Algorithms: Total Energy Minimization

Jörg Neugebauer
Fritz-Haber-Institut, Berlin

Goal: Calculate total energy

$$E_{tot}^{BO}\left(\left\{\vec{R}_i\right\}\right) = \min_{\{\varphi_i\}} E_{tot}[\{\vec{R}_I\}, \{\varphi_i\}]$$

Questions:
- **How to construct E_{tot}?**
- **How to find (efficiently) the electronic ground state?**
  $\Rightarrow$ **Solve Kohn-Sham Equations**

$$H[n]\varphi_i = \varepsilon_i \varphi_i \qquad \textbf{with} \qquad n(\vec{r}) = \sum_i \left|\Psi_i(\vec{r})\right|^2$$

Outline

How to get to the Born-Oppenheimer surface?

- **Direct diagonalization**
- **Iterative diagonalization**
    - **Method**
    - **Scaling**
    - **Improvements**
    - **Self-consistency**

How to move on the Born-Oppenheimer surface?

- **Equilibrium geometry (T=0K)**
- **Molecular dynamics**

Basis set representation

Basis set expansion:
$$|\varphi_i\rangle = \sum_\mu c_{i\mu}|\mu\rangle \qquad \langle\mu|\nu\rangle = \delta_{\mu\nu} \quad \text{orthogonal}$$

$$\sum_\mu |\mu\rangle\langle\mu| = \hat{1} \quad \text{complete}$$

KS-equations:
$$H\,\varphi_i = \varepsilon_i\varphi_i \implies \sum_\mu \langle\nu|H|\mu\rangle\langle\mu|\varphi_i\rangle = \varepsilon_i\langle\nu|\varphi_i\rangle$$

$\Rightarrow$ **Solve eigenvalue problem:** $\qquad \hat{H}\,\vec{Y} = \varepsilon\vec{Y}$

direct diagonalization: $\hat{U}^T\hat{H}\hat{U} = \hat{1}\vec{\varepsilon}$
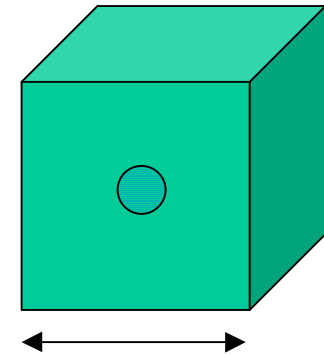
$$
\begin{pmatrix}
U_{11} & \cdots & \cdots & U_{N1} \\
\vdots & \ddots & & \vdots \\
\vdots & & \ddots & \vdots \\
U_{1N} & \cdots & \cdots & U_{NN}
\end{pmatrix}
\begin{pmatrix}
H_{11} & \cdots & \cdots & H_{1N} \\
\vdots & \ddots & & \vdots \\
\vdots & & \ddots & \vdots \\
H_{N1} & \cdots & \cdots & H_{NN}
\end{pmatrix}
\begin{pmatrix}
U_{11} & \cdots & \cdots & U_{1N} \\
\vdots & \ddots & & \vdots \\
\vdots & & \ddots & \vdots \\
U_{N1} & \cdots & \cdots & U_{NN}
\end{pmatrix}
=
\begin{pmatrix}
\varepsilon_{11} & \cdots & \cdots & 0 \\
\vdots & \ddots & & \vdots \\
\vdots & & \ddots & \vdots \\
0 & \cdots & \cdots & \varepsilon_{NN}
\end{pmatrix}
$$

Problems:
- **N^3 scaling**
- **complete Hamiltonian has to be saved**

Discussion: Plane wave basis

Example: cubic cell with a=5Å

E_{cut}	n_{max}	N_{PW}
5	14	534
10	20	1510
15	25	2774
20	28	4271
40	40	12081
100	64	47752

Realistic systems have basis sets with $10^{4...6}$ functions!

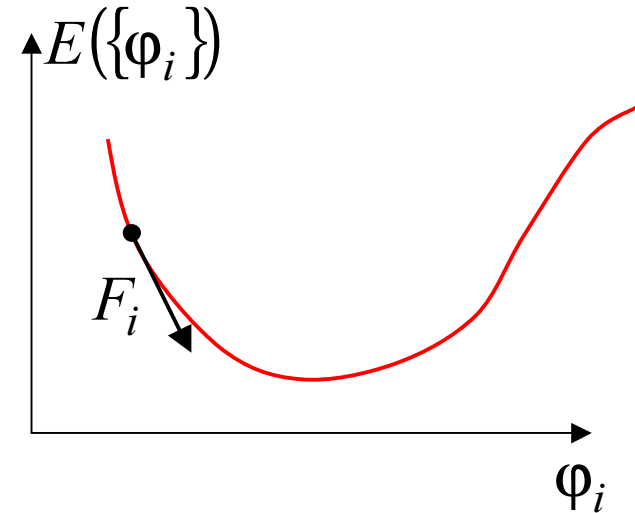$\Rightarrow$ **direct diagonalization inefficient/impossible**

Iterative diagonalization

Basic idea:

Total energy gradient:

$$F_i = -\frac{\delta E(\{\varphi_i\})}{\delta \varphi_i} \quad \text{with} \quad \langle \varphi_i | \varphi_j \rangle = \delta_{ij}$$

Residuum: $\quad |F_i\rangle = (H - \varepsilon_i)|\varphi_i\rangle$

$\Rightarrow$ **Find equation of motion (EOM) which minimizes total energy!**

First order EOM: $\quad -|F_i\rangle = \alpha|\dot{\varphi}_i\rangle \qquad \qquad \alpha$ **"friction" parameter**

or: $\quad (H - \varepsilon)|\varphi_i\rangle = -\alpha|\dot{\varphi}_i\rangle$

interesting aspect: time dependent Schrödinger equation for $\quad \alpha \rightarrow i\hbar$

Scaling of the iterative algorithm

Equation of motion: $\left(H - \varepsilon\right)\left|\varphi_i\right\rangle = -\alpha\left|\dot{\varphi}_i\right\rangle = -\alpha\dfrac{\varphi_i^{(n+1)} - \varphi_i^{(n)}}{\Delta t} = -\tilde{\alpha}\Delta\varphi_i^{(n)}$

Plane wave basis set: $\displaystyle\sum_{G'}\left\langle G\left|\left(H - \varepsilon\right)\right|G'\right\rangle\left\langle G'\middle|\varphi_i\right\rangle = -\tilde{\alpha}\left\langle G\middle|\Delta\varphi_i^{(n)}\right\rangle$

O(N²M) operations
M number of states (<<N)

Trick: Use locality of the contributions of the Hamiltonian in real and reciprocal space!

Hamilton Operator: $H = -\nabla^2 + V^{eff}\left(\vec{r}\right)$ **real space** $\overset{\textbf{FFT}}{\longleftrightarrow}$ **reciprocal space**

O(N ln(N)) operations

local in reciprocal space

$$\begin{pmatrix} -G_1^2 & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & -G_N^2 \end{pmatrix}\begin{pmatrix} \varphi_i(G_1) \\ \vdots \\ \vdots \\ \varphi_i(G_N) \end{pmatrix}$$

O(N) operations

local in real space

$$\begin{pmatrix} V^{eff}(\vec{r}_1) & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & V^{eff}(\vec{r}_N) \end{pmatrix}\begin{pmatrix} \varphi_i(\vec{r}_1) \\ \vdots \\ \vdots \\ \varphi_i(\vec{r}_N) \end{pmatrix}$$

O(N) operations

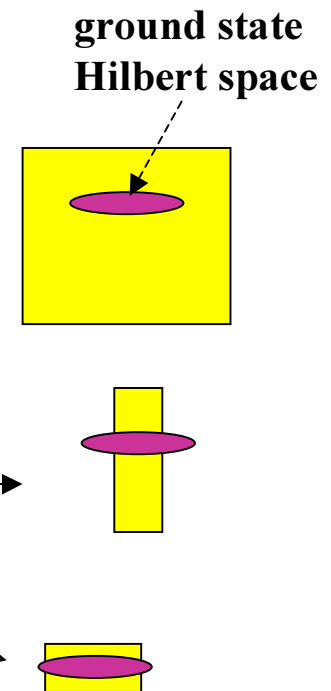# Construction of the initial wave functions

Iterative schema: $\varphi_i^{(n+1)} = F\left(\left\{\varphi_i^{(n)}\right\}\right)$

Problem: How to construct $\left\{\varphi_i^{(0)}\right\}$ **?**

$$\Psi_{\vec{k}}(\vec{r}) = \sum_{\vec{G}} c_{\vec{k}}(\vec{G})\, e^{i(\vec{G}+\vec{k})\vec{r}}$$

ground state Hilbert space

Solutions:
- **Direct diagonalization (not possible for large systems)**
- **Random numbers**
- **Direct diagonalization in PW subset**
- **Direct diagonalization in LCAO basis set**

Diagonalization in LCAO basis set

Expand atomic orbitals in plane waves:
$$\mu_{\vec{k}}(\vec{r}) = \sum_{\vec{G}} \mu_{\vec{k}}(\vec{G}) \, e^{i(\vec{G}+\vec{k})\vec{r}}$$

μ ... **atomic s, p, d orbitals**

Formally fully equivalent to wave functions:
$$\Psi_{\vec{k}}(\vec{r}) = \sum_{\vec{G}} c_{\vec{k}}(\vec{G}) \, e^{i(\vec{G}+\vec{k})\vec{r}}$$

Construct hamiltonian and overlap matrix:
$$H_{\mu\nu}(\vec{k}) = \left\langle \mu_{\vec{k}} \left| \hat{H} \right| \nu_{\vec{k}} \right\rangle$$

$$S_{\mu\nu}(\vec{k}) = \left\langle \mu_{\vec{k}} \middle| \nu_{\vec{k}} \right\rangle$$

Solve generalized eigenvalue problem:
$$\left\{ H_{\mu\nu}(\vec{k}) - \varepsilon_i \, S_{\mu\nu}(\vec{k}) \right\} \left\langle \mu_{\vec{k}} \middle| \psi_i \right\rangle = 0$$

Number of atomic orbitals << Number of plane waves (approx. 1:100)

➡ **Eigenvalue problem can be easily solved!**

Convergence Criteria (I)

iterative solution: $\varphi^{(n+1)} = \varphi^{(n)} - \left(H - \varepsilon\right)\left|\varphi^{(n)}\right\rangle \Delta t$ $\varepsilon = \left\langle \varphi^{(n)}\left|H\right|\varphi^{(n)}\right\rangle$

exact solution: $\left(H - \varepsilon\right)\left|\chi_i\right\rangle = 0$

representation of the trial wave function in the χ_i set:

$$\left|\varphi^{(n)}\right\rangle = \sum_i \alpha_i \left|\chi_i\right\rangle \qquad \left|\varphi^{(n+1)}\right\rangle = \sum_i \beta_i \left|\chi_i\right\rangle$$

$$\Longrightarrow \quad \beta_i = \alpha_i - \left(\varepsilon - \varepsilon_i\right)\Delta t$$

Before iteration step:
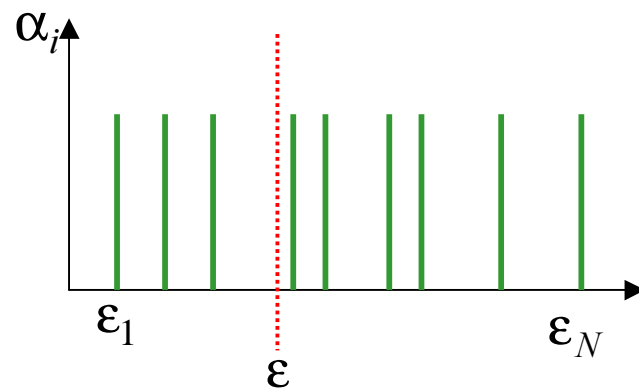
After "ideal" iteration step:
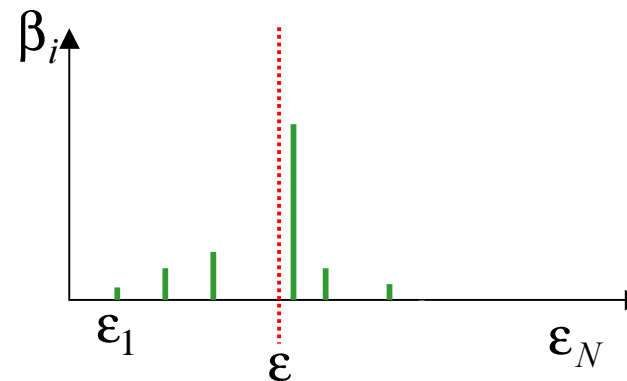
Convergence Criteria (II)

Before iteration step:

α_i

ε_1 ε_N

After iteration step:

β_i

ε_1 ε_N

After orthogonalization: β_i

ε_1 ε ε_N
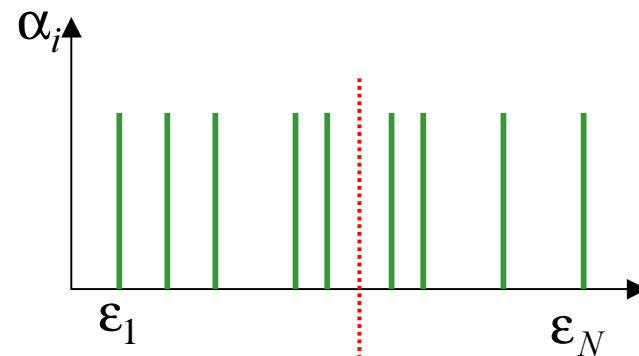
EOM: $\beta_i = \alpha_i - (\varepsilon - \varepsilon_i)\Delta t$
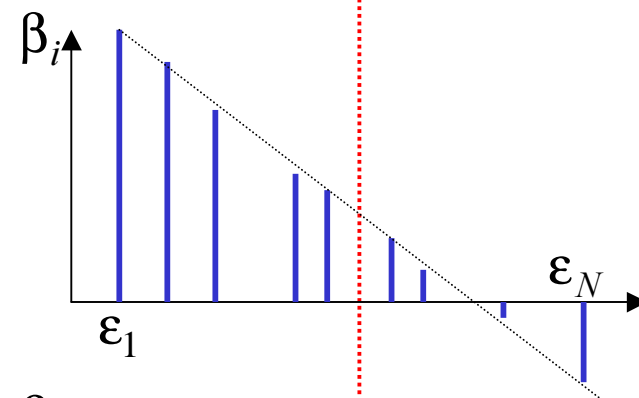
optimum time step:

$$0 \le \beta_N = \alpha_N - (\varepsilon - \varepsilon_N)\Delta t_{max}$$

$$\Delta t_{max} = \frac{1}{\varepsilon_N - \varepsilon_1} = \frac{1}{\underbrace{\varepsilon_{max} - \varepsilon_{min}}_{\textbf{spectral radius}}}$$
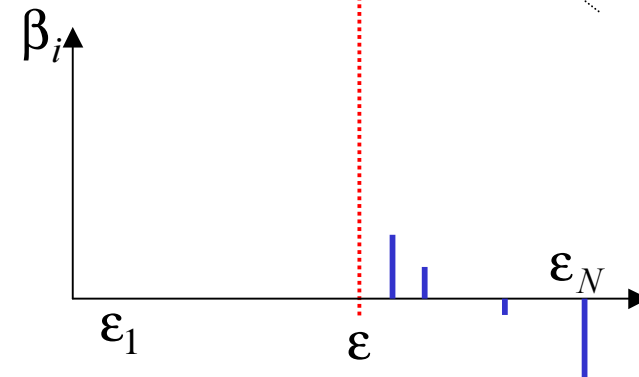
- **Orthogonalization essential for each iteration step**
- **Convergence rate decreases with increasing number of PW**

Preconditioning

Residual error: $\underbrace{(H-\varepsilon)\big|\chi + \Delta\varphi\big\rangle}_{=:|\varphi\rangle} = (H-\varepsilon)\big|\Delta\varphi\big\rangle = \big|\Delta\tilde{\varphi}\big\rangle$
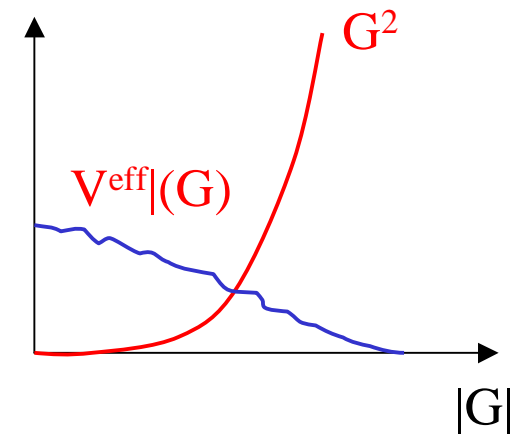
we need this   **we get this**

$$\Longrightarrow \quad (H-\varepsilon)^{-1}\big|\Delta\tilde{\varphi}\big\rangle = \big|\Delta\varphi\big\rangle$$

Hamiltonian in PW basis:

$$H_{GG'} = -G^2\delta_{GG'} + V^{eff}(G-G')$$

Kinetic energy dominates for high wave numbers!:

G^2

$V^{eff}|(G)$

$|G|$

Partition Hamiltonian: $\hat{H} = \hat{D} + \hat{L}$
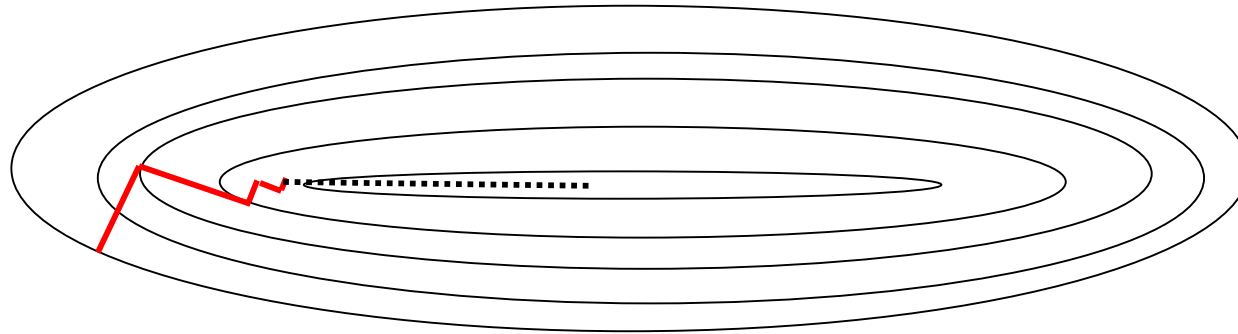
diagonal matrix   **all diagonal elements are zero**

For high wave numbers: $\hat{H} \approx \hat{D}$ $\Longrightarrow$ $\left(\hat{D}-\varepsilon\right)^{-1}\Delta\tilde{\varphi} = \Delta\varphi$

Preconditioning:Geometric interpretation

Without preconditioning:

With preconditioning:
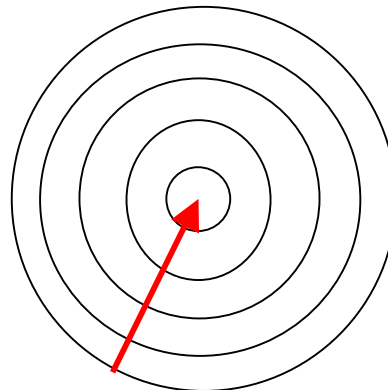
Williams-Soler Algorithm

Equation of motion: $\left(H - \varepsilon_i\right)\left|\varphi_i\right\rangle = -\alpha\left|\dot{\varphi}_i\right\rangle$

In plane-wave basis set:

$$\alpha\,\dot{c}_i(G) = \underbrace{\left[G^2 + V^{eff}(G-G)\right]c_i(G)}_{\text{diagonal part}} - \sum_{\substack{G' \\ G' \neq G}} \underbrace{V^{eff}(G-G')c_i(G')}_{\text{off-diagonal part}}$$

EOM for a single state: $\dot{c}_i(G) = \omega_G^2\, c_i(G) - B_G$

Assumption: B_G **not time dependent**

$$\Rightarrow \quad c_i(G, t = \Delta t) = -\frac{B_G}{\omega_G^2} + \left[c_i(G, t = 0) + \frac{B_G}{\omega_G^2}\right]\exp\left(-\omega_G^2\Delta t\right)$$
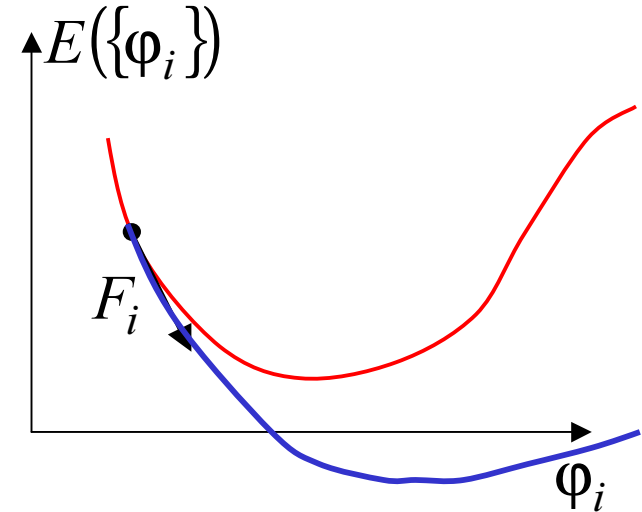
Improves significantly the convergence rate for high PW energy cutoffs!

Higher order EOM's

Goal: Find total-energy minimum more efficiently

Strategies:
- **minimize number of iterations**
- **minimize computational effort for each step**

Two approaches

Use higher derivatives	Use higher order in time
$$D_{ij} = \frac{\partial^2 E(\{\varphi_i\})}{\partial \varphi_i \partial \varphi_j} \quad , \quad \ldots$$	$$\varphi_i = F(\dot{\varphi}_i, \ddot{\varphi}_i, \ldots)$$ **equivalent:** $$\varphi_i^{(n+1)} = \widetilde{F}\left(\varphi_i^{(n)}, \varphi_i^{(n-1)}, \ldots\right)$$

high computational effort!   **no additional computational effort!**
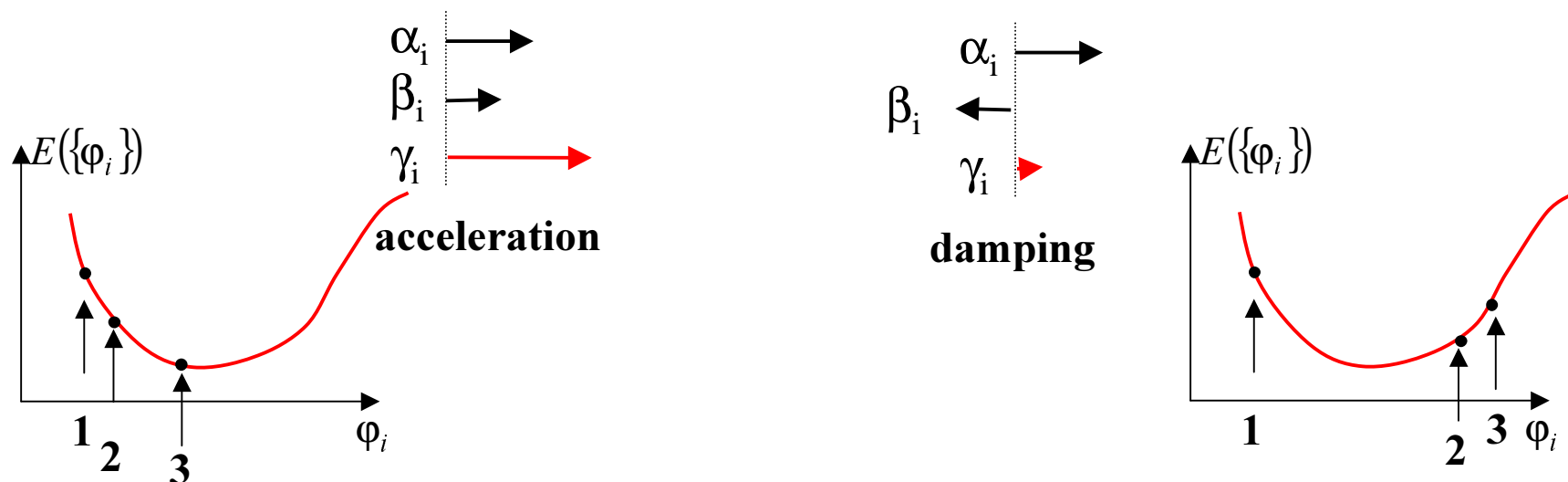
Second order equation of motion

Finite differences:

$$\varphi^{(n+1)} = \varphi^{(n)} - (H - \varepsilon)\left|\varphi^{(n)}\right\rangle \Delta t_1 - (H - \varepsilon)\left|\varphi^{(n-1)}\right\rangle \Delta t_2$$

Expansion with respect to eigenfunctions χ_i:

$$\left|\varphi^{(n-1)}\right\rangle = \sum_i \alpha_i \left|\chi_i\right\rangle \qquad \left|\varphi^{(n)}\right\rangle = \sum_i \beta_i \left|\chi_i\right\rangle \qquad \left|\varphi^{(n+1)}\right\rangle = \sum_i \gamma_i \left|\chi_i\right\rangle$$
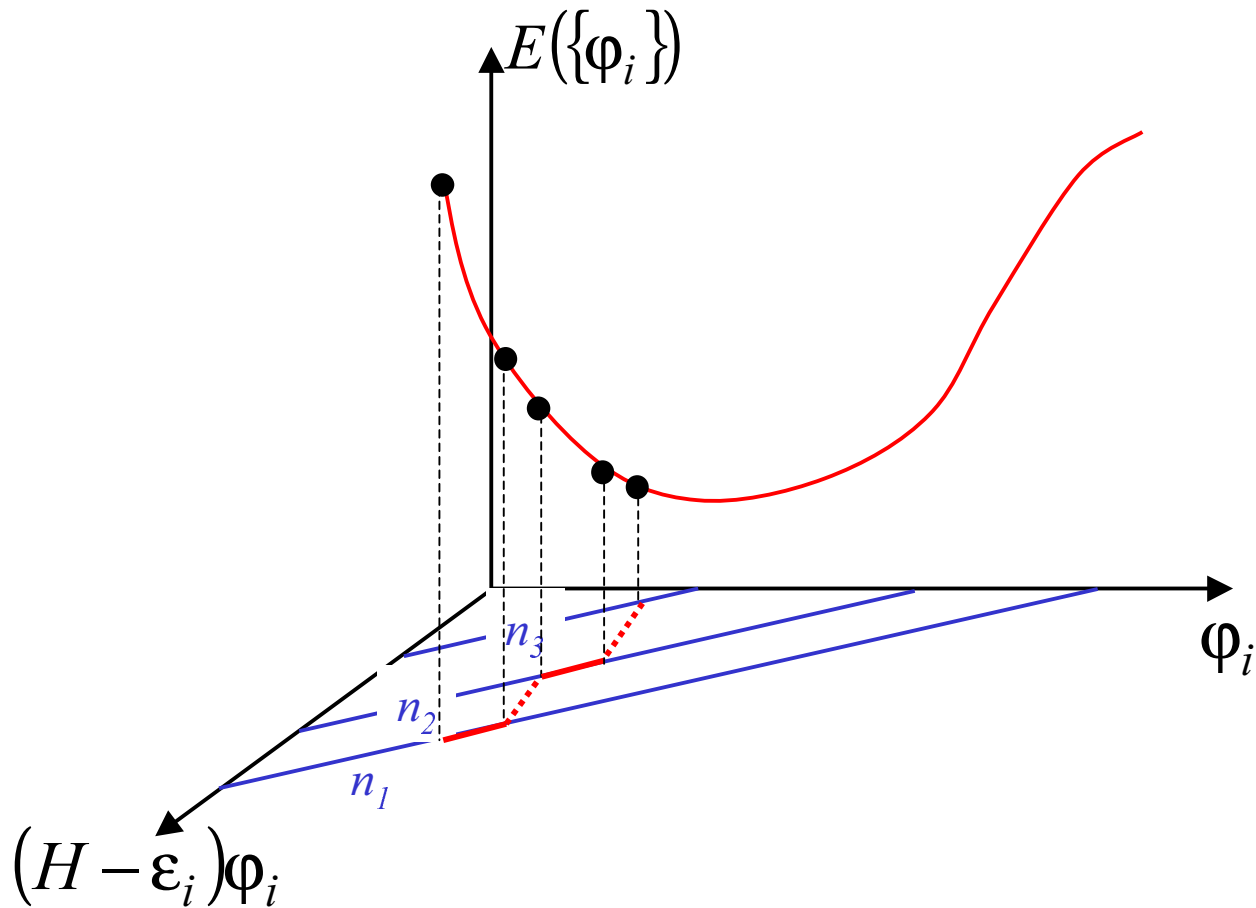
$$\Longrightarrow \quad \gamma_i = \beta_i - \beta_i(\varepsilon - \varepsilon_i)\Delta t_1 - \alpha_i(\varepsilon - \varepsilon_i)\Delta t_2$$

Two possible scenarios

$\alpha_i \rightarrow$
$\beta_i \rightarrow$
$\gamma_i \rightarrow$

acceleration

$\alpha_i \rightarrow$
$\beta_i \leftarrow$
$\gamma_i \rightarrow$

damping

$E(\{\varphi_i\})$ φ_i

1   2   3

$E(\{\varphi_i\})$ φ_i

1   2   3

Direct vs iterative methods

Self-consistency: $\left(H - \varepsilon_i\right)\left|\varphi_i\right\rangle = 0$ **with** $H = H\left(\left\{\varphi_i\right\}\right)$

Improved convergence rate by enforcing self-consistency at each iteration step!
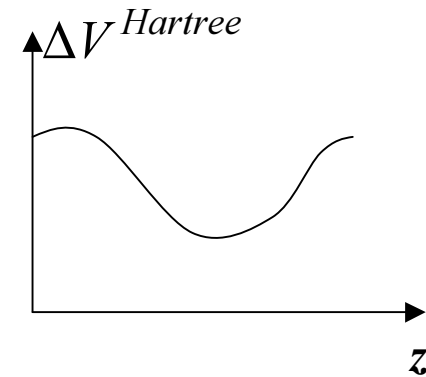
Instabilities: Charge sloshing

Hartree potential:
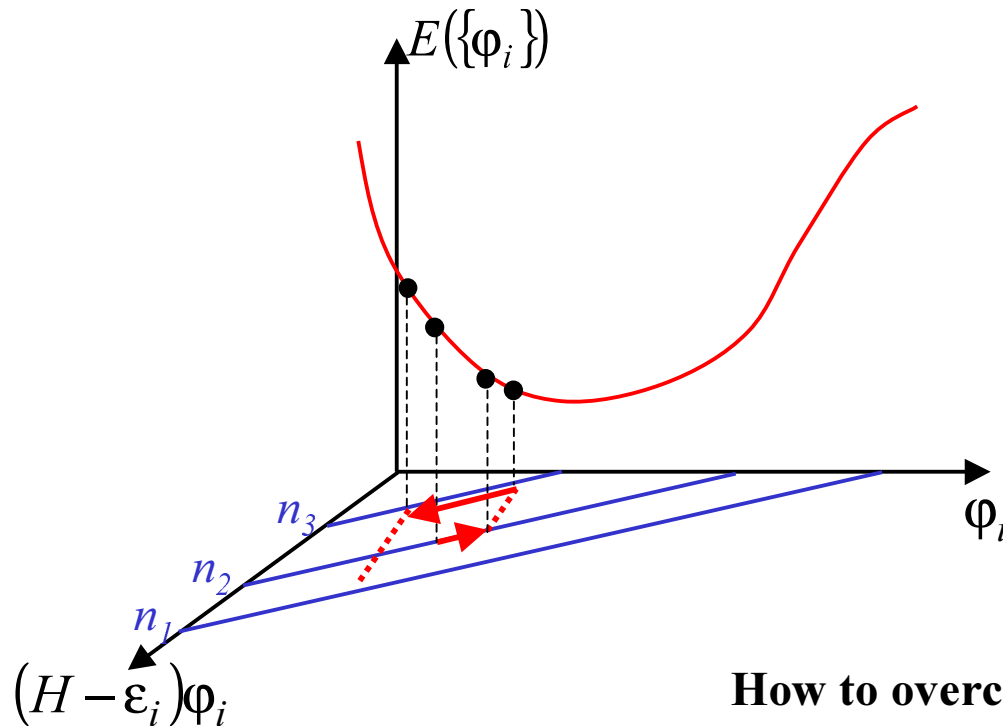
$$V^{Hartree}(G) = \frac{4\pi e \rho(G)}{|G|^2}$$

Change in Hartree potential:

$$\Delta V^{Hartree}(G) = \frac{4\pi e \Delta\rho(G)}{|G|^2}$$

$\Delta V^{Hartree}$ vs z

Instability increases with increasing system size:

$$G_{\min} \propto \frac{1}{l_{\max}}$$

$E(\{\varphi_i\})$ vs φ_i, $(H - \varepsilon_i)\varphi_i$, n_1, n_2, n_3
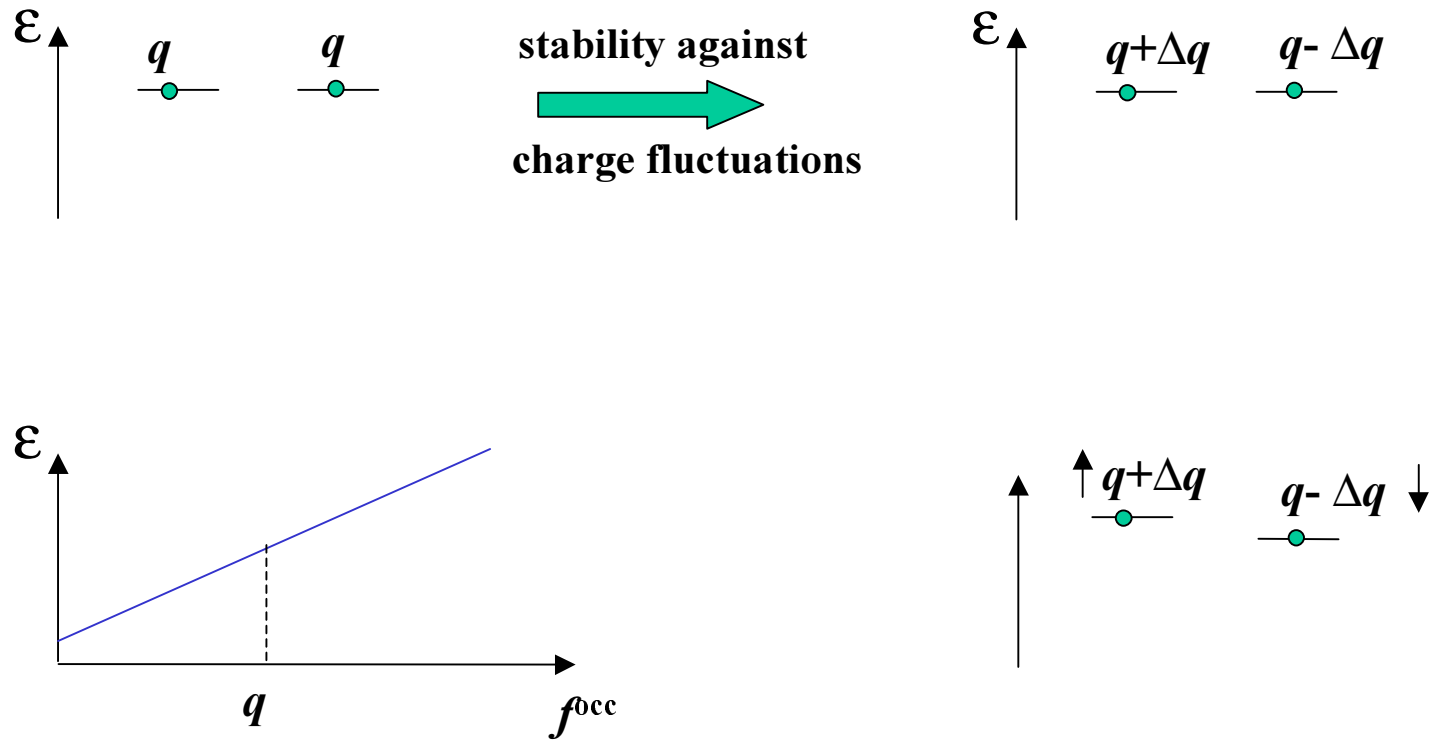
How to overcome the instabilities?
- **Use higher order algorithms**
- **Mixing of the charge density, e.g.**

$$\rho(\vec{r}) = \lambda_{mix}\rho^{old}(\vec{r}) + (1 - \lambda_{mix})\rho^{new}(\vec{r})$$
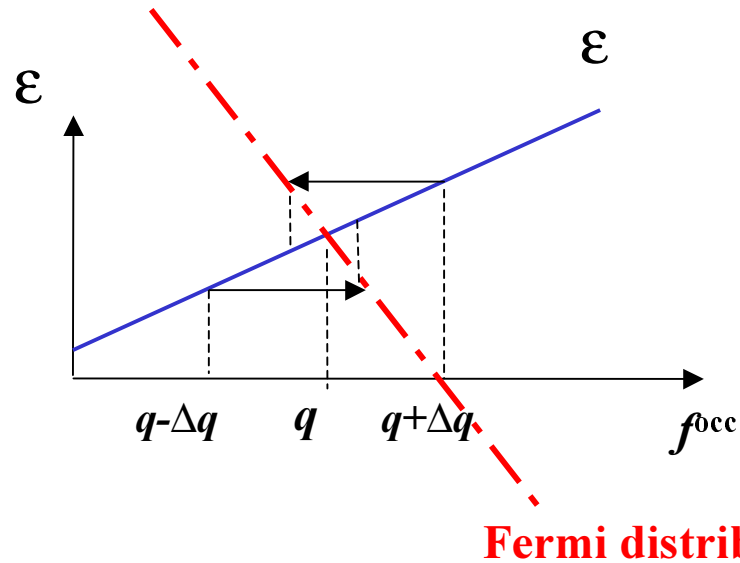
Instabilities: Metallic systems (I)

Fractionally occupied states:
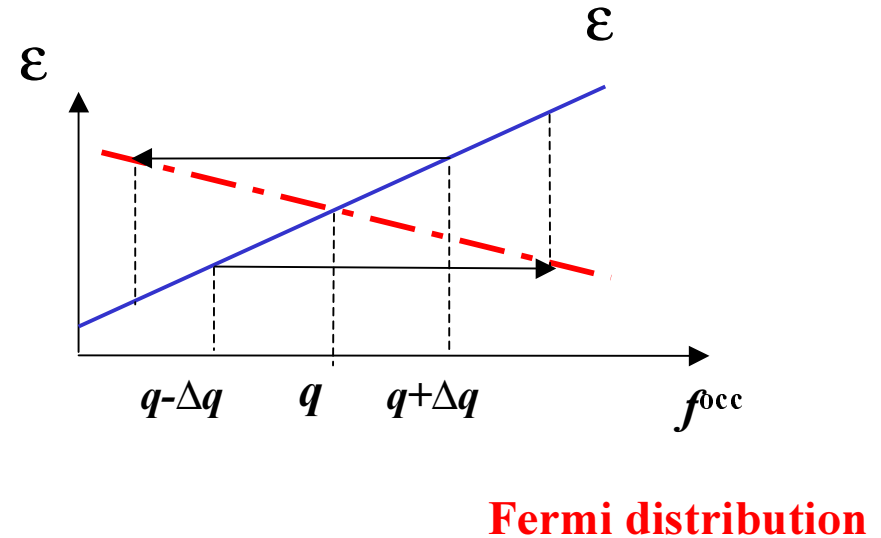
Consider two degenerate states close to the Fermi level:

Instabilities: Metallic systems (II)

High electronic temperatures:

Fermi distribution

Low electronic temperatures:

Fermi distribution
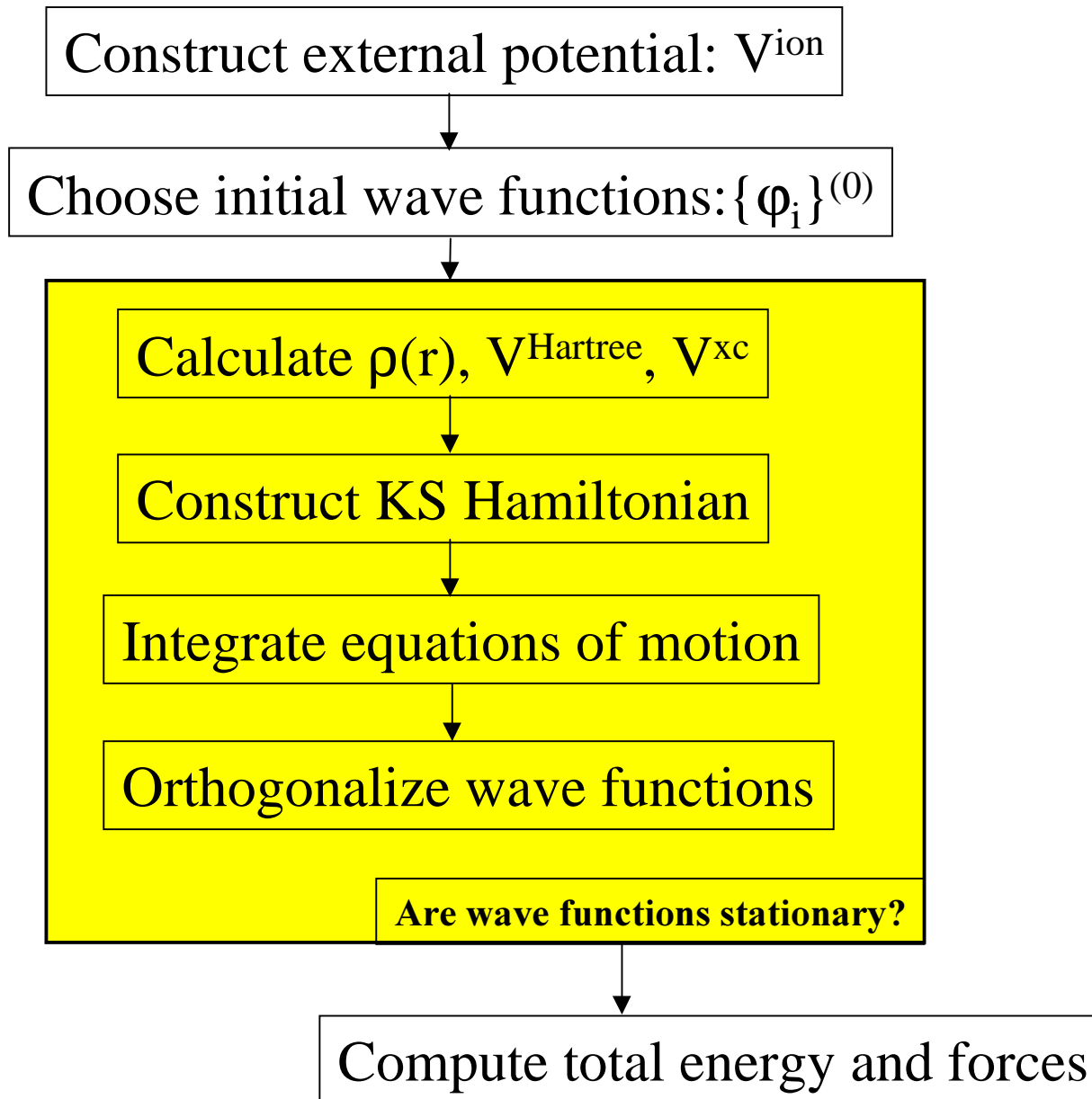
Stable against charge fluctuations!

Unstable against charge fluctuations!

How can we avoid these instabilities?
- **Improve k-point sampling**
- **damp occupation numbers**

How to get down to the BO surface

Construct external potential: V^{ion}

Choose initial wave functions: $\{\varphi_i\}^{(0)}$

Calculate $\rho(r)$, $V^{Hartree}$, V^{xc}

Construct KS Hamiltonian

Integrate equations of motion

Orthogonalize wave functions

Are wave functions stationary?

Compute total energy and forces

Hellmann-Feynman forces

Hellmann-Feynman theorem:

$$\vec{F}_I = -\frac{\partial E\left(\left\{\vec{R}_I\right\}\right)}{\partial \vec{R}_I}$$

$$= -\sum_i \left\langle \varphi_i \left| \frac{\partial H}{\partial \vec{R}_I} \right| \varphi_i \right\rangle$$

This theorem is valid only when: $\left(H - \varepsilon_i\right)\left|\varphi_i\right\rangle = 0$

Forces are not variational with respect to the wave functions
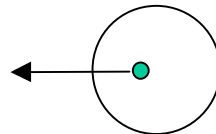$\Rightarrow$ **good convergence required**

Ionic motion

Basic idea: Simultaneous motion of electrons <u>and</u> ions per iteration step

   fictitious dynamics for electrons keeps wave functions on BO surface

Car-Parrinello Lagrangian:

$$L = \sum_i \frac{1}{2}\mu\langle\dot{\varphi}_i|\dot{\varphi}_i\rangle + \sum_I \frac{1}{2}M_I\dot{\vec{R}}_I^2 - E\big[\{\varphi_i\},\{\vec{R}_I\}\big]$$

Equations of motions:

$$M_I\ddot{\vec{R}}_I = -\frac{\partial E\big[\{\varphi_i\},\{\vec{R}_I\}\big]}{\partial\vec{R}_I} \qquad \mu\ddot{\varphi}_i = -\frac{\partial E\big[\{\varphi_i\},\{\vec{R}_I\}\big]}{\partial\varphi_i}$$

First order EOM:

Second order EOM:

How do we move the atoms

Two schemes

Molecular dynamics

Structure optimization

Integrate equations of motion:

$$M_I \ddot{\vec{R}}_I = \vec{F}_I$$

2nd order EOM with damping

$$\mu_I \ddot{\vec{R}}_I + \lambda_I \dot{\vec{R}}_I = \vec{F}_I$$

effective mass   damping/friction