

1 Introduction to Molecular Dynamics

Ralf Schneider, Amit Raj Sharma, and Abha Rai

Max-Planck-Institut für Plasmaphysik, Teilinstitut Greifswald,
17491 Greifswald, Germany

Molecular dynamics is the science of simulating the time dependent behavior of a system of particles. The time evolution of the set of interacting atoms is followed by integrating their equation of motion with boundary conditions appropriate for the geometry or symmetry of the system. Molecular dynamics generate information at the microscopic level, which are: atomic positions, velocities. In order to calculate the microscopic behavior of a system from the laws of classical mechanics, MD requires, as an input, a description of the interaction potential (or force field). The quality of the results of an MD simulation depends on the accuracy of the description of inter-particle interaction potential. This choice depends very strongly on application. Thus MD technique acts as a *computational microscope*. This microscopic information is then converted to the macroscopic observable like pressure, temperature, heat capacity and stress tensor etc. using statistical mechanics. Molecular dynamic techniques have been widely used by almost all the branches of science. Namely, determination of reaction rates in chemistry, solid state structures, surfaces and defects formation in material science, protein folding in biochemistry and so on. Recent applications employing common force fields include an exploration of protein folding pathways in solution [1], structural and dynamical properties of ion channels [2, 3]. The disadvantage of a model force-field is that a system is restricted to a single molecular connectivity. This prohibits force field models from describing chemical processes involving bond breaking and forming. An alternative approach is the combination of classical dynamics with electronic structure: internuclear forces are computed *on the fly* from an electronic structure calculation as a MD simulation proceeds [4, 5]. This method, known as *ab initio* molecular dynamics, requires no input potential model and is capable of describing chemical events, although it has high computational overhead.

1.1 Basic Approach

The essential elements for a molecular dynamics simulation are (i) the interaction potential (i.e., potential energy) for the particles, from which the forces can be calculated, and (ii) the equations of motion governing the dynamics of the particles. We follow the laws of classical mechanics, mainly Newton's law

$$\mathbf{F}_i = m_i \mathbf{a}_i, \quad (1.1)$$

for each atom i in a system constituted by N atoms. Here, m_i is the atom mass, \mathbf{a}_i its acceleration and \mathbf{F}_i the force acting upon it due to the interactions with the other atoms. Equivalently one can solve classical Hamiltonian equation of motion

$$\dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{r}_i}, \quad (1.2)$$

$$\dot{\mathbf{r}}_i = \frac{\partial H}{\partial \mathbf{p}_i}, \quad (1.3)$$

where \mathbf{p}_i and \mathbf{r}_i are the momentum and position co-ordinates for the i^{th} atom. H , the Hamiltonian, which is defined as a function of position and momenta, is given by

$$H(\mathbf{p}_i, \mathbf{r}_i) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{r}_i). \quad (1.4)$$

The force on an atom can be calculated as the derivative of energy with respect to the change in the atom's position

$$\mathbf{F}_i = m_i \mathbf{a}_i = -\nabla_i V = -\frac{dE}{d\mathbf{r}_i}. \quad (1.5)$$

Knowledge of the atomic forces and masses can then be used to solve for the positions of each atom along a series of extremely small time steps (on the order of femtoseconds). The velocities are calculated from the accelerations

$$\mathbf{a}_i = \frac{d\mathbf{v}_i}{dt}. \quad (1.6)$$

Finally, the positions are calculated from the velocities

$$\mathbf{v}_i = \frac{d\mathbf{r}_i}{dt}. \quad (1.7)$$

To summarize the procedure, at each step, the forces on the atoms are computed and combined with the current positions and velocities to generate new positions and velocities a short time ahead. The force acting on each atom is assumed to be constant during the time interval. The atoms are then moved to the new positions, an updated set of forces is computed and new dynamics cycle goes on.

Usually molecular dynamics simulations scale by either $O(N \log N)$ or $O(N)$, with N as the number of atoms. This makes simulations with macroscopic number of atoms or molecules ($\sim 10^{23}$) impossible to handle with MD. Therefore, statistical mechanics is used to extract the macroscopic information from the microscopic information provided by MD.

Two important properties of the equations of motion should be noted. One is that they are time reversible, i.e., they take the same form when the transformation $t \rightarrow -t$ is made. The consequence of time reversal symmetry is that the microscopic physics is independent of the direction of the flow of time. Therefore, in contrast to

the Monte Carlo method, molecular dynamics is a *deterministic* technique: Given an initial set of positions and velocities, the subsequent time evolution is *in principle* [6] completely determined from its current state. Molecular dynamics calculates the real dynamics, i.e. behavior of the system, from which the time average of the system's properties can be calculated. The second important property of the equations of motion is that they conserve the Hamiltonian. This can be easily seen by computing the time derivative of H and substituting (1.2) and (1.3) for the time derivatives of position and momentum

$$\frac{dH}{dt} = \sum_{i=1}^N \left[\frac{\partial H}{\partial \mathbf{r}_i} \dot{\mathbf{r}}_i + \frac{\partial H}{\partial \mathbf{p}_i} \dot{\mathbf{p}}_i \right] = \sum_{i=1}^N \left[\frac{\partial H}{\partial \mathbf{r}_i} \frac{\partial H}{\partial \mathbf{p}_i} - \frac{\partial H}{\partial \mathbf{p}_i} \frac{\partial H}{\partial \mathbf{r}_i} \right] = 0. \quad (1.8)$$

The conservation of the Hamiltonian is equivalent to the conservation of the total energy of the system and provides an important link between molecular dynamics and statistical mechanics.

1.1.1 Statistical Ensemble

Statistical mechanics connects the microscopic details of a system to physical observables such as equilibrium thermodynamic properties, transport coefficients, and spectra. Statistical mechanics is based on the Gibbs ensemble concept. That is, many individual microscopic configurations of a very large system lead to the same macroscopic properties, implying that it is not necessary to know the precise detailed motion of every particle in a system in order to predict its properties. It is sufficient to simply average over a large number of identical systems, each in a different microscopic configuration; i.e., the macroscopic observables of a system are formulated in terms of ensemble averages. Statistical ensembles are usually characterized by fixed values of thermodynamic variables such as energy E , temperature T , pressure P , volume V , particle number N or chemical potential μ . One fundamental ensemble is called the *micro-canonical* ensemble and is characterized by constant particle number N , constant volume V and constant total energy E , and is denoted as the NVE ensemble. Other examples include the *canonical* or NVT ensemble, the *isothermal-isobaric* or NPT ensemble, and the *grand-canonical* or μVT ensemble. The thermodynamic variables that characterize an ensemble can be regarded as experimental control parameters that specify the conditions under which an experiment is performed.

Now consider a system of N particles occupying a container of volume V and evolving under Hamilton's equations of motion. According to (1.8), the Hamiltonian will be a constant E , equal to the total energy of the system. In addition, the number of particles and the volume are assumed to be fixed. Therefore, a dynamical trajectory of this system will generate a series of classical states having constant N , V , and E , corresponding to a micro-canonical ensemble. If the dynamics generates all possible states having a fixed N , V , and E , then an average over this trajectory will yield the same result as an average in a micro-canonical ensemble. The energy

conservation condition, $H(\mathbf{p}, \mathbf{r}) = E$, which imposes a restriction on the classical microscopic states accessible to the system, defines a hyper-surface in the phase space called the constant energy surface. A system evolving according to Hamilton's equations of motion will remain on this surface. The assumption that a system, given an infinite amount of time, will cover the entire constant energy hyper-surface is known as the ergodic hypothesis. Thus, under the ergodic hypothesis, averages over a trajectory of a system obeying Hamilton's equations are equivalent to averages over the micro-canonical ensemble.

1.2 Macroscopic Parameters

Statistical mechanics provides a link between the macroscopic properties of matter (like temperature, pressure, etc.) and the microscopic properties (like positions, velocities, individual kinetic and potential energies) of atoms and molecules that constitute it. These macroscopic properties reflect the time average behavior of the atoms at equilibrium (i.e. in one of the many possible degenerate minimum energy states accessible to the system). Often even in an NVE simulation one does some simple tricks to control temperature and/or pressure. This give something of a NVT or NVP and NVE hybrid. However temperature and pressure fluctuate, and the system does not behave as a true NVT or NVP ensemble in the thermodynamic sense. But on average temperature and pressure have the desired value. In true NVT or NPT (non-Hamiltonian) algorithms it is possible to have T and P have exactly the desired value, and the simulation directly corresponds to the thermodynamic ensembles.

At the start of the MD simulation the atomic positions and velocities have to be initialized. In the case of crystalline solids the starting positions will be defined by the crystal symmetry and positions of atoms within the unit cell of the crystal. The unit cell is then repeated to fill up the desired dimensions of the system. Realistic atomic displacements from crystal lattice sites can also be derived using the Debye model. For amorphous solids the particles can be randomly distributed within the desired dimensions making sure that there exists a minimum distance between the atoms so that strong local forces do not exist in the system.

The initial velocities are set by assuming a Maxwell-Boltzmann distribution for velocities along the three dimensions. This is done by using Gaussian distributed random numbers multiplied by a mean square velocity given by $\sqrt{2k_B T/m}$ in each of the three directions and making sure that the system has total momentum equal to zero. Generally speaking, if sensible (tailored to avoid large impulsive forces) position and velocity distributions are chosen, particle positions at equilibrium relax to oscillating around the minimum energy locations of the potential Φ . A Maxwellian distribution of velocities is naturally obtained in the simulation.

Therefore the initial temperature and total energy of the system has been fixed. The temperature is fixed by the velocity distribution. The total energy of the system is given by

$$E_{\text{tot}} = (KE)_{\text{tot}} + (PE)_{\text{tot}} , \quad (1.9)$$

where $(KE)_{\text{tot}}$ is the total kinetic energy in the system given by

$$(KE)_{\text{tot}} = \sum_{i=1}^N \frac{1}{2} m (v_{x,i}^2 + v_{y,i}^2 + v_{z,i}^2) \quad (1.10)$$

and $(PE)_{\text{tot}}$ is the total potential energy of the system given by

$$(PE)_{\text{tot}} = \sum_{i=1}^N \Phi_i(r_i) \quad (1.11)$$

with $v_{x,y,z}$ being the velocities, r being the positions of atoms, and i being the index that sums over all the atoms N in the system. $\Phi_i(r_i)$ is the potential energy of the i^{th} atom due to all other atoms in the system.

1.2.1 Temperature Scaling

In equilibrium simulations, especially if long-range interactions are involved and a potential truncated at a cut-off radius is used, an unavoidable slow drift occur that need correction. A possible trivial temperature scaling is to force the system temperature to be exactly T during every time step. This can be a rather severe perturbation of the atom motion especially if there are only a few atoms. Better methods to control temperature and pressure are discussed in [7, 8, 9] and will be shortly summarized in the following.

The Berendsen method [7] is essentially a direct scaling, but softened with a time constant. Let T_0 be the desired temperature, Δt is the time step of the simulation and τ_T be the time constant for temperature control. In the Berendsen temperature control scheme, all velocities are scaled at each time step by a factor λ given by

$$\lambda = \sqrt{1 + \frac{\Delta t}{\tau_T} \left(\frac{T_0}{T} - 1 \right)} , \quad (1.12)$$

τ_T has to be greater than Δt . According to Berendsen [7] if $\tau_T > 100\Delta t$ then the system has natural fluctuations about the average.

1.2.2 Pressure Scaling

The Berendsen pressure control is implemented by changing all atom positions, and the system cell size during the simulation. If the desired pressure is P_0 and τ_P is the time constant for pressure control, which should be typically greater than $100\Delta t$, the scaling factor μ is given by:

$$\mu = \left[1 - \frac{\beta \Delta t}{\tau_P} (P_0 - P) \right]^{1/3} , \quad (1.13)$$

where β is the isothermal compressibility of the system ($= 1/\text{bulk modulus}$) and P is the current pressure. The change in all atom positions and the system size is given by

$$\mathbf{r}(t + \delta t) = \mu \mathbf{r}(t) , \quad (1.14)$$

$$\mathbf{S}(t + \delta t) = \mu \mathbf{S}(t) , \quad (1.15)$$

and the volume of the system also changes by

$$V(t + \delta t) = \mu^3 V(t) . \quad (1.16)$$

This type of temperature and pressure scaling should be done after the solution of the equations of motions gives realistic fluctuations in temperature and pressure for a system in equilibrium and when large values of τ_T and τ_P are chosen.

1.2.3 Time Scale Dilemma

Design of a molecular dynamics simulation can often encounter limits of computational power. The simulation's time duration is dependent on the time length of each time-step, between which forces are recalculated. The time-step must be chosen small enough to avoid discretization errors, and the number of time-steps, and thus simulation time, must be chosen large enough to capture the effect being modeled without taking an extraordinary period of time i.e. smaller than the vibrational frequency of the system. The length of the simulation should be large enough that the system goes through all possible phase space points in the ensemble. As a rule of thumb: the atoms should not move more than $1/20$ of the nearest neighbor distance in the chosen time step. There exists a wide range of time scales over which specific processes occur and one need to resolve vibrations at these scales, for example, bond vibrations (femtosecond), collective vibrations (picosecond) and protein foldings (millisecond to microsecond). The integration time step which is determined by the fastest varying force is of the order femtoseconds. This limits the accessible time scale by MD simulations from picoseconds to several nanoseconds. So, no matter how many processors (how powerful the computer is) one can only reach several picoseconds in time because time cannot be parallelize [10]. As a consequence of time scale dilemma, slower mechanisms like MD has limited accessibility to handle diffusion. This can only be overcome using multi-scale models.

1.3 Inter-Atomic Potentials

1.3.1 Pair Potentials

For pair potentials, the total potential energy of a system can be calculated from the sum of energy contributions from pairs of atoms and it depends only on the distance between atoms. One example of a pair potential is the *Lennard-Jones potential* [11]

(also known as the 6–12 potential). Other examples of pair potential are Coulomb potential, Morse potential [12] etc. Lennard-Jones potential is the most commonly used form

$$V(r)^{\text{LJ}} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right], \quad (1.17)$$

where ϵ is the cohesive energy well depth and σ is the equilibrium distance. The $(\sigma/r)^{12}$ term describes the repulsive force due to overlapping of electron orbitals (Pauli repulsion) and does not have a true physical motivation, other than that the exponent must be larger than 6 to get a potential well. One often uses 12 because it can be calculated efficiently (square of 6). The term $(\sigma/r)^6$ describes the attractive force (Van der Waals) and can be derived classically by considering how two charged spheres induce dipole-dipole interactions into each other. This potential was used in the earliest studies of the properties of liquid argon [13, 14]. LJ potentials are not a good choice for very small r ($r \lesssim 0.1$ nm) since the true interaction is $\sim (1/r)\exp(-r)$ and not $1/r^{12}$.

Typical simulation sizes in molecular dynamics simulation are very small up to 1000 atoms. As a consequence, most of the extensive quantities are small in magnitude when measured in macroscopic units. There are two possibilities to overcome this problem: Either one should work with atomic-scale units (ps, amu, nm) or to make all the observable quantities dimensionless with respect to their characteristic values. The second approach is more popular. The scaling is done with the model parameters e.g size σ , energy ϵ , mass m . So the common recipe is, one chooses a value for one atom/molecule pair potential arbitrarily (ϵ) and then other model parameters (say energy E) are given in terms of this reference value ($E^* = E/\epsilon$). The other parameters are also calculated similarly. For example, dimensionless distance ($r^* = r/\sigma$), energy ($E^* = E/\epsilon$), temperature ($T^* = kT/\epsilon$), time ($t^* = t/[\sigma(m/\epsilon)^{1/2}]$), force ($F^* = F\sigma/\epsilon$), diffusion coefficient ($D^* = D/[\sigma(\epsilon/m)^{1/2}]$) and so on.

Now if we write the LJ potential in dimensionless form

$$V^*(r^*)^{\text{LJ}} = 4 \left[\left(\frac{1}{r^*} \right)^{12} - \left(\frac{1}{r^*} \right)^6 \right]. \quad (1.18)$$

We see that it is parameter independent, consequently all the properties must also be parameter independent. If a potential only has a couple of parameters then this scaling has a lot of advantages. Namely, potential evaluation can be really efficient in reduced units and as the results are always the same, so the results can be transferred to different systems with straight forward scaling by using the model parameters σ , ϵ and m . This is equivalent to selecting unit value for the parameters and it is convenient to report system properties in this form e.g $P^*(\rho^*)$.

1.3.2 Molecular Interaction Models

To describe atomic interactions in molecules more complex than a dimer a pair potential is not enough. Since molecules are bonded by covalent bonds, at least

angular terms are needed, and in many cases many more complicated terms as well. For instance, in carbon chains the difference between *single* and *double* bonds often is important, and for this at least a four-body term is needed.

To describe complex molecules a large set of inter-atomic potentials (often also called force fields) have been developed by chemists, physicists and biochemists. At least when force fields are used to describe atom motion inside molecules and interactions between molecules (but not actual chemical reactions) the term molecular mechanics is often used.

The total energy of a molecule can be given as

$$E = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsion}} + E_{\text{oop}} + E_{\text{cross}} + E_{\text{nonbond}} . \quad (1.19)$$

Where:

- E_{bond} describes the energy change related to a change of bond length, and thus is simply a pair potential V_2 .
- E_{angle} describes the energy change associated with a change in the bond angle, i.e. is a three-body potential V_3 .
- E_{torsion} describes the torsion, i.e. energy associated with the rotation between two parts of a molecule relative to each other.
- E_{oop} describes *out-of-plane* interactions, i.e. the energy change when one part of a molecule is out of the plane with another.
- E_{cross} are cross terms between the other interaction terms.
- E_{nonbond} describes interaction energies which are not associated with covalent bonding. Could be e.g. ionic or van-der-Waal-terms.

In the following we describe the terms, using notation more common on chemistry rather than the physics notation used earlier.

1.3.2.1 The Term E_{bond}

This term describes the energy change associated with the bond length. It is a simple pair potential, and could be e.g. a Morse or LJ potential. At its simplest, it is purely harmonic, i.e.

$$E_{\text{bond}} = \sum_{\text{bonds}} \frac{1}{2} k_b (b - b_0)^2 , \quad (1.20)$$

where b is the bond length. If we write this term instead as

$$E_i = \sum_j \frac{1}{2} k (r_{ij} - r_0)^2 , \quad (1.21)$$

we see that this is essentially the same thing as the pair potentials dealt with earlier. So this is essentially the same thing as approximating the bond as a string with the string constant k . Although the approximation is very simple, it can be good enough in problems where we are always close to equilibrium, since any smooth potential well can always be to the first order approximated by a harmonic well. But harmonic

potentials obviously can not describe large displacements of atoms or bond breaking reasonably. In solids, the harmonic approximation corresponds to the elastic regime, i.e. the one where stress is directly proportional to the strain (Hooke's law).

To improve on the bond model beyond the elastic regime, one can add higher-order terms to it, e.g.

$$E_{\text{bond}} = \sum_{\text{bonds}} K_2(b - b_0)^2 + K_3(b - b_0)^3 + K_4(b - b_0)^4 . \quad (1.22)$$

This way also larger strain can be described, but this still does not describe bond breaking (dissociation).

Also the Morse potential

$$E_{\text{bond}} = \sum_{\text{bonds}} D_b \{1 - e^{-a(b-b_0)}\}^2 \quad (1.23)$$

is much used to describe bond energies. It is good in that it tends to zero when b tends to infinity so it can describe bond breaking. But on the other hand it never goes fully to zero, which is not quite realistic either as in reality a covalent bond does break essentially completely at some inter-atomic distance.

1.3.2.2 Angular Terms E_{angle}

The angular terms describe the energy change associated with two bonds forming an angle with each other. Most kinds of covalent bonds have some angle which is most favored by them – for sp^3 hybridized bonds it is $\sim 109^\circ$, for sp^2 120° and so on. Like for bond lengths, the easiest way to describe bond angles is to use a harmonic term like

$$E_{\text{angle}} = \sum_{\theta} H_{\theta} (\theta - \theta_0)^2 , \quad (1.24)$$

where θ_0 is the equilibrium angle and H_{θ} a constant which describes the angular dependence well.

This may work well up to 10° or so, but for larger angles additional terms are needed. A typical means for improvement is the third-order terms and so forth, for instance

$$E_{\text{angle}} = \sum_{\theta} H_2(\theta - \theta_0)^2 + H_3(\theta - \theta_0)^3 . \quad (1.25)$$

1.3.2.3 Torsional Terms E_{torsion}

The bond and angular terms were already familiar from the potentials for solids. In the physics and chemistry of molecules there are many important effects which can not be described solely with these terms. The most fundamental of these is probably

torsion. By this, the rotations of one part of a molecule with respect to another is meant. A simple example is the rotation of two parts of the ethane molecule C_2H_6 around the central C-C carbon bond.

Torsional forces can be caused by e.g. dipole-dipole-interactions and bond conjugation. If the angle between two parts is described by an angle ϕ , it is clear that the function f which describes the rotation should have the property $f(\phi) = f(\phi + 2\pi)$, because it is possible to do a full rotation around the central bond and return to the initial state. The trigonometric functions sine and cosine of course fulfill this requirement, so it is natural to describe the torsional energy with a few terms in a Fourier series

$$E_{\text{torsion}} = V_1(1 + \cos(\phi)) + V_2(1 + \cos(2\phi)) + V_3(1 + \cos(3\phi)) . \quad (1.26)$$

The first part of the torsional term V_1 is often interpreted to be related to dipole-dipole interactions, V_2 to bond conjugation and V_3 to steric energy.

1.3.2.4 Out-of-Plane Terms E_{oop}

With the out-of-plane-terms one describes the energy which in (some cases) is associated with the displacement of atoms out of the plane in which they should be. This is relevant in some (parts of) molecules where atoms are known to lie all in the same plane. The functional form can be rather simple

$$E_{\text{oop}} = \sum_{\chi} H_{\chi} \chi^2 , \quad (1.27)$$

where χ is the displacement out of the plane.

1.3.2.5 Cross Terms E_{cross}

The cross-terms are functions which contain several of the above-mentioned quantities. They could e.g. describe how a stretched bond has a weaker angular dependence than a normal one. Or they can describe the relations between two displacements, an angle and a torsion and so on.

1.3.2.6 Non-Bonding Terms E_{nonbond}

With the non-bonding terms all effects which affect the energy of a molecule but are not covalent bonds are meant. These are e.g. van-der-Waals-terms, electrostatic Coulomb interactions and hydrogen bonds. For this terms one could thus further divide

$$E_{\text{nonbond}} = E_{\text{vdW}} + E_{\text{Coulomb}} + E_{\text{hbond}} . \quad (1.28)$$

The van der Waals term is often a simple Lennard-Jones-potential, and E_{Coulomb} a Coulomb potential for some, usually fractional, charges q_i .

1.3.3 Reactive Potentials

Most of the potential functions used in MD simulations are intended for modeling physical processes, not chemical reactions. The formation and breaking of chemical bonds are inherently quantum mechanical processes, and are often studied using first-principles methods. Nevertheless, classical potentials do exist that can empirically model changes in covalent bonding.

One successful method for treating covalent bonding interactions in computer simulations is the Tersoff-type potential [15, 16, 17, 18]. Unlike traditional molecular mechanics force fields [19, 20, 21, 22, 23, 24, 25, 26], the Tersoff model allows for the formation and dissociation of covalent chemical bonds during a simulation. Many-body terms reflecting the local coordination environment of each atom are used to modify the strength of more conventional pairwise terms. With this approach, individual atoms are not constrained to remain attached to specific neighbors, or to maintain a particular hybridization state or coordination number. Models of this sort, despite being purely classical, can provide a realistic description of covalent bonding processes in non-electrostatic systems. Potentials of this type have been developed to treat systems containing silicon [16], carbon [17, 27], germanium [18], oxygen [27], or hydrogen [27], as well as heterogeneous systems containing various combinations of these species [18, 28, 29, 30, 31].

One particularly successful example of a Tersoff potential is the reactive empirical bond-order (REBO) potential developed by Brenner [30, 31, 32, 33]. This model uses a Tersoff-style potential to describe the covalent bonding interactions in carbon and hydrocarbon systems. Originally developed for use in simulating the chemical vapor deposition of diamond [30], the REBO potential has been extended to provide more accurate treatment of the energetic, elastic, and vibrational properties of solid carbon and small hydrocarbons [33]. This potential has been used to model many different materials and processes, including fullerenes [32], carbon nanotubes [34], amorphous carbon [35], and the tribology and tribochemistry of diamond interfaces [36, 37, 38, 39, 40, 41, 42].

The REBO potential is not appropriate for studying every hydrocarbon system, however. In particular, the absence of dispersion and non-bonded repulsion terms makes the potential poorly suited for any system with significant intermolecular interactions. This is the case for many important hydrocarbon systems, including liquids and thin films, as well as some solid-state materials such as graphite and fullerenes. Even covalent materials such as diamond can benefit from a treatment including non-bonded interactions. The bulk phase is dominated by covalent interactions, but longer-range forces become quite important when studying interfacial systems [27].

Various attempts have been made previously to combine non-bonded interactions with the Tersoff or REBO potentials in a way that preserves the reactive capabilities of the model [43, 44, 45]. One such improvement of the Tersoff potential was presented by Kai Nordlund et al. [46] which retains the good description of the covalent bonding and yet also describes accurately both the short-range repulsive part of the potential and the long-range bonding between graphite planes. One

way to do this is to simply reduce the repulsive barrier associated with the Lennard-Jones or other potential [47], although this results in barriers which are too large for radical species and too small for saturated compounds. Another alternative, taken by Nyden et al. [44], is to allow bonds to dissociate with a Morse potential [12], and explicitly check for recombination reactions between dissociated radicals. This approach has been used to model thermal decomposition of polymers [44], but is not general enough to treat arbitrary reactions in hydrocarbons, such as addition across unsaturated bonds. Another method, used by Che et al. [45] is to reduce the repulsive non-bonded interactions based on the covalent interaction energy, rather than the distance. This method can help eliminate non-bonded interactions during bond dissociations, but will again tend to overestimate barriers in association reactions.

1.4 Numerical Integration Techniques

The potential energy is a function of the atomic positions ($3N$) of all the atoms in the system. Due to the complicated nature of this function, there is no analytical solution to the equations of motion and these equations must be solved numerically.

Numerous numerical algorithms have been developed for integrating the equations of motion. We list several here.

- (i) Verlet algorithm [14],
- (ii) Leap-frog algorithm [48],
- (iii) Velocity Verlet [49],
- (iv) Beeman's algorithm [50] and
- (v) Symplectic reversible integrators [51, 52].

In choosing which algorithm to use, one considers the following criteria:

- (i) The algorithm should conserve energy and momentum and is reversible. When $\delta t \rightarrow -\delta t$ the system should go back to original state.
- (ii) It should be computationally efficient.
- (iii) It should permit a long time step for integration.
- (iv) Only one force evaluation per time step (important for complex potential).

1.4.1 Verlet's Algorithm

The most widely used finite-difference method is a third-order Störmer algorithm first used by Verlet [14] and widely known as the Verlet's method. It is derived from the two Taylor expansion

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2}\delta t^2 \mathbf{a}(t) + \frac{1}{3!}\delta t^3 \dot{\mathbf{a}}(t) + O(\delta t^4), \quad (1.29)$$

$$\mathbf{r}(t - \delta t) = \mathbf{r}(t) - \delta t \mathbf{v}(t) + \frac{1}{2}\delta t^2 \mathbf{a}(t) - \frac{1}{3!}\delta t^3 \dot{\mathbf{a}}(t) + O(\delta t^4), \quad (1.30)$$

summing the above two equations eliminates the odd-order terms. Rearranging gives

$$\mathbf{r}(t + \delta t) + \mathbf{r}(t - \delta t) = 2\mathbf{r}(t) + \delta t^2 \mathbf{a}(t) , \quad (1.31)$$

$$\mathbf{r}(t + \delta t) = 2\mathbf{r}(t) - \mathbf{r}(t - \delta t) + \delta t^2 \mathbf{a}(t) + O(\delta t^4) . \quad (1.32)$$

Notice that the position vector \mathbf{r} at time $t + \delta t$ is calculated from position vector at time t and $t - \delta t$, this makes the Verlet's algorithm a two-step method. Therefore it is not self-starting, initial positions $\mathbf{r}(0)$ and velocities $\mathbf{v}(0)$ are not sufficient to begin a calculation. Also the velocities are missing from the above equation and can be calculated from

$$\mathbf{v}(t) = \frac{\mathbf{r}(t + \delta t) - \mathbf{r}(t - \delta t)}{2\delta t} . \quad (1.33)$$

In its original form it treats velocity as less important than positions. This approach is conflicting for ergodic system. The phase space trajectory depends equally on positions and velocities.

The local error (error per iteration) in position of the Verlet integrator is $O(\delta t^4)$ and local error in velocity is $O(\delta t^2)$. However the global error in position is $O(\delta t^2)$ and the global error in velocity is $O(\delta t^2)$.

Because the velocity is determined in a non-cumulative way from the positions in the Verlet integrator, the global error in velocity is also $O(\delta t^2)$. In molecular dynamics simulations, the global error is typically far more important than the local error, and the Verlet integrator is therefore known as a second-order integrator.

1.4.2 General Predictor-Corrector Algorithms

Predictor-corrector methods are composed of three steps: prediction, evaluation and correction. Starting from the current position $\mathbf{r}(t)$ and velocity $\mathbf{v}(t)$, the numerical steps are as follows.

- (i) Predict the position $\mathbf{r}(t + \delta t)$ and velocity $\mathbf{v}(t + \delta t)$ at the end of the next step.
- (ii) Evaluate the forces by taking the gradient of the potential at $\delta t + t$ using the predicted position. The difference in the calculated acceleration (this step) and the predicted acceleration (step 1) constitutes an error signal.
- (iii) The error signal is used to correct the predictions using some combination of the predicted and previous values of position and velocity.

Using a Taylor series expansion to predict the system configuration at time $(t + \delta t)$ one gets

$$\begin{aligned} \mathbf{r}(t + \delta t) &= \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t) + \frac{1}{3!} \delta t^3 \mathbf{b}(t) + \dots , \\ \mathbf{v}(t + \delta t) &= \mathbf{v}(t) + \delta t \mathbf{a}(t) + \frac{1}{2} \delta t^2 \mathbf{b}(t) + \dots , \\ \mathbf{a}(t + \delta t) &= \mathbf{a}(t) + \delta t \mathbf{b}(t) + \dots , \\ \mathbf{b}(t + \delta t) &= \mathbf{b}(t) + \dots , \end{aligned} \quad (1.34)$$

where \mathbf{b} is the time derivative of the acceleration \mathbf{a} and is known at time t .

If the Taylor expansions are truncated, so that only the terms shown explicitly in (1.34) are left, then the quantities can be called the predicted values \mathbf{r}^p , \mathbf{v}^p , \mathbf{a}^p and \mathbf{b}^p . The force is computed by taking the gradient of potential at the predicted position \mathbf{r}^p , and new acceleration value is computed. Since the predicted values are not based on physics the re-calculated acceleration is different from the predicted acceleration \mathbf{a}^p (acceleration in (1.34)). The difference between the two values is called the error signal or error

$$\Delta\mathbf{a}(t + \delta t) = \mathbf{a}^c(t + \delta t) - \mathbf{a}^p(t + \delta t). \quad (1.35)$$

This error signal is used to correct all predicted quantities in (1.34)

$$\begin{aligned} \mathbf{r}^c(t + \delta t) &= \mathbf{r}^p(t + \delta t) + c_0\Delta\mathbf{a}(t + \delta t), \\ \mathbf{v}^c(t + \delta t) &= \mathbf{v}^p(t + \delta t) + c_1\Delta\mathbf{a}(t + \delta t), \\ \mathbf{a}^c(t + \delta t) &= \mathbf{a}^p(t + \delta t) + c_2\Delta\mathbf{a}(t + \delta t), \\ \mathbf{b}^c(t + \delta t) &= \mathbf{b}^p(t + \delta t) + c_3\Delta\mathbf{a}(t + \delta t). \end{aligned} \quad (1.36)$$

All the corrected quantities are proportional to the error signal, and the proportional coefficients are determined to maximize the stability of the calculation. These corrected values are now better approximations of the true quantities, and are used to predict the quantities in the next iteration. The best choice for these coefficients depends on the order of both the differential equations and the Taylor series [53]. These coefficients are computed based on the order of the algorithm being used in the simulation. In addition, the accuracy of the numerical integrator algorithms also depends on the time step size, which is typically on the order of fractions of femto-seconds (10^{-15} s). Thus, the simulation as a whole is able to describe only short-time scale phenomena that last on the order of pico- (10^{-12}) up to nano-seconds (10^{-9} s).

1.4.3 Leap-Frog

In this algorithm, the velocities are first calculated at time $t + 1/2\delta t$; these are used to calculate the positions, \mathbf{r} , at time $t + \delta t$. In this way, the velocities leap over the positions, then the positions leap over the velocities. The advantage of this algorithm is that the velocities are explicitly calculated, however, the disadvantage is that they are not calculated at the same time as the positions. The velocities at time t can be approximated by the relationship

$$\mathbf{v}(t) = \frac{1}{2} \left[\mathbf{v} \left(t - \frac{1}{2}\delta t \right) + \mathbf{v} \left(t + \frac{1}{2}\delta t \right) \right]. \quad (1.37)$$

Therefore:

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \mathbf{v} \left(t + \frac{1}{2}\delta t \right) \delta t, \quad (1.38)$$

$$\mathbf{v} \left(t + \frac{1}{2}\delta t \right) = \mathbf{v} \left(t - \frac{1}{2}\delta t \right) + \mathbf{a}(t)\delta t. \quad (1.39)$$

1.4.4 Velocity Verlet

One starts with the following equations

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t) + \dots, \quad (1.40)$$

$$\mathbf{v}(t + \delta t) = \mathbf{v}(t) + \frac{1}{2} \delta t [\mathbf{a}(t) + \mathbf{a}(t + \delta t)]. \quad (1.41)$$

Each integration cycle consists of the following step:

(i) Calculate the velocities at mid-step using

$$\mathbf{v}\left(t + \frac{\delta t}{2}\right) = \mathbf{v}(t) + \frac{1}{2} \delta t \mathbf{a}(t). \quad (1.42)$$

(ii) Calculate $\mathbf{r}(t + \delta t)$

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \mathbf{v}\left(t + \frac{\delta t}{2}\right) \delta t. \quad (1.43)$$

(iii) Calculate $\mathbf{a}(t + \delta t)$ from the potential.

(iv) Update the velocity using

$$\mathbf{v}(t + \delta t) = \mathbf{v}\left(t + \frac{\delta t}{2}\right) + \frac{1}{2} \delta t \mathbf{a}(t + \delta t). \quad (1.44)$$

1.4.5 Beeman's Algorithm

The advantage of this algorithm is that it provides a more accurate expression for the velocities and better energy conservation. The disadvantage is that the more complex expressions make the calculation more expensive

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{2}{3} \delta t^2 \mathbf{a}(t) - \frac{1}{6} \delta t^2 \mathbf{a}(t - \delta t). \quad (1.45)$$

The predicted velocity is given by

$$\mathbf{v}(t + \delta t) = \mathbf{v}(t) + \frac{3}{2} \delta t \mathbf{a}(t) - \frac{1}{2} \delta t \mathbf{a}(t - \delta t). \quad (1.46)$$

The acceleration is based on the predicted velocity

$$\mathbf{a}(t + \delta t) = F(\{\mathbf{r}_i(t + \delta t), \mathbf{v}_i(t + \delta t)\}, i = 1, 2, \dots, n), \quad (1.47)$$

where \mathbf{v}_i is the predicted velocity from the previous equation. The corrected velocity is given by

$$\mathbf{v}(t + \delta t) = \mathbf{v}(t) + \frac{1}{3} \delta t \mathbf{a}(t + \delta t) + \frac{5}{6} \delta t \mathbf{a}(t) - \frac{1}{6} \delta t \mathbf{a}(t - \delta t). \quad (1.48)$$

1.4.6 Gear Algorithm

The fifth-order Gear predictor-corrector method [53] predicts the molecular position \mathbf{r}_i at time $t + \delta t$ using fifth-order Taylor series based on position and their derivatives at time t . It is particularly useful for stiff differential equations.

1.4.7 Symplectic Integrators

Symplectic integrators are designed for the numerical solution of Hamiltonian's equation of motion. They preserve Poincaré invariants when integrating classical trajectories (see [54] and earlier references therein). The Hamiltonian which is slightly perturbed from the original value is conserved. This approach has the big advantage, that it guarantees and preserves conservation laws.

1.5 Analysis of MD Runs

In this section we will describe how the output of MD simulations (positions and velocities) are analysed to get the physical quantities of interest.

1.5.1 Ergodic Hypothesis

To calculate a physical quantity A in molecular dynamics, it is calculated as the time average of A

$$\langle A \rangle_{\text{time}} = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(\mathbf{p}^N(t), \mathbf{r}^N(t)) dt \approx \frac{1}{M} \sum_{t=1}^M A(\mathbf{p}^N, \mathbf{r}^N) , \quad (1.49)$$

where t is the simulation time, M is the number of time steps in the simulation and $A(\mathbf{p}^N, \mathbf{r}^N)$ is the instantaneous value of A . This integral is generally extremely difficult to calculate because one must calculate all possible states of the system.

In statistical mechanics experimental observables are assumed to be ensemble averages

$$\langle A \rangle_{\text{ensemble}} = \iiint d\mathbf{p}^N d\mathbf{r}^N A(\mathbf{p}^N, \mathbf{r}^N) \rho(\mathbf{p}^N, \mathbf{r}^N) , \quad (1.50)$$

where $A(\mathbf{p}^N, \mathbf{r}^N)$ is the observable of interest, $\rho(\mathbf{p}^N, \mathbf{r}^N)$ is the probability density of the ensemble. The integration is carried over all possible values of position \mathbf{r} and momenta \mathbf{p} . The *ergodic hypothesis*, which states that the time average equals the ensemble average

$$\langle A \rangle_{\text{time}} = \langle A \rangle_{\text{ensemble}} , \quad (1.51)$$

The basic idea is that if one allows the system to evolve in time indefinitely, then the system will eventually pass through all possible states. One goal, therefore, of a

molecular dynamics simulation is to generate enough representative conformations such that this equality is satisfied. If this is the case, experimentally relevant information concerning structural, dynamic and thermodynamic properties may then be calculated using a feasible amount of computer resources. Because the simulations are of fixed duration, one must be certain to sample a sufficient amount of phase space.

1.5.2 Standard Diagnostics

There are a number of different physical quantities which one may be interested in. For a liquid, these may be liquid structure factors, transport coefficients (eg. diffusion coefficient, viscosity or thermal conductivity) etc. For solids, these may be crystal structure, adsorption of molecules on surface, melting behaviour etc. Here, we will consider the diagnostics methods to calculate internal energy, pressure tensor, self-diffusion coefficient and pair distribution function. More details are described in [55, 56].

1.5.2.1 Energy

The energy is the simplest and most straightforward quantity to calculate. From all pair of atoms (i, j) , one calculates their separation r_{ij} . These are then substituted into the chosen form of potential $U(r)$. The energy has contributions from both potential and kinetic terms. The kinetic energy should be calculated after the momenta \mathbf{p} have been updated, i.e., after the force routine has been called. The kinetic energy can then be calculated, and then added to the potential energy

$$\langle E \rangle = \langle H \rangle = \langle K \rangle + \langle U \rangle = \left\langle \sum_i \frac{|p_i|^2}{2m_i} \right\rangle + \langle U(r) \rangle. \quad (1.52)$$

$U(r)$ is obtained directly from the potential energy calculations. For calculating average temperature

$$E_{\text{kin}} = \langle K \rangle = \frac{3}{2} N k_B T \Rightarrow T = \frac{1}{3Nk_B} \left\langle \sum_{i=1}^N \frac{|p_i|^2}{m_i} \right\rangle. \quad (1.53)$$

1.5.2.2 Pressure

Pressure is a second rank tensor. For inhomogeneous systems, one calculates this tensor by finding the force across potential surfaces [57]. However, for homogeneous systems, it is not the most efficient method and one uses the virial theorem to calculate the configurational part of the pressure tensor, and then add that to the kinetic part. For the derivation of the virial theorem one can refer to [58]. The full expression for the pressure tensor of a homogeneous system of particles is given as

$$\mathbf{P}(\mathbf{r}, t) = \frac{1}{V} \left[\sum_{i=1}^N m_i \mathbf{v}_i(t) \mathbf{v}_i(t) + \sum_{i=1}^N \sum_{j>i}^N \mathbf{r}_{ij}(t) \mathbf{F}_{ij}(t) |_{\mathbf{r}_i(t)=\mathbf{r}} \right], \quad (1.54)$$

where V is the volume, m_i , v_i are the mass and velocity of particle i respectively. The first term represents the kinetic contribution and the second term represents the configurational part of the pressure tensor. It is clear that the interaction between the pairs is calculated just once. Note that the above equation is valid for atomic systems at equilibrium, system of molecules require some modifications to be made, as do non-equilibrium systems.

1.5.2.3 Pair Correlation Function

The static properties of the system e.g. structure, energy, pressure etc. are obtained from the pair (or radial) correlation function. Pair correlation function, $g(r)$, gives the information on the structure of the material. It gives the probability of locating pairs of atoms separated by a distance r , relative to that for a completely random distribution at the same density (i.e. the ideal gas). For a crystal, it exhibits a sequence of peaks at positions corresponding to shells around a given system. For amorphous materials and liquid, $g(r)$ exhibits its major peak close to the average atomic separation of neighboring atoms, and oscillates with less pronounced peaks at larger distances. The magnitude of the peaks usually decays exponentially with distance as $g(r) \rightarrow 1$. In most cases, $g(r)$ vanishes below a certain distance where atomic repulsion is strong enough to prevent pairs of atoms from getting too close.

It is defined as

$$g(r) = \frac{V}{N^2} \left\langle \sum_{i=1}^N \sum_{j \neq i}^N \delta(\mathbf{r} - \mathbf{r}_{ij}) \right\rangle. \quad (1.55)$$

In a computer simulation, the delta function is replaced by a function that is finite (say, given a value 1) over a small range of separations, and a histogram is accumulated over time of all pair separations that fall within this range. $g(r)$ is effectively a measure of structural properties, but is particularly important because all thermodynamic quantities may be expressed as some function of it [56, 59].

1.5.2.4 Time Correlation Function

The dynamic and transport properties of the system are obtained from time correlation functions. Any Transport coefficient K can be calculated using generalized Einstein and Green-Kubo Formulas [60]

$$K(t) = \lim_{t \rightarrow \infty} \frac{\langle [A(t) - A(0)]^2 \rangle}{2t} = \int_0^\infty d\tau \langle \dot{A}(\tau) \dot{A}(0) \rangle. \quad (1.56)$$

If one wants to calculate the self diffusion coefficient then $A(t) = \mathbf{r}_i(t)$ is the atom position at time t and $\dot{A} = \mathbf{v}_i(t)$ is the velocity of the atom. For calculating the shear viscosity, $A(t) = \sum m_i \mathbf{r}_i(t) \mathbf{v}_i(t)$ and $\dot{A} = \boldsymbol{\sigma}_{\alpha\beta}$. Other transport quantities can also be calculated similarly. If we compare the value of $A(t)$ with its value at zero time, $A(0)$ the two values will be correlated at sufficiently short times, but at longer times the value of $A(t)$ will have no correlation with its value at $t = 0$. Information on relevant dynamical processes is contained in the time decay of $K(t)$. Time correlation function can be related to the experimental spectra by a fourier transformation.

1.5.2.5 Diffusion Coefficients

As discussed above, we obtain diffusion coefficient using the Einstein relation

$$D = \lim_{t \rightarrow \infty} \frac{\langle [r(t) - r(0)]^2 \rangle}{2dt}, \quad (1.57)$$

where D is the diffusion coefficient, d is the dimensionality of the system and $r(t)$ is the position of atom at time t . Angle brackets represents averaging over all possible time origins (see [56] for more information). This is proportional to the slope of the mean square displacement of a single particle undergoing Brownian motion at the long time limit.

Warrier et al. [61] analysed the diffusion of hydrogen atoms in porous graphite. They found, that different length scales for jumps are present in the system. J. Klafter et al. [62] talk about random walk that are *sub-diffusive* (wherein the trajectory results in a mean square displacement that shows slower-than-linear growth with time), and *super-diffusive* (wherein the trajectory results in a mean square displacement that shows faster-than-linear growth with time). Such random walks are called Lévy flights and can show up super-diffusive behaviour with infinite variance and their trajectories show self-similar patterns characteristics of fractals.

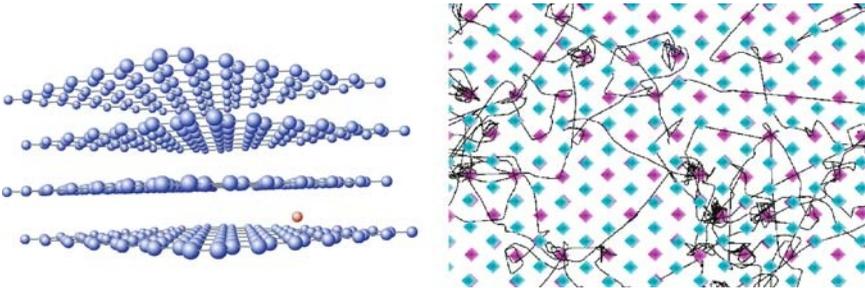


Fig. 1.1. **Left:** One hydrogen atom in a carbon lattice. **Right:** Diffusion paths at 900 K for a hydrogen atom in graphite. Small frequent jumps and rare large jumps are visible

1.5.3 Multi-Scale Modeling

Multi-Scale modeling is the field of solving physical problems which have important features at multiple scales, particularly multiple spatial and temporal scales. As an example, the problem of protein folding has multiple time scales. While the time scale for the vibration of the covalent bonds is of the order of femtoseconds (10^{-15} s), folding time for proteins may very well be of the order of seconds. Well-known examples of problems with multiple length scales include turbulent flows, mass distribution in the universe, and vortical structures on the weather map [63]. In addition, different physical laws may be required to describe the system at different scales. Take the example of fluids. At the macroscale (meters or millimeters), fluids are accurately described by the density, velocity and temperature fields, which obey the continuum Navier-Stokes equations. On the scale of mean free path, it is necessary to use kinetic theory (Boltzmann equations) to get a more detailed description in the terms of the one-particle phase-space distribution function. At the nanometer scale, molecular dynamics in the form of Newton's law has to be used to give the actual position and velocity of each individual atom that makes up the fluid. If a liquid such as water is used as the solvent for protein folding, then the electronic structure of the water molecules becomes important and these are described by Schrödinger's equation in quantum mechanics. The boundaries between different levels of theories may vary, depending on the system being studied, but the overall trend described above is generally valid. At each finer scale a more detailed theory has to be used, giving rise to more detailed information on the system. Warrior et al.

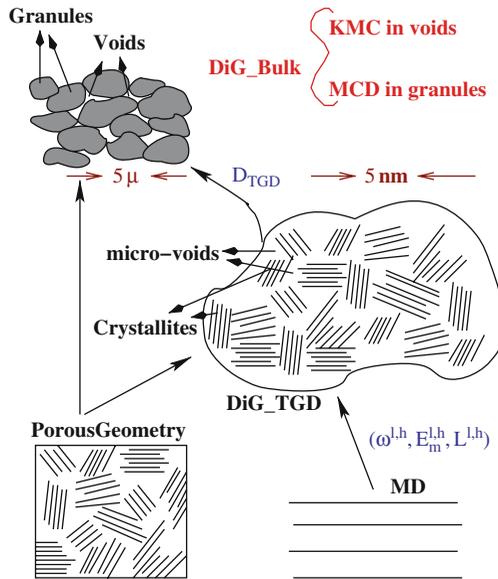


Fig. 1.2. Multi-scale modeling approach for diffusion of hydrogen in porous graphite

[61] have done a multi-scale analysis of the diffusion of hydrogen isotope in porous graphite. They used the insight gained from microscopic models (consisting of a few hundreds of atoms over a time scale of a few picoseconds and length scale of nanometers using MD) into modeling the hydrogen isotope reactions and transports at meso-scale (trans-granular diffusion, with length scales of few microns) and further into the macro-scale (typically a centimeter over a time scale of milliseconds). Therefore a multi-scale (both in length and time) approach to modeling plasma surface interaction is necessary. The figure below explains the multi-scale modeling approach clearly.

1.6 From Classical to Quantum-Mechanical MD

Classical molecular dynamics using predefined potentials is well established as a powerful tool to investigate many-body condensed matter systems. The broadness, diversity, and level of sophistication of this technique is documented in several monographs as well as proceedings of conferences and scientific schools [56, 64, 65, 66, 67, 68, 69]. At the very heart of any molecular dynamics scheme is the question of how to describe, that is in practice how to approximate, the interatomic interactions. The traditional route followed in molecular dynamics is to determine these potentials in advance. Typically, the full interaction is broken up into two-body, three-body and many-body contributions, long-range and short-range terms etc., which have to be represented by suitable functional forms, discussed under the inter-atomic potentials section of this article. After decades of intense research, very elaborate interaction models including the non-trivial aspect to represent them analytically were devised [70, 71, 72].

Despite overwhelming success of the pre-calculated potentials, the *fixed model potential* implies serious drawbacks. Among the most delicate ones are systems where

- (i) many different atom or molecule types give rise to a myriad of different interatomic interactions that have to be parameterized and/or
- (ii) the electronic structure and thus the bonding pattern changes qualitatively in the course of the simulation.

These systems can be called *chemically complex*.

The reign of traditional molecular dynamics and electronic structure methods was greatly extended by the family of techniques that is called here *ab initio* molecular dynamics. Other names that are currently in use are for instance Car-Parrinello, Hellmann-Feynman, First principles, quantum chemical, on-the-fly, direct, potential-free, quantum, etc. molecular dynamics. The basic idea underlying every *ab initio* molecular dynamics method is to compute the forces acting on the nuclei from electronic structure calculations that are performed *on-the-fly* as the molecular dynamics trajectory is generated. In this way, the electronic variables are not integrated out before-hand, but are considered as active degrees of freedom. This implies that, given a suitable approximate solution of the many-electron

problem, also *chemically complex* systems can be handled by molecular dynamics. But this also implies that the approximation is shifted from the level of selecting the model potential to the level of selecting a particular approximation for solving the Schrödinger equation.

1.7 Ab Initio MD

In this approach, a global potential energy surface is constructed in a first step either empirically or based on electronic structure calculations. In a second step, the dynamical evolution of the nuclei is generated by using classical mechanics, quantum mechanics or semi/quasiclassical approximations of various sorts.

Suppose that a useful trajectory consists of about 10^M molecular dynamics steps, i.e. 10^M electronic structure calculations are needed to generate one trajectory. Furthermore, it is assumed that 10^n independent trajectories are necessary in order to average over different initial conditions so that 10^{M+n} *ab initio* molecular dynamics steps are required in total. Finally, it is assumed that each single-point electronic structure calculation needed to devise the global potential energy surface and one *ab initio* molecular dynamics time step requires roughly the same amount of CPU time. Based on this truly simplistic order of magnitude estimate, the advantage of *ab initio* molecular dynamics vs. calculations relying on the computation of a global potential energy surface amounts to about $10^{3N+6+M+n}$. The crucial point is that for a given statistical accuracy (that is for M and n fixed and independent on N) and for a given electronic structure method, the computational advantage of *on-the-fly* approaches grows like 10^N with system size. Of course, considerable progress has been achieved in trajectory calculations by carefully selecting the discretization points and reducing their number, choosing sophisticated representations and internal coordinates, exploiting symmetry etc. but basically the scaling 10^N with the number of nuclei remains a problem. Other strategies consist for instance in reducing the number of active degrees of freedom by constraining certain internal coordinates, representing less important ones by a (harmonic) bath or friction, or building up the global potential energy surface in terms of few-body fragments. All these approaches, however, invoke approximations beyond the ones of the electronic structure method itself. Finally, it is evident that the computational advantage of the *on-the-fly* approaches diminish as more and more trajectories are needed for a given (small) system. For instance extensive averaging over many different initial conditions is required in order to calculate quantitatively scattering or reactive cross sections.

A variety of powerful *ab initio* molecular dynamics codes have been developed, few of them listed here are CASTEP [73], CP-PAW [74], fhi98md [75], NWChem [76], VASP [77], GAUSSIAN [78], MOLPRO [79] and ABINIT [80, 81].

1.8 Car-Parrinello Molecular Dynamics

The basic idea of the Car-Parrinello [4] approach can be viewed to exploit the quantum-mechanical adiabatic time-scale separation of fast electronic and slow nuclear motion by transforming that into classical-mechanical adiabatic energy-scale separation in the framework of dynamical systems theory. In order to achieve this goal the two-component quantum/classical problem is mapped onto a two-component purely classical problem with two separate energy scales at the expense of loosing the explicit time-dependence of the quantum subsystem dynamics.

Car and Parrinello postulated the following class of Lagrangians [4] to serve this purpose

$$L_{CP} = \underbrace{\sum_I \frac{1}{2} M_I \dot{R}_I^2}_{\text{normal kinetic energy}} + \underbrace{\sum_i \frac{1}{2} \mu_i \langle \dot{\psi}_i | \dot{\psi}_i \rangle}_{\text{potential energy}} - \underbrace{\langle \Psi_0 | H_e | \Psi_0 \rangle}_{\text{orthonormality}} + \text{constraints} \quad , \quad (1.58)$$

where μ_i ($= \mu$) are the *fictitious masses* or inertia parameters assigned to the orbital degrees of freedom; the units of the mass parameter μ are energy times a squared time for reasons of dimensionality. ψ_i are regarded as classical fields, M_I are the ionic masses. The potential energy in the Car-Parrinello Lagrangian can be written as

$$\langle \Psi_0 | H_e | \Psi_0 \rangle = E_{KS} [\{\psi_i\}, \mathbf{R}_I] \quad , \quad (1.59)$$

E_{KS} is the LDA-KS energy functional. Within the pseudopotential implementation of the local density approximation (LDA) in the Kohn-Sham (KS) scheme, the ionic potential energy corresponding to the electron in the ground state can be found by minimizing the KS total-energy functional $E_{KS} [\{\psi_i\}, \mathbf{R}_I]$ with respect to the one-particle wavefunction $\psi_i(\mathbf{r})$ describing the valence-electron density subject to orthonormalization constraints. The explicit expression of E_{KS} in terms of orthonormal one-particle orbitals $\psi_i(\mathbf{r})$ is

$$\begin{aligned} E_{KS} [\{\psi_i(\mathbf{r})\}, \{\mathbf{R}_I\}] &= \sum_i f_i \int \psi_i^*(\mathbf{r}) \left(-\frac{1}{2} \nabla^2 \right) \psi_i(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \\ &+ \int \epsilon_{XC}(\rho(\mathbf{r})) \rho(\mathbf{r}) d\mathbf{r} + E_{el}([\psi_i(\mathbf{r})], \{\mathbf{R}_I\}) + U_I^0(\{\mathbf{R}_I\}) \quad . \quad (1.60) \end{aligned}$$

The terms on the right-hand side of the previous equation are, respectively, the electronic kinetic energy, the electrostatic Hartree term, the integral of the LDA exchange and correlation energy density ϵ_{XC} , the electron-ion pseudopotential interaction, and the ion-ion interaction potential energy. The electronic density $\rho(\mathbf{r})$ is given by

$$\rho(\mathbf{r}) = \sum_i f_i |\psi_i(\mathbf{r})|^2 \quad , \quad (1.61)$$

where f_i are occupation numbers.

The corresponding Newtonian equations of motion are obtained from the associated Euler-Lagrange equations

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{R}}_I} = \frac{\partial L}{\partial \mathbf{R}_I}, \quad (1.62)$$

$$\frac{d}{dt} \frac{\delta L}{\delta \dot{\psi}_i^*} = \frac{\delta L}{\delta \psi_i^*} \quad (1.63)$$

like in classical mechanics, but here for both the nuclear positions and the orbitals; note $\psi_i^* = \langle \psi_i |$ and that the constraints are holonomic (which can be expressed in the form $f(\mathbf{r}_1, \mathbf{r}_2, \dots, t) = 0$). Following this route of ideas, generic Car-Parrinello equations of motion are found to be of the form

$$M_I \ddot{\mathbf{R}}_I(t) = -\frac{\partial}{\partial \mathbf{R}_I} \langle \Psi_0 | H_e | \Psi_0 \rangle + \frac{\partial}{\partial \mathbf{R}_I} \{\text{constraints}\}, \quad (1.64)$$

$$\mu_i \ddot{\psi}_i(t) = -\frac{\delta}{\delta \psi_i^*} \langle \Psi_0 | H_e | \Psi_0 \rangle + \frac{\delta}{\delta \psi_i^*} \{\text{constraints}\}. \quad (1.65)$$

Note that the constraints within the total wavefunction lead to *constraint forces* in the equations of motion. Note also that these constraints might be a function of both the set of orbitals $\{\psi_i\}$ and the nuclear positions $\{\mathbf{R}_I\}$. These dependencies have to be taken into account properly in deriving the Car-Parrinello equations following from (1.58) using (1.62) and (1.63).

According to the Car-Parrinello equations of motion, the nuclei evolve in time at a certain (instantaneous) physical temperature $\propto \sum_I M_I \dot{\mathbf{R}}_I^2$, whereas a *fictitious temperature* $\propto \sum_i \mu_i \langle \psi_i | \dot{\psi}_i \rangle$ is associated to the electronic degrees of freedom. In this terminology, *low electronic temperature* or *cold electrons* means that the electronic subsystem is close to its instantaneous minimum energy $\min_{\{\psi_i\}} \langle \Psi_0 | H_e | \Psi_0 \rangle$ i.e. close to the exact Born-Oppenheimer (BO) surface. Thus, a ground-state wavefunction optimized for the initial configuration of the nuclei will stay close to its ground state also during time evolution if it is kept at a sufficiently low temperature. The remaining task is to separate in practice nuclear and electronic motion such that the fast electronic subsystem stays cold also for long times but still follows the slow nuclear motion adiabatically (or instantaneously). Simultaneously, the nuclei are nevertheless kept at a much higher temperature. This can be achieved in nonlinear classical dynamics via decoupling of the two subsystems and (quasi-)adiabatic time evolution. This is possible if the power spectra stemming from both dynamics do not have substantial overlap in the frequency domain so that energy transfer from the *hot nuclei* to the *cold electrons* becomes practically impossible on the relevant time scales. This amounts in other words to imposing and maintaining a metastability condition in a complex dynamical system for sufficiently long times.

The Hamiltonian or conserved energy is the constant of motion (like classical MD, with relative variations smaller than 10^{-6} and with no drift), which serves as an extremely sensitive check of the molecular dynamics algorithm. Contrary to that the electronic energy displays a simple oscillation pattern due to the simplicity of

the phonon modes. Most importantly, the fictitious kinetic energy of the electrons is found to perform bound oscillations around a constant, i.e. the electrons *do not heat up* systematically in the presence of the hot nuclei.

As we have seen above, Car-Parrinello method gives physical results even if the orbitals are not at the BO surface, provided that the electronic and ionic degrees of freedom remain adiabatically separated and electrons remain close to the BO surface. Loss of adiabaticity would mean that there is transfer of energy from *hot nuclei* to *cold electron* and Car-Parrinello MD deviates from BO surface.

1.8.1 Adiabaticity

The metastable two-temperature regime setup in the CP dynamics is extremely efficient at approximating the constraints of maintaining the electronic energy functional at the minimum without explicit minimization. At the beginning of the numerical simulation, the electronic subsystem is in an initial state which is very close to the minimum of the energy surface. When the ions start moving, their motion causes a change in the instantaneous position of the minimum in the electronic parameter space. The electrons experience restoring forces and start moving. If they start from a neighborhood of a stable equilibrium position, there will be range of initial velocities such that a regime of small oscillations is originated.

A simple harmonic analysis of the frequency spectrum of the orbital classical fields close to the minimum defining the ground state yields [82]

$$\omega_{ij} = \left(\frac{2(\epsilon_i - \epsilon_j)}{\mu} \right)^{1/2}, \quad (1.66)$$

where ϵ_j and ϵ_i are the eigen values of occupied and unoccupied orbitals, respectively. The analytic estimate for the lowest possible electronic frequency

$$\omega_e^{\min} \propto \left(\frac{E_{\text{gap}}}{\mu} \right)^{1/2} \quad (1.67)$$

shows that this frequency increases like the square root of the electronic energy difference E_{gap} between the lowest unoccupied and the highest occupied orbital. On the other hand it increases similarly for a decreasing fictitious mass parameter μ . Since the parameters E_{gap} and the maximum phonon frequency (ω_n^{\max}) are dictated by physics, the only parameter in our hands to control adiabatic separation is the fictitious mass, which is therefore also called *adiabaticity parameter*. However, decreasing μ not only shifts the electronic spectrum upwards on the frequency scale, but also stretches the entire frequency spectrum according to (1.66). This leads to an increase of the maximum frequency according to

$$\omega_e^{\max} \propto \left(\frac{E_{\text{cut}}}{\mu} \right)^{1/2}, \quad (1.68)$$

where E_{cut} is the largest kinetic energy in an expansion of the wavefunction in terms of a plane wave basis set. Limitation to decrease arbitrarily kicks in due to the

maximum length of the molecular dynamics time step Δt^{\max} that can be used. The time step is inversely proportional to the highest frequency in the system, which is ω_e^{\max} and thus the relation

$$\Delta t^{\max} \propto \left(\frac{\mu}{E_{\text{cut}}} \right)^{1/2}. \quad (1.69)$$

In the limit, when, electronic gap is very small or even vanishes $E_{\text{gap}} \rightarrow 0$ as is the case for metallic systems, all the above-given arguments break down due to the occurrence of zero-frequency electronic modes in the power spectrum according to (1.67), which necessarily overlap with the phonon spectrum. It has been shown that the coupling of separate Nosé-Hoover thermostats [68, 69, 83] to the nuclear and electronic subsystem can maintain adiabaticity by counterbalancing the energy flow from ions to electrons so that the electrons stay *cool* [84]; see [85] for a similar idea to restore adiabaticity. Although this method is demonstrated to work in practice [86], this ad hoc cure is not entirely satisfactory from both a theoretical and practical point of view so that the well-controlled Born-Oppenheimer approach is recommended for strongly metallic systems.

1.9 Potential Energy Surface

In the past two decades, or so, there have been dramatic improvements in both the accuracy and efficiency of high-level electronic structure calculations [87, 88, 89, 90]. These advances, along with the increasing speed of modern computers have made possible very high-quality ab initio calculations for small polyatomic systems [91, 92]. For three- and four-atom systems, calculations with errors less than 1 kcal/mol are feasible. Gradients and Hessians are also becoming widely available. However, many uses of this vast supply of data require that it be re-expressed with a suitable local or global representation as a potential energy surface (PES). Since the inception of quantum mechanics, considerable effort has been devoted to finding better ways of utilizing ab initio data and/or experimental data to construct PES. The earliest and most common methods involve least-squares fitting to empirical or semi-empirical functional forms [71, 93, 94]. This approach is mature and well understood, although sophisticated schemes involving complex functional forms continue to evolve. During the past decade, generic multivariate interpolation techniques have gathered attention as alternatives to complicated functional forms [95, 96, 97, 98, 99]. The goal of these methods is to produce a general framework for constructing PESs that will reduce the effort and expertise required to turn high-level calculations into usable surfaces.

Another solution is to skip the surface construction step entirely and to use the ab initio results directly in dynamical studies [100, 101, 102]. However, such direct dynamics techniques are inherently classical trajectory approaches and require tens of ab initio calculations for dynamically significant trajectories, and thus, this approach is limited by the available electronic structure calculation techniques. Its application

has been restricted to cases in which modest potential quality seems sufficient and in which discrete spectral lines or state-selected dynamics are not required, as in rate constant calculations based on classical trajectories [103] or in transition state theory [104, 105]. In contrast, the highest-accuracy *ab initio* calculations can take hours or more of computer time, even for small systems. Another obstacle for *on-the-fly* calculations of *ab initio* energies is the failure or non-convergence of the *ab initio* method. One frequently comes across this problem when the nuclear configurations are in a state for which the selected *ab initio* method fails. This is seen in particular for dissociating molecules. The absence of *ab initio* energy on the surface can be treated as *hole* in the surface and can be corrected on the pre-calculated surface. Moreover, carefully adding the *ab initio* fragment data for the dissociating molecule allows to study reaction dynamics on high quality surface. Thus, the construction of accurate analytic representations of PES is a necessary step in full quantum spectroscopic and dynamics studies.

The number of high-level *ab initio* data points currently needed for adequate sampling of dynamically significant regions typically ranges from several hundred to several thousand points for tri- and tetra-atomic systems. Methods that use derivatives typically use fewer configurations; however, the number of pieces of information is typically in the same range [106, 107, 108, 109, 110, 111, 112, 113].

In constructing the PES the prescribed functional form must be carefully crafted so that it

- (i) does not introduce arbitrary features,
- (ii) achieves the required smoothness,
- (iii) preserves any necessary permutation symmetry, and
- (iv) agrees with any known asymptotic form of the underlying PES.

An analytic fit that has a residual significantly larger than the error in the high level *ab initio* calculations is only marginally more useful than if a lower-level calculation is employed. High-quality *ab initio* calculations demand representations that preserve their level of accuracy.

One such method named, Reproducing kernel Hilbert space (RKHS) was introduced by Hollebeek et al. [114]. Several other examples of carefully crafted analytic representations are listed in [114].

1.10 Advanced Numerical Methods

A system consisting of N particles in which the particles interact through forces with a cutoff distance R_c , each particle feels the forces from $N_c \propto \rho R_c^3$ neighbors. CPU time required to advance the system one time step δt is proportional to the number of forces calculated, $NN_c/2$. Clearly the simulation time grows as the cube of the cutoff distance. A frequently encountered problem in molecular dynamics is how to treat the long times that are required to simulate condensed systems consisting of particles interacting through long range forces. Standard methods require the calculation of the forces at every time step. Because each particle interacts with

all particles within the interaction range of the potential the longer the range of the potential the larger the number forces that must be calculated at each time step.

1.10.1 Ewald Summation Method

The Ewald summation is the method of choice to compute electrostatic interactions in systems with periodic boundary conditions [56]. It avoids all problems associated with the use of a cut-off radius and there is no need for switching or shifting functions. Lennard-Jones interactions are calculated normally; due to their shorter range the errors are normally negligible. The Ewald sum consists of a short-range term that is computed in normal space (r -part) and a second term, the k -sum, that is calculated in Fourier-space (k -space). A parameter, usually labeled κ or η , controls the relationship between the two parts. Its value should be chosen so that the r -part interaction between a pair of particles is zero at the cut-off distance, which is still used although it is more a formal parameter in Ewald summation. The more one dampens the r -part (and thus shortens the computer time required for its calculation), the more costly the calculation of the k -sum becomes. Even highly optimized computer codes for the Ewald sum are, therefore, slower than cut-off based methods. If one does not make an error in the choice of η (κ) vs. the cut-off distance and includes enough terms in the k -sum, the calculation of the electrostatic energy using the Ewald summation is exact.

1.10.1.1 Minimum Image

The simulation region or *cell* is effectively replicated in all spatial directions, so that particles leaving the cell reappear at the opposite boundary. For systems governed by a short-ranged potential – say Lennard-Jones or hard spheres – it is sufficient to take just the neighbouring simulation volumes into account, leading to the *minimum-image* configuration shown in Fig. 1.3.

The potential seen by the particle at \mathbf{r}_i is summed over all other particles \mathbf{r}_j , or their periodic images ($\mathbf{r}_j \pm \mathbf{n}$), where $\mathbf{n} = (i_{\hat{x}}, i_{\hat{y}}, i_{\hat{z}})L$, with $i_{\alpha} = 0, \pm 1, \pm 2, \pm 3, \dots \pm \infty$ whichever is closest. L denotes the length of the simulation box. More typically, this list is further restricted to particles lying within a sphere centred on \mathbf{r}_i^6 . For long-range potentials, this arrangement is inadequate because the contributions from more distant images at $2L, 3L$ etc., are no longer negligible.

1.10.1.2 Ewald Summation

One is faced with the challenge of arranging the terms in the potential energy equation so that the contribution from oppositely charged pairs of charges cancel and the summation series converges, and preferably as fast as possible.

A way to achieve this is to add image cells radially outwards from the origin as shown in Fig. 1.4 (this is to build up sets of images contained within successively larger spheres surrounding the simulation region).

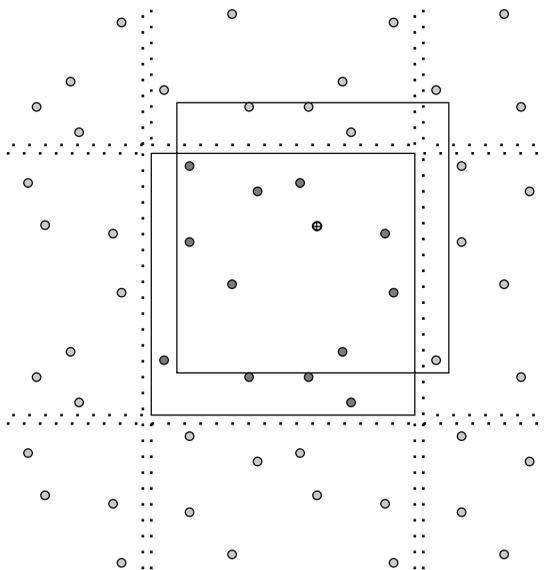


Fig. 1.3. Periodic boundary conditions for simulation region (centre, *dark-shaded particles* at positions \mathbf{r}_j), showing *minimum-image* box for reference ion \oplus at position \mathbf{r}_i containing nearest periodic images (*lightshaded particles* at positions $\mathbf{r}_j \pm \mathbf{n}$)

For the above scheme the potential at r_i due to charges at r_j and image cells is

$$V_s(\mathbf{r}_i) = \sum_n' \sum_{j=1}^N \frac{q_j}{|\mathbf{r}_{ij} + \mathbf{n}|}, \tag{1.70}$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, \mathbf{n} and i_α is same as above. The prime in the summation over \mathbf{n} indicates that the term $j = i$ is omitted for the primary cell $\mathbf{n} = 0$. Taking the image cells in the order perscribed by Fig. 1.4 ensures that the sum in (1.70) converges

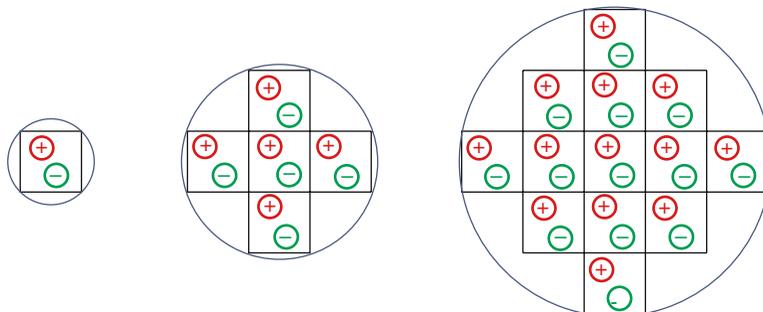


Fig. 1.4. Constructing a convergent sum over periodic images (adapted from Allen & Tildesley)

to the correct value, but only slowly. The summation over the boxes as shown in (1.70) is computationally expensive for N -body problem. The $O(N^2)$ is turned into a $N_{\text{box}} \times N^2$ operation problem.

Ewald's idea got around this problem by recasting the potential equation into sum of two rapidly converging series, one in real space and one in the reciprocal k -space. Consider the simple Gaussian distribution originally used by Ewald himself

$$\sigma(r) = \frac{\alpha^3}{\pi^{3/2}} e^{-\alpha^2 r^2} , \tag{1.71}$$

which is normalized such that

$$\int_0^\infty \sigma(r) dr = 1 . \tag{1.72}$$

Note that α determines the height and width of the effective size of the charges (called spreading function). To obtain the real-space term depicted in Fig. 1.5, we just subtract the lattice sum for the smeared out charges from the original point-charge sum, thus

$$\begin{aligned} V_r(\mathbf{r}_i) &= \sum'_{\mathbf{n}} \sum_{j=1}^N \frac{q_j}{|\mathbf{r}_{ij} + \mathbf{n}|} \left[1 - \int_0^\infty \sigma(r - r_{ij}) d^3r \right] \\ &= \sum'_{\mathbf{n}} \sum_j q_j \left[\frac{1}{|\mathbf{r}_{ij} + \mathbf{n}|} - \frac{4\alpha^3}{\pi^{1/2} |\mathbf{r}_{ij} + \mathbf{n}|} \int_0^{|\mathbf{r}_{ij} + \mathbf{n}|} r^2 e^{-\alpha^2 r^2} dr \right. \\ &\quad \left. - \frac{4\alpha^3}{\pi^{1/2}} \int_{|\mathbf{r}_{ij} + \mathbf{n}|}^\infty r e^{-\alpha^2 r^2} dr \right] . \end{aligned} \tag{1.73}$$

The second term in the above equation can be integrated by parts to give an error function

$$\text{erfc}(x) = 1 - \frac{2}{\pi^{1/2}} \int_0^x e^{-t^2} dt , \tag{1.74}$$

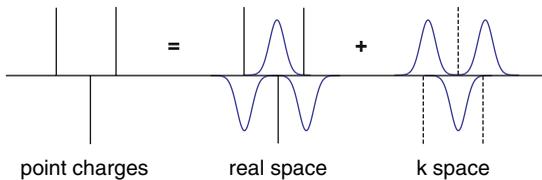


Fig. 1.5. Splitting the sum for point charges into two rapidly convergent series for Gaussian-shaped charges

plus a term which exactly cancels the third term. This gives

$$V_s(\mathbf{r}_i) = \sum_{\mathbf{n}}' \sum_{j=1}^N q_j \frac{\operatorname{erfc}(\alpha|\mathbf{r}_{ij} + \mathbf{n}|)}{|\mathbf{r}_{ij} + \mathbf{n}|} . \quad (1.75)$$

Now for the reciprocal-space sum, consider the charge density of the whole lattice at some arbitrary position r

$$\rho(r) = \sum_j q_j \delta(r - r_j) . \quad (1.76)$$

Since the lattice is periodic, we can express this equivalently as a Fourier sum

$$\rho(r) = L^{-3} \sum_j \sum_{\mathbf{k}} f(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}} , \quad (1.77)$$

where $\mathbf{k} = 2\pi/(L(i_{\hat{x}}, i_{\hat{y}}, i_{\hat{z}}))$; $i_{\alpha} = 0, 1, 2, \dots$ etc. and

$$f(\mathbf{k}) = \int_{L^3} \rho(r) e^{i\mathbf{k}\cdot\mathbf{r}} d^3r , \quad (1.78)$$

where the integration is restricted to the unit cell volume $V = L^3$. Substituting $\rho(r)$ from (1.76) into (1.78) and making use of standard identity picks out modes corresponding to the point charges

$$f(\mathbf{k}) = \sum_j q_j e^{i\mathbf{k}\cdot\mathbf{r}_j} . \quad (1.79)$$

The smeared charge distribution is

$$\rho'(r) = \sum_j q_j \sigma(r - r_j) = \int_{L^3} \rho(r - r') \sigma(r') d^3r' . \quad (1.80)$$

This is the convolution of function $\rho(r)$ with function $\sigma(r)$, which can be expressed in Fourier space as

$$\rho'(r) = \frac{1}{L^3} \sum_{\mathbf{k}}' f(\mathbf{k}) \phi(\mathbf{k}, \alpha) e^{-i\mathbf{k}\cdot\mathbf{r}} , \quad (1.81)$$

where $\phi(\mathbf{k}, \alpha)$ is the Fourier transform of the charge-smearing function $\sigma(r)$, i.e.

$$\phi(\mathbf{k}, \alpha) = e^{-|\mathbf{k}|^2/(4\alpha^2)} . \quad (1.82)$$

The potential due to the smeared charges in k -space at the reference position r_i is

$$V_k(r_i) = \int_0^{\infty} \frac{\rho'(r_i + r)}{r} dr = \frac{1}{L^3} \sum_{\mathbf{k}}' f(\mathbf{k}) \phi(\mathbf{k}, \alpha) e^{-i\mathbf{k}\cdot\mathbf{r}} \int_0^{\infty} \frac{e^{-i\mathbf{k}\cdot\mathbf{r}}}{r} d^3r . \quad (1.83)$$

The integral on the right of the above expression is $4\pi/k^2$. Combining this with earlier results from (1.79) and (1.82) we get

$$V_k(r_i) = \frac{4\pi}{L^3} \sum'_{\mathbf{k}} \sum_j q_j e^{i\mathbf{k}\cdot(\mathbf{r}_j - \mathbf{r}_i)} \frac{e^{-|\mathbf{k}|^2/(4\alpha^2)}}{|\mathbf{k}|^2}. \quad (1.84)$$

This potential includes an unphysical *self-term* corresponding to a smeared out charge centered at r_i , which needs to be subtracted off:

$$\begin{aligned} V_s(r_i) &= q_i \int_0^\infty \sigma(r) d^3r \\ &= \frac{4\pi q_i \alpha^3}{\pi^{3/2}} i \int_0^\infty r e^{-\alpha^2 r^2} d^3r \\ &= \frac{2\alpha}{\pi^{1/2}} q_i. \end{aligned} \quad (1.85)$$

Adding the partial sum given by (1.75), (1.84) and (1.85) we obtain the Ewald sum

$$\begin{aligned} V_E(\mathbf{r}_i) &= \sum'_{\mathbf{n}} \sum_{j=1}^N q_j \frac{\operatorname{erfc}(\alpha|\mathbf{r}_{ij} + \mathbf{n}|)}{|\mathbf{r}_{ij} + \mathbf{n}|} \\ &\quad + \frac{4\pi}{L^3} \sum_{\mathbf{k} \neq 0} \sum_j q_j e^{-|\mathbf{k}|^2/(4\alpha^2)} e^{i\mathbf{k}\cdot(\mathbf{r}_j - \mathbf{r}_i)} - \frac{2\alpha}{\pi^{1/2}} q_i \end{aligned} \quad (1.86)$$

and the force on charge i is given by

$$\begin{aligned} \mathbf{f}_i &= -\nabla_{\mathbf{r}_i} U \\ &= \underbrace{\frac{q_i}{4\pi\epsilon_0} \sum_{\mathbf{n}} \sum_{j=1, j \neq i}^N q_j \left[\frac{\operatorname{erfc}(\alpha|\mathbf{r}_{ij} + \mathbf{n}|)}{|\mathbf{r}_{ij} + \mathbf{n}|} + \frac{2\alpha}{\sqrt{\pi}} e^{-\alpha^2|\mathbf{r}_{ij} + \mathbf{n}|^2} \right] \frac{\mathbf{r}_{ij} + \mathbf{n}}{|\mathbf{r}_{ij} + \mathbf{n}|}}_{\text{Real-space term}} \\ &\quad + \underbrace{\frac{2}{\epsilon_0 V} \sum_{\mathbf{k} > 0} q_i \frac{\mathbf{k}}{k^2} e^{-k^2/(4\alpha^2)} \left[\sin(\mathbf{k} \cdot \mathbf{r}_i) \sum_{j=1}^N q_j \cos(\mathbf{k} \cdot \mathbf{r}_j) \right. \\ &\quad \quad \left. - \cos(\mathbf{k} \cdot \mathbf{r}_i) \sum_{j=1}^N q_j \sin(\mathbf{k} \cdot \mathbf{r}_j) \right]}_{\text{Reciprocal-space term}} \\ &\quad - \underbrace{\frac{q_i}{6\epsilon_0 V} \sum_{j=1}^N q_j \mathbf{r}_j}_{\text{Surface dipole term}}. \end{aligned} \quad (1.87)$$

One needs an additional correction for the intra-molecular self-energy

$$-\frac{1}{4\pi\epsilon_0} \sum_{n=1}^M \sum_{\kappa=1}^{N_m} \sum_{\lambda=\kappa+1}^{N_m} q_{n\kappa} q_{n\lambda} \frac{\text{erf}(\alpha|\mathbf{r}_{\kappa\lambda}|)}{|\mathbf{r}_{\kappa\lambda}|}, \quad (1.88)$$

whose derivative is absent from the equation for the forces (1.87). This term corrects for interactions between charges on the same molecule which are implicitly included in the reciprocal space sum, but are not required in the rigid-molecule model. Although the site forces \mathbf{f}_i , do include unwanted terms, these sum to zero in the evaluation of the molecular center-of-mass forces and torques (by the conservation laws for linear and angular momentum).

Both, the real- and reciprocal-space series (the sums over \mathbf{n} and \mathbf{k}) converge fairly rapidly so that only a few terms are need to be evaluated. One defines the cut-off distances r_c and k_c so that only terms with $|\mathbf{r}_{ij} + \mathbf{n}| < r_c$ and $|\mathbf{k}| < k_c$ are included. The parameter α determines how rapidly the terms decrease and the values of r_c and k_c needed to achieve a given accuracy.

For a fixed α and accuracy the number of terms in the real-space sum is proportional to the total number of sites, N but the cost of the reciprocal-space sum increases as N^2 . An overall scaling of $N^{3/2}$ may be achieved if α varies with N . This is discussed in detail in an excellent article by D. Fincham [115]. The optimal value of α is

$$\alpha = \sqrt{\pi} \left(\frac{t_R}{t_F} \frac{N}{V^2} \right)^{\frac{1}{6}}, \quad (1.89)$$

where t_R and t_F are the execution times needed to evaluate a single term in the real- and reciprocal-space sums respectively. If we require that the sums converge to an accuracy of $\epsilon = \exp(-p)$ the cutoffs are then given by

$$r_c = \frac{\sqrt{p}}{\alpha}, \quad (1.90)$$

$$k_c = 2\alpha\sqrt{p}. \quad (1.91)$$

A representative value of t_R/t_F has been established as 5.5. Though this will vary on different processors and for different potentials its value is not critical since it enters the equations as a sixth root.

It must be emphasized that the r_c is used as a cutoff for the short-ranged potentials as well as for the electrostatic part. The value chosen above does not take the nature of the non-electrostatic part of the potential into account.

1.10.1.3 Uniform Sheet Correction

In a periodic system the electrostatic energy is finite only if the total electric charge of the MD cell is zero. The reciprocal space sum for $\mathbf{k} = 0$ takes the form

$$\frac{1}{k^2} e^{-k^2/(4\alpha^2)} \left| \sum_{i=1}^N q_i \right|^2, \quad (1.92)$$

which is zero in the case of electro-neutrality but infinite otherwise. Its omission is physically equivalent to adding a uniform jelly of charge which exactly neutralizes the unbalanced point charges. But though the form of the reciprocal space sum is unaffected by the uniform charge jelly the real-space sum is not. The real-space part of the interaction of the jelly with each point charge as well as the self-energy of the jelly itself must be included giving

$$- \frac{1}{8\epsilon_0 V \alpha^2} \left| \sum_{i=1}^N q_i \right|^2. \quad (1.93)$$

1.10.1.4 Surface Dipole Term

This term accounts for different periodic boundary conditions. It was suggested by De Leeuw, Perram and Smith [116, 117, 118] in order to accurately model dipolar systems and is necessary in any calculation of a dielectric constant

$$+ \left[\frac{1}{6\epsilon_0 V} \left| \sum_{i=1}^N q_i \mathbf{r}_i \right|^2 \right]. \quad (1.94)$$

Consider a near-spherical cluster of MD cells. The *infinite* result for any property is the limit of its *cluster* value as the size of the cluster tends to infinity. However, this value is non-unique and depends on the dielectric constant, ϵ_s of the physical medium surrounding the cluster. If this medium is conductive ($\epsilon_s = \infty$) the dipole moment of the cluster is neutralized by image charges, whereas in a vacuum ($\epsilon_s = 1$) it remains. It is trivial to show that in that case the dipole moment per unit volume (or per MD cell) does not decrease with the size of the cluster. This term is then just the dipole energy, and ought to be used in any calculation of the dielectric constant of a dipolar molecular system.

1.10.2 Multipole Methods

There is a large number of N -body problems for which periodic boundaries are completely inappropriate, for example: galaxy dynamics, electron-beam transport, large proteins [119], and any number of problems with complex geometries. Two new approaches were put forward in the mid-1980's, the first from Appel [120] and Barnes & Hut [121], who proposed $O(N \log N)$ -schemes based on hierarchical grouping of distant particles; the second from Greengard & Rohklin [122] with an $O(N)$ (better than $O(N \log N)$) solution with rounding-error accuracy. These two methods are known today as the *hierarchical tree algorithm* and the *Fast Multipole Method* (FMM) respectively – have revolutionized N -body simulation in a much

broader sense than the specialized periodic methods discussed earlier. They offer a generic means of accelerating the computation of many-particle systems governed by central, long-range potentials.

References

1. Y. Duan, L. Wang, P. Kollman, P. Natl. Acad. Sci. USA **95**, 9897 (1998) 3
2. Q. Zhong, P. Moore, D. News, M. Klein, FEBS Lett. **427**, 267 (1998) 3
3. Q. Zhong, Q. Jiang, P. Moore, D. News, M. Klein, Biophys. J. **74**, 3 (1998) 3
4. R. Car, M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985) 3, 25
5. G. Galli, M. Parrinello, in *Proceedings of the NATO Advanced Study Institute on Computer Simulation in Material Science: Interatomic Potentials, Simulation Techniques and Applications, Aussois, France, 25 March - 5 April 1991*, Vol. 3, ed. by M. Meyer, V. Pontikis (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991), Vol. 3, pp. 283–304 3
6. D. Heermann, *Computer Simulation Methods* (Springer, Berlin Heidelberg New York, 1986) 5
7. H. Berendsen, J. Postma, W. van Gunsteren, A. DiNola, J. Haak, J. Chem. Phys. **81**, 3684 (1984) 7
8. H. Andersen, J. Chem. Phys. **72**, 2384 (1980) 7
9. W. Hoover, Phys. Rev. A **31**, 1695 (1985) 7
10. A. Voter, F. Montalenti, T. Germann, Annu. Rev. Mater. Res. **32**, 321 (2002) 8
11. J. Lennard-Jones, P. Roy. Soc. Lond. **43**, 461 (1931) 8
12. P. Morse, Phys. Rev. **34**, 57 (1929) 9, 14
13. A. Rahman, Phys. Rev. **136**, A405 (1964) 9
14. L. Verlet, Phys. Rev. **159**, 98 (1967) 9, 14
15. J. Tersoff, Phys. Rev. Lett. **56**, 632 (1986) 13
16. J. Tersoff, Phys. Rev. B **37**, 6991 (1988) 13
17. J. Tersoff, Phys. Rev. Lett. **61**, 2879 (1988) 13
18. J. Tersoff, Phys. Rev. B **39**, 5566 (1989) 13
19. W. Jorgensen, J. Madura, C. Swenson, J. Am. Chem. Soc. **106**, 6638 (1984) 13
20. N. Allinger, K. Chen, J. Lii, J. Comput. Chem. **17**, 642 (1996) 13
21. W. Jorgensen, D. Maxwell, J. Tiradorives, J. Am. Chem. Soc. **118**, 11225 (1996) 13
22. W. Cornell, P. Cieplak, C. Bayly, I. Gould, K. Merz, D. Ferguson, D. Spellmeyer, T. Fox, J. Caldwell, P. Kollman, J. Am. Chem. Soc. **118**, 2309 (1996) 13
23. T. Halgren, J. Comput. Chem. **17**, 490 (1996) 13
24. S. Nath, F. Escobedo, J. de Pablo, J. Chem. Phys. **108**, 9905 (1998) 13
25. M. Martin, J. Siepmann, J. Phys. Chem. B **102**, 2569 (1998) 13
26. H. Sun, J. Phys. Chem. B **102**, 7338 (1998) 13
27. D. Brenner, Mat. Res. Soc. Symp. Proc. **141**, 59 (1989) 13
28. M. Ramana Murty, H. Atwater, Phys. Rev. B **51**, 4889 (1995) 13
29. A. Dyson, P. Smith, Surf. Sci. **355**, 140 (1996) 13
30. D. Brenner, Phys. Rev. B **42**, 9458 (1990) 13
31. D. Brenner, Phys. Rev. B **46**, 1948 (1992) 13
32. D. Brenner, J. Harrison, C. White, R. Colton, Thin Solid Films **206**, 220 (1991) 13
33. D. Brenner, K. Tupper, S. Sinnott, R. Colton, J. Harrison, Abstr. Pap. Am. Chem. S. **207**, 166 (1994) 13

34. J. Harrison, S. Stuart, D. Robertson, C. White, *J. Phys. Chem. B* **101**, 9682 (1997) 13
35. S. Sinnott, R. Colton, C. White, O. Shenderova, D. Brenner, J. Harrison, *J. Vac. Sci. Technol. A* **15**, 936 (1997) 13
36. J. Harrison, C. White, R. Colton, D. Brenner, *Phys. Rev. B* **46**, 9700 (1992) 13
37. J. Harrison, R. Colton, C. White, D. Brenner, *Wear* **168**, 127 (1993) 13
38. J. Harrison, C. White, R. Colton, D. Brenner, *J. Phys. Chem.* **97**, 6573 (1993) 13
39. J. Harrison, D. Brenner, *J. Am. Chem. Soc.* **116**, 10399 (1994) 13
40. J. Harrison, C. White, R. Colton, D. Brenner, *Thin Solid Films* **260**, 205 (1995) 13
41. M. Perry, J. Harrison, *Langmuir* **12**, 4552 (1996) 13
42. D. Allara, A. Parikh, E. Judge, *J. Chem. Phys.* **100**, 1761 (1994) 13
43. R. Smith, K. Beardmore, *Thin Solid Films* **272**, 255 (1996) 13
44. M. Nyden, T. Coley, S. Mumby, *Polym. Eng. Sci* **37**, 1496 (1997) 13, 14
45. J. Che, T. Cagin, W. Goddard, *Theor. Chem. Acc.* **102**, 346 (1999) 13, 14
46. K. Nordlund, J. Keinonen, T. Mattila, *Phys. Rev. Lett.* **77**, 699 (1996) 13
47. S. Stuart, B. Berne, *J. Phys. Chem.* **100**, 11934 (1996) 14
48. R. Hockney, J. Eastwood, *Computer Simulation Using Particles* (McGraw-Hill, New-York, USA, 1981) 14
49. W. Swope, H. Andersen, P. Berens, K. Wilson, *J. Chem. Phys.* **76**, 637 (1982) 14
50. D. Beeman, *J. Comput. Phys.* **20**, 130 (1976) 14
51. G. Martyna, M. Tuckerman, *J. Chem. Phys.* **102**, 8071 (1995) 14
52. M. Tuckerman, B. Berne, G. Martyna, *J. Chem. Phys.* **97**, 1990 (1992) 14
53. C. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations (Chap. 9)* (Prentice Hall, Englewood Cliffs, NJ, USA, 1971) 16, 18
54. H. Yoshida, *Phys. Lett. A* **150**, 262 (1990) 18
55. D. Frenkel, B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, San Diego, 1996) 19
56. M. Allen, D. Tildesley, *Computer simulation of liquids* (Clarendon Press, Oxford, 1987) 19, 20, 21, 23, 30
57. B. Todd, D. Evans, P. Daivis, *Phys. Rev. E* **52**, 1627 (1995) 19
58. J. Irving, J. Kirkwood, *J. Chem. Phys.* **18**, 817 (1950) 19
59. D. McQuarrie, *Statistical Mechanics* (Harper and Row, New York, 1976) 20
60. D. Frenkel, B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, San Diego, 2002) 20
61. M. Warrior, R. Schneider, E. Salonen, K. Nordlund, *Contrib. Plasma Phys.* **44**, 307 (2004) 21, 23
62. J. Klafter, M. Shlesinger, G. Zumofen, *Phys. Today* **2**, 33 (1996) 21
63. E. Weinan, B. Engquist, *Not. Am. Math. Soc* **50**, 1062 (2003) 22
64. B. Berne, G. Ciccotti, C. D.F. (eds.), *Classical and Quantum Dynamics in Condensed Phase Simulations* (World Scientific Publishing Company, Singapore, Singapore, 1998) 23
65. K. Binder, G. Ciccotti (eds.), *Monte Carlo and Molecular Dynamics of Condensed Matter Systems* (Editrice Compositori, Bologna, Italy, 1996) 23
66. G. Ciccotti, D. Frenkel, I. McDonald, *Simulation of Liquids and Solids* (North Holland, Amsterdam, 1987) 23
67. R. Esser, P. Grassberger, J. Grotendorst, M. Lewerenz (eds.), *Molecular Dynamics on Parallel Computers* (World Scientific Publishing Company, Singapore, Singapore, 1999) 23
68. D. Frenkel, B. Smit, *Understanding Molecular Simulations: From Algorithms to Applications* (Academic Press, San Diego, 2005) 23, 28

69. R. Haberlandt, S. Fritzsche, G. Peinel, K. Heinzinger, *Molekulardynamik - Grundlagen und Anwendungen* (H.-L.Vörtlter, Lehrbuch, Vieweg, Wiesbaden, 1995) 23, 28
70. G. Gray, K. Gubbins, *Theory of Molecular Fluids* (Clarendon Press, Oxford, 1984) 23
71. G. Schatz, *Rev. Mod. Phys.* **61**, 669 (1989) 23, 28
72. M. Sprik, in *NATO ASI Series C*, Vol. 397, ed. by M. Allen, D. Tildesley (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993), Vol. 397, pp. 211–259 23
73. M. Segall, P. Lindan, M. Probert, C. Pickard, P. Hasnip, S. Clark, M. Payne, *J. Phys-Condens. Mat.* **14**, 2717 (2002) 24
74. P. Blöchl, *Phys. Rev. B* **50**, 17953 (1994) 24
75. M. Bockstedte, A. Kley, J. Neugebauer, M. Scheffler, *Comput. Phys. Commun.* **107**, 187 (1997) 24
76. R. Kendall, E. Apra, D. Bernholdt, E. Bylaska, M. Dupuis, G. Fann, R. Harrison, J. Ju, J. Nichols, J. Nieplocha, T. Straatsma, T. Windus, A. Wong, *Comput. Phys. Commun.* **128**, 260 (2000) 24
77. G. Kresse, J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996) 24
78. M. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, J. Montgomery, Jr., T. Vreven, K. Kudin, J. Burant, J. Millam, S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. Knox, H. Hratchian, J. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. Stratmann, O. Yazyev, A. Austin, R. Cammi, C. Pomelli, J. Ochterski, P. Ayala, K. Morokuma, G. Voth, P. Salvador, J. Dannenberg, V. Zakrzewski, S. Dapprich, A. Daniels, M. Strain, O. Farkas, D. Malick, A. Rabuck, K. Raghavachari, J. Foresman, J. Ortiz, Q. Cui, A. Baboul, S. Clifford, J. Cioslowski, B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Martin, D. Fox, T. Keith, C. Al-Laham, M.A. and Peng, A. Nanayakkara, M. Challacombe, P. Gill, B. Johnson, W. Chen, M. Wong, C. Gonzalez, J. Pople, *Gaussian 03 User's Reference, Revision C.02*. Gaussian, Inc., Wallingford, CT, 2004 24
79. H.J. Werner, P. Knowles, R. Lindh, R. Manby, M. Schütz, P. Celani, T. Korona, G. Rauhut, R. Amos, A. Bernhardsson, A. Berning, D. Cooper, M. Deegan, A. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, A. Lloyd, S. McNicholas, W. Meyer, M. Mura, A. Nicklass, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A. Stone, R. Tarroni, T. Thorsteinsson, *MOLPRO, Version 2006.1*. Cardiff, UK (2006). A package of ab initio programs, see <http://www.molpro.net> 24
80. X. Gonze, J.M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, P. Ghosez, J. Raty, D. Allan, *Comp. Mater. Sci.* **25**, 478 (2002) 24
81. X. Gonze, G. Rignanese, M. Verstraete, J. Beuken, Y. Pouillon, R. Caracas, F. Jollet, M. Torrent, G. Zerah, M. Mikami, P. Ghosez, M. Veithen, J. Raty, V. Olevanov, F. Bruneval, L. Reining, R. Godby, G. Onida, D. Hamann, D. Allan, *Z. Kristallogr.* **220**, 558 (2005) 24
82. G. Pastore, E. Smargiassi, F. Buda, *Phys. Rev. A* **44**, 6334 (1991) 27
83. M. Allen, D. Tildesley, *Computer simulation of liquids* (Clarendon Press: Oxford, 1990) 28
84. P.E. Blöchl, M. Parrinello, *Phys. Rev. B* **45**, 9413 (1992) 28
85. E. Fois, A. Selloni, M. Parrinello, R. Car, *J. Phys. Chem.* **92**, 3268 (1988) 28
86. A. Pasquarello, K. Laasonen, R. Car, C. Lee, D. Vanderbilt, *Phys. Rev. Lett.* **69**, 1982 (1992) 28

87. Y. Yamaguchi, Y. Osamura, J. Goddard, H. Schaefer, *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory* (Oxford University Press, New York, 1994) 28
88. W. Hehre, L. Radom, P. Schleyer, J. Pople, *Ab initio molecular orbital theory* (Wiley, New York, 1986) 28
89. M. Headgordon, *J. Phys. Chem.* **100**, 13213 (1996) 28
90. W. Kohn, A. Becke, R. Parr, *J. Phys. Chem.* **100**, 12974 (1996) 28
91. T. Dunning, *Advances in Molecular and Electronic Structure Theory*, Vol. 1 (Jai Press, Greenwich, CT, 1990) 28
92. B. Jeziorski, R. Moszynski, K. Szalewicz, *Chem. Rev.* **94**, 1887 (1994) 28
93. J. Murrell, S. Carter, S. Farantos, P. Huxley, A. Varandas, *Molecular Potential Energy Functions* (John Wiley and Sons Ltd, New York, 1984) 28
94. D. Truhlar, R. Steckler, M. Gordon, *Chem. Rev.* **87**, 217 (1987) 28
95. J. Ischtwan, M. Collins, *J. Chem. Phys.* **100**, 8080 (1994) 28
96. M. Collins, *Adv. Chem. Phys.* **93**, 389 (1996) 28
97. T.S. Ho, H. Rabitz, *J. Chem. Phys.* **104**, 2584 (1996) 28
98. T. Hollebeck, T.S. Ho, H. Rabitz, *J. Chem. Phys.* **106**, 7223 (1997) 28
99. T.S. Ho, H. Rabitz, in *Fashioning a Model: Optimization Methods in Chemical Physics*, ed. by A. Ernesti, J. Hutson, N. Wright (1998), pp. 28–34 28
100. T. Helgaker, E. Uggerud, H. Jensen, *Chem. Phys. Lett.* **173**, 145 (1990) 28
101. W. Chen, W. Hase, H. Schlegel, *Chem. Phys. Lett.* **228**, 436 (1994) 28
102. R. Steckler, G. Thurman, J. Watts, R. Bartlett, *J. Chem. Phys.* **106**, 3926 (1997) 28
103. A. Varandas, P. Abreu, *Chem. Phys. Lett.* **293**, 261 (1998) 29
104. Y. Chuang, D. Truhlar, *J. Phys. Chem. A* **101**, 3808 (1997) 29
105. J. Corchado, J. Espinosa-Garcia, O. Roberto-Neto, Y. Chuang, D. Truhlar, *J. Phys. Chem. A* **102**, 4899 (1998) 29
106. M. Jordan, K. Thompson, M. Collins, *J. Chem. Phys.* **102**, 5647 (1995) 29
107. M. Jordan, K. Thompson, M. Collins, *J. Chem. Phys.* **103**, 9669 (1995) 29
108. M. Jordan, M. Collins, *J. Chem. Phys.* **104**, 4600 (1996) 29
109. K. Thompson, M. Collins, *J. Chem. Soc. Faraday T.* **93**, 871 (1997) 29
110. K. Thompson, M. Jordan, M. Collins, *J. Chem. Phys.* **108**, 564 (1998) 29
111. K. Thompson, M. Jordan, M. Collins, *J. Chem. Phys.* **108**, 8302 (1998) 29
112. T. Ishida, G. Schatz, *J. Chem. Phys.* **107**, 3558 (1997) 29
113. I. Takata, T. Taketsugu, K. Hirao, M. Gordon, *J. Chem. Phys.* **109**, 4281 (1998) 29
114. T. Hollebeck, T.S. Ho, H. Rabitz, *Annu. Rev. Phys. Chem.* **50**, 537 (1999) 29
115. D. Fincham, *Mol. Simulat.* **13**, 1 (1994) 35
116. S. Deleeuw, J. Perram, E. Smith, *P. Roy. Soc. Lond. A Mat.* **373**, 27 (1980) 36
117. S. Deleeuw, J. Perram, E. Smith, *P. Roy. Soc. Lond. A Mat.* **373**, 57 (1980) 36
118. S. Deleeuw, J. Perram, E. Smith, *P. Roy. Soc. Lond. A Mat.* **388**, 177 (1983) 36
119. T. Schlick, R. Skeel, A. Brunger, L. Kale, J. Board, J. Hermans, K. Schulten, *J. Comput. Phys.* **151**, 9 (1999) 36
120. A. Appel, *Siam J. Sci. Stat. Comp.* **6**, 85 (1985) 36
121. J. Barnes, P. Hut, *Nature* **324**, 446 (1986) 36
122. L. Greengard, V. Rokhlin, *J. Comput. Phys.* **73**, 325 (1987) 36