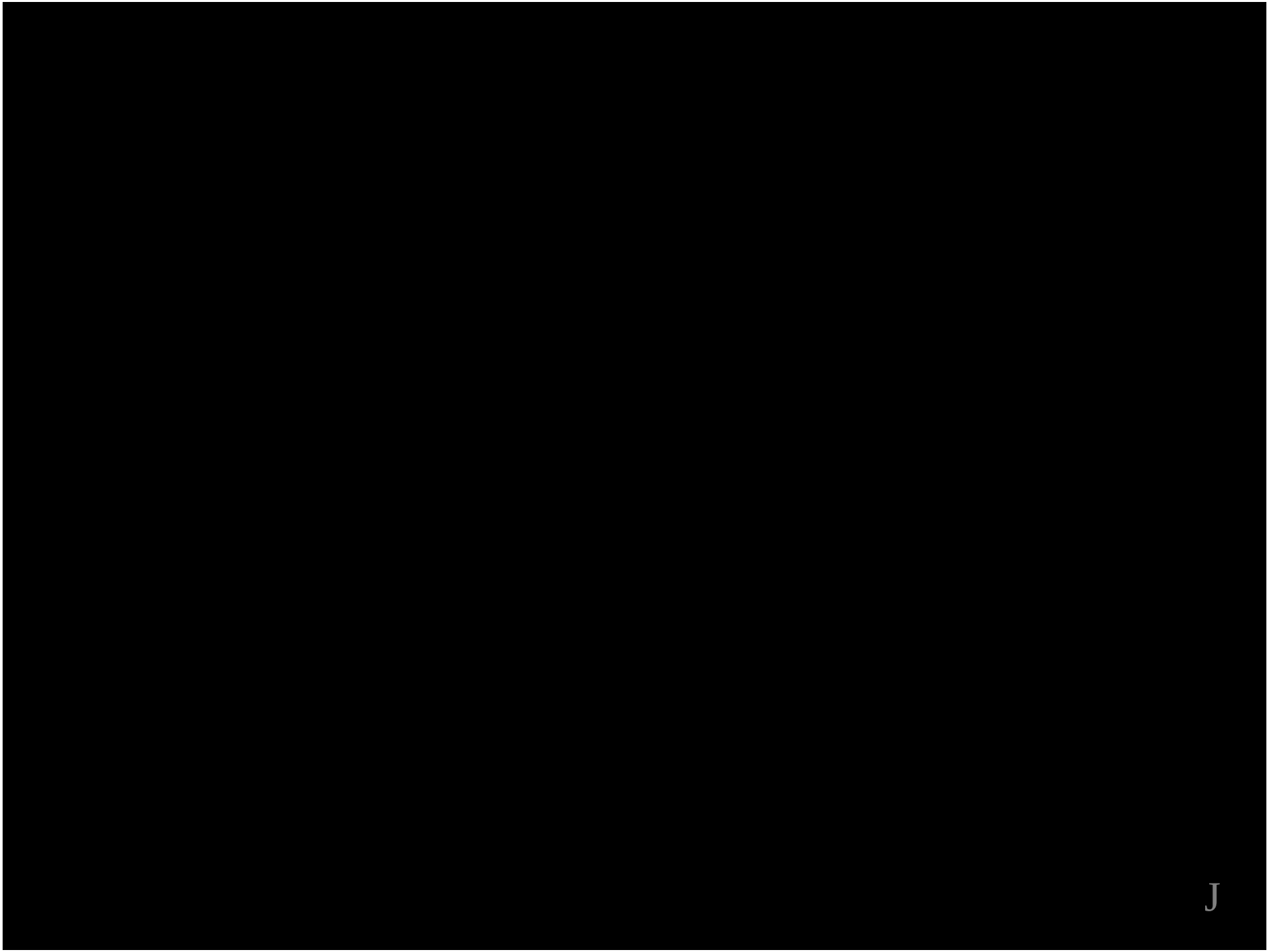

J

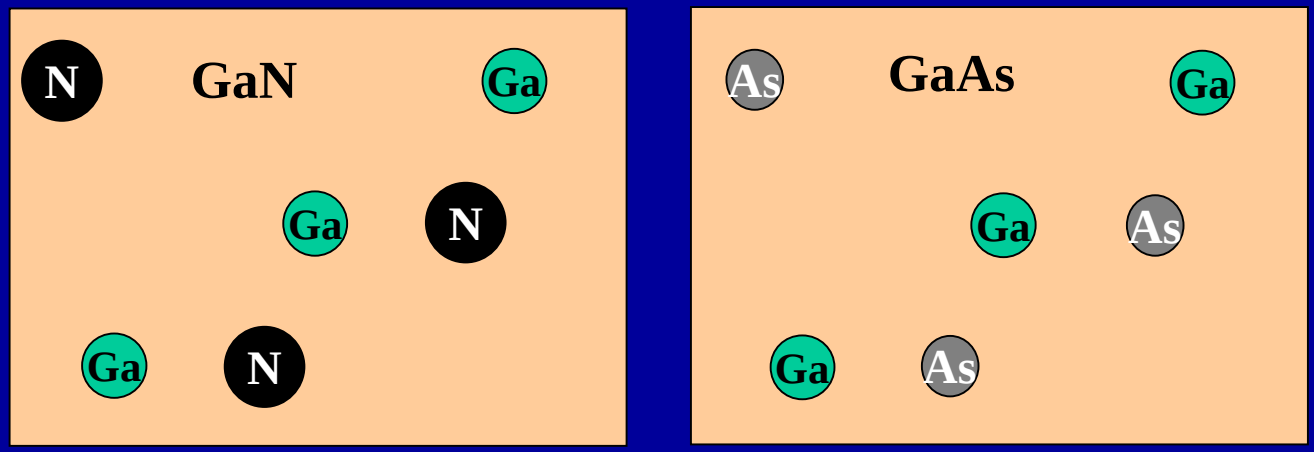
Algorithms for Total Energy Minimization

Jörg Neugebauer
Fritz-Haber-Institut, Berlin

Previous talks:

$$E_{tot}\left[\{\vec{R}_I\},\{\varphi_i\}\right]$$

Example:



Step 1:
provide H

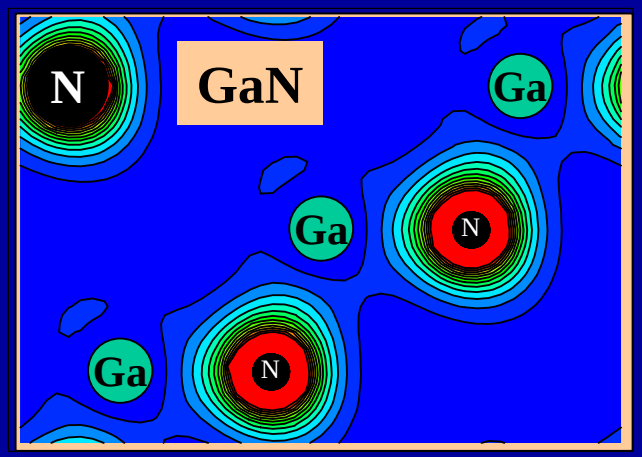
Algorithms for Total Energy Minimization

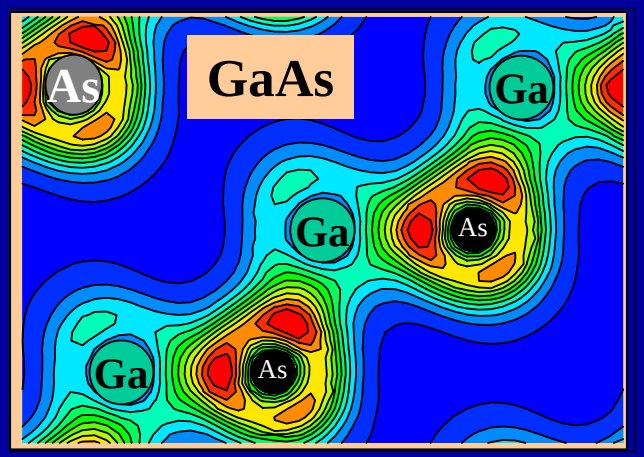
Jörg Neugebauer
Fritz-Haber-Institut, Berlin

Goal: Minimize total energy

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

Example:





Step 1:
provide H

Step 2:
Solve KS-Eq.

Questions/Problem

How to find (efficiently) the electronic ground state?

$\Rightarrow$ Solve Kohn-Sham Equations

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

How to find (efficiently) the self-consistent charge density?

Outline

How to get to the Born-Oppenheimer surface?

(1) Diagonalize the Hamiltonian

$$H[n]\varphi_i = \varepsilon_i \varphi_i$$

- Direct diagonalization
- Iterative diagonalization

(2) Perform self consistency (find n_{SCF})

- direct methods
- iterative methods [combine step (1) and (2)]

How to move on the Born-Oppenheimer surface?

- Equilibrium geometry (T=0K)
- Molecular dynamics ($\Rightarrow$ Session L16)

Basis set representation

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

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

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

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

direct diagonalization: $\hat{U}^T\hat{H}\hat{U} = \hat{1}\vec{\varepsilon}$ **[see e.g. Numerical recipes]**

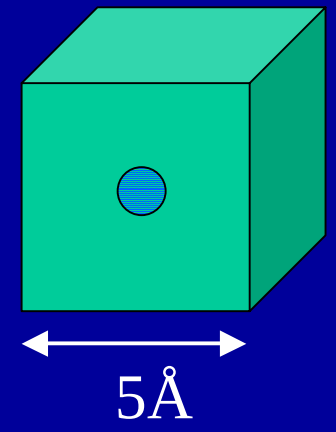
$$\begin{pmatrix} U_{11} & \cdots & \cdots & U_{N1} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ U_{1N} & \cdots & \cdots & U_{NN} \end{pmatrix} \begin{pmatrix} H_{11} & \cdots & \cdots & H_{1N} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ H_{N1} & \cdots & \cdots & H_{NN} \end{pmatrix} \begin{pmatrix} U_{11} & \cdots & \cdots & U_{1N} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ U_{N1} & \cdots & \cdots & U_{NN} \end{pmatrix} = \begin{pmatrix} \varepsilon_{11} & \cdots & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \varepsilon_{NN} \end{pmatrix}$$

Problems: • **N^3 scaling**

• **complete Hamiltonian has to be saved**

Discussion: Plane wave basis

Example: cubic cell with a=5Å



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

$\Rightarrow$ direct diagonalization inefficient/impossible

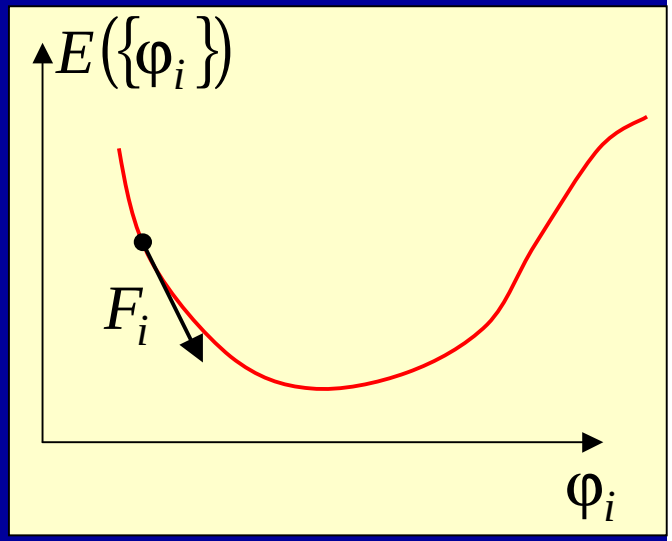
Iterative diagonalization

Basic idea:

Total energy gradient:

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

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



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

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

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

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

Scaling of the iterative algorithm

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

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

$O(N^2M)$ 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 $\overset{\text{FFT}}{\longleftrightarrow}$ **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

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

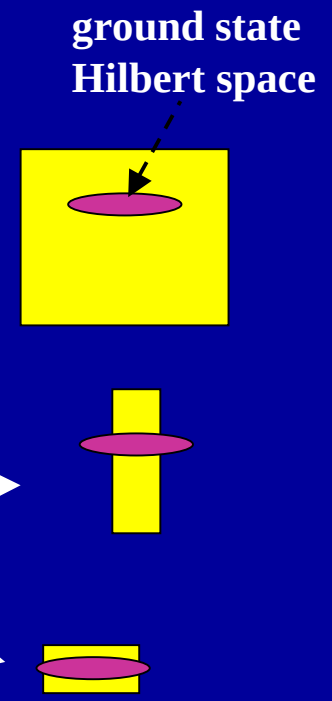
O(N) operations

Construction of the initial wave functions

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

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

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



Solutions:

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

J. Neugebauer and C. Van de Walle, Mat. Res. Soc. Symp. Proc. 408, 43-48 (1996).

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:

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

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

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

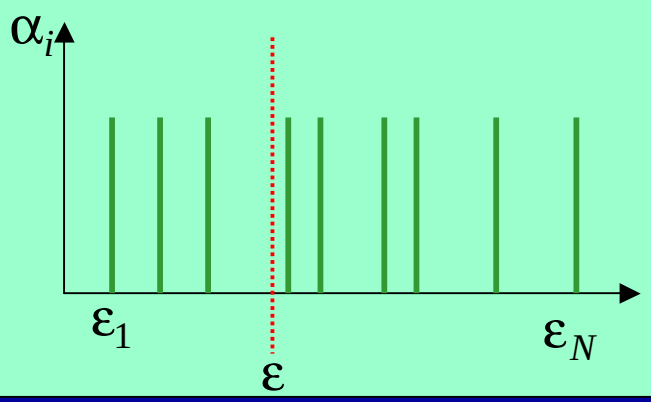
exact solution: $(H - \varepsilon)\left|\chi_i\right\rangle = 0$

representation of the trial wave function in the χ_i set:

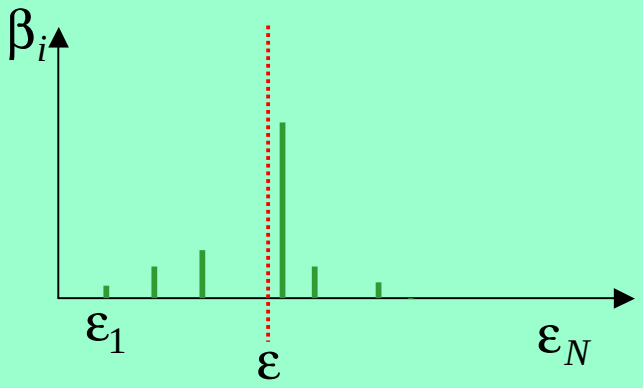
$$\left|\varphi^{(n)}\right\rangle = \sum_i \alpha_i \left|\chi_i\right\rangle \qquad \left|\varphi^{(n+1)}\right\rangle = \sum_i \beta_i \left|\chi_i\right\rangle$$

$$\Rightarrow \beta_i = \alpha_i - (\varepsilon_i - \varepsilon)\Delta t$$

Before iteration step:

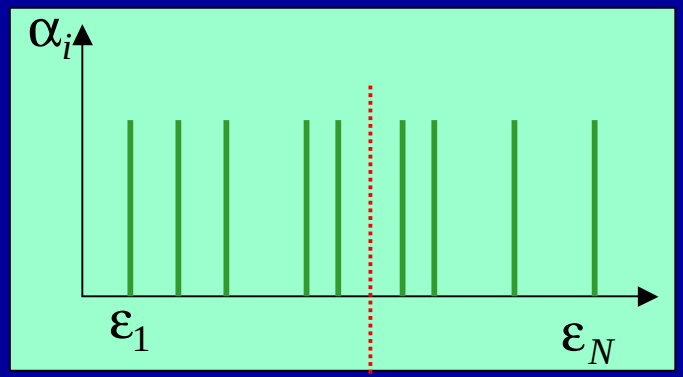


After "ideal" iteration step:



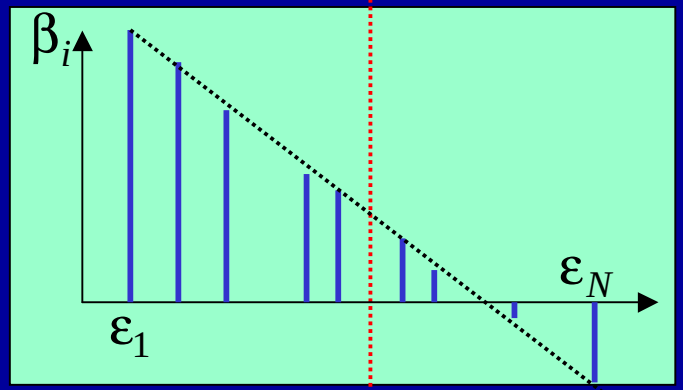
Convergence Criteria (II)

Before iteration step:



EOM: $\beta_i = \alpha_i - (\varepsilon - \varepsilon_i)\Delta t$

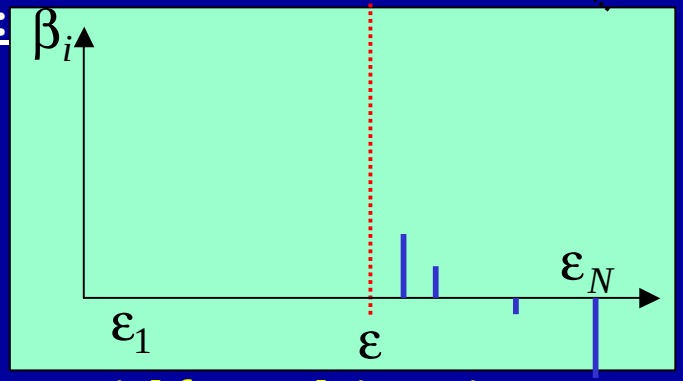
After iteration step:



optimum time step:

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

After orthogonalization:



$$\Delta t_{\max} = \frac{1}{\varepsilon_N - \varepsilon_1} = \frac{1}{\underbrace{\varepsilon_{\max} - \varepsilon_{\min}}_{\text{spectral radius}}}$$

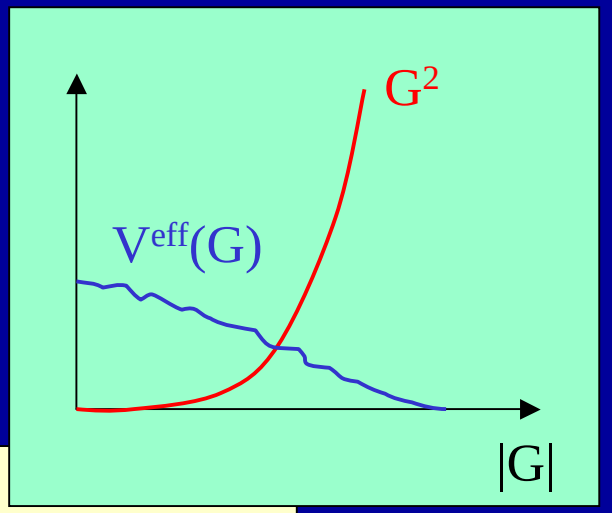
- **Orthogonalization essential for each iteration step**
- **Convergence rate decreases with increasing number of PW**

Preconditioning

Residual error: $$(H-\varepsilon)\underbrace{|\chi+\Delta\varphi\rangle}_{=:|\varphi\rangle} = (H-\varepsilon)|\Delta\varphi\rangle = |\Delta\tilde{\varphi}\rangle$$

we need this — we get this

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



Hamiltonian in PW basis:

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

Kinetic energy dominates for high wave numbers!:

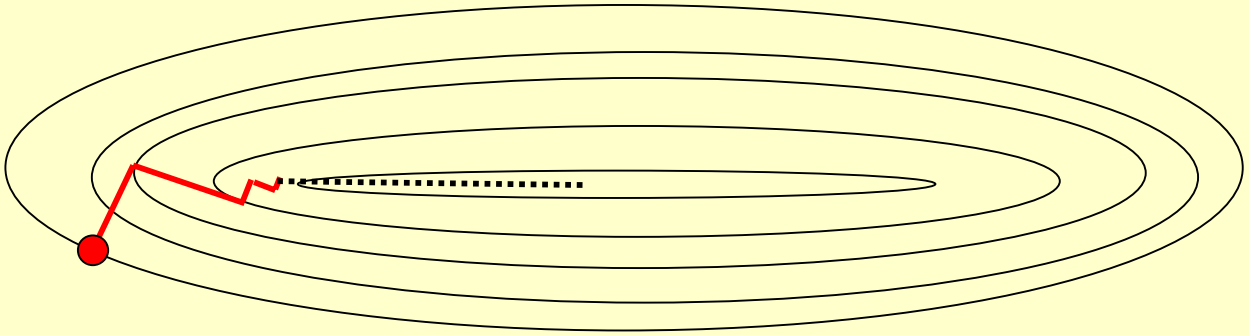
Partition Hamiltonian: $$\hat{H} = \hat{D} + \hat{L}$$

diagonal matrix — all diagonal elements are zero

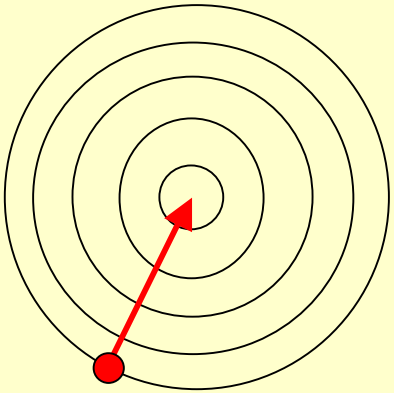
For high wave numbers: $\hat{H} \approx \hat{D}$ $\Rightarrow$ $(\hat{D}-\varepsilon)^{-1}\Delta\tilde{\varphi} = \Delta\varphi$

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:

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

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

Assumption: B_G **not time dependent**

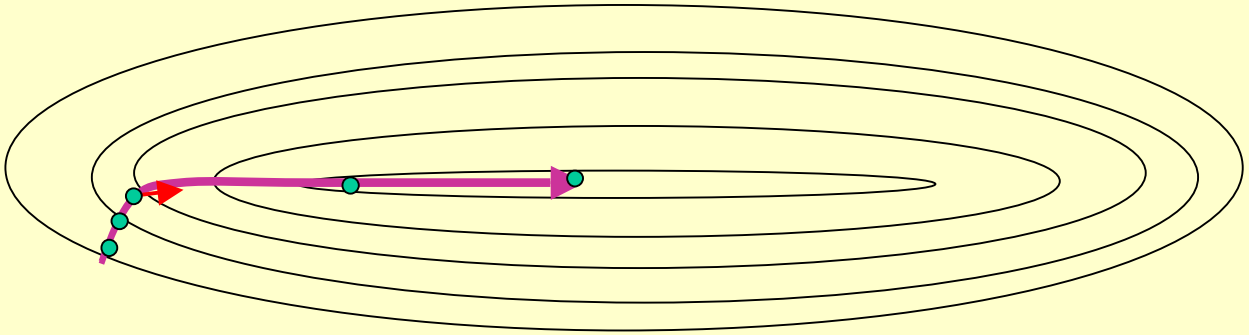
$$\Rightarrow\quad c_i(G,t=\Delta t)=-\frac{B_G}{\omega_G^2}+\left[c_i(G,t=0)+\frac{B_G}{\omega_G^2}\right]\exp\left(-\omega_G^2\Delta t\right)$$

Improves significantly the convergence rate for high PW energy cutoffs!

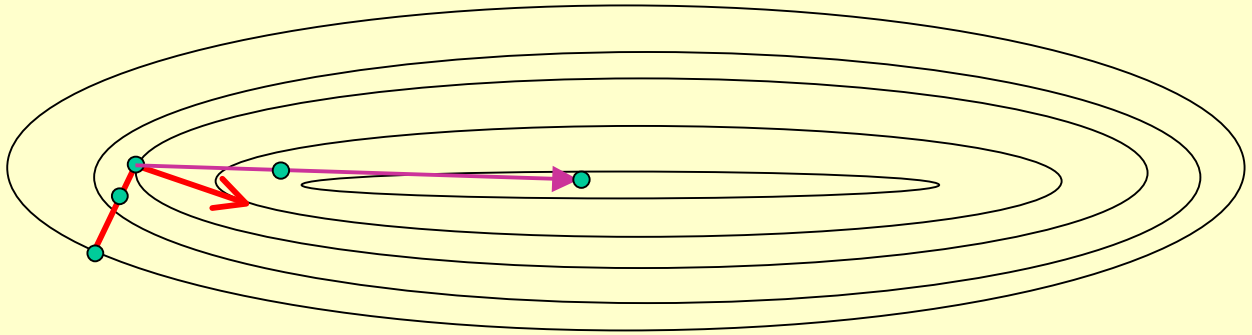
A. Williams and J. Soler, Bull. Am. Phys. Soc. 32, 562 (1987).

How to obtain more efficient schemes?

(1) Use higher order equations of motion:



(2) Perform accurate line minimization along the search direction:

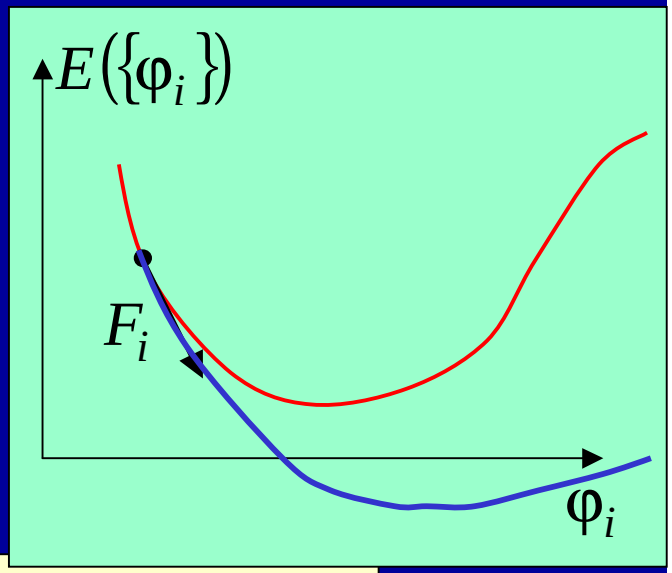


Higher order EOM's

Goal: Find total-energy minimum more efficiently

Strategies:

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



Two approaches

Use higher derivatives

$$D_{ij} = \frac{\partial^2 E(\{\varphi_i\})}{\partial\varphi_i \partial\varphi_j}, \quad \cdots$$

high computational effort!

Use higher order in time

$$\varphi_i = F(\dot{\varphi}_i, \ddot{\varphi}_i, \ldots)$$

equivalent:

$$\varphi_i^{(n+1)} = \tilde{F}(\varphi_i^{(n)}, \varphi_i^{(n-1)}, \ldots)$$

no additional computational effort!

Second order equation of motion

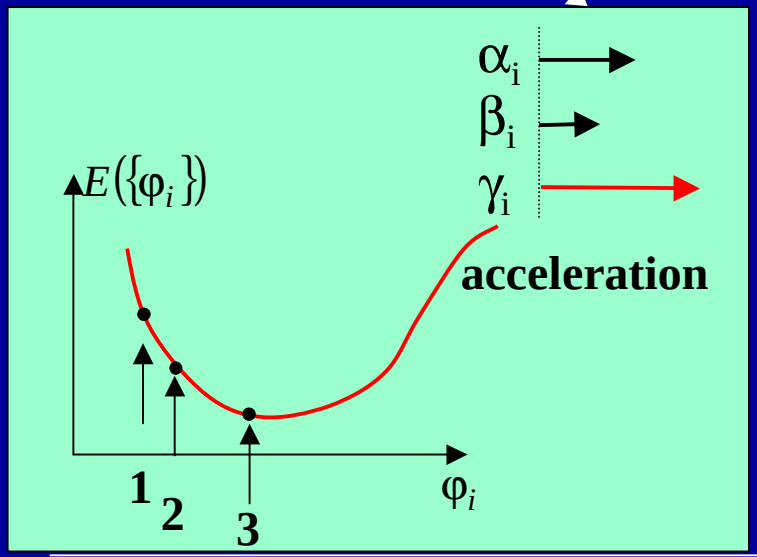
Finite differences: $\varphi^{(n+1)} = \varphi^{(n)} - (H-\varepsilon)\left|\varphi^{(n)}\right\rangle \Delta t_1 - (H-\varepsilon)\left|\varphi^{(n-1)}\right\rangle \Delta t_2$

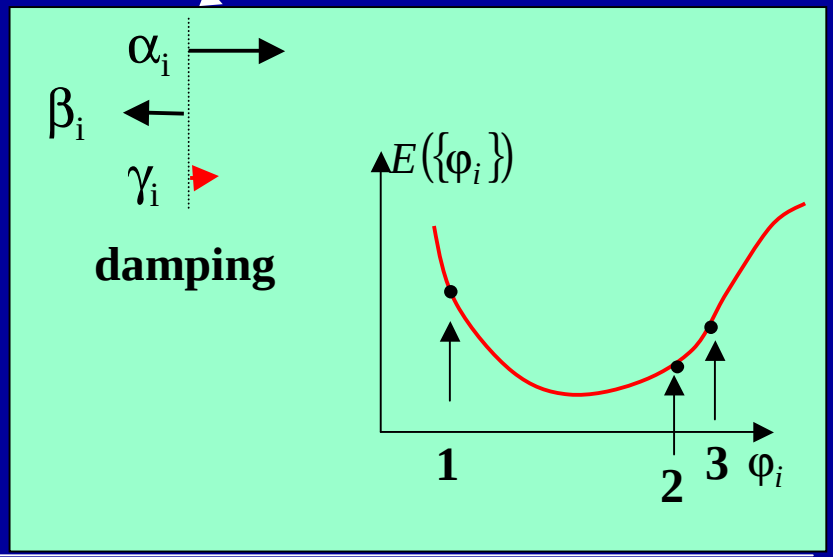
Expansion with respect to eigenfunctions χ_i:

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

$$\Rightarrow \quad \gamma_i = \beta_i - \beta_i(\varepsilon - \varepsilon_i)\Delta t_1 - \alpha_i(\varepsilon - \varepsilon_i)\Delta t_2$$

Two possible scenarios

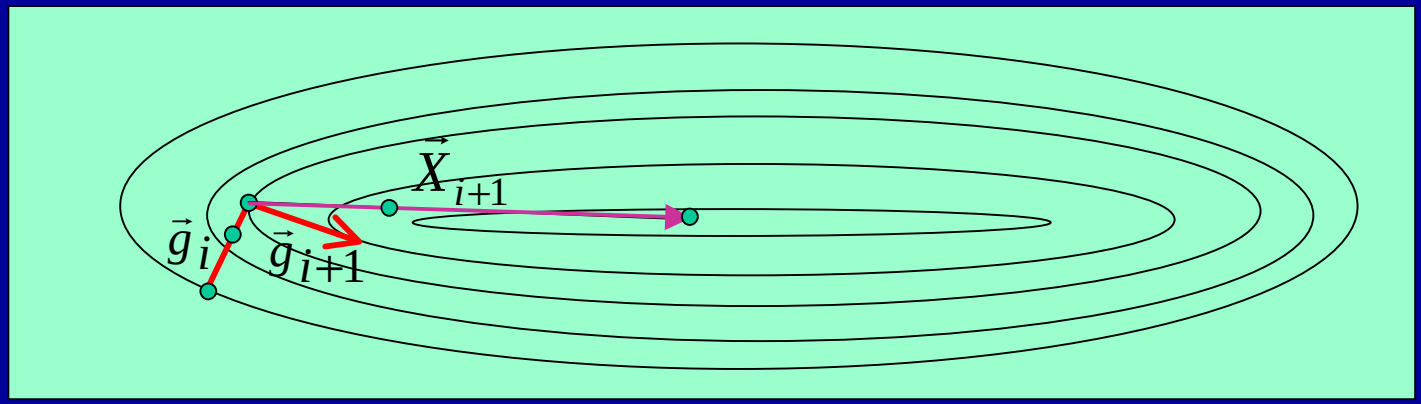




F. Tassone et al., Phys. Rev. B50, 10561 (1994).

Conjugate Gradient Schemes

Construct the best possible search direction based on the history of gradients:



This direction can be analytically calculated assuming a harmonic total energy surface:

$$|X_{i+1}\rangle = |g_i\rangle + \lambda_i |X_i\rangle \qquad \lambda_i = \frac{\langle g_{i+1}|g_{i+1}\rangle}{\langle g_i|g_i\rangle}$$

in praxis replaced by preconditioned gradient!

Properties:

- **+ very efficient: number of iterations ≤ dimensionality of problem**
- **+ search directions are completely decoupled (conjugated)**
- **(-) accurate line minimization crucial (i.e., one CCG-step needs at least two electronic steps)**

see e.g.: M.C. Payne et al., Phys. Rev. B56, 2656 (1986).

Charge self-consistency

What did we do so far?

efficient methods to calculate one-particle energies and eigenfunctions

$$H\left[n^{(\alpha)}\right]\varphi_i^{(\alpha+1)} = \varepsilon_i \varphi_i^{(\alpha+1)}$$

provided

What remains to be done?

calculate new charge density:

$$n^{(\alpha+1)}(\vec{r}) = \sum_i \left|\varphi_i^{(\alpha+1)}(\vec{r})\right|^2$$

Goal:

self-consistent charge density

$$n^{(\alpha+1)}(\vec{r}) = n^{(\alpha)}(\vec{r})$$

output density **input density**

How to achieve charge self-consistency?

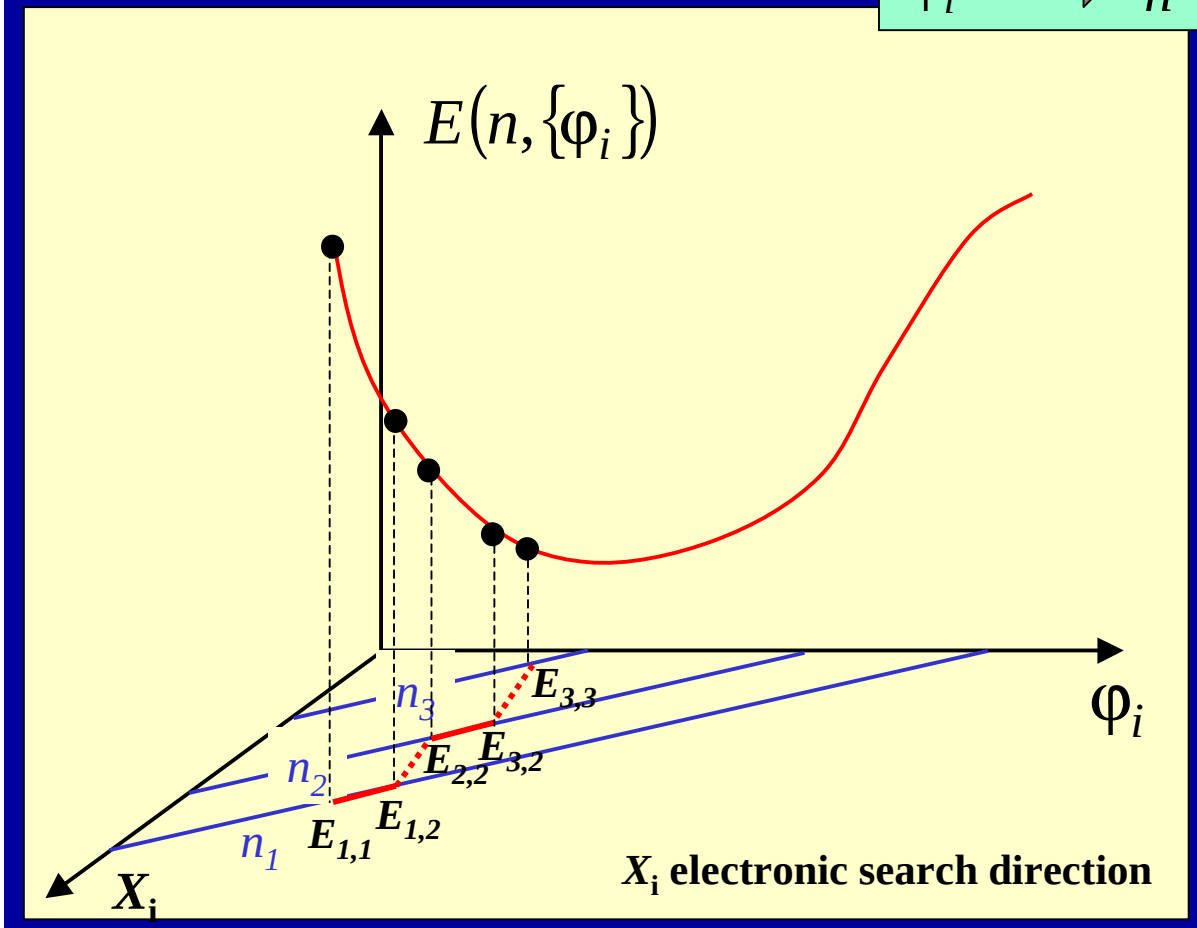
Indirect Methods

Update after each electronic step:

$$\varphi_i^{(\alpha)} \Rightarrow n^{(\alpha)} \Rightarrow H[n^{(\alpha)}] \Rightarrow \varphi_i^{(\alpha+1)}$$

$$E\left(n^{(\alpha)}, \{\varphi_i^{(\alpha)}\}\right) := E_{\alpha,\alpha}$$

$$E\left(n^{(\alpha)}, \{\varphi_i^{(\alpha+1)}\}\right) := E_{\alpha,\alpha+1}$$



Approach works for

- **Steepest Descent**
- **Williams Soler**
- **Higher order schemes**

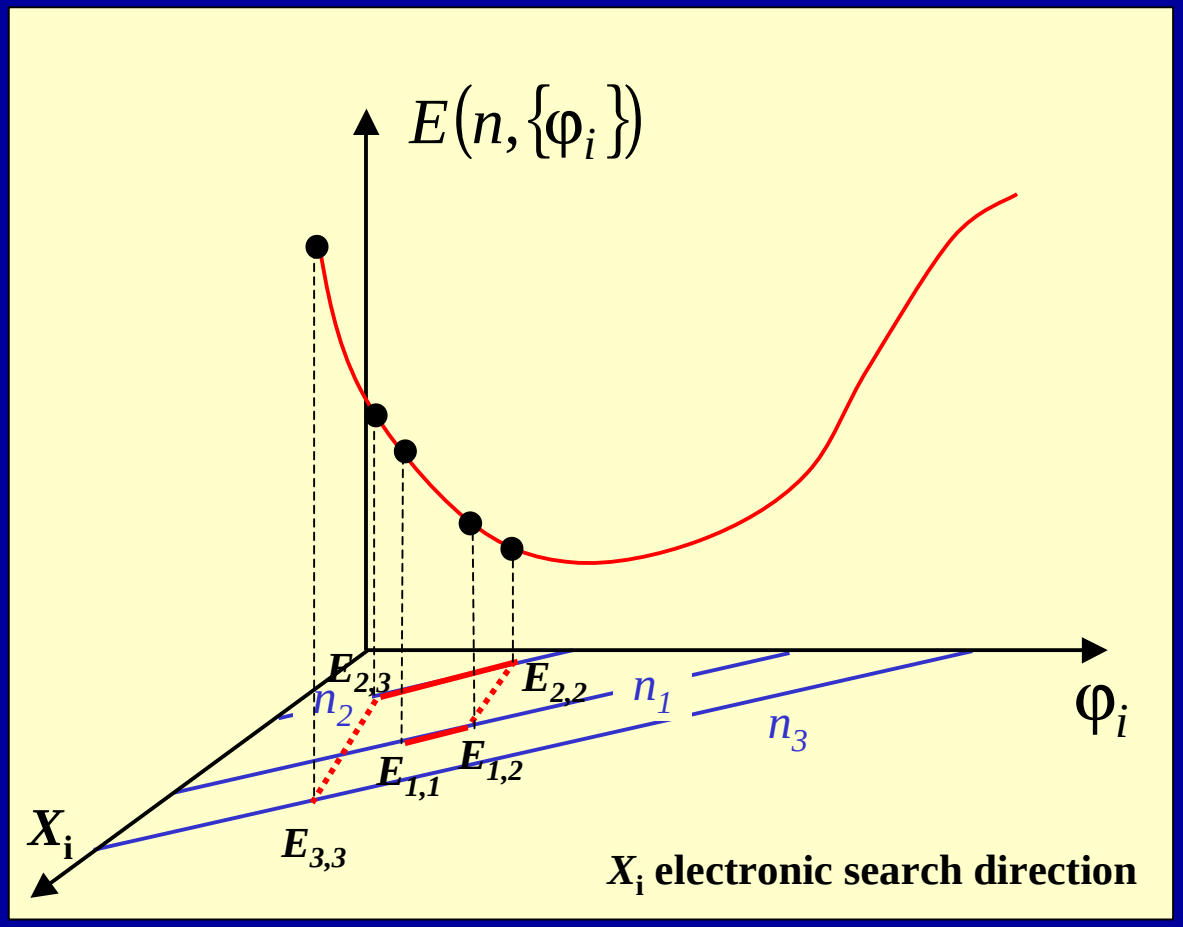
but not for

- **Conjugate gradient**
  → **line minimization not possible**

Self-consistency is enforced at each iteration step!

Occurrence of Instabilities

Behavior for badly converging systems:



Instabilities may occur ⇒ convergence in total energy destroyed

Origin of the Instabilities (I)

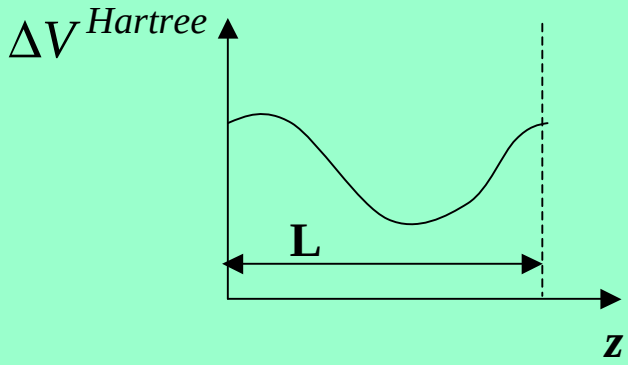
Behavior of the electrostatic potential:

Hartree potential:

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

Change in Hartree potential:

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



Instability increases with increasing system size:

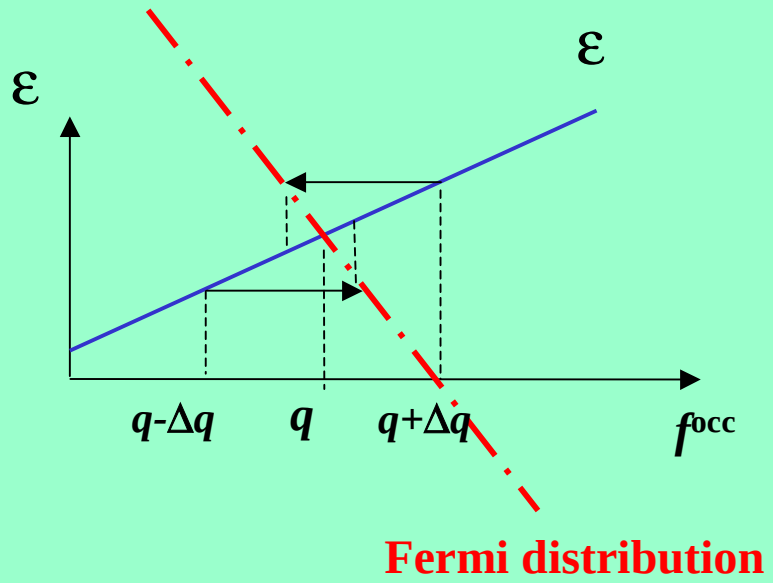
$$G_{\min} \propto \frac{1}{l_{\max}}$$

With increasing system size even small perturbations in the charge density may have a huge effect on the electrostatic potential!

Origin of the Instabilities (II)

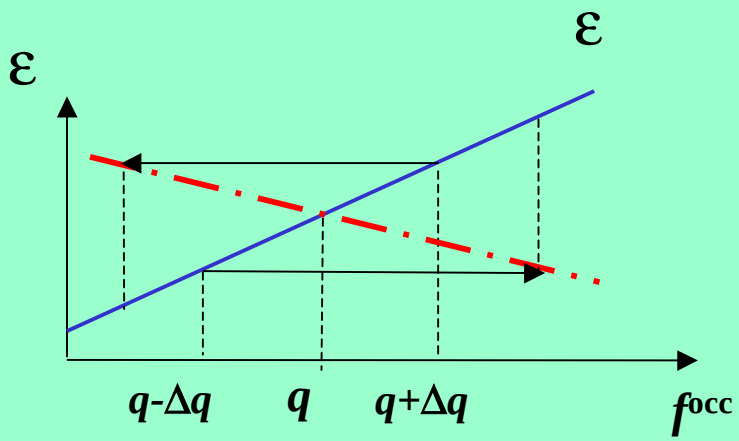
Metallic systems:

High electronic temperatures:



Stable against charge fluctuations!

Low electronic temperatures:



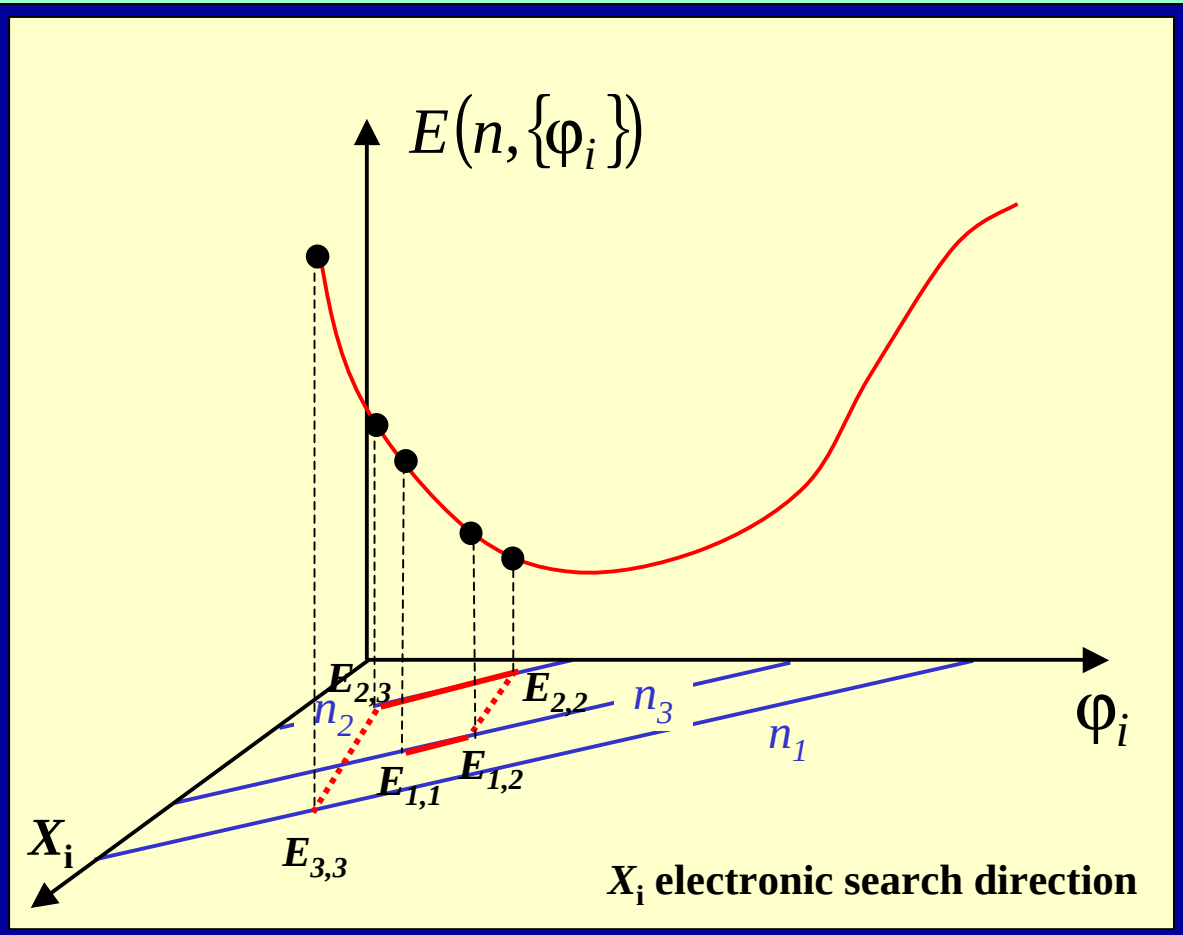
Unstable against charge fluctuations!

How to avoid these instabilities?

Instabilities: Solutions (I)

Solution: Make self-consistency step (dashed line) shorter

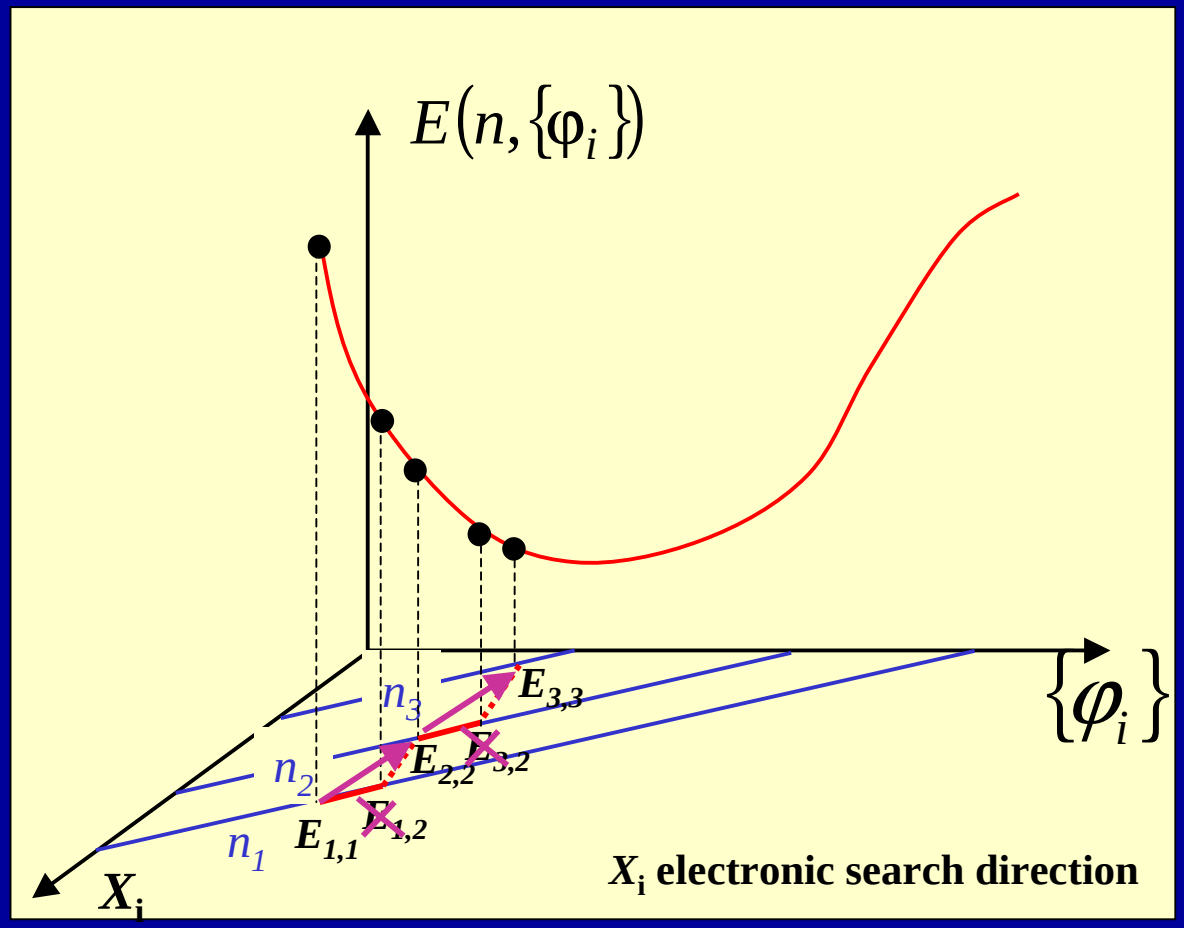
- **use charge density mixer:** $n_{mix}^{(\alpha+1)} = (1-\lambda)n^{(\alpha+1)} + \lambda n^{(\alpha)}$
- **use damping in the occupation numbers**
- **use smaller time step**



Instabilities: Solutions (II)

Solution:

- Construct a search direction which minimizes energy with respect to both self-consistency and diagonalization
- Apply a CCG minimizer



Advantages:

- very stable
- monotonous behavior
- fast convergence

Note:

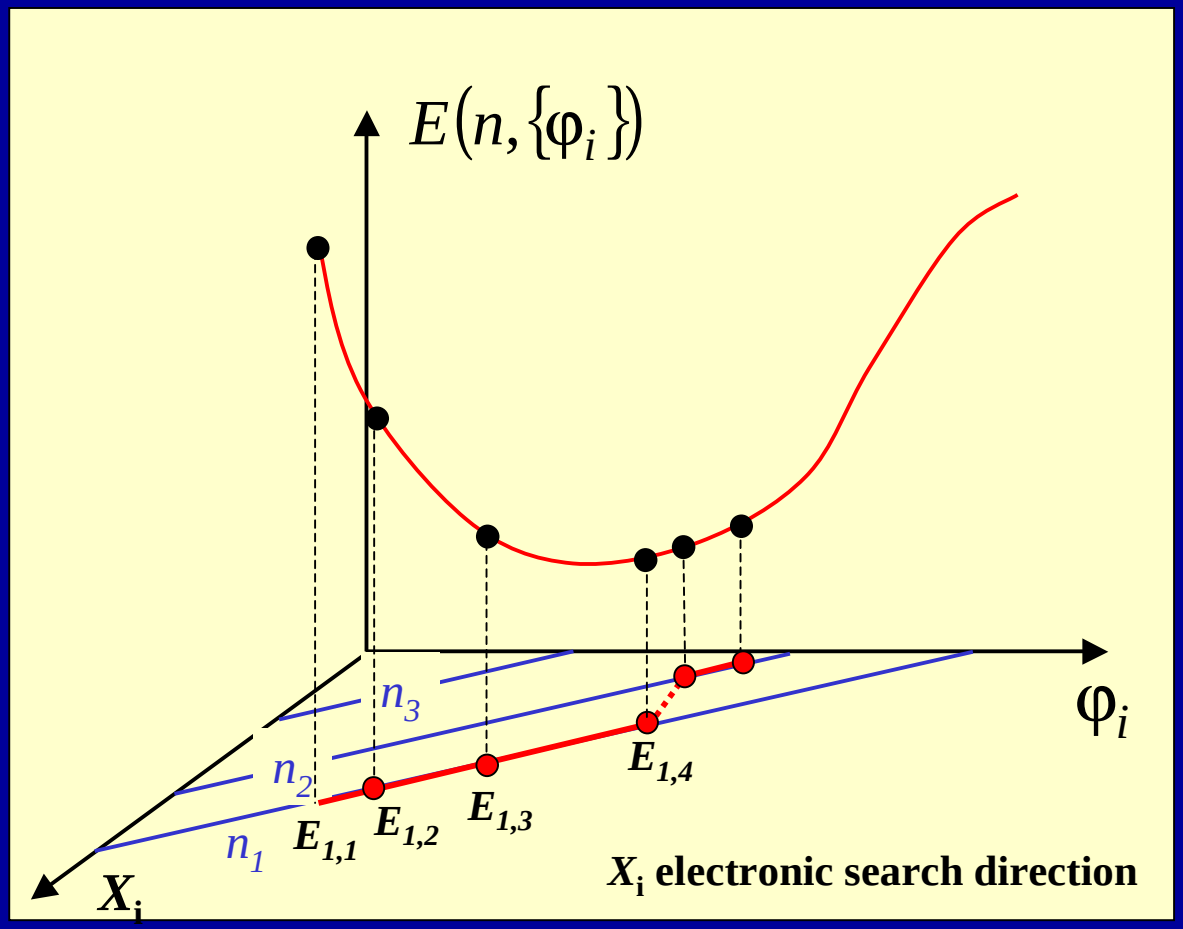
- can be used only for non-metallic systems

S. Ismail-Beigi and T.A. Arias, Comp. Phys. Comm. 128, 1 (2000).

Instabilities: Solutions (III)

Solution: Direct Methods

- **Fully diagonalize the Hamiltonian**
- **Perform than an efficient mixing scheme**



G. Kresse, J. Furthmüller, Phys. Rev. B54, 11169 (1996).

Efficient Mixing Schemes

indirect Methods: $n_{out} = f(n_{in}, \{\varphi_i\})$

direct Methods: $n_{out} = f(n_{in})$

Residuum: $R = n_{out} - n_{in}$ **Self-consistency:** $\langle R | R \rangle = 0$

$R = \tilde{f}(n_{in})$

Idea: Assume linearity of $\tilde{f}$ **⇒ Pulay mixing**

⇒ use information from previous steps $[R(n^{(i)})]$

Minimize residuum: $\langle R^{(m+1)} | R^{(m+1)} \rangle \rightarrow \min$

with: $R^{(m+1)} = \sum_{i=1}^{m} \alpha_i R^{(i)}$ **determines** α_i

Construct optimum new input density: $n_{in}^{(m+1)} = \sum_{i=1}^{m} \alpha_i n_{in}^{(i)}$

Practical realization: use $R^{(m)}$-$R^{(m-1)}$ **rather then** $R^{(m)}$ **(DIIS)**

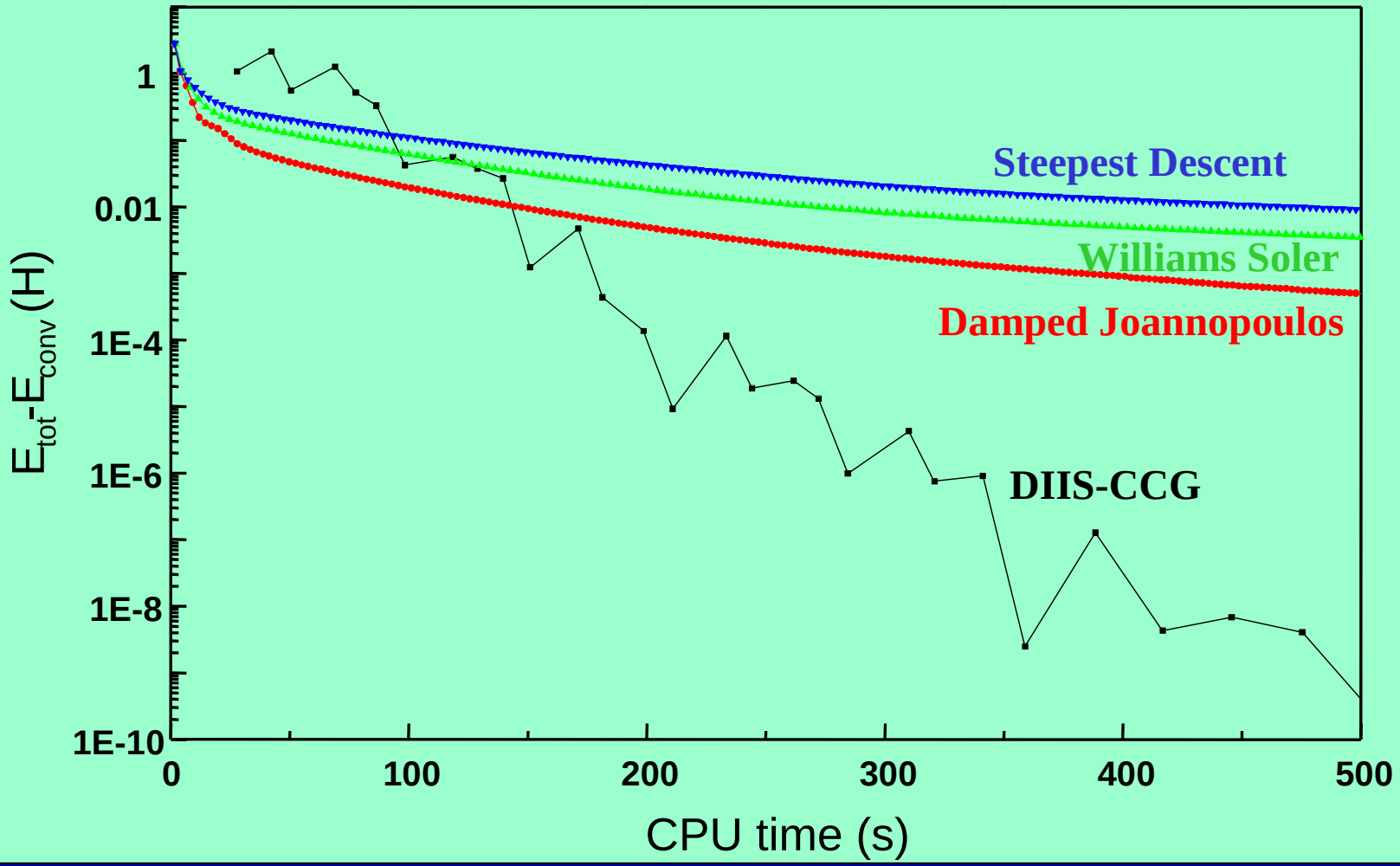
precondition $R^{(m)}$ **(Kerker mixing)**

Comparison of the Different Methods

Method	diag.	charge density	# internal steps	# steps	systems
Steepest Descent	single it.	it.	1	-	all
Williams Soler	single it.	it.	1	+	all
Damped Joannopoulos	single it.	it.	1	++	all
All band Conj. Grad.	single it.	it.	2	+++	non metallic
State by State Conj. Grad.	full it.	DIIS	2...10	++	all

Example I: Metallic Systems

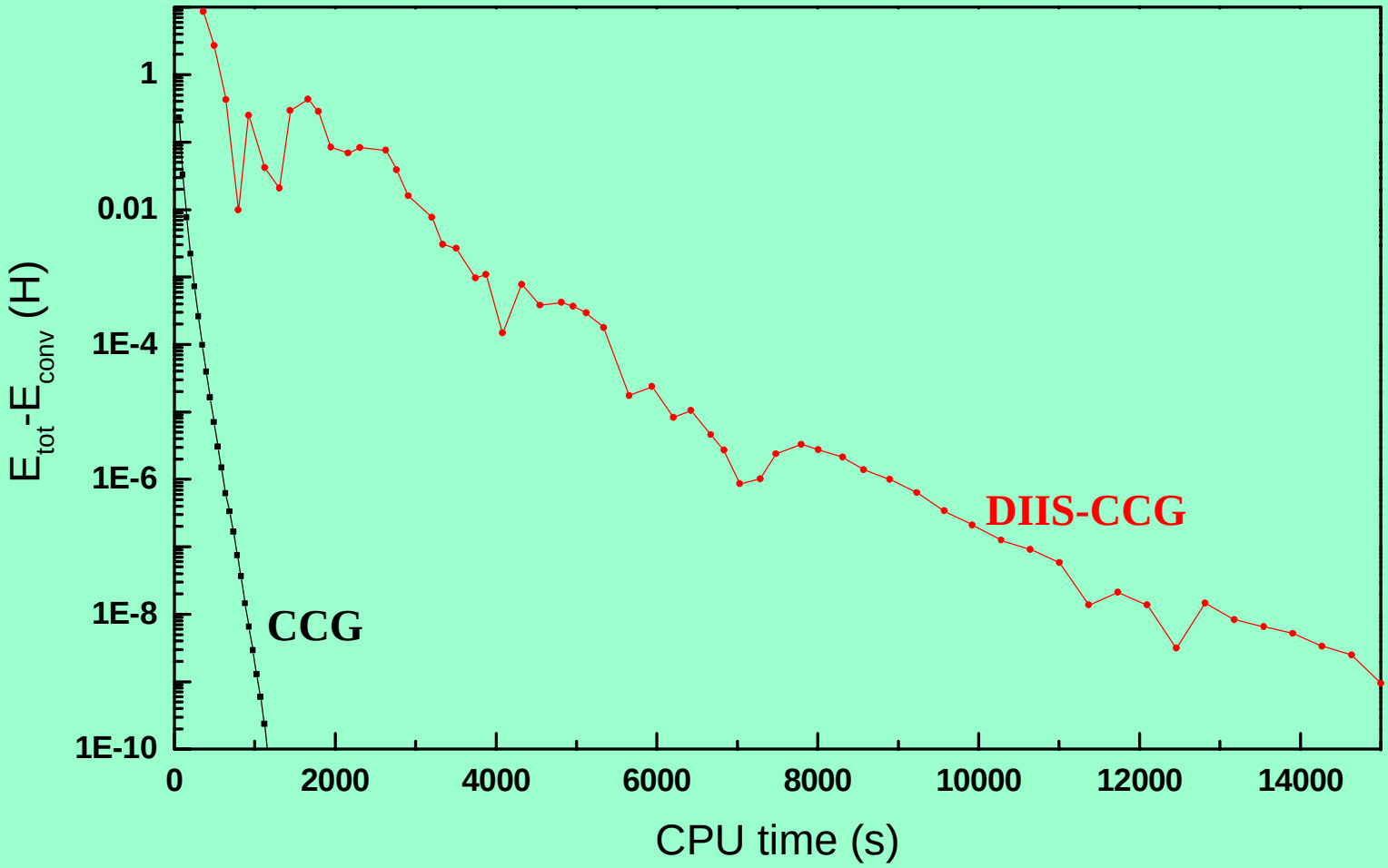
System: Spin-polarized cubic MnAs



DIIS-CCG has fastest convergence rate (but convergence step rather expensive)

Example II: Non-metallic Systems

System: 3-10 helix (30 atoms, E_{cut}=40 Ry)



CCG decreases monotonously and shows much faster convergence speed!

Conclusions

SFHIngS provides a variety of efficient minimization techniques

Conjugate gradient based schemes are most stable ones:

- **All band conjugate gradient**
  - **⇒ Optimum choice for non-metallic systems**
  - **⇒ no input parameters**
- **Band-by-band conjugate gradient:**
  - **⇒ Excellent convergence but significantly slower**
  - **⇒ parameters for preconditioner in density mixing needed (Kerker mixing)**

First and second order schemes:

- **⇒ may be faster than CG-based methods for simple metals such as e.g. Al**
- **⇒ require careful choice of convergence parameters such as time step and damping**

To obtain maximum performance (e.g. for large/complex systems) you may combine the various methods. This can be easily set up in the SFIngX input!