Academic year 2018/2019

TYC Materials Modelling Course

Ab-initio Molecular Dynamics Simulation

Prof. Jochen Blumberger

Contents

Recommended book: D. Marx and J. Hutter, *Ab initio Molecular Dynamics: Basic theory and Advanced Methods*, Cambridge University Press, 2009.

1 Introducing Molecular Dynamics Simulation

1.1 The Born-Oppenheimer approximation

The stationary Schrödinger equation of the coupled nuclear-electronic problem reads

$$[\mathbf{T}_n(\mathbf{R}) + \mathbf{T}_e(\mathbf{r}) + V_{en}(\mathbf{R},\mathbf{r}) + V_{ee}(\mathbf{r}) + V_{nn}(\mathbf{R})]\Psi(\mathbf{R},\mathbf{r}) = E_{\text{tot}}\Psi(\mathbf{R},\mathbf{r}) \quad (1)$$

where $\mathbf{T}_n = \sum_I -\frac{1}{2M_I}\nabla^2_{\mathbf{R}_I}$ and $\mathbf{T}_e = \sum_i -\frac{1}{2m}\nabla^2_{\mathbf{r}_i}$ denote the total kinetic energy operator for the nuclei (index I, mass M_I) and electrons (index i, mass m), respectively, and V denotes the usual nuclei-electron (ne), electron-electron (ee) and nuclear-nuclear (nn) interactions. The cartesian positions of the nuclei and electrons are denoted by the vectors $\mathbf{R}$ and $\mathbf{r}$, respectively. The total wavefunction $\Psi(\mathbf{R},\mathbf{r})$ can be expanded in a product of electronic wavefunctions $\Psi^i_{\mathbf{R}}(\mathbf{r})$ and nuclear wavefunctions $\Phi^i(\mathbf{R})$ (Born-Ansatz),

$$\Psi(\mathbf{R},\mathbf{r}) = \sum_i \Psi^i_{\mathbf{R}}(\mathbf{r})\Phi^i(\mathbf{R}), \quad (2)$$

where the $\Psi^i_{\mathbf{R}}(\mathbf{r})$ are solutions of the electronic Schrödinger equation for a fixed set of nuclear coordinates $\mathbf{R}$,

$$[\mathbf{T}_e(\mathbf{r}) + V_{en}(\mathbf{R},\mathbf{r}) + V_{ee}(\mathbf{r}) + V_{nn}(\mathbf{R})]\Psi^i_{\mathbf{R}}(\mathbf{r}) = E^i_{\mathbf{R}}\Psi^i_{\mathbf{R}}(\mathbf{r}). \quad (3)$$

Note that $\Psi^i_{\mathbf{R}}$ and the corresponding eigenvalues $E^i_{\mathbf{R}}$ depend parametrically on the nuclear positions which is indicated by the subscript $\mathbf{R}$.

The Born-Oppenheimer approximation (contd)

In the following we neglect all electronic excitations in Eq. 2 and assume $\Psi(\mathbf{R},\mathbf{r}) = \Psi^0_{\mathbf{R}}(\mathbf{r})\Phi^0(\mathbf{R})$ where $\Psi^0_{\mathbf{R}}$ is the electronic ground state with energy $E^0_{\mathbf{R}}$, and $\Phi^0(\mathbf{R})$ is the corresponding nuclear wavefunction. As we deal only with electronic ground states in the following, we drop the superscript "0".

To proceed we consider the action of the nuclear and electronic kinetic energy operators in Eq. 1 on the product $\Psi_{\mathbf{R}}(\mathbf{r})\Phi(\mathbf{R})$ and note that the mass of the nuclei (M_I) is at least 3 orders of magnitude larger than the mass of the electron (m). This implies that $\mathbf{T}_n(\mathbf{R})\Psi_{\mathbf{R}}(\mathbf{r}) << \mathbf{T}_e(\mathbf{r})\Psi_{\mathbf{R}}(\mathbf{r})$ may be a good approximation. Assuming that $\mathbf{T}_n(\mathbf{R})\Psi_{\mathbf{R}}(\mathbf{r}) = 0$ is called the *Born-Oppenheimer (BO) approximation.*

Inserting Eq. 3 for $i = 0$ in Eq. 1 and assuming the BO-approximation one obtains

$$[\mathbf{T}_n(\mathbf{R}) + E_{\mathbf{R}}]\Phi(\mathbf{R}) = E_{\mathrm{tot}}\Phi(\mathbf{R}). \tag{4}$$

This equation is the time-independent Schrödinger equation for nuclear motion in the BO-approximation. Note, this equation does no longer depend explicitly on the electronic wavefunction. The electronic terms of Eq. 1 are condensed into an effective electronic potential energy $E_{\mathbf{R}}$, that depends parametrically on the nuclear coordinates. $E_{\mathbf{R}}$ is also termed ground state potential energy hyper surface or short, ground state potential energy surface (PES). The above arguments imply that the BO-approximation is exact in the limit of infinitely heavy nuclei, but it is usually a good approximation for physically relevant systems in the electronic ground state as a consequence of the large proton to electron mass ratio.

1.2 Classical approximation of nuclear motion

In the BO approximation the quantum mechanical motion of the nuclei is given by the following time-dependent Schrödinger Equation:

$$i\hbar\frac{\partial}{\partial t}\Phi(\mathbf{R},t) = [\mathbf{T}_n(\mathbf{R}) + E_{\mathbf{R}}]\Phi(\mathbf{R}) \tag{5}$$

This equation can be solved numerically only for a few degrees of freedom. In many situations it is sufficient to treat the nuclei as classical degrees of freedom, e.g. at high temperature and/or heavy nuclei. Classical dynamics can be formally derived from Eq. 5. Writing Φ in polar representation, $\Phi(\mathbf{R},t) = A(\mathbf{R},t)\exp(iS(\mathbf{R},t)/\hbar)$, $A \in R > 0, S \in R$, and inserting into Eq. 5 one obtains after separation of real and imaginary parts

$$\frac{\partial S}{\partial t} + \sum_I \frac{1}{2M_I}\left(\nabla_I S\right)^2 + E_{\mathbf{R}} - \sum_I \frac{\hbar^2}{2M_I}\frac{\nabla_I^2 A}{A} = 0 \tag{6}$$

$$\frac{\partial A}{\partial t} + \sum_I \frac{1}{M_I}\left(\nabla_I S\right)\left(\nabla_I A\right) + \sum_I \frac{1}{2M_I} A\nabla_I^2 S = 0 \tag{7}$$

Taking the limit $\hbar \to 0$ in Eq. 6 one obtains an expression for S that is isomorph to the Hamilton-Jacobi Equation of classical mechanics,

$$\frac{\partial S^{\mathrm{cl}}}{\partial t} + \sum_I \frac{\mathbf{p}_I^2}{2M_I} + E_{\mathbf{R}} = 0 \tag{8}$$

where S^{cl} is the classical action and p_I the classical momentum of particle I. The classical equations of motions are obtained from the derivatives of the Hamiltonian function $H = \sum_I \mathbf{p}_I^2/(2M_I) + E_{\mathbf{R}}$, $\dot{\mathbf{R}}_I = \partial H/\partial \mathbf{p}_I$ and $\dot{\mathbf{p}}_I = -\partial H/\partial \mathbf{R}_I$. The latter equation is Newton's second law of motion,

$$\mathbf{f}_I = M_I\ddot{\mathbf{R}}_I = -\nabla_I E_{\mathbf{R}}. \tag{9}$$

1.3 Molecular dynamics

Assuming classical nuclear motion on a potential energy surface $E_{\mathbf{R}}$, the time evolution or trajectory of the nuclei can be solved using Newton's equation of motion,

$$\mathbf{f}_I = M_I \ddot{\mathbf{R}}_I \tag{10}$$

$$\mathbf{f}_I = -\frac{\partial E_{\mathbf{R}}}{\partial \mathbf{R}_I} \tag{11}$$

where $\mathbf{f}_I = \mathrm{d}\mathbf{p}_I/\mathrm{d}t$ is the force on atom I. Solving the coupled equation for all atoms I in the system is generally referred to as *molecular dynamics* (MD) calculation.

If the potential energy $E_{\mathbf{R}}$ is given by an empirical function of the nuclear coordinates it is termed *classical molecular dynamics*. If $E_{\mathbf{R}}$ is obtained from ab initio or density functional theory (DFT) the calculations are termed *ab initio molecular dynamics* (AIMD) or *density functional based molecular dynamics* (DFTMD).

There are two ways of carrying out AIMD or DFMD simulations in practice. One option is to carry out ab initio or DFT energy calculations for many different configurations of the system followed by a fit of the points to an analytic representation of $E_{\mathbf{R}}$ *prior* to the MD simulation. Once $E_{\mathbf{R}}$ is obtained Eqs. 10-11 are solved. The other option is to solve the electronic structure problem and Eqs. 10-11 simultaneously. This latter option is termed *on the fly* AIMD or DFMD, which is nowadays the standard.

1.4 Time stepping algorithms

The classical dynamics of interacting many-particle systems as given by Equations 10-11 cannot be solved analytically. Iterative numerical schemes have to be used.

The first step is a discretization of time in terms of small increments called *time steps* which we assume to be of equal length δt. Counting the successive equidistant points on the time axis by the index m, $t_m = m\delta t$ with $t_0 = 0$, the evolution of the system is described by a series of the coordinate values

$$\mathbf{R}(t_0) = \mathbf{R}(0), \ldots, \mathbf{R}(t_{m-1}) = \mathbf{R}(t_m - \delta t), \mathbf{R}(t_m), \ldots, \mathbf{R}(t_{m+1}) = \mathbf{R}(t_m + \delta t) \tag{12}$$

plus a similar series for the velocities $\dot{\mathbf{R}}$. In the following we will discuss three popular algorithms for computation of the discrete series of positions and velocities, the Verlet, Velocity Verlet and leap-frog algorithms.

1.5 The Verlet algorithm

The algorithm is based on a Taylor expansion of the coordinates around t forward and backward in time,

$$\mathbf{R}_I(t+\delta t) = \mathbf{R}_I(t) + \dot{\mathbf{R}}_I(t)\delta t + \frac{\mathbf{f}_I(t)}{2M_I}\delta t^2 + \frac{\mathbf{b}_I(t)}{6}\delta t^3 + O(\delta t^4) \quad (13)$$

$$\mathbf{R}_I(t-\delta t) = \mathbf{R}_I(t) - \dot{\mathbf{R}}_I(t)\delta t + \frac{\mathbf{f}_I(t)}{2M_I}\delta t^2 - \frac{\mathbf{b}_I(t)}{6}\delta t^3 + O(\delta t^4) \quad (14)$$

where we have used Eq. 10 to express the second derivative in terms of force and mass. Adding Eqs. 13 and 14 we obtain

$$\mathbf{R}_I(t+\delta t) = 2\mathbf{R}_I(t) - \mathbf{R}_I(t-\delta t) + \frac{\mathbf{f}_I(t)}{M_I}\delta t^2 + O(\delta t^4) \quad (15)$$

Note that the accuracy of Eq. 15 is fourth order in time, one order of magnitude better than Eqs. 13 and 14. The velocities are obtained by subtracting Eq. 14 from Eq. 13.

$$\dot{\mathbf{R}}_I(t) = \frac{1}{2\delta t}[\mathbf{R}_I(t+\delta t) - \mathbf{R}_I(t-\delta t)] + O(\delta t^3) \quad (16)$$

In the Verlet algorithm, Eqs. 15-16, one needs to know the positions at time $t+\delta t$ and $t-\delta t$ to know the velocity at time t. That is, the velocity update is one step behind the position update. This is in principle not a problem but can be inconvenient when one wants to calculate velocity dependent quantities such as kinetic energy.

1.6 The Velocity Verlet algorithm

Velocity Verlet is a modification of the Verlet algorithm where the update of the positions and velocities occur at the same time step. The positions are given by Eq. 13 neglecting terms higher than second order.

$$\mathbf{R}_I(t+\delta t) = \mathbf{R}_I(t) + \dot{\mathbf{R}}_I(t)\delta t + \frac{\mathbf{f}_I(t)}{2M_I}\delta t^2 + O(t^3) \tag{17}$$

To obtain the velocities at time t we first calculate the force at the updated positions.

$$\mathbf{f}_I(t+\delta t) = \mathbf{f}_I(\mathbf{R}_I(t+\delta t)) \tag{18}$$

Substituting Eq. 18 into the Taylor expansion Eq. 14 for going back in time from $t+\delta t$ to t we find

$$\mathbf{R}_I(t) = \mathbf{R}_I(t+\delta t) - \dot{\mathbf{R}}(t+\delta t)\delta t + \frac{\mathbf{f}_I(t+\delta t)}{2M_I}\delta t^2 + O(t^3) \tag{19}$$

Adding Eq. 19 to the forward expansion Eq. 17 we obtain

$$\dot{\mathbf{R}}_I(t+\delta t) = \dot{\mathbf{R}}_I(t) + \frac{1}{2M_I}[\mathbf{f}_I(t) + \mathbf{f}_I(t+\delta t)]\delta t + O(t^3) \tag{20}$$

Although they seem dissimilar, one can easily show that the Verlet and Velocity Verlet algorithms are equivalent and produce exactly the same trajectory.

1.7 The leap-frog algorithm

Another modification of the Verlet algorithm is the leap-frog algorithm. In this scheme the positions and velocities are half a time step out of step. The velocities at half integer times are defined as

$$\dot{\mathbf{R}}_I(t - \delta t/2) = \frac{1}{\delta t}[\mathbf{R}_I(t) - \mathbf{R}_I(t - \delta t)] \tag{21}$$

$$\dot{\mathbf{R}}_I(t + \delta t/2) = \frac{1}{\delta t}[\mathbf{R}_I(t + \delta t) - \mathbf{R}_I(t)] \tag{22}$$

The following sequence of updates is used to propagate position and velocity, one 'leaping' over the other with a full step step,

$$\dot{\mathbf{R}}_I(t + \delta t/2) = \dot{\mathbf{R}}_I(t - \delta t/2) + \frac{\mathbf{f}_I(t)}{M_I}\delta t \tag{23}$$

$$\mathbf{R}_I(t + \delta t) = \mathbf{R}_I(t) + \dot{\mathbf{R}}_I(t + \delta t/2)\delta t \tag{24}$$

with the velocity at time t obtained as

$$\dot{\mathbf{R}}_I(t) = \frac{1}{2}[\dot{\mathbf{R}}_I(t - \delta t/2) + \dot{\mathbf{R}}_I(t + \delta t/2)] \tag{25}$$

The leap-frog algorithm is equivalent to the Verlet and Velocity Verlet algorithm.

2 Born-Oppenheimer molecular dynamics (BOMD)

In Born-Oppenheimer molecular dynamics (BOMD) Newton's Equations of motions are solved using ionic forces that are obtained from static electronic structure calculations. Every time step the ground state wavefunction $\Psi_{\mathbf{R}}$, the corresponding ground state energy $E_{\mathbf{R}}$ of the electronic Hamiltonian $\mathbf{H}_{\mathrm{e}}$ and the ionic forces $\mathbf{f}_I$ are calculated at the instantaneous nuclear configuration $\mathbf{R}$. Using the varational principle, we obtain

$$E_{\mathbf{R}} = \min_{\{\Psi_{\mathbf{R}}\}} \langle \Psi_{\mathbf{R}} | \mathbf{H}_{\mathrm{e}} | \Psi_{\mathbf{R}} \rangle = \langle \Psi_{\mathbf{R}} | \mathbf{H}_{\mathrm{e}} | \Psi_{\mathbf{R}} \rangle \tag{26}$$

$$\mathbf{H}_{\mathrm{e}} = \mathbf{T}_e(\mathbf{r}) + V_{en}(\mathbf{R}, \mathbf{r}) + V_{ee}(\mathbf{r}) + V_{nn}(\mathbf{R}). \tag{27}$$

Note that Eq. 26 is equivalent with solving the ground state of the Schrödinger equation $\mathbf{H}_{\mathrm{e}}\Psi_{\mathbf{R}} = E_{\mathbf{R}}\Psi_{\mathbf{R}}$. The forces on the nuclei are equal to minus the gradient of the ground state energy with respect to the nuclear coordinates,

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_{\mathbf{R}_I} \langle \Psi_{\mathbf{R}} | \mathbf{H}_{\mathrm{e}} | \Psi_{\mathbf{R}} \rangle. \tag{28}$$

The dynamics of the nuclei is computed according to one of the previously discussed molecular dynamics algorithms, identifying the force on ion I, $\mathbf{f}_I$, with the expression on the rhs of Eq. 28.

2.1 Brief Review: Kohn-Sham Density Functional Theory (KS-DFT)

In Kohn-Sham (KS) theory the electron density $\rho(\mathbf{r})$ is expressed in terms of orbitals $(\phi_1, \ldots, \phi_n)$, termed KS-orbitals,

$$\rho(\mathbf{r}) = \sum_i |\phi_i|^2, \tag{29}$$

and the ground state energy at a given nuclear configuration $\mathbf{R}$, $E_{\mathbf{R}}^{\mathrm{KS}}$, is obtained by minimisation of the energy functional,

$$\begin{aligned} E_{\mathbf{R}}[\rho(\{\phi_i\})] &= -\frac{1}{2}\sum_i \langle\phi_i|\nabla^2|\phi_i\rangle + \int \mathrm{d}\mathbf{r} V_{en}(\mathbf{R})\rho(\mathbf{r}) && (30) \\ &\quad + \frac{1}{2}\int \mathrm{d}\mathbf{r_1}\mathrm{d}\mathbf{r_2}\frac{\rho(\mathbf{r_1})\rho(\mathbf{r_2})}{\mathbf{r_{12}}} + E_{\mathrm{xc}}[\rho] + V_{nn}(\mathbf{R}) && (31) \end{aligned}$$

with respect to all electron densities ρ that integrate to N electrons,

$$E_{\mathbf{R}}^{\mathrm{KS}} = \min_{\rho(\{\phi_i\})} E_{\mathbf{R}}[\rho(\{\phi_i\})], \tag{32}$$

$E_{\mathbf{R}}^{\mathrm{KS}} = E_{\mathbf{R}}$ if the exact exchange correlation functional (E_{xc}) was known. The minimisation in Eq. 32 can be carried out with respect to the KS-orbitals under the constraint that they remain orthonormal,

$$0 = -\frac{\delta}{\delta\phi_i^*} E_{\mathbf{R}}\left[\rho(\{\phi_i\})\right] + \sum_j \Lambda_{ij}\phi_j \quad \forall i, \tag{33}$$

where Λ_{ij} is a matrix of Lagrange undetermined multipliers. Equation 33 is the minimisation condition from which the Kohn-Sham equations are derived,

$$\mathbf{H}^{\mathrm{KS}}\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}). \tag{34}$$

2.2 BOMD using KS-DFT (DFMD)

If Kohn-Sham DFT (KS-DFT) is used for the electronic structure problem in BOMD, one solves at each time step, or ionic configuration $\mathbf{R}$, the equations

$$0 = -\frac{\delta}{\delta\phi_i^*} E_{\mathbf{R}}\left[\rho(\{\phi_i\})\right] + \sum_j \Lambda_{ij}\phi_j \tag{35}$$

$$M_I\ddot{\mathbf{R}}_I(t) = -\nabla_{\mathbf{R}_I} \min_{\rho(\{\phi_i\})} E_{\mathbf{R}}[\rho(\{\phi_i\})] = -\nabla_{\mathbf{R}_I}\langle\Psi_{\mathbf{R}}^{\mathrm{KS}}|\mathbf{H}^{\mathrm{KS}}|\Psi_{\mathbf{R}}^{\mathrm{KS}}\rangle, \tag{36}$$

where Eq. 35 is identical with Eq. 33. This means that at each time step the ground state energy and orbitals have to be computed by solving Eq. 35 (or equivalently, by solving the KS equations Eq. 34), and by calculation of the ionic forces according to Eq. 36. $\Psi_{\mathbf{R}}^{\mathrm{KS}}$ on the rhs of Eq. 36 is the KS determinant obtained from the N KS orbitals that satisfy Eq. 35 (or the KS equations Eq. 34).

The dynamics of the nuclei is computed according to one of the previously discussed molecular dynamics algorithms, identifying the force on ion I, $\mathbf{f}_I$, with the expression on the rhs of Eq. 36.

2.3 Hellmann-Feynman theorem

The forces in Eqs. 28 and 36 can be evaluated using the Hellmann-Feynman theorem.

Consider a Hamiltonian that depends on a parameter λ, $\mathbf{H}(\lambda)$, an eigenfunction of this Hamiltonian, Ψ_λ (not necessarily the ground state), and the corresponding eigenvalue E_λ. Then

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle\Psi_\lambda|\mathbf{H}(\lambda)|\Psi_\lambda\rangle = \langle\Psi_\lambda|\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathbf{H}(\lambda)|\Psi_\lambda\rangle \tag{37}$$

Proof:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle\Psi_\lambda|\mathbf{H}(\lambda)|\Psi_\lambda\rangle &= \langle\frac{\mathrm{d}}{\mathrm{d}\lambda}\Psi_\lambda|\mathbf{H}(\lambda)|\Psi_\lambda\rangle + \langle\Psi_\lambda|\mathbf{H}(\lambda)|\frac{\mathrm{d}}{\mathrm{d}\lambda}\Psi_\lambda\rangle && (38)\\
&\quad +\langle\Psi_\lambda|\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathbf{H}(\lambda)|\Psi_\lambda\rangle && (39)\\
&= E_\lambda\langle\frac{\mathrm{d}}{\mathrm{d}\lambda}\Psi_\lambda|\Psi_\lambda\rangle + E_\lambda\langle\Psi_\lambda|\frac{\mathrm{d}}{\mathrm{d}\lambda}\Psi_\lambda\rangle && (40)\\
&\quad +\langle\Psi_\lambda|\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathbf{H}(\lambda)|\Psi_\lambda\rangle && (41)\\
&= E_\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle\Psi_\lambda|\Psi_\lambda\rangle + \langle\Psi_\lambda|\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathbf{H}(\lambda)|\Psi_\lambda\rangle && (42)\\
&= \langle\Psi_\lambda|\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathbf{H}(\lambda)|\Psi_\lambda\rangle && (43)
\end{aligned}$$

since the norm of Ψ_λ is invariant with respect to a change in λ, $\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle\Psi_\lambda|\Psi_\lambda\rangle = 0$. Note that the Hellmann-Feynman theorem is valid for any Hamiltonian (e.g. Hartree-Fock, Kohn-Sham) as long as Ψ_λ is an exact eigenfunction of this Hamiltonian.

Identifying λ with the components of the nuclear position vector $\mathbf{R}$, we find

$$\nabla_{\mathbf{R}_I}\langle\Psi^{\mathrm{KS}}_{\mathbf{R}}|\mathbf{H}^{\mathrm{KS}}|\Psi^{\mathrm{KS}}_{\mathbf{R}}\rangle = \langle\Psi^{\mathrm{KS}}_{\mathbf{R}}|\nabla_{\mathbf{R}_I}\mathbf{H}^{\mathrm{KS}}|\Psi^{\mathrm{KS}}_{\mathbf{R}}\rangle. \tag{44}$$

3 Car-Parrinello molecular dynamics

In the Car-Parrinello (CP) approach both the nuclear positions $\mathbf{R}_I$ and the electronic orbitals $\phi_i(\mathbf{r}_i)$ are considered as dynamical variables that obey classical equations of motion. This is accomplished by introducing an orbital velocity $\dot{\phi}_i(\mathbf{r}_i) = \mathrm{d}\phi_i(\mathbf{r}_i)/\mathrm{d}t$ and a fictitious mass μ associated with the orbital dynamics. The Lagrangian for the coupled nuclear and orbital degree of freedom reads

$$\mathcal{L}^{\mathrm{CP}} = \sum_I \frac{M_I}{2}|\dot{\mathbf{R}}_I|^2 + \sum_i \mu\langle\dot{\phi}_i|\dot{\phi}_i\rangle - E_{\mathbf{R}}\left[\{\phi_i\}\right] + \sum_{ij}\Lambda_{ij}(\langle\phi_i|\phi_j\rangle - \delta_{ij}) \quad (45)$$

where Λ_{ij} is a matrix of Lagrange multiplier imposing the orthonormality condition $\langle\phi_i|\phi_j\rangle = \delta_{ij}$ on the orbitals and δ_{ij} is the Kronecker symbol.

The introduction of a fictitious motion for the electronic orbitals allows for an analytic but approximate computation of the electronic potential energy $E_{\mathbf{R}}\left[\{\phi_i\}\right]$ on the fly, that is, it allows for a simultaneous propagation of the orbitals and nuclear positions in time. In this way the major bottleneck of BOMD, namely minimizing $E_{\mathbf{R}}\left[\{\phi_i\}\right]$ or, equivalently, solving the Kohn-Sham equation self consistently at each molecular dynamics step, is avoided.

3.1 CP equations of motion

The CP equations of motion are obtained by applying the associated Euler-Lagrange equations on the CP Lagrangian $\mathcal{L}^{\mathrm{CP}}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}^{\mathrm{CP}}}{\partial \dot{\mathbf{R}}_I} = \frac{\partial \mathcal{L}^{\mathrm{CP}}}{\partial \mathbf{R}_I} \tag{46}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta \mathcal{L}^{\mathrm{CP}}}{\delta \dot{\phi}_i^*} = \frac{\delta \mathcal{L}^{\mathrm{CP}}}{\delta \phi_i^*} \tag{47}$$

and differentiation of Eqs. 46 and 47,

$$M_I \ddot{\mathbf{R}}_I(t) = -\frac{\partial}{\partial \mathbf{R}_I} E_{\mathbf{R}}\left[\{\phi_i\}\right] \tag{48}$$

$$\mu \ddot{\phi}_i(t) = -\frac{\delta}{\delta \phi_i^*} E_{\mathbf{R}}\left[\{\phi_i\}\right] + \sum_j \Lambda_{ij}\phi_j. \tag{49}$$

In practice the electronic orbitals $\phi_i(t)$ are expanded in a set of time-independent basis functions φ_k with time-dependent coefficients $c_{ik}(t)$, $\phi_i = \sum_k c_{ik}(t)\varphi_k(\mathbf{r}_i)$. Substituting into Eq. 49 we obtain the equation of motion for the orbitals in terms of the orbital expansion coefficients.

$$\mu \ddot{c}_{ik} = \frac{\partial}{\partial c_{ik}^*} E_{\mathbf{R}}\left[\{c_{ik}\}\right] + \sum_j \Lambda_{ij} c_{jk}. \tag{50}$$

Eqs. 48 and 49 (or Eq. 50, respectively) couple all nuclear and electronic degrees of freedom and therefore have to be solved numerically.

3.2 Why does the Car-Parrinello algorithm work ?

If the fictitious mass of the orbitals is chosen significantly smaller than the mass of the lightest nucleus of the system, the electron dynamics (Eq. 49) is *adiabatically* separated from the nuclear dynamics. This situation is shown in Figure 1 for a simple but realistic model system. The highest nuclear frequency (or phonon mode) is indicated by a triangle, and the frequencies of the electronic motion in solid lines. The latter are obtained from the Fourier transform of the average velocity autocorrelation function of the orbitals,

$$f(\omega) = \int_0^\infty \cos(\omega t) \sum_i \langle \dot{\phi}_i; t | \dot{\phi}_i; 0 \rangle \mathrm{d}t \tag{51}$$

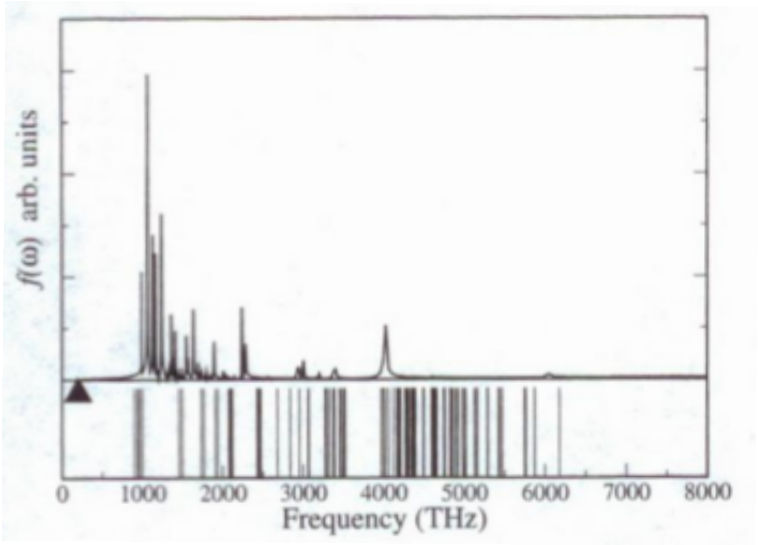

Figure 1: $f(\omega)$ (upper part) and harmonic approximation thereof (lower part) for two silicon atoms on a periodic diamond lattice. $\mu = 300 a.u.$ and time step = 0.3 fs.

If the mass or frequency gap is large the electrons will follow the nuclear motion adiabatically. This behavior is shown in Figure 2 where the fictitious kinetic energy of the orbitals, $T_{\rm e}$, and the potential energy $V_{\rm e}=E_{\mathbf{R}}[\{\phi_i(\mathbf{R}(t))\}]$ is plotted against time. Indeed, the fictitious electron dynamics tracks the change in nuclear configurations as represented by the change in potential energy. This is a consequence of the coupling of Eq. 49 with Eq. 48. However, in CP dynamics the orbitals are not the ones that minimize the electronic Hamiltonian (as in BOMD), but oscillate around them. These intrinsic oscillations can be seen after closer inspection of Figure 2, where they manifest as very fast oscillations of $T_{\rm e}$, superimposed on the slower oscillations due to the nuclei. Thus, since the orbitals oscillate on the subfemtosecond time scale around the minimum energy orbitals, the nuclei feel an average electronic force on the relevant femtosecond time scale.

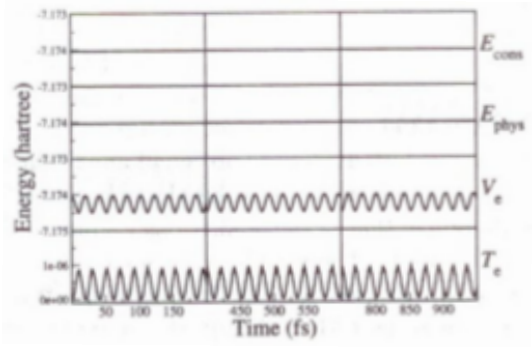

Figure 2: Time series of energies along a CPMD run.

$$E_{\rm cons} = T_{\rm e} + V_{\rm e} + \sum_j \frac{1}{2} M_I |\dot{\mathbf{R}}_I|^2 \quad (52)$$

$$E_{\rm phys} = E_{\rm cons} - T_{\rm e} \quad (53)$$

$$T_{\rm e} = \sum_i \mu \langle \dot{\phi}_i | \dot{\phi}_i \rangle \quad (54)$$

$$V_{\rm e} = E_{\mathbf{R}}[\{\phi_i\}] \quad (55)$$

$E_{\rm cons}$ defined in Eq. 52 is the conserved energy of motion in CP dynamics and $E_{\rm phys}$ defined in Eq. 53 is the physical energy. 'Physical' because it does not contain the fictitious orbital kinetic energy $T_{\rm e}$.

Naturally, the question arises as to how CPMD compares to 'exact' BOMD (exact in the sense that in BOMD the nuclear forces are derived from the true ground state wavefunction at each MD step). To investigate this question we reconsider the equations of motions in CPMD

$$M_I \ddot{\mathbf{R}}_I(t) = -\frac{\partial}{\partial \mathbf{R}_I} E_{\mathbf{R}}\left[\{\phi_i\}\right] \tag{56}$$

$$\mu \ddot{\phi}_i(t) = -\frac{\delta}{\delta \phi_i^*} E_{\mathbf{R}}\left[\{\phi_i\}\right] + \sum_j \Lambda_{ij}\phi_j, \tag{57}$$

and in BOMD,

$$M_I \ddot{\mathbf{R}}_I(t) = -\frac{\partial}{\partial \mathbf{R}_I} \min_{\{\phi_i\}} E_{\mathbf{R}}\left[\{\phi_i\}\right] \tag{58}$$

$$0 = -\frac{\delta}{\delta \phi_i^*} E_{\mathbf{R}}\left[\{\phi_i\}\right] + \sum_j \Lambda_{ij}\phi_j. \tag{59}$$

If on the left hand side of Eq. 57 μ is set equal to zero one obtains the minimization condition from which the Kohn-Sham equations are derived. In this case Eq. 57 is identical with Eq. 59 and Eq. 56 is identical with Eq. 58. CPMD has been reduced to BOMD. Thus, the fictitious kinetic energy of the orbitals introduced in CPMD, $\sum_i \mu \langle \dot{\phi}_i | \dot{\phi}_i \rangle$, can be interpreted as a measure for the deviation from the exact BO PES. For typical values of μ (300-500 a.u.) the fictitious energy is orders of magnitude smaller than the potential energy (see Figure 2). In such cases the nuclear forces in CPMD reproduce the exact BO forces very well as is illustrated in Figure 3.

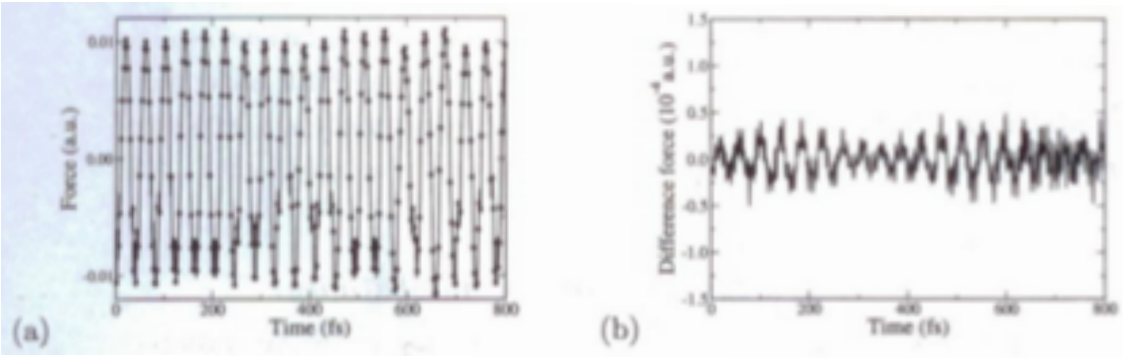

Figure 3: (a) Comparison of the x-component of the force acting on one atom from BOMD (dots) and CPMD (solid lines). (b) Enlarged view of the difference.

3.3 Comparing CPMD and BOMD

	CPMD	BOMD
forces	approximate but can be made arbitrarily accurate	exact if well converged
MD time step	determined by electron dynamics, typically 0.1-0.2 fs	determined by nuclear dynamics, typically 0.5-1 fs
energy conservation	very good, energy drift typically $< 1 \times 10^{-6}$ a.u./ps/atom	depends on quality of wavefunction optimization, tight: $< 1 \times 10^{-5}$ a.u./ps/atom, standard: $< 1 \times 10^{-4}$ a.u./ps/atom
efficiency	NO wavefunction optimization $\rightarrow$ gain of $\approx$ factor $\geq 5 - 10$. BUT smaller time step $\rightarrow$ loss of $\approx$ factor 3-5	wavefunction optimization every MD step
	CPMD is usually more efficient than BOMD, although with recent advances in wavefunction optimization techniques the difference is not as large any more as it used to be !	
metals	not possible in a straightforward manner (zero gap $\rightarrow$ no adiabatic separation)	method of choice

4 Ensembles and Temperature in MD

4.1 Time averages

What one typically would like to do in a molecular dynamics simulation is to compute the time average of some observable A that is a function of all classical phase space variables $\mathbf{R}$ and $\mathbf{p}$.

$$A(t) = A(\mathbf{R}(t), \mathbf{p}(t)) \tag{60}$$

The phase space variables change in time, that is, along the MD trajectory, and so does A. Moreover, the time evolution of $A(t)$ will be different for different trajectories, i.e. for different initial values of the phase space variables ($\mathbf{R}(0)$, $\mathbf{p}(0)$)

The time average of A over a continuous trajectory $\mathbf{R}(t)$ of finite length Δt is given by

$$\bar{A}_{\Delta t} = \frac{1}{\Delta t} \int_0^{\Delta t} \mathrm{d}t A(t). \tag{61}$$

In practical calculations the trajectories are discrete and the continuous time average Eq. 61 is approximated by the discrete time average,

$$\bar{A}_{\Delta t} \approx \frac{1}{M} \sum_{m=1}^{M} A(t_m) \tag{62}$$

where $\Delta t = M \delta t$ and δt is the time step.

4.2 Microcanonical ensemble

Consider a large number of copies of an isolated system that all have the same constant particle number N, constant volume V and constant total energy E. The collection of states in phase space $(\mathbf{R}, \mathbf{p})$ that have the same total energy E is called the microcanonical or NVE ensemble. One of the fundamental assumptions of statistical mechanics is that the probability of a state, $\rho_{\mathrm{NVE}}(\mathbf{R}, \mathbf{p})\, \mathrm{d}\mathbf{R}\mathrm{d}\mathbf{p}$, is the same for all states of the NVE ensemble. The condition that only states of energy E are occupied and that all of them have the same probabilty can be expressed in terms of a delta function δ, which restricts the manifold of accessible phase space points to a hypersurface of constant E only.

$$\rho_{\mathrm{NVE}}(\mathbf{R}, \mathbf{p}) = \frac{(h^{3N} N!)^{-1}}{\Omega_{\mathrm{NVE}}} \delta\left[\mathbf{H}(\mathbf{R}, \mathbf{p}) - E\right] \tag{63}$$

$$\Omega_{\mathrm{NVE}} = (h^{3N} N!)^{-1} \int \mathrm{d}\mathbf{R}\, \mathrm{d}\mathbf{p}\, \delta\left[\mathbf{H}(\mathbf{R}, \mathbf{p}) - E\right] \tag{64}$$

Here ρ_{NVE} is the probability density, $\mathbf{H} = \sum_{I=1}^{N} \mathbf{p}_I^2/(2M_I) + E_{\mathbf{R}}$ is the Hamiltonian, $E_{\mathbf{R}}$ is the potential energy, h is Planck's constant, and Ω_{NVE} is the microcanonical partition function. The denominator of the probability density contains the partition function which normalizes the probability to unity, $\int \mathrm{d}\mathbf{R}\, d\mathbf{p}\, \rho_{\mathrm{NVE}} = 1$. It also contains the unit volume in phase space, $h^{3N} N!$, which makes the probability dimensionless.

4.3 Ergodic hypothesis

In order to obtain the average or the expectation value of an observable $A(\mathbf{R}, \mathbf{p})$ in the NVE ensemble, $\langle A \rangle_{\mathrm{NVE}}$, one has to integrate A over the phase space variables, where each point in phase space is weighted with the probability density ρ_{NVE}.

$$\langle A \rangle_{\mathrm{NVE}} = \int \mathrm{d}\mathbf{R}\, \mathrm{d}\mathbf{p}\, \rho_{\mathrm{NVE}}(\mathbf{R}, \mathbf{p}) A(\mathbf{R}, \mathbf{p}) \tag{65}$$

The ergodic hypothesis establishes a connection between the time averages obtained from molecular dynamics and the ensemble averages used in statistical mechanics. It states that for ergodic systems, the time average of an observable A, Eq. 61, is equal to the ensemble average Eq. 65 in the limit of an infinitely long trajectory.

$$\lim_{\Delta t \to \infty} \bar{A}_{\Delta t} = \langle A \rangle_{\mathrm{NVE}} \tag{66}$$

This means that the number of times a point in phase space is visited along an (infinitely long) MD trajectory is proportional to its statistical weight in the NVE ensemble. The equivalence of time and and ensemble averages is an assumption valid for stable many-body systems. However, there are systems for which this condition is not satisfied, such as glasses.

4.4 Canonical ensemble

Consider a large number of copies of a system that is connected to a heat bath at temperature T. All copies have the same constant particle number N, constant volume V, but they can exchange heat with the bath and thus can have a different total system energy E. The collection of these states is called the canonical or NVT ensemble. From arguments based on the maximization of possibilities to realize a given total energy $E_{\rm tot} = E + E_{\rm bath}$, one can derive that each state in the canonical ensemble has a probability density

$$\rho_{\rm NVT}(\mathbf{R}, \mathbf{p}) = \frac{(h^{3N} N!)^{-1}}{\Omega_{\rm NVT}} \exp\left[-\beta \mathbf{H}(\mathbf{R}, \mathbf{p})\right] \quad (67)$$

$$\Omega_{\rm NVT} = (h^{3N} N!)^{-1} \int \mathrm{d}\mathbf{R}\, \mathrm{d}\mathbf{p}\, \exp\left[-\beta \mathbf{H}(\mathbf{R}, \mathbf{p})\right] \quad (68)$$

where $\beta = 1/(k_{\rm B} T)$ and $k_{\rm B}$ is the Boltzmann constant. The distribution Eq. 67 is also called Boltzmann distribution, and $\Omega_{\rm NVT}$ is the canonical partition function.

The expectation value for an observable A in the NVT ensemble is calculated similarly as for the NVE ensemble (Eq. 65), replacing $\rho_{\rm NVE}$ by the probability density $\rho_{\rm NVT}$ of Eq. 67.

$$\langle A \rangle_{\rm NVT} = \int \mathrm{d}\mathbf{R}\, \mathrm{d}\mathbf{p}\, \rho_{\rm NVT}(\mathbf{R}, \mathbf{p}) A(\mathbf{R}, \mathbf{p}) \quad (69)$$

In analogy to Eq. 66 there exists an ergodic hypothesis also for NVT ensembles. However, before we state this explicitly we have to define temperature in molecular dynamics simulation.

4.5 Temperature in the NVT ensemble

Temperature was introduced previously as a parameter in the exponent of the canonical ensemble distribution function, Eq. 67. Comparison of the expressions of state functions in statistical mechanics and classical thermodynamics shows that the statistical temperature is indeed equivalent with the empirical temperature measured in experiments.

One can easily show that in the canonical ensemble temperature is proportional to the expectation value of the total kinetic energy. Insertion of $A = \sum_{I=1}^{N} \frac{\mathbf{p}_I^2}{2\,M_I}$ in Eq. 69 we obtain

$$\begin{aligned}
\left\langle \sum_{I=1}^{N} \frac{\mathbf{p}_I^2}{2\,M_I} \right\rangle_{\mathrm{NVT}} &= \frac{(h^{3N}N!)^{-1} \int \mathrm{d}\mathbf{R}\,\mathrm{d}\mathbf{p}\,(\sum_{I=1}^{N} \frac{\mathbf{p}_I^2}{2\,M_I}) \exp\left[-\beta \sum_{J=1}^{N} \frac{\mathbf{p}_J^2}{2\,M_J} + E_{\mathbf{R}}\right]}{(h^{3N}N!)^{-1} \int \mathrm{d}\mathbf{R}\,\mathrm{d}\mathbf{p}\, \exp\left[-\beta \sum_{I=1}^{N} \frac{\mathbf{p}_I^2}{2\,M_I} + E_{\mathbf{R}}\right]} \\
&= \frac{3N}{2} k_{\mathrm{B}} T, \qquad (70)
\end{aligned}$$

where we have used $\int_{-\infty}^{\infty} dx\, x^2 \exp(-ax^2) = \frac{1}{2a}\sqrt{\pi/a}$. Thus, temperature can be defined by the average kinetic energy, which is perhaps a more intuitive quantity than the ensemble distribution function.

4.6 Temperature in the NVE ensemble

It is not obvious how to define temperature in the microcanonical ensemble. Here we rely on a proof in statistical mechanics showing that the fractional difference between the microcanonical average and the canonical average of properties such as kinetic energy vanishes in the thermodynamic limit of very large N. Hence, $\lim_{N\to\infty} \frac{1}{N} \left\langle \sum_{I=1}^{N} \frac{\mathbf{p}_I^2}{2\,M_I} \right\rangle_{\mathrm{NVE}} = \frac{3}{2} k_{\mathrm{B}} T$. Thus, we define temperature for the NVE ensemble similarly as for the NVT ensemble,

$$T = \frac{2}{3N k_{\mathrm{B}}} \left\langle \sum_{I=1}^{N} \frac{\mathbf{p}_I^2}{2\,M_I} \right\rangle_{\mathrm{NVE}} \tag{71}$$

and note that the two temperatures in Eqs. 70 and 71 will be equal for large N. Assuming ergodicity, Eq. 66, we can calculate the temperature from molecular dynamics simulation as the time average of the instantaneous temperature $\mathcal{T}$, which is proportional to the instantaneous kinetic energy,

$$T = \lim_{\Delta t\to\infty} \frac{1}{\Delta t} \int_0^{\Delta t} \mathrm{d}t\, \mathcal{T}(t) \;\approx\; \frac{1}{M} \sum_{m=1}^{M} \mathcal{T}_m \tag{72}$$

$$\mathcal{T}(t) \;=\; \frac{2}{3N k_{\mathrm{B}}} \sum_{I=1}^{N} \frac{\mathbf{p}_I^2(t)}{2\,M_I}. \tag{73}$$

4.7 The Nosé thermostat

The MD algorithms presented in the first lecture are numerical schemes that solve Newton's equation of motion. As the total energy is conserved during the motion, Newton Dynamics samples the NVE ensemble (assuming ergodicity). On the other hand, experiments are usually carried out at constant temperature. Thus, in order to allow for a consistent comparison between simulation and experiment, it would be desirable to carry out MD simulations in the NVT ensemble. The famous Nosé constant temperature MD algorithm has been developed for this purpose. In this method Newton's equations are extended by a friction term, i.e. a term proportional to the velocity, $\zeta\dot{\mathbf{R}}_I$,

$$\ddot{\mathbf{R}}_I = \frac{\mathbf{f}_I}{M_I} - \zeta\dot{\mathbf{R}}_I \tag{74}$$

$$\dot{\zeta} = \frac{1}{Q}\left(\sum_{I=1}^{N}\frac{M_I\dot{\mathbf{R}}_I^2}{2} - 3Nk_{\mathrm{B}}T\right), \tag{75}$$

where ζ is the friction constant, Q is the fictitious mass and T is the target temperature.

During the dynamics defined by Eqs. 74-75 the total energy is not conserved any more due to the dissipation of heat caused by the friction term. This can be shown by taking the time derivative of the Hamiltonian $\mathbf{H}=\sum_{I=1}^{N}\frac{\mathbf{p}_I^2}{2\,M_I}+E_{\mathbf{R}}$ using the chain rule as before, and substituting $\dot{\mathbf{p}}=M_I\ddot{\mathbf{R}}$ by Eq. 74. The result is

$$\frac{\mathrm{d}\mathbf{H}}{\mathrm{d}t} = -\zeta\sum_{I=1}^{N}M_I\dot{\mathbf{R}}_I^2. \tag{76}$$

The dissipation by the friction term can be positive or negative, thus leading either to an accelaration or deccelaration of the particles in addition to the accelaration or deccelaration due to the atomistic potential energy $E_{\mathbf{R}}$.

4.8 How the Nose-Hoover thermostat works

There is an energy quantity $\tilde{\mathbf{H}}$ that is conserved during the Nosé dynamics, Eqs. 74-75,

$$\tilde{\mathbf{H}} = \mathbf{H} + \frac{Q}{2}\zeta^2 + 3Nk_{\mathrm{B}}T\int_0^t \mathrm{d}t'\zeta(t') \tag{77}$$

This can be shown by taking the time derivative of the extended Hamiltonian $\tilde{\mathbf{H}}$,

$$\frac{\mathrm{d}\tilde{\mathbf{H}}}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{H}}{\mathrm{d}t} + Q\zeta\dot{\zeta} + 3Nk_{\mathrm{B}}T\zeta, \tag{78}$$

substituting $\dot{\zeta}$ by Eq. 75, and inserting Eq. 76.

$$\frac{\mathrm{d}\tilde{\mathbf{H}}}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{H}}{\mathrm{d}t} + \zeta\sum_{I=1}^{N} M_I\dot{\mathbf{R}}_I^2 = 0 \tag{79}$$

Equations 77 and 79 can be used for a qualitative explanation of the Nosé thermostat. Assume that Q is sufficiently large so that for a small time interval $\Delta t = t_2 - t_1$ the change in the friction constant ζ is negligible. Then $\tilde{\mathbf{H}}(t_2) - \tilde{\mathbf{H}}(t_1) \approx \mathbf{H}(t_2) - \mathbf{H}(t_1) + 3Nk_{\mathrm{B}}T\int_{t_1}^{t_2}\mathrm{d}t\,\zeta = 0$ and the change in the total atomic energy in the time interval Δt is

$$\mathbf{H}(t_2) - \mathbf{H}(t_1) = -3Nk_{\mathrm{B}}T\int_{t_1}^{t_2}\mathrm{d}t\zeta. \tag{80}$$

Thus, for $\zeta > 0$ the total atomic energy decreases (= 'Cooling') and for $\zeta < 0$ the total energy increases (= 'Heating').

Assuming that the Nosé dynamics Eqs. 74-75 is ergodic, one can in fact show that each state is sampled according to the canonical probability distribution Eq. 67, that is, the Nosé dynamics samples the NVT ensemble.