbf{A} x_{k-1}; x_k = \frac{\tilde{x}_k}{\|\tilde{x}_k\|} \\ k \rightarrow \infty & \quad x_k \rightarrow u_1 \end{aligned} \quad (18)$$

In such iteration procedures we never need the full matrix \mathbf{A} , but only the product vector $\mathbf{A} x_k$. The special properties of the single particle Hamiltonian \mathcal{H} , eq. (5), allow us to calculate this product vector without ever calculating the N^2 matrix elements. It turns out that we have to perform only $\mathcal{O}(N \log_2 N)$ operations. This becomes clear if one inspects the Hamiltonian matrix in k -space:

$$\begin{aligned} H_{\vec{k}, \nu}(\vec{G}^n, \vec{G}^m) &= \delta_{nm} \frac{(\hbar \vec{G}^n)^2}{2m} + \tilde{V}_{loc}^{ps}(\vec{G}^n - \vec{G}^m) + \tilde{\phi}(\vec{G}^n - \vec{G}^m) + \tilde{V}_{xc}(\vec{G}^n - \vec{G}^m) \\ &+ \sum_{\ell} \tilde{S}_{\ell}^*(\vec{k} + \vec{G}^n) \tilde{S}_{\ell}(\vec{k} + \vec{G}^m) \end{aligned} \quad (19)$$

$$\text{with } \tilde{S}_{\ell}(\vec{k} + \vec{G}^n) = 4\pi \sum_{\mu} Y_{\ell\mu}(\Omega_{\vec{k}+\vec{G}^n}) \tilde{T}_{\ell}^{ps}(|\vec{k} + \vec{G}^n|)$$

where the $Y_{\ell\mu}(\Omega_{\vec{k}+\vec{G}^n})$ ($\mu = -\ell, \dots, +\ell$) are the spherical harmonics for angular momentum ℓ , depending on the angles of $\vec{k} + \vec{G}^n$.

This k -space representation of the Hamiltonian can be compared to the real space operator in eq. (5) with the effective potential from eq. (7) and the external pseudopotential from eq. (15). The kinetic energy is a differential operator in r -space, and a diagonal matrix in k -space. The potential terms (local pseudopotential, Hartree potential and LDA exchange correlation potential) which depend on $(\vec{G}^n - \vec{G}^m)$ are convolution operations in k -space, and they transform into simple multiplication operators in r -space. The separable (projection operator) parts of the pseudopotentials are integration operators in r -space, and vector multiplications in k -space. This means that parts of the Hamilton operators are “quasi-diagonal” in conjugated spaces. An operator with such properties is called “pseudo-spectral”. When calculating the product vector $\mathcal{H} \psi$ in the iterative procedure, we can use each of the terms in the space where it requires the least effort, e.g the kinetic energy and the separable pseudopotential in k -space, and the local parts of the potential energy operator in r -space. This can be written symbolically:

$$\mathcal{H} \psi = (\mathcal{H}_{kin} + V_{sep}^{ps}) \psi_{\vec{G}} + (V_{loc}^{ps} + \phi + V_{xc}) \psi_{\vec{r}} \quad (20)$$

Before we can apply the r -space part of the Hamiltonian we have to Fourier transform the wave functions into r -space, and the resulting products back into k -space, where the two terms of

eq. (20) are summed up. The wavefunctions are kept in k -space for further iterations. By means of Fast-Fourier-transform (FFT) we can do the transformations between the conjugated spaces with an effort of $\mathcal{O}(N \log_2 N)$ multiplications (see appendix A and Ref. [31]).

Following Davidson [32] and Kosugi [33], the iteration procedure described above for one eigenvector can be generalized to determine a group of eigenvectors, e.g. the ones with the lowest M eigenvalues which determine the ground state electron density according to eq. (6). Typically, M (= number of electrons/2) is a few percent of the number of basis functions, N . In this method one generates a subspace of dimension $N_{sub} \gtrsim 2M$, which contains the trial vectors for the wanted eigenvectors. One iterates this subspace by applying the Hamiltonian in the spirit of eq. (20), and filters out the eigenvectors with the lowest eigenvalues in every step. During the scheme one has to diagonalize $N_{sub} \times N_{sub}$ -matrices in the subspace, which requires much less effort than for the full $N \times N$ -matrix. In appendix B we give a short account of this method.

The “pseudo-spectral” properties of the single particle Hamiltonian can also be used to calculate the “forces” on the wavefunctions in the dynamical approach, governed by eq. (13), which we repeat here for convenience:

$$\mu \ddot{\psi}_i(\vec{r}, t) = F_i + F_{i,ortho} = -\frac{\delta E}{\delta \psi_i^*(\vec{r}, t)} + \sum_j \Lambda_{ij} \psi_j(\vec{r}, t) = -\mathcal{H} \psi_i(\vec{r}, t) + \sum_j \Lambda_{ij} \psi_j(\vec{r}, t) \quad (21)$$

As can be seen, the forces F_i are given by the application of the single particle Hamiltonian \mathcal{H} on the wavefunction ψ_i . As shown in eq. (20) this can be performed partly in real space and in reciprocal space, depending on convenience. Thus, the close analogy between an iteration step in the diagonalization methods and the force calculation in dynamical methods is obvious. The orthogonalization forces $F_{i,ortho}$ have to be applied, if many eigenfunctions and eigenvalues should be calculated simultaneously. To approach the point in the $M \times N$ -dimensional space of the coefficients $c_{\vec{k},\nu}(\vec{G}^n)$, where all $\ddot{\psi}_i = 0$, and where the electronic energy, eq. (1), is minimal, we can use a simple “steepest descent” method, i.e. change the time dependent Fourier coefficients of the wavefunctions along the forces:

$$c_{\vec{k},\nu}(\vec{G}^n, t + \Delta t) = c_{\vec{k},\nu}(\vec{G}^n, t) + \underbrace{\left(-\sum_m H_{\vec{k},\nu}(\vec{G}^n, \vec{G}^m) c_{\vec{k},\nu}(\vec{G}^m, t) + F_{\vec{k},\nu,ortho} \right)}_{F_{\vec{k},\nu}(\vec{G}^n, t)} \frac{(\Delta t)^2}{\mu} \quad (22)$$

In spite of the high dimension of the vector space $\{\vec{G}^n\}$ it turns out that there are no relative side minima but only **one global** minimum for the electronic energy as function of the $c_{\vec{k},\nu}(\vec{G}^n)$. Thus, we do not have to fight the difficulties described in Section 2.2. To reach the point of vanishing forces we can use very efficient so-called “quasi-Newton” methods [34, 35] which are described in Appendix C. In Ref. [34] you can also find a description of the “conjugate gradient” method which is another well-known alternative to find the extrema of a function(al).

5 Dynamical properties of amorphous Si

Car and Parrinello [21] have first applied the DFT/Molecular Dynamics scheme according to eqs.(12) to amorphous Si, using the “correct” (LDA-) interatomic potentials. Since amorphous

materials do not have translation symmetry, one has to use large (periodically repeated) supercells to model their behavior. If the supercell is large enough, the artificial periodicity should not matter. Car and Parrinello have used an fcc supercell with a volume of 1033 \AA^3 , containing 54 Si-atoms. This setup allows to study the short range atomic order up to a distance of $\approx 6 \text{ \AA}$ (to be compared to the nearest-neighbor distance in Si-bulk of $\approx 2.35 \text{ \AA}$) without perturbation by boundary effects. The aim of the study was the calculation of the microscopic structure of an amorphous glass-like material without empirical parameters.

To represent the electronic wavefunctions a planewave basis-set with $E_{cut} = 5.5 \text{ Ry}$ and one k -point ($\vec{k} = 0$) was used. This is sufficient because Si has a well-behaved pseudopotential and the supercell is rather large. The time step to solve the discretized coupled equations of motion, eqs. (12), was $\Delta t = 1.7 \times 10^{-16}$, one order of magnitude smaller than for usual MD simulations. Because of the small (fictitious) “electronic” mass, μ , the electronic degrees of freedom show a much faster variation than the ions. After the electronic ground state has been determined for the starting configuration of the ions (see below), the equations of motion, eqs. (12), were integrated with the Verlet algorithm [36] (see also contribution by R.G.Winkler). For the electronic part (in reciprocal space) this reads:

$$c_{\vec{k},\nu}(\vec{G}^n, t + \Delta t) = 2c_{\vec{k},\nu}(\vec{G}^n, t) - c_{\vec{k},\nu}(\vec{G}^n, t - \Delta t) + (F_{\vec{k},\nu}(\vec{G}^n, t) + F_{\vec{k},\nu,ortho}) \frac{(\Delta t)^2}{\mu} + \mathcal{O}(\Delta t^4) \quad (23)$$

As mentioned above, due to the small fictitious mass, $\mu = 300 \text{ a.u.}$, the electronic system de-

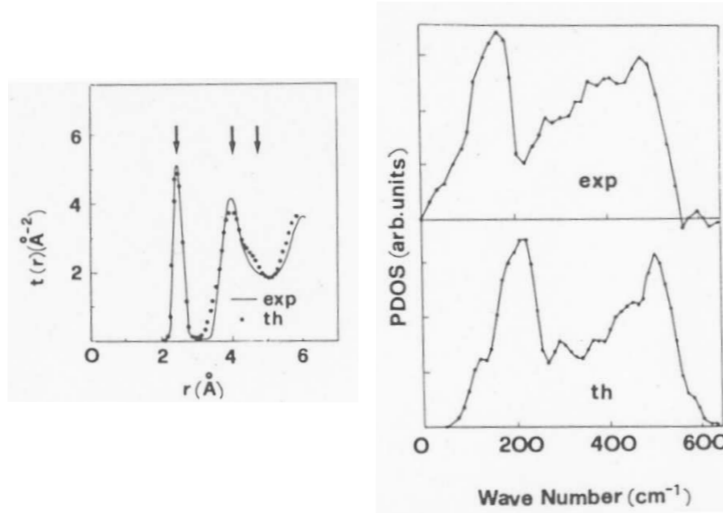


Fig. 6: Comparison of theory and experiment for amorphous Si: (a) pair correlation function $t(r) = 4\pi\bar{\rho}rg(r)$ (with $\bar{\rho}$ the average particle density), and (b) phonon density of states (PDOS).

viates from the ground state by a very small energy (a few tenths of eV) over several 1000 time steps. Thus the dynamics of the 54 Si atoms is realistically modeled for an appreciable time without forcing the electrons to the ground state (Born-Oppenheimer-surface) by the complicated and time consuming selfconsistency procedure described in Section 4.2. This “energetic decoupling” is the main advantage of the combined DFT/MD method. With the solution of the equations of motion the “diagonalization” of the electronic Hamiltonian, the selfconsistency, and the ionic motion are handled simultaneously.

Car and Parrinello started with a diamond-like structure for the 54 Si atoms, for which the electronic ground state was determined with the methods discussed in Section 4.2. Then the system was brought to $T = 2200\text{K}$, where it melted. As in classical MD, the temperature is determined by the average kinetic energy of the ions, which can be controlled by a scaling of

the velocities (see also H.G.Winkler). Then the temperature was reduced to 300 K by reducing the kinetic energy during several thousand time steps to obtain an amorphous solid system. At this temperature the evolution of the system was followed for a total time of $t = 2 \times 10^{-12}$ s. To characterize the structure the time averages of the pair correlation function and of the velocity-autocorrelation function (equivalent to the phonon density of states) were calculated. In Fig. 6 the results are compared to experiments. The agreement is quite remarkable. This demonstrates that thermodynamic properties of real materials can be calculated with parameter-free (“ab initio”) quantum methods.

6 Simulated annealing studies of Se clusters

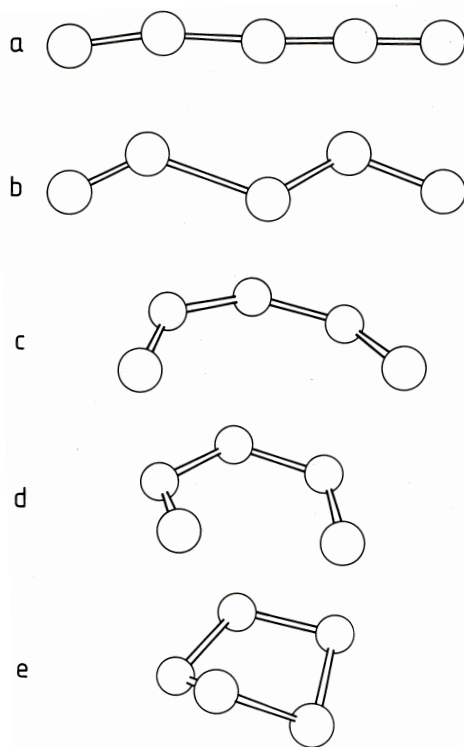


Fig. 7: Time evolution of a Se_5 cluster calculated by DFT/MD, starting from the linear chain shown in (a). The end configuration (e) is the lowest energy configuration of Se_5 .

Another example of the application of the DFT/MD method is the study of low energy configurations of small atom clusters. In particular, the clusters of (partially) covalently bonding atoms show a large variety of configurations with very different bonding characteristics, but similar energies. Selenium is a specific example [22], which forms chains in the stable solid phase. The simulation for Se_n -clusters ($n = 3, \dots, 8$) was performed in a supercell with a volume of 1000 \AA^3 . A planewave basis-set with one k -point ($\vec{k} = 0$) and $E_{\text{cut}} = 14 \text{ Ry}$ was used. The fictitious mass was set to $\mu = 1800 \text{ a.u.}$, and the time step $\Delta t = 3.5 \times 10^{-16}$. In an “arbitrary” starting geometry the electronic degrees of freedom were relaxed to selfconsistency, i.e., the system was brought to the Born-Oppenheimer surface. During the time evolution the system stayed there very accurately. Finite temperatures were used to overcome small energy barriers between relative minima. It might be interesting to know that for Se_5 one time step required 1 CPU-s on the CRAY-XMP (with 100 MFLOPS), and ≈ 10.000 time steps were needed to come to the final configuration. To probe different minima of the energy surface the temperature was

varied and slowly reduced to $T = 0$ (simulated annealing). Depending on n , the Se_n -clusters showed very different stable structures: Up to $n = 4$, planar configurations were found to be most stable. Starting with $n = 5$, three-dimensional structures were obtained. For all structures the bond lengths and angles are analyzed. Excellent agreement is found for the cluster geometries which have been measured (Se_6 and Se_8). Also, the electronic structure can be analyzed for the stable configurations, and a good understanding of the electronic distribution for the highest occupied (HOMO) and lowest unoccupied (LUMO) orbitals is obtained.

Fig. 7 shows a sequence of configurations obtained by following the time evolution of a Se_5 cluster by DFT/MD, starting from the linear chain of Fig. 7a. Several side minima were found (b - d) during the time development, and the final most stable configuration is a closed ring (three-dimensional) with C_S -symmetry.

7 Surfactant mediated growth: Kinetics of adatoms

As a third example we discuss the possible reactions on a semiconductor surface during epitaxial growth, and how to calculate the kinetic parameters for adatoms. The design of modern electronic devices often requires new materials combinations and the control of interfaces down to the atomic scale. A prominent example is the Ge/Si-system. One would like to combine the perfection of large Si wafers with the larger electron mobility in Ge. To do this one has to deposit a defect free Ge film on a Si substrate. This is not an easy task because due to the larger lattice constant ($a_{\text{Ge}}/a_{\text{Si}} = 1.04$) Ge grows on clean Si surfaces in the form of large three-dimensional islands. However, if a single-atom layer of group-V atoms (As, Sb, Bi) is placed on top of the Si substrate before the Ge atoms are deposited, one can obtain atomically flat Ge films. The group-V atom layer (called “surfactant” layer, short for surface active layer) floats on top of the growing Ge film, and enforces a layer-by-layer growth of Ge on Si. This means that all deposited Ge(Si) atoms are incorporated under the surfactant layer. The group-V atoms prevent the complicated reconstruction of the clean group-IV surfaces, and they lower the surface energy. As shown in Fig. 8 the equilibrium structure of As-covered Si(111)

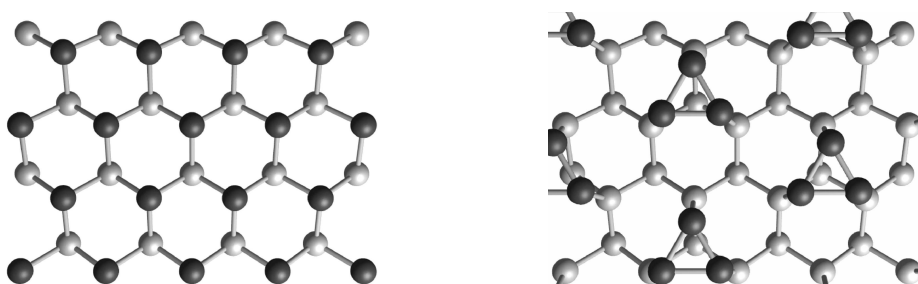


Fig. 8: Equilibrium surface structures of surfactant-covered Si(111): (a) with As-coverage the (1×1) -structure is found, the As atoms constitute the upper layer of the terminating double layer; (b) with Sb (or Bi)-coverage the $(\sqrt{3} \times \sqrt{3})$ -structure is found, the Sb (Bi) atoms form T4-centered trimers on top of a Si double layer. Si atoms are represented by light spheres, As, Sb, Bi atoms by dark spheres.

is the unreconstructed (1×1) surface whereas the Sb and Bi-covered Si(111) surfaces show a $(\sqrt{3} \times \sqrt{3})$ surface structure instead of the (7×7) -reconstruction of the clean Si(111) surface. Due to the extra electron the group-V atoms saturate their valence shell with three bonds. On

the surface they have an occupied lone pair instead of the energetically costly single electron dangling bond of the group-IV atoms at the surface (which are saturated by the reconstruction of clean surfaces).

In addition, the adatom kinetics is changed by the surfactants. The quantitative analysis of experiments [37, 38] reveals the following facts for surfactant-mediated epitaxy: (i) Si homoepitaxy proceeds via nucleation of islands on the flat terraces, and the density of nucleated islands for Si homoepitaxy is much higher with surfactants than without [37]. (ii) Ge heteroepitaxy proceeds with a competition between step flow growth and island nucleation on the flat terraces. (iii) The strain due to the 4% lattice mismatch is released by a dislocation network at the Ge/Si-interface after a Ge film of ≈ 10 layers has been grown [38].

The different growth behavior of Si and Ge can be understood on the basis of a different effective diffusion barrier for the two atom species: For Si adatoms the effective barrier should be larger with surfactants than without, resulting in a shorter effective diffusion path, which in turn leads to a high nucleation density of islands. On the contrary, for Ge the effective barrier should be small. Then the Ge adatoms can reach the (very distant) terrace steps. To substantiate these ideas and understand the experimental results we have calculated the kinetic parameters for deposited adatoms (Ge and Si) on As- and Sb-covered Si(111) [3, 39, 40, 41, 42]. Since the surface with an adatom is not periodic anymore, we have to use a supercell approach. All calculation were performed in a (3×3) surface supercell. For the plane wave basis we used 4 k -points out of the surface Brillouin-zone and a cut-off $E_{cut} = 13.69$ eV. As shown in Fig. 9, we used iterative methods to obtain minimum energy configurations.

For the surfactant-covered surfaces we have to consider two reaction paths: (i) the adatom diffusion step on top of the surfactant layer, and (ii) the incorporation step, which brings the group-IV atoms under the surfactant layer and lifts the surfactant atoms to a new on-top position.

Search for a reaction path

The calculation of a reaction path and the reaction barrier requires to determine the starting configuration, $\{\vec{R}_n\}^0$, and the end configuration, $\{\vec{R}_n\}^1$, of the path, and the saddle point, $\{\vec{R}_n\}^S$, for the reaction. The saddle point is the configuration with the lowest maximal energy of all possible paths connecting $\{\vec{R}_n\}^0$ and $\{\vec{R}_n\}^1$, and the reaction barrier for the reaction from $0 \rightarrow 1$ is the energy difference between starting configuration and saddle point

$$E_r(0 \rightarrow 1) = E_{pot}(\{\vec{R}_n\}^S) - E_{pot}(\{\vec{R}_n\}^0),$$

with $E_{pot}(\{\vec{R}_n\}) = E[\{n_{GS}(\vec{r})\}, \{\vec{R}_n\}] + E_{ion}(\{\vec{R}_n\})$ (24)

the potential energy of the system. The starting and end configurations are (meta)stable positions, i.e. (relative) minima of the potential energy contour with respect to the ionic degrees of freedom. We can determine these minima in the way discussed in Sections 3 and 4. The full procedure is summarized in the hierarchy of iterative loops depicted in Fig. 9.

Of course, we have to make sure that we find the most significant (lowest) minimum for the adatom configurations. Since the surfactant covered surfaces still show a high symmetry with only few special points (see Fig. 8), we can construct possible diffusion paths almost by inspection. The actual equilibrium (minimum energy) configuration can then be found by starting from a few “most asymmetric” configurations in the surface unit cell. If all starting configurations yield the same minimum, we can be pretty sure that we have found the correct one. As the equilibrium site for Si and Ge adatoms on top of the As-covered Si(111) surface we find the “H3”-position (in the middle of the hexagon formed by the outer double layer, see Figs. 8 and 1a).

- (1) **start:** choose atomic configuration
- (2) **start:** calculate selfconsistent electron density with DFT methods
by solving the Kohn-Sham equations with quasi-Newton methods
- (3) solve eigenvalue problem with Davidson-Kosugi method
- (3) **end:** eigenvalue problem ($\delta\varepsilon < tol_\varepsilon$)
- (2) **end:** electron density ($\delta n(r) < tol_n$) \rightarrow selfconsistency reached
- (1) calculate forces on atoms
and find new atomic configuration with quasi-Newton methods
- (1) **end:** atomic configuration ($\max(\vec{F}_n) < tol_F$) \rightarrow minimum reached

Fig. 9: Flow chart with iteration loops to determine minima of $E_{pot}(\{\vec{R}_n\})$

The diffusion path is a connection between two equivalent equilibrium configurations in neighboring unit cells. Several configurations on the Si-adatom diffusion-path on As-covered Si(111) are shown in Fig. 1. When calculating the actual transition path between two minima, we have to prevent the system from relaxing to the minima at either end of the path. This can be achieved by introducing the condition (“constraint”), that the system is allowed to vary the configuration only in a hyperplane perpendicular to the connecting vector between the path terminating minima, e.g. by projecting the forces on the atoms to this plane. Introducing the hypervectors for the ionic degrees of freedom

$$\begin{aligned} \text{coordinates} \quad |\vec{X}\rangle &= |\vec{R}_1, \vec{R}_2, \dots, \vec{R}_N\rangle, \\ \text{forces} \quad |\vec{F}\rangle &= |\vec{F}_1, \vec{F}_2, \dots, \vec{F}_N\rangle, \end{aligned} \quad (25)$$

we can formulate this constraint as

$$\begin{aligned} |\vec{F}\rangle_{plane} &= |\vec{F}\rangle - \frac{\langle \Delta \vec{X} | \vec{F} \rangle}{\langle \Delta \vec{X} | \Delta \vec{X} \rangle} |\Delta \vec{X}\rangle \\ \text{with} \quad |\Delta \vec{X}\rangle &= |\vec{X}^1 - \vec{X}^0\rangle. \end{aligned} \quad (26)$$

This reduces the $3N$ -dimensional configuration space to a $(3N - 1)$ -dimensional hyperplane. To find the saddle point we have to determine the plane with the energetically highest minimum under the constraint formulated in eq. (26). Using the projected forces, eq. (26), we can follow the iteration procedure in Fig. 9 also for the minimization of the (constrained) energy on the hyperplanes. For the diffusion path the constraint, eq. (26), mostly concerns the moving adatom, since the surfactant-terminated surface is almost inert and the atoms do not move a lot during the diffusion jump of the adatom.

We define equidistant points on the straight connection between neighboring H3-positions and minimize the potential energy on the orthogonal hyperplanes going through these points. The configuration with the highest minimal energy is an approximation to the saddle point of the diffusion path. The residual forces perpendicular to the plane give an estimate of the quality of the approximation. We can improve the quality by applying the same procedure to the interval between the two points with the two highest minimal energies. To check that the calculated saddle point is indeed dividing the attraction basins of the two minima, we can back-relax the system from neighboring configurations *without constraint*. If it always ends up in one of the minima, we have indeed found a valid saddle point. As a saddle point for the diffusion of Si and Ge adatoms on As-terminated Si(111) we find the “T4”-position (adatom above the Si atoms of the second surface layer, see Fig. 1c). The diffusion activation energies are very similar for Si, Ge (and also Sn) adatoms, we find $E_D = 0.25$ eV [40].

The calculations are much more complicated for the incorporation of the adatoms under the surfactant layer. In this case, we have to consider a two-atom reaction. The starting configuration for the reaction on the As-covered Si(111) surface again is the H3-adatom configuration $\text{Si}_{\text{add}}(\text{H3})$. In the final configuration the adatom replaces a surfactant atom which has to move away. We can determine this final configuration again with unconstrained minimization by putting the extra Si (or Ge) atom to a substitutional position in the first surface layer, and locate the replaced As atom in an on-top position in the same surface unit cell. The minimization of the energy yields indeed an almost ideal substitutional site, Si_{sub} , for both group-IV atoms and an adatom bridge-position for the As atom with one bond to the group-IV atom and a second one to a neighboring As atom (see Fig. 10d). Both Si and Ge atoms gain energy when incorporated (however Sn does not!), the binding energies are 0.80 eV and 0.20 eV for Si and Ge, respectively [40].

But how does the adatom get to the substitutional position, i.e. what is the reaction path? The first guess is to start from the straight connection path between the H3-position for the Si adatom (plus substitutional As) and the substitutional Si (plus bridge As-adatom) and apply the energy minimization with constraints. Here the path is determined by the coordinated motion of two atoms (Si adatom and replaced As atom), and the projection of the forces, eq. (26), effects mainly these two atoms. The hyperplane effectively remains a 5-dimensional space (plus small displacements of the neighboring atoms to adjust bonds), and one can imagine that complications might arise. Indeed, when we check the approximately calculated saddle point configuration starting from the straight path by back-relaxation without constraints, we find new (relative) minimum configurations of the potential energy. Thus, we have to apply the search for the saddle points for all possible paths between these minima and find the path with the lowest saddle points connecting the adatom configuration Si_{add} to the substitutional configuration Si_{sub} . It turns out that the optimal path goes via four extra minima (see Fig. 11b). The reason is that on this path the system can avoid to break more than one bond at a time, and thus the energy barriers are much lower than close to the straight path.

In such complicated energy landscapes it is not easy to guarantee that the path found by the end points of the minimization on the hyperplanes with the constraints, eq. (26), is continuous. One has to determine the relevant minima carefully, and use rather closely neighboring hyperplanes, especially close to the saddle points. Other people try to overcome this problem by the “nudged band method” [43], which uses an additional constraint by introducing an energy term, which depends on the distance of the discrete points visited on the path.

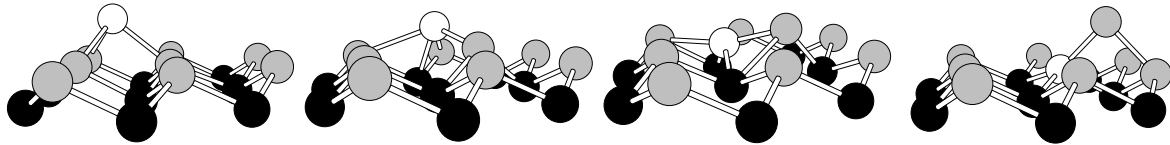


Fig. 10: Exchange path of Si-As pair on Si(111):As in perspective view (Si adatom white, As atoms gray, bulk Si atoms black); (a) on top equilibrium configuration $\text{Si}_{\text{add}}\text{-As}_{\text{sub}}$; (b) Si-As-dimer parallel to surface (side minimum); (c) saddle point configuration of the Si-As-pair before full re-bonding has taken place; (d) substitutional equilibrium configuration $\text{Si}_{\text{sub}}\text{-As}_{\text{add}}$.

In Fig. 10 three-dimensional plots of some configurations encountered on the exchange path from Si_{add} to Si_{sub} are shown. Although the exchange path is the same for Ge adatoms, the energies on the path differ substantially, as can be seen in Fig. 12. For Si the highest barrier

is found in the first section, and $E_{ex}(Si) = 0.27$ eV is comparable to $E_D(Si)$. However, when the Ge atom approaches the Si layer, the energy rises above that of the Si adatom. On this last section of the exchange path the energy for the Ge-As-exchange is shifted upward almost rigidly by ~ 0.5 eV compared to the Si-As-exchange, nearly the same amount by which the binding energies of the incorporated atoms differ. For Ge adatoms $E_{ex}(Ge) = 0.71$ eV is much larger than $E_D(Ge)$.

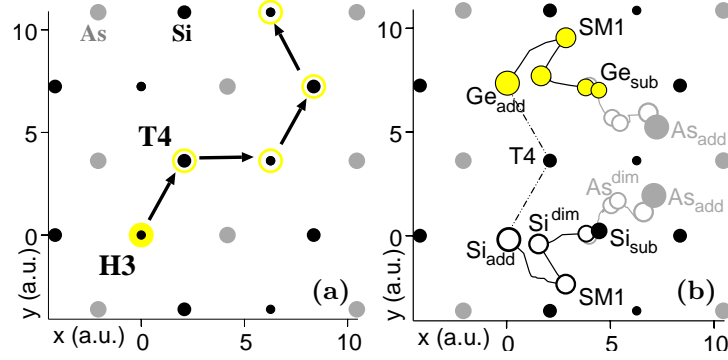


Fig. 11: Topview of the adatom diffusion and exchange paths for group-IV atoms on Si(111):As [40]. (a) The diffusion path starts at the equilibrium H3-site (small black dots), goes via the T4 saddle point (medium black dots) and ends at the H3-site in a neighboring cell. As atoms are large gray dots.- (b) The exchange paths for Si (lower half of figure) and Ge (upper half) adatoms are very similar. The exchanging pairs visit several minima (open circles, whose sizes indicate the height above the surface).

The results of the calculations are summarized in the Figs. 11 and 12, and in Table 2. The kinetic behavior of the adatoms is determined by the competition of on-top diffusion and exchange reaction, whose temperature dependence is governed by the activation energies listed in Table 2.

Table 2: Energies (in eV; relative to H3 adatom position) for group-IV adatoms on Si(111):As and Si(111):Sb in different configurations [39, 40]

surface struct	Si(111):As			Si(111):Sb	
	(1×1)			(1×1)	$\sqrt{3} \times \sqrt{3}$
Conf.	Si	Ge	Sn	Si	Si
E_D	0.25	0.25	0.23	0.21	0.55
E_{ex}	0.27	0.71		0.60	0.20
$E_{sub} = -E_B$	-0.80	-0.20	0.44	-1.50	-0.80
E_{reex}	1.07	0.91		2.10	1.00
$E_{D,eff}$	1.07	0.91		0.21	1.00

For Si the exchange barrier is as small as the diffusion barrier, and from each H3 equilibrium site the Si adatom can undergo either reaction. Due to the large energy gain for the substitutional configuration, it will spend most of the time there. The effective diffusion step is thus the re-exchange step with an activation energy of ≈ 1 eV (compared to 0.6 eV on clean Si(111) [37]). This explains the effective slowing down of the Si adatom diffusion due to the As coverage

and the high nucleation density during homoepitaxy on Si(111):As. On the contrary, for Ge the exchange barrier is much larger than the diffusion barrier. One can estimate, that at typical growth temperatures ($T \approx 500^\circ\text{C}$) a Ge adatom on the average makes about 500 jumps before it is incorporated (and even more at lower temperatures). Thus it has a good chance to visit terrace steps, where it also can be incorporated. This explains the experimentally found mixed growth mode for Ge on Si(111):As by step flow and island nucleation. A similar argument holds for Sb-covered Si(111).

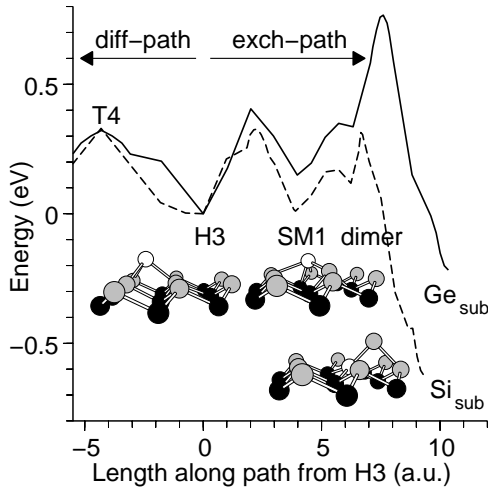


Fig. 12: Energy (relative to H3 position of adatoms) along the diffusion path (to the left of H3) and along the exchange path to the substitutional position (to the right). Compared are the energies for Si (dashed line) and Ge (full line) on Si(111) at a_{Si} [40]. The insets show configurations for important minima: H3: adatom equilibrium; dimer: Si(Ge)-As-dimer on top of As-layer; Si(Ge)_{sub}: substitutional Si(Ge) and As adatom (Si-bulk black, As gray, adatom white).

In Table 2 we have included results for diffusion of Si on Si(111):Sb [39]. The different values found for the diffusion and exchange barriers on the different surface structures ((1×1) and $(\sqrt{3} \times \sqrt{3})$) help to understand that during homoepitaxial growth on Si(111):Sb islands with both types of surface structures are found [39, 41].

Appendices

A Fast Fourier Transform

Fast-Fourier-Transform [31] finds applications in many numerical procedures, e.g. solving linear differential equations. In principle, it was already familiar to C.F. Gauss, but has not been used extensively up to 1965, since for one-dimensional transforms with $N = 12, 24$, or 36 planewaves it does not have big advantages. But with the onset of broad use of high-performance computers, FFT was re-discovered by Cooley and Tukey, and has revolutionized the numerical world. Today, vectorized and parallel FFT program packages can be found in every good on-line computer library.

How many multiplications are needed for the discrete Fourier transform of N data points f_j ?

$$F_k = \sum_{j=0}^{N-1} e^{2\pi jk/N} f_j = \sum_{j=0}^{N-1} W^{kj} f_j \quad (27)$$

At first sight it seems to be $\mathcal{O}(N^2)$ as for each matrix multiplication. But, if N can be factorized as a product of natural numbers, this is not so! As an example we will demonstrate this for

$N = 2^l$. Then F_k can be written as a sum of two Fourier transforms of length $N/2$, one with the even-indexed points and one with the odd indexes:

$$\begin{aligned}
 F_k &= \sum_{j=0}^{N-1} e^{2\pi i j k / N} f_j \\
 &= \sum_{j=0}^{N/2-1} e^{2\pi i k (2j) / N} f_{2j} + \sum_{j=0}^{N/2-1} e^{2\pi i k (2j+1) / N} f_{2j+1} \\
 &= \sum_{j=0}^{N/2-1} e^{2\pi i k j / (N/2)} f_{2j} + W^{k,1} \sum_{j=0}^{N/2-1} e^{2\pi i k j / (N/2)} f_{2j+1} \\
 &= F_k^g + W^{k,1} F_k^u
 \end{aligned} \tag{28}$$

One thus has reduced the N -point Fourier transform to two $(N/2)$ -point transforms (and a multiplication). This reduction can be used recursively for each of the terms of eq. (28) until one ends up with transformations of length unity (the identity). With every further reduction the number of terms doubles and an additional multiplication has to be done. Finally, one arrives at an effort of $N \log_2 N$ multiplications. We leave the proof (e.g. by induction) to the reader. The difference between $N \log_2 N$ and N^2 multiplications is immense for large numbers: for $N = 10^6$ we find 10^8 vs. 10^{12} multiplications, which translates to a fraction of a second vs. an hour of CPU-time on a fast processor!

B Davidson Kosugi method

Davidson and Kosugi [32, 33] have suggested an iterative procedure to determine a group of eigenvalues and the corresponding eigenvectors of a real symmetric matrix \mathbf{A} simultaneously. It can also be applied to hermitian matrices like the Hamiltonian matrix, eq. (19). The method is basically the Ritz iteration method to correct trial vectors simultaneously using correction vectors in a subspace of fixed dimension N_{sub} . We describe the method used to determine the M lowest eigenvalues:

1. Initialization ($m = 0$) of N_{sub} vectors:

- (a) Choose M vectors and eigenvalues $\{|\psi_\nu^{[0]}\rangle, \epsilon_\nu^{[0]}\}$, ($\nu = 1, M$) as starting vectors; e.g. the eigenvectors corresponding to the lowest eigenvalues of a submatrix \mathbf{A}_s of \mathbf{A} with dimension $> N_{sub}$, calculated by diagonalization.

2. Iterations, ($m \geq 0$)

- (a) calculate the residues: $|R[\psi_\nu]\rangle^{[m]} = (\mathbf{A} - \epsilon_\nu^{[m]})|\psi_\nu^{[m]}\rangle$
- (b) $\|R[\psi_\nu]^{[m]}\| < \delta$: convergence reached, leave iteration cycle, go to (3).
- (c) Condition the residue vectors with a suitable diagonal matrix \mathbf{D} , e.g. the diagonal elements of \mathbf{A} : $|\psi_\nu^{corr,[m]}\rangle = \frac{1}{\mathbf{D}_{-\epsilon^{[m]}}} |R[\psi_\nu]\rangle^{[m]}$.
- (d) Orthonormalize $|\psi_\nu^{[m]}\rangle, |\psi_\nu^{corr,[m]}\rangle$, and $N_{sub} - 2M$ unit vectors. These vectors define the basis vectors $|e^\alpha\rangle$, ($\alpha = 1, N_{sub}$) for the m th iteration of the subspace.

- (e) Diagonalize $\langle e^\beta | \mathbf{A} | e^\alpha \rangle$. The eigenvectors corresponding to the M lowest eigenvalues are the starting vectors for the next iteration step: $\{\psi_\nu^{[m+1]}, \epsilon_\nu^{[m+1]}\}, (\nu = 1, M)$. Go to (a).

3. **Orthonormalization.** After convergence orthonormalize the converged wavefunctions: $\langle \psi_\nu, \psi_\mu \rangle = \delta_{\nu\mu}, (\nu, \mu = 1, M)$.

During the iteration the basis vectors of the subspace are adjusted in every step by adding orthogonal correction vectors to the ones corresponding to the lowest eigenvalues. In this way the full N -dimensional vector space is sampled successively by “rotating” the subspace.

C Quasi-Newton methods

“Quasi”-Newton methods are used for minimization or optimization problems when the Hessian matrix of second derivatives of the energy function cannot be used (e.g. because it would be too costly to calculate it). Then one has to find a way to iterate the Hessian matrix along the path sampled by just using the calculated values of the energy function and its first derivatives. An account of quasi-Newton methods and the possible ways to iterate the Hessian can be found in *Numerical Recipes* [34] and more explicit in *Computational Methods for Optimization* [35]. Here we will just give a short description of the ideas behind the quasi-Newton methods.

We first shortly repeat the regular Newton-Raphson scheme. Suppose one can expand the energy function $E[\vec{x}]$ where \vec{x} is a vector with many components $\vec{x} = (x_1, x_2, \dots, x_N)$ around a point $\vec{x}^{(m)}$ (close to the minimum \vec{x}^*) to second order:

$$E[\vec{x}] = E[\vec{x}^{(m)}] - \langle \vec{f}^{(m)} | \vec{x} - \vec{x}^{(m)} \rangle + \frac{1}{2} \langle \vec{x} - \vec{x}^{(m)} | \mathbf{H}^{(m)} | \vec{x} - \vec{x}^{(m)} \rangle \quad (29)$$

with $\vec{f}^{(m)} = - \frac{\delta E}{\delta \vec{x}} \big|_{\vec{x}^{(m)}} \quad \text{and} \quad \mathbf{H}^{(m)} = \frac{\delta^2 E}{\delta \vec{x}^2} \big|_{\vec{x}^{(m)}}$

To estimate the minimum we use the condition that the first derivative of the energy has to be zero. This yields a next estimate \vec{x} of the minimum point.

$$\begin{aligned} \frac{\delta E}{\delta \vec{x}} &= 0 = -|\vec{f}^{(m)}\rangle + \mathbf{H}^{(m)}|\vec{x} - \vec{x}^{(m)}\rangle \\ |\vec{x}\rangle &= |\vec{x}^{(m)}\rangle + (\mathbf{H}^{(m)})^{-1} |\vec{f}^{(m)}\rangle \end{aligned} \quad (30)$$

Of course, eq. (30) yields the correct minimum only if the starting point is so close to the minimum, that the derivatives of the energy function at $\vec{x}^{(m)}$ are exactly the ones at \vec{x}^* , especially $\mathbf{H}^{(m)} = \mathbf{H}^*$. (In principle, this is true if we expand around the minimum). Otherwise, we could proceed with an iteration by expanding the energy at point $\vec{x}^{(m+1)} = \vec{x}$, and so on.

For each iteration step we have to make sure that the energy decreases, which means that the Hessian matrix has to be positive-definite. If we start far from a minimum, this cannot be ensured for the actual second derivative matrix of the energy function. Thus it would be desirable to find a way to construct approximations to the Hessian with the required positive definiteness. Close to the minimum the updating formula should approach the true Hessian and we thus can enjoy the quadratic convergence of the Newton-Raphson scheme. In addition, in many cases the Hessian matrix cannot be calculated explicitly because the CPU and/or the memory requirements are too high. Then one has to iterate the Hessian by only using the information on the

energy function and its first derivative at the points visited. This explains the name “quasi”-Newton methods because we do not use the actual Hessian calculated from the energy function to find the next iteration point, but an updated approximation of it. Such a procedure can be established in the following way: Suppose we write down eq. (30) for two consecutive iteration points, $\vec{x}^{(m)}$ and $\vec{x}^{(m+1)}$, assuming that we have found the correct Hessian \mathbf{H}^* already, and then subtract the two equations. This yields

$$|\vec{x}^{(m+1)}\rangle - |\vec{x}^{(m)}\rangle = (\mathbf{H}^*)^{-1} \left(|\vec{f}^{(m+1)}\rangle - |\vec{f}^{(m)}\rangle \right) \quad (31)$$

We now require that the new update of the Hessian matrix, $(\mathbf{H}^{(m+1)})^{-1}$ (calculated after point $\vec{x}^{(m+1)}$ has been visited) obeys eq. (31), and can be written as a correction to the previous approximation $(\mathbf{H}^{(m)})^{-1}$. One also has to avoid storing N^2 elements for the Hessian, and thus a correction in the form of a dyadic (or “outer” or “direct”) product of two vectors (which requires to store $\mathcal{O}(N)$ elements for each iteration step) would be most convenient:

$$(\mathbf{H}^{(m+1)})^{-1} = (\mathbf{H}^{(m)})^{-1} + \sum_k g_k |\vec{u}_k^{(m+1)}\rangle \otimes \langle \vec{u}_k^{(m+1)}| \quad (32)$$

There are several ways to construct the update vectors $\vec{u}^{(m+1)}$ and the weights g_k from the difference vectors $\Delta\vec{x}^{(m+1)} = \vec{x}^{(m+1)} - \vec{x}^{(m)}$ and $\Delta\vec{f}^{(m+1)} = \vec{f}^{(m+1)} - \vec{f}^{(m)}$. They are named after the authors. The most commonly used ones are due to Davidon-Fletcher-Powell (DFP) and Broyden-Fletcher-Goldfarb-Shanno (BFGS). The DFP updating formula reads

$$\begin{aligned} \text{DFP: } (\mathbf{H}^{(m+1)})^{-1} &= (\mathbf{H}^{(m)})^{-1} + \frac{|\Delta\vec{x}^{(m+1)}\rangle \otimes \langle \Delta\vec{x}^{(m+1)}|}{\langle \Delta\vec{x}^{(m+1)} | \Delta\vec{f}^{(m+1)} \rangle} \\ &\quad - \frac{|\mathbf{H}^{(m)} \Delta\vec{f}^{(m+1)}\rangle \otimes \langle \mathbf{H}^{(m)} \Delta\vec{f}^{(m+1)}|}{\langle \Delta\vec{f}^{(m+1)} | \mathbf{H}^{(m)} \Delta\vec{f}^{(m+1)} \rangle}, \end{aligned} \quad (33)$$

whereas the BFGS updating can be written as

$$\begin{aligned} \text{BFGS: } (\mathbf{H}^{(m+1)})^{-1} &= (\mathbf{H}^{(m)})^{-1} \\ &+ \frac{|\Delta\vec{x}^{(m+1)}\rangle \otimes \langle \Delta\vec{x}^{(m+1)}|}{\langle \Delta\vec{x}^{(m+1)} | \Delta\vec{f}^{(m+1)} \rangle} \left(1 + \frac{\langle \Delta\vec{f}^{(m+1)} | \mathbf{H}^{(m)} \Delta\vec{f}^{(m+1)} \rangle}{\langle \Delta\vec{x}^{(m+1)} | \Delta\vec{f}^{(m+1)} \rangle} \right) \\ &- \frac{|\Delta\vec{x}^{(m+1)}\rangle \otimes \langle \mathbf{H}^{(m)} \Delta\vec{f}^{(m+1)}| + |\mathbf{H}^{(m)} \Delta\vec{f}^{(m+1)}\rangle \otimes \langle \Delta\vec{x}^{(m+1)}|}{\langle \Delta\vec{f}^{(m+1)} | \mathbf{H}^{(m)} \Delta\vec{f}^{(m+1)} \rangle} \end{aligned} \quad (34)$$

We leave it to the reader to show that both forms indeed satisfy eq. (31). From eq. (32) one can see immediately that the approximated Hessian stays symmetric if one starts with a symmetric form. Usually, one starts with a term proportional to the unit matrix

$$(\mathbf{H}^{(0)})^{-1} = \alpha \mathbf{1}. \quad (35)$$

Inserting (35) into eq. (30) one can see that with this form the configuration change is strictly along the force and the parameter α sets the scale for the change. If no update of the Hessian is made, this is the “steepest descent” method. In any case, α has to be adjusted for each problem to get optimal performance of the iteration procedure. For further discussions see Refs. [35, 34].

References

- [1] S.Yip (Ed.), *Handbook of Materials Modeling*, (Springer, Berlin 2005).
- [2] R.O.Jones and O.Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
- [3] K. Schroeder, S. Blügel, B. Engels, P. Richard, Phys. Rev. Lett. **80**, 2873 (1998).
- [4] R. Jones and D. Hohl, Chapter 24 in 20th IFF Spring School *Computersimulation in der Physik*, Forschungszentrum Jülich, 1989.
- [5] P.Hohenberg and W.Kohn, Phys. Rev. **136**, B864 (1964).
- [6] W.Kohn and L.J.Sham, Phys. Rev. **140**, A1170 (1965).
- [7] J.Harris and R.O.Jones, J. Phys. F **4**, 1170 (1974).
- [8] O.Gunnarsson and B.I.Lundquist, Phys. Rev. B **13**, 4274 (1976).
- [9] D.M.Ceperley and B.J.Alder, Phys. Rev. Lett. **45**, 566 (1980)
- [10] S.H.Vosko, L.Wilk, and M.Nusair, Can. J. Phys. **58**, 1200 (1980).
- [11] R.O.Jones, J. Chem. Phys. **82**, 325 (1985).
- [12] L.P.Perdew and Y.Wang, Phys. Rev. B **33**, 13244 (1992).
- [13] J.P.Perdew, K.Burke, and M.Ernzerhof
Phys. Rev. Lett. **77**, 3865 (1996); **78**, 1396 (1997); **80**, 891 (1998).
- [14] N. Atodiresei, Ph.D.-thesis, RWTH Aachen, 2004.
- [15] S.J.McGlone, P.S.Elmes, R.D.Brown, and P.D.Godfrey, J. Mol. Struc. **225**, 485 (1999).
- [16] M.R.Hoar and J.A.McInnes, Adv. Phys. **32**, 791 (1983).
- [17] L.T.Wille and J.Vennik, J. Phys. A **18**, L419, L1113 (1985).
- [18] M.R.Garey and D.S.Johnson *Computers and Intractability:
A Guide to the Theory of NP-Completeness*, (Freeman, San Francisco, 1979).
- [19] S.Kirkpatrick, C.D.Gelatt, and M.P.Vecchi, Science **220**, 671 (1983).
- [20] R.Car and M.Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
- [21] R.Car and M.Parrinello, Phys. Rev. Lett. **60**, 204 (1988).
- [22] D.Hohl, R.O.Jones, R.Car, and M.Parrinello, Chem. Phys. Lett. **139**, 540 (1987).
- [23] M.L.Cohen, Science **234**, 549 (1986).
- [24] D.R.Hamann, W.Schlüter, and C.Chiang, Phys. Rev. Lett. **43**, 1494 (1979).
- [25] G.B.Bachelet, D.R.Hamann, and W.Schlüter, Phys. Rev. B **26**, 4199 (1982).

- [26] L.Kleinman and D.M.Bylander, Phys.. Rev. Lett. **48**, 1425 (1982).
- [27] P.E.Blöchl, Phys. Rev. **41**, 5414 (1990).
- [28] R.Car, F.de Angelis, P.Giannozzi, and N.Marzari, in Ref. [1], p. 59
- [29] P.E.Blöchl, J.Kästner, C.J.Först, in Ref. [1], p. 93
- [30] J.H.Wilkinson and C.Reinsch, *Linear Algebra*,
Vol.2 of *Handbook for Automatic Computation*, (Springer, New York 1971).
- [31] E.O.Brigham, *The Fast Fourier Transform*, (Prentice-Hall, Englewood Cliffs N.J. 1974).
- [32] W.E.Davidson, J. Comput. Phys. **17**, 87 (1975).
- [33] N.Kosugi, J. Comput. Phys. **55**, 426 (1984).
- [34] W.H.Press, S.A.Teukolsky, W.T.Vetterling, and B.P.Flannery, *Numerical Recipes in C++*,
(Cambridge University Press, Cambridge, UK, 2nd ed. 2002).
Earlier editions show programs in C and FORTRAN.
- [35] E.Polak, *Computational Methods for Optimization*, (Academic Press, New York, 1971).
- [36] L.Verlet, Phys. Rev. **159**, 98 (1967).
- [37] B.Voigtländer, A.Zinner, T.Weber, and H.P.Bonzel, Phys. Rev. B **51**, 7583 (1995).
- [38] M. von Hoegen, Z. Kristallogr. **214**, 591, 614 (1999).
- [39] K.Schroeder, A. Antons, R.Berger, S. Blügel: Phase Transitions **75**, 91 (2002).
- [40] K.Schroeder, A. Antons, R.Berger, S. Blügel: Phys. Rev. Lett. **88**, 046101 (2002).
- [41] A. Antons, K.Schroeder, B. Voigtländer, V. Cherepanov, R. Berger, S. Blügel,
Phys. Rev. Lett. **89**, 236101 (2002).
- [42] A. Antons, Y. Cao, B. Voigtländer, K. Schroeder, R. Berger, S. Blügel,
Europhys. Lett., **62** 547-553 (2003).
- [43] J.J.Mortensen, Y.Morikawa, B.Hammer, and J.K.Norskov, J. of Catalysis **169**, 85 (1997).