# *Ab initio* Molecular Dynamics

Jürg Hutter

Physical Chemistry University of Zurich

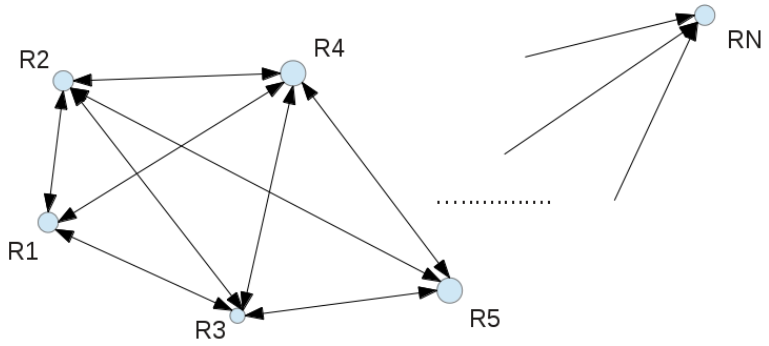
August 25, 2015

# Overview

- Equations of motion
- Integrators, Sources of error
- General Lagrangian for AIMD
- Car–Parrinello MD
- Born–Oppenheimer MD
- *ab initio* Langevin MD
- Stability and efficiency

# Equations of Motion (EOM)

Newton's EOM for a set of classical point particles in a potential.




$$M_I \ddot{\boldsymbol{R}}_I = -\frac{\mathrm{d}V(\boldsymbol{R})}{\mathrm{d}\boldsymbol{R}_I}$$

These EOM generate for a given number of particles $N$ in a volume $V$ the micro canonical ensemble (NVE ensemble).

The total energy $E$ is a constant of motion!

# Total Energy

Total Energy = Kinetic energy + Potential energy

$$\text{Kinetic energy} = T(\dot{\boldsymbol{R}}) = \sum_{I=1}^{N} \frac{M_I}{2} \dot{\boldsymbol{R}}^2$$

$$\text{Potential energy} = V(\boldsymbol{R})$$

We will use the total energy as an indicator for the numerical accuracy of simulations.

# Lagrange Equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{R}}} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{R}}$$

with

$$\mathcal{L}(\boldsymbol{R}, \dot{\boldsymbol{R}}) = T(\dot{\boldsymbol{R}}) - V(\boldsymbol{R})$$

Equivalent to Newton's EOM in Cartesian coordinates, but is more general and flexible.

- Extended systems
- Constraints

# Integration of EOM

Discretization of time

$$\boldsymbol{R}(t) \rightarrow \boldsymbol{R}(t+\tau) \rightarrow \boldsymbol{R}(t+2\tau) \rightarrow \cdots \rightarrow \boldsymbol{R}(t+m\tau)$$
$$\boldsymbol{V}(t) \rightarrow \boldsymbol{V}(t+\tau) \rightarrow \boldsymbol{V}(t+2\tau) \rightarrow \cdots \rightarrow \boldsymbol{V}(t+m\tau)$$

- Efficiency: minimal number of force evaluations, minimal number of stored quantities
- Stability: minimal drift in constant of motion (energy)
- Accuracy: minimal distance to exact trajectory

# Sources of Errors

- Type of integrator
  predictor-corrector, time-reversible, symplectic
- Time step $\tau$
  short time accuracy measured as $\mathcal{O}(\tau^n)$
- Consistency of forces and energy
  e.g. cutoffs leading to non-smooth energy surfaces
- Accuracy of forces
  e.g. convergence of iterative force calculations (SCF, constraints)

# Velocity Verlet Integrator

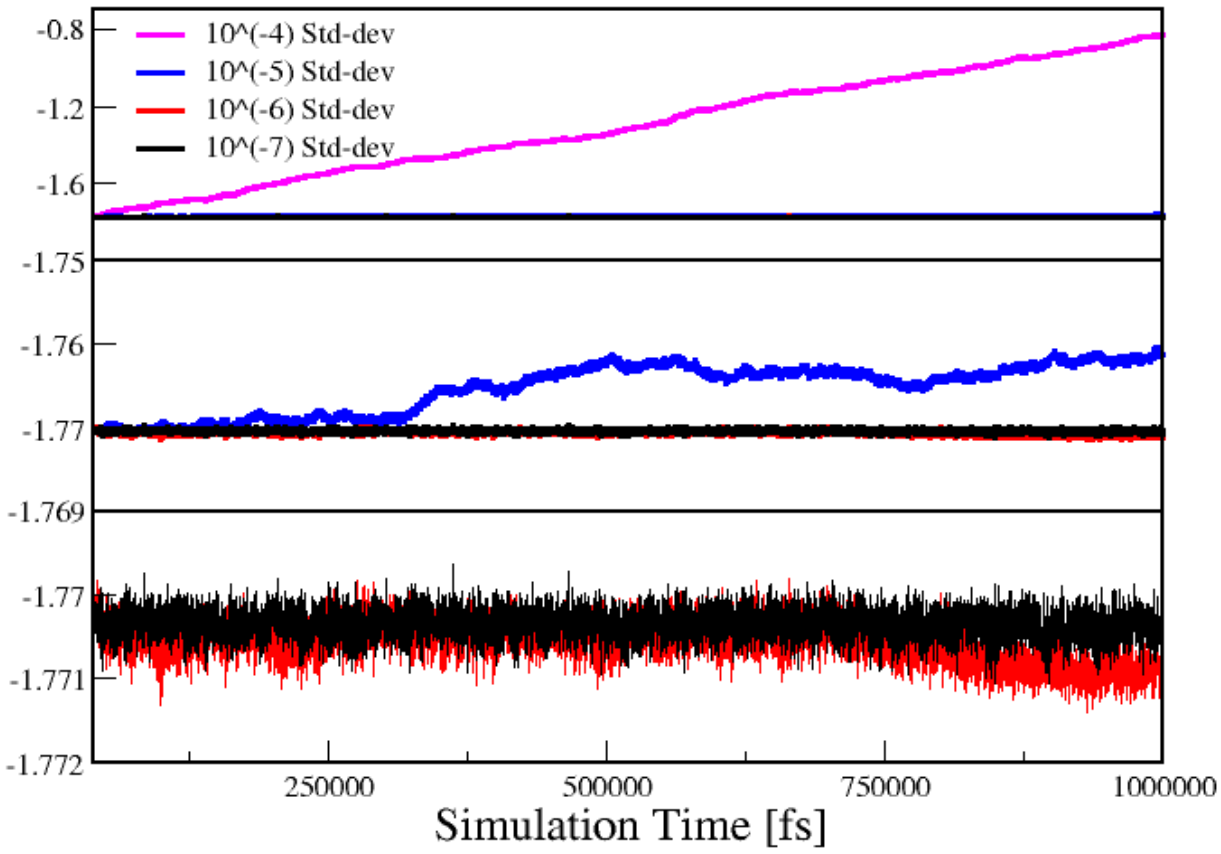
$$\boldsymbol{R}(t+\tau) = \boldsymbol{R}(t) + \tau \boldsymbol{V}(t) + \frac{\tau^2}{2M}\boldsymbol{f}(t)$$

$$\boldsymbol{V}(t+\tau) = \boldsymbol{V}(t) + \frac{\tau}{2M}[\boldsymbol{f}(t) + \boldsymbol{f}(t+\tau)$$

- Efficiency: 1 force evaluation, 3 storage vectors
- Stability: time reversible
- Accuracy: $\mathcal{O}(\tau^2)$
- Simple adaptation for constraints (shake, rattle, roll)
- Simple adaptation for multiple time steps and thermostats

# Test on Required Accuracy of Forces

Classical Force Field Calculations, 64 molecules, 330 K
TIP3P (flexible), SPME ($\alpha = 0.44$, GMAX $= 25$),



# Stability: Accuracy of Forces

| Stdev. $\Delta$f Hartree/Bohr | Stdev. Energy $\mu$Hartree | Drift $\mu$Hartree/ns | Drift Kelvin/ns |
|---|---|---|---|
| – | 170.35 | 35.9 | 0.06 |
| $10^{-10}$ | 179.55 | -85.7 | -0.14 |
| $10^{-08}$ | 173.68 | 6.5 | 0.01 |
| $10^{-07}$ | 177.83 | -58.2 | -0.10 |
| $10^{-06}$ | — | -385.4 | -0.63 |
| $10^{-05}$ | — | 9255.8 | 15.21 |
| $10^{-04}$ | — | 972810.0 | 1599.31 |

# Born–Oppenheimer MD: The Easy Way

$$M_I \ddot{\boldsymbol{R}}_I = -\frac{\mathrm{d}V(\boldsymbol{R})}{\mathrm{d}\boldsymbol{R}_I} \qquad \text{EOM}$$

$$V(\boldsymbol{R}) = \min_{\Phi} \left[E_{\mathrm{KS}}(\{\Phi(\boldsymbol{r})\}; \boldsymbol{R}) + \text{const.}\right] \qquad \text{Kohn–Sham BO potential}$$

## Forces

$$\begin{aligned} f_{\mathrm{KS}}(\boldsymbol{R}) &= \frac{\mathrm{d}\min_{\Phi} E_{\mathrm{KS}}(\{\Phi(\boldsymbol{r})\}; \boldsymbol{R})}{\mathrm{d}\boldsymbol{R}} \\ &= \frac{\partial E_{\mathrm{KS}}}{\partial \boldsymbol{R}} + \frac{\partial \text{const.}}{\partial \boldsymbol{R}} + \sum_i \underbrace{\frac{\partial (E_{\mathrm{KS}} + \text{const.})}{\partial \Phi_i}}_{=0} \frac{\partial \Phi_i}{\partial \boldsymbol{R}} \end{aligned}$$

# Computational Details

- System
  - 64 water molecules
  - density $1 gcm^{-3}$
  - Temperature $\approx 330K$
  - Timestep $0.5fs$

- DFT Calculations
  - GPW, TZV2P basis (2560 bsf), PBE functional
  - Cutoff 280 Rydberg, $\epsilon_{\text{default}} = 10^{-12}$
  - OT-DIIS, Preconditioner FULL_SINGLE_INVERSE
  - Reference trajectory ($1ps$), $\epsilon_{\text{SCF}} = 10^{-10}$

# Stability in BOMD

Unbiased initial guess; $\Phi(t) = \Phi_0(\boldsymbol{R}(t))$

| $\epsilon_{\mathrm{SCF}}$ | MAE $E_{\mathrm{KS}}$ Hartree | MAE f Hartree/Bohr | Drift Kelvin/ns |
|---|---|---|---|
| $10^{-08}$ | $1.2 \cdot 10^{-11}$ | $5.1 \cdot 10^{-09}$ | 0.0 |
| $10^{-07}$ | $9.5 \cdot 10^{-10}$ | $5.6 \cdot 10^{-08}$ | 0.1 |
| $10^{-06}$ | $6.9 \cdot 10^{-08}$ | $4.8 \cdot 10^{-07}$ | 0.4 |
| $10^{-05}$ | $7.4 \cdot 10^{-06}$ | $5.6 \cdot 10^{-06}$ | 2.3 |
| $10^{-04}$ | $3.3 \cdot 10^{-04}$ | $5.9 \cdot 10^{-05}$ | 50 |

**Consistent with results from classical MD**
**Note accuracy of forces!**

# Efficiency: Initial Guess of Wavefunction

4th order Gear predictor (PS extrapolation in CP2K)

| Method | $\epsilon_{\mathrm{SCF}}$ | Iterations | Drift (Kelvin/ns) |
|---|---|---|---|
| Guess | $10^{-06}$ | 14.38 | 0.4 |
| Gear(4) | $10^{-07}$ | 6.47 | 5.7 |
| Gear(4) | $10^{-06}$ | 5.22 | 11.8 |
| Gear(4) | $10^{-05}$ | 4.60 | 86.8 |

What is the problem?

Time reversibility has been broken!

# Generalized Lagrangian

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} M \dot{\mathbf{q}}^2 + \frac{1}{2} \mu \dot{\mathbf{x}}^2 - E(\mathbf{q}, \mathbf{y}) + k\mu G(\|\mathbf{x} - \mathbf{y}\|)$$

| | |
|---|---|
| $\mathbf{y} = F(\mathbf{q}, \mathbf{x})$ | wavefunction optimization |
| $G(\|\mathbf{x} - \mathbf{y}\|)$ | wavefunction retention potential |

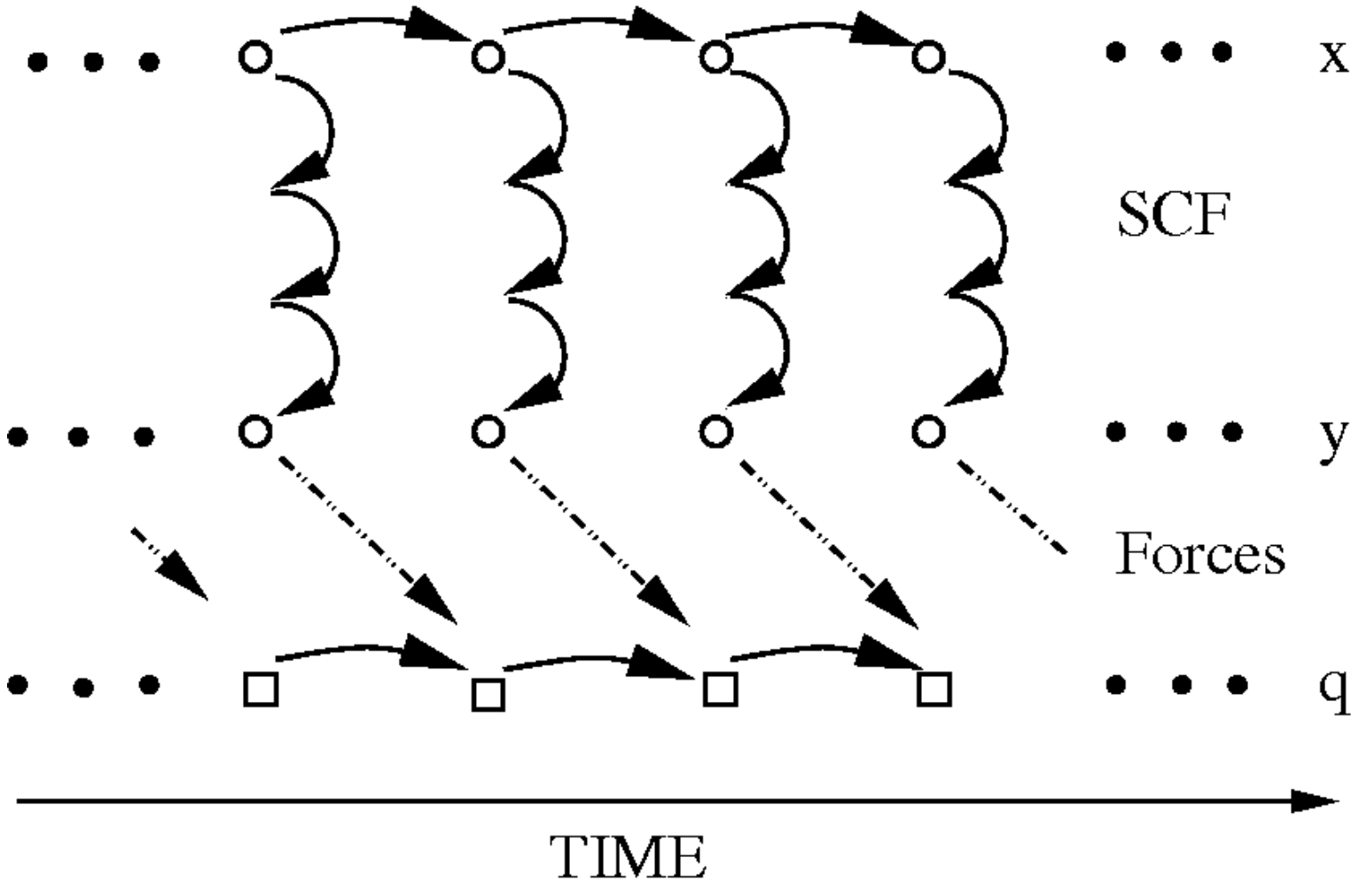
**Equations of motion**

$$M\ddot{\mathbf{q}} = -\frac{\partial E}{\partial \mathbf{q}} - \frac{\partial E}{\partial \mathbf{y}} \frac{\partial F}{\partial \mathbf{q}} + k\mu \frac{\partial G}{\partial \mathbf{y}} \frac{\partial F}{\partial \mathbf{q}}$$

$$\mu\ddot{\mathbf{x}} = -\frac{\partial E}{\partial \mathbf{y}} \frac{\partial F}{\partial \mathbf{x}} + k\mu \left[ \frac{\partial G}{\partial \mathbf{x}} + \frac{\partial G}{\partial \mathbf{y}} \frac{\partial F}{\partial \mathbf{x}} \right]$$

Extension of Niklasson Lagrangian, PRL **100** 123004 (2008)

# Dynamical System



# Car–Parrinello Molecular Dynamics

$$\mathbf{y} = \mathbf{x} \quad \Rightarrow \quad G(\|\mathbf{x} - \mathbf{y}\|) = 0$$

## Lagrangian

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} M \dot{\mathbf{q}}^2 + \frac{1}{2} \mu \dot{\mathbf{x}}^2 - E(\mathbf{q}, \mathbf{x})$$

## Equations of motion

$$M\ddot{\mathbf{q}} = -\frac{\partial E}{\partial \mathbf{q}}$$

$$\mu\ddot{\mathbf{x}} = -\frac{\partial E}{\partial \mathbf{x}}$$

# Properties of CPMD

- **Accuracy:** Medium
  Distance from BO surface controlled by mass $\mu$
  Requires renormalization of dynamic quantities (e.g. vibrational spectra)

- **Stability:** Excellent
  All forces can be calculated to machine precision easily

- **Efficiency:** Good
  Efficiency is strongly system dependent (electronic gap)
  Requires many nuclear gradient calculations

**Not implemented in CP2K!**

# BOMD

$$\mathbf{y} = \mathrm{Min}_{\mathbf{x}} E(\mathbf{q}, \mathbf{x}) \quad \text{and} \quad \mu = 0$$

## Lagrangian

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} M \dot{\mathbf{q}}^2 + E(\mathbf{q}, \mathbf{y})$$

$$\mathcal{L}(\mathbf{x}, \dot{\mathbf{x}}) = \frac{1}{2} \dot{\mathbf{x}}^2 + kG(\|\mathbf{x} - \mathbf{y}\|)$$

## Equations of motion

$$\left.\begin{aligned} M\ddot{\mathbf{q}} &= -\frac{\partial E}{\partial \mathbf{q}} \\ \ddot{\mathbf{x}} &= -k\frac{\partial G}{\partial \mathbf{x}} \end{aligned}\right\} \text{decoupled equations}$$

# BOMD with Incomplete SCF Convergence

$$\mathbf{y} = F(\mathbf{q}, \mathbf{x}) \approx \mathrm{Min}_{\mathbf{x}} E(\mathbf{q}, \mathbf{x}) \quad \text{and} \quad \mu = 0$$

**Equations of motion**

$$M\ddot{\mathbf{q}} = -\frac{\partial E}{\partial \mathbf{q}} - \frac{\partial E}{\partial \mathbf{y}} \frac{\partial F}{\partial \mathbf{q}}$$

$$\ddot{\mathbf{x}} = -k\frac{\partial G}{\partial \mathbf{x}} - \frac{\partial G}{\partial \mathbf{y}} \frac{\partial F}{\partial \mathbf{x}}$$

SCF Error: Neglect of force terms $\frac{\partial E}{\partial \mathbf{y}} \frac{\partial F}{\partial \mathbf{q}}$ and $\frac{\partial G}{\partial \mathbf{y}} \frac{\partial F}{\partial \mathbf{x}}$.

EOM are coupled through terms neglected!

# ASPC Integrator

Integration of electronic DOF ($\mathbf{x}$) has to be

- accurate: good wavefunction guess gives improved efficiency
- stable: do not destroy time-reversibility of nuclear trajectory

## ASPC: Always Stable Predictor Corrector

- J. Kolafa, J. Comput Chem. **25**: 335–342 (2004)
- ASPC(k): time-reversible to order $2k + 1$

# Orthogonality Constraint

Wavefunction extrapolation for non-orthogonal basis sets

$$\mathbf{C}_{\text{init}} = \sum_{j=0}^{K} B_{j+1} \underbrace{\mathbf{C}_{t-j\tau}\mathbf{C}_{t-j\tau}^{T}\mathbf{S}_{t-j\tau}}_{\mathbf{PS}} \mathbf{C}_{t-\tau}$$

Coefficients $B_i$ are given by ASPC algorithm

$$\mathbf{x} = \mathbf{C}_{\text{init}}$$
$$\mathbf{y}[t] = \mathbf{C}_t$$
$$\mathbf{S} \quad \text{overlap matrix}$$

J. VandeVondele et al. Comp. Phys. Comm. **167**: 103–128 (2005)

# Importance of Time-Reversibility

| Method | $\epsilon_{\mathrm{SCF}}$ | Iterations | Drift (Kelvin/ns) |
|---|---|---|---|
| Guess | $10^{-06}$ | 14.38 | 0.4 |
| ASPC(3) | $10^{-06}$ | 5.01 | 0.2 |
| ASPC(3) | $10^{-05}$ | 3.02 | 4.5 |
| Gear(4) | $10^{-07}$ | 6.47 | 5.7 |
| Gear(4) | $10^{-06}$ | 5.22 | 11.8 |
| Gear(4) | $10^{-05}$ | 4.60 | 86.8 |

# Efficiency and Drift

| Method | $\epsilon_{\mathrm{SCF}}$ | Iterations | Drift (Kelvin/ns) |
|---|---|---|---|
| Guess | $10^{-06}$ | 14.38 | 0.4 |
| ASPC(4) | $10^{-06}$ | 5.01 | 0.2 |
| ASPC(4) | $10^{-05}$ | 3.02 | 4.5 |
| ASPC(4) | $10^{-04}$ | 1.62 | 1742.4 |
| ASPC(4) | $10^{-02}$ | 1.03 | 21733.2 |

# Efficiency and Drift

| Method | $\epsilon_{\mathrm{SCF}}$ | Iterations | Drift (Kelvin/ns) |
|---|---|---|---|
| ASPC(4) | $10^{-04}$ | 1.62 | 1742.4 |
| ASPC(5) | $10^{-04}$ | 1.63 | 1094.0 |
| ASPC(6) | $10^{-04}$ | 1.79 | 397.4 |
| ASPC(7) | $10^{-04}$ | 1.97 | 445.8 |
| ASPC(8) | $10^{-04}$ | 2.06 | 24.1 |

# Langevin BOMD

Starting point: BOMD with ASPC(k) extrapolation

Analysis of forces

$$f_{\mathrm{BO}}(\boldsymbol{R}) = \underbrace{f_{\mathrm{HF}}(\boldsymbol{R}) + f_{\mathrm{Pulay}}(\boldsymbol{R})}_{f(\boldsymbol{R})} + f_{\mathrm{nsc}}(\boldsymbol{R})$$

- $f_{\mathrm{BO}}$ : correct BO force
- $f_{\mathrm{HF}}$ : Hellmann–Feynman force
- $f_{\mathrm{Pulay}}$ : Pulay force
- $f_{\mathrm{nsc}}$ : non-self consistency error force
- $f$ : approximate BO force

# Forces in Approximate BOMD

Approximate $f_{\mathrm{nsc}}$ by

$$\tilde{f}_{\mathrm{nsc}} = -\int \mathrm{d}\boldsymbol{r} \left[ \left( \frac{\partial V_{\mathrm{xc}}(\rho^i)}{\partial \rho^i} \right) \Delta\rho + V_{\mathrm{H}}(\Delta\rho) \right] \nabla_I \rho^i$$

with $\Delta\rho = \rho^o - \rho^i$, $\rho^o$ final (output) density, $\rho^i$ initial (predicted) density.

Now assume

$$f(\boldsymbol{R}) + \tilde{f}_{\mathrm{nsc}}(\boldsymbol{R}) = f_{\mathrm{BO}}(\boldsymbol{R}) - \gamma_D \dot{\boldsymbol{R}}$$

where $\gamma_D$ is a constant friction parameter.

# Langevin EOM

$$M\ddot{\boldsymbol{R}} = f_{\mathrm{BO}}(\boldsymbol{R}) - (\gamma_D + \gamma_L)\dot{\boldsymbol{R}} + \Theta$$

with $\Theta$ a Gaussian random noise term and

$$\langle \Theta(0)\Theta(t) \rangle = 6(\gamma_D + \gamma_L) M k_B T \delta(t)$$

Given temperature $T = \frac{k_B}{3}\langle M\dot{\boldsymbol{R}}^2 \rangle$ and an arbitrary $\gamma_L$ this determines $\gamma_D$

# Langevin BOMD

T. D. Kühne et al., Phys. Rev. Lett. **98** 066441 (2007)

- single SCF step plus force correction needed
  extremely efficient for systems with slow SCF convergence
- $\gamma_D$ is small: correct statistics and dynamics
- difficult to stabilize in complex systems

Advanced Review

# Car–Parrinello molecular dynamics

Jürg Hutter*

The Car–Parrinello (CP) method made molecular dynamics simulation with on-the-fly computation of interaction potentials from electronic structure theory computationally feasible. The method reformulates *ab initio* molecular dynamics (AIMD) as a two-component classical dynamical system. This approach proved to be valuable far beyond the original CP molecular dynamics method. The modern formulation of Born–Oppenheimer (BO) dynamics is based on the same basic principles and can be derived from the same Lagrange function as the CP method. These time-reversible BO molecular dynamics methods allow higher accuracy and efficiency while providing similar longtime stability as the CP method. AIMD is used in many fields of computational physics and chemistry. Its applications are instrumental in fields as divers as enzymatic catalysis and the study of the interior of planets. With its versatility and predictive power, AIMD has become a major approach in atomistic simulations. 

**How to cite this article:**
*WIREs Comput Mol Sci* 2012, 2: 604–612 doi: 10.1002/wcms.90

# Summary: BOMD in CP2K

- Born–Oppenheimer MD with ASPC(3) is default in CP2K
- SCF convergence criteria depends on system
  $10^{-5} - 10^{-6}$ is a reasonable starting guess
- Best used together with OT and the
  FULL_SINGLE_INVERSE preconditioner
- Langevin dynamics is an option
  needs some special care, mostly for uniform systems