Efficient iterative methods for electronic structure computation

Emphasis on Plane Wave Methods

African School on Electronic Structure Methods and Applications
AIMS, Capetown, July 14-25, 2008

Richard M. Martin
Department of Physics and Materials Computation Center
University of Illinois at Urbana-Champaign
Department of Applied Physics, Stanford University

Reference:
Electronic Structure: Basic Theory and Practical Methods,
Outline

• **Plane waves solution for electronic states in crystals**
  – Hamiltonian matrix given simply in terms involving Fourier components of the potential

• **Clever ideas!**
  – New approaches, algorithms, . . .

• **Efficient iterative methods**
  – Well developed in plane waves because of the simplicity of the operations
  – Uniform grid operations

• **Some version use in ALL the modern plane wave codes**
  – Different codes use different methods
  – But all can be understood in a general framework
Independent electrons – Plane wave methods

The independent-particle Schrödinger equation

The eigenstates of any independent particle Schrödinger-like equation in which each electron moves in an effective potential $V_{\text{eff}}(\mathbf{r})$, such as the Kohn–Sham equations, satisfy the eigenvalue equation

$$\hat{H}_{\text{eff}}(\mathbf{r})\psi_i(\mathbf{r}) = \left[ -\frac{\hbar^2}{2m_e} \nabla^2 + V_{\text{eff}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}). \quad (12.1)$$

Using the fact that any periodic function can be expanded in the complete set of Fourier components, an eigenfunction can be written

$$\psi_i(\mathbf{r}) = \sum_q c_{i,q} \times \frac{1}{\sqrt{\Omega}} \exp(i\mathbf{q} \cdot \mathbf{r}) \equiv \sum_q c_{i,q} \times |\mathbf{q}\rangle, \quad (12.2)$$

where $c_{i,q}$ are the expansion coefficients of the wavefunction in the basis of orthonormal plane waves $|\mathbf{q}\rangle$ satisfying

$$\langle q'|q \rangle \equiv \frac{1}{\Omega} \int_{\Omega} d\mathbf{r} \exp(-i\mathbf{q}' \cdot \mathbf{r}) \exp(i\mathbf{q} \cdot \mathbf{r}) = \delta_{q,q'}. \quad (12.3)$$

Plane wave basis – always possible – Fourier analysis
Independent electrons – Plane wave methods

The hamiltonian matrix in a plane wave basis

**Definition:** \( q = k + G \) where \( q \) is a general wavevector. \( k \) is defined to be in the BZ, \( G \) is a reciprocal lattice vector.

Finally, if we define \( q = k + G_m \) and \( q' = k + G_{m'} \) (which differ by a reciprocal lattice vector \( G''_m = G_m - G_{m'} \)), then the Schrödinger equation for any given \( k \) can be written as the matrix equation

\[
\sum_{m'} H_{m,m'}(k)c_{i,m'}(k) = \varepsilon_i(k)c_{i,m}(k),
\]

where

\[
H_{m,m'}(k) = \langle k + G_m|\hat{H}_{eff}|k + G_{m'} \rangle = \frac{\hbar^2}{2m_e}|k + G_m|^2 \delta_{m,m'} + V_{eff}(G_m - G_{m'}). \]

**Note that** \( k \) **only appears in the kinetic energy!**
*(Actually V depends on k in pseudopotentials)*

Matrix eigenvalue equation with size (\# plane waves) x (\# plane waves)
Solving Kohn-Sham Equations

- Structure, types of atoms

- Guess for input

- Solve KS Eqs.

- New Density and Potential

- Self-consistent?

- Output:
  - Total energy, force, stress, ...
  - Eigenvalues

Self-Consistent Kohn-Sham Equations

\[
V_{\text{eff}}^{\sigma}(r) = V_{\text{ext}}(r) + V_{\text{Hart}}[n] + V_{\text{xc}}^{\sigma}[n^\uparrow, n^\downarrow]
\]

\[
\left(-\frac{1}{2} \nabla^2 + V_{\text{eff}}^{\sigma}(r)\right)\psi_i^{\sigma}(r) = \varepsilon_i^{\sigma}\psi_i^{\sigma}(r)
\]

\[
n^{\sigma}(r) = \sum_i f_i^{\sigma}|\psi_i^{\sigma}(r)|^2
\]
Solving Kohn-Sham Equations

- Critical steps in the algorithm
  - Solving the KS equations
    - By far the greatest computer time is spent in the operations involving the hamiltonian and the eigenstates
  - Achieving self-consistency
    - Iterative approaches to the solution
  - Important point to notice
    - As one approaches self-consistency, the $V_{\text{eff}}$ changes only a little from one iteration to the next
    - If we move atoms a small amount, $V_{\text{eff}}$ changes only a little from one calculation to the next

Self-Consistent Kohn-Sham Equations

1. Initial Guess
   \[ n^+(r), n^-(r) \]
2. Calculate Effective Potential
   \[ V_{\text{eff}}^\sigma(r) = V_{\text{ext}}(r) + V_{\text{Hart}}[n] + V_{\text{xc}}[n^+, n^-] \]
3. Solve KS Equation
   \[ \left( -\frac{1}{2} \nabla^2 + V_{\text{eff}}^\sigma(r) \right) \psi_i^\sigma(r) = \epsilon_i^\sigma \psi_i^\sigma(r) \]
4. Calculate Electron Density
   \[ n^\sigma(r) = \sum_i f_i^\sigma |\psi_i^\sigma(r)|^2 \]
5. Self-consistent?
   - Yes
     - Output Quantities
       - Energy, Forces, Stresses, Eigenvalues, …
   - No
     - Go back to step 1
Solving the matrix equations

• Standard diagonalization methods to find eigenvalues and eigenvectors
  • Lapack, . . .
  • Excellent codes – the way to go if one has a dense matrix

• Computational time
  • In general scales as (# Basis functions)^3
  • If one wants only a few eigenvectors –
    (# eigenfunctions) x (# Basis functions)^2

• Electronic structure problem
  • Very appropriate for small # Basis functions
    Local orbitals, LMTO, . . .
  • But for plane waves -- # Basis functions is very large
    # Basis functions >> # eigenfunctions

• Can we find more efficient methods for plane waves?
  • How can we use the fact that the operations are simple?
  • How can we use the fact that one needs to solve problems that
    are only slightly changed from a previous problem
Efficient iterative methods for plane waves

• From “Electronic Structure”, R. M. Martin, – Summary for Chapter 18

Of all the recent methods for computing the properties of materials from electronic equations, one stands out: i.e. the quantum molecular dynamics (QMD) simulations pioneered by Car and Parrinello in 1985 [156]. This work and subsequent developments have led to a revolution in the capabilities of theory to treat real, complex molecules, solids, and liquids including thermal motion (molecular dynamics), with the forces derived from the electrons treated by (quantum) density functional methods. Altogether, four advances create the new approach to electronic structure. These comprise:

• optimization methods (instead of variational equations),
• equations of motion (instead of matrix diagonalization),
• fast Fourier transforms (FFTs) – (instead of matrix operations), and
• a trace of occupied subspace (instead of eigenvector operations).

Car and Parrinello combined these features into one unified algorithm for electronic states, self-consistency, and nuclear movement. There has also been an explosion of alternative approaches that utilize the force theorem, together with efficient iterative methods described in App. M or simpler tight-binding-type methods. These are described in the present chapter as well as the Car–Parrinello method per se.
Review: the Kohn Sham approach

- Start again from the original Kohn-Sham problem

\[ n_0(r) = \sum_\sigma \sum_{i=1} \left| \psi_i^\sigma (r) \right|^2 , \]

\[ E_{KS} = \frac{1}{2} \sum_\sigma \sum_{i=1} |\nabla \psi_i^\sigma|^2 + \int dV_{ext}(r)n(r) + E_{Hartree}[n] + E_{II} + E_{xc}[n]. \]

Equations for independent particles - **soluble**

Exchange-Correlation Functional – Exact theory but **unknown** functional!

- We will proceed assuming that we have some approximate practical from for \( E_{xc}[n] \)
Review: the Kohn-Sham Equations

- Assuming a form for $E_{xc}[n]$
- Minimizing energy (with constraints) $\rightarrow$ Kohn-Sham Eqs.

$$n_0(r) = \sum_{\sigma} \sum_{i=1} |\psi_i^\sigma(r)|^2,$$

$$E_{KS} = \frac{1}{2} \sum_{\sigma} \sum_{i=1} |\nabla \psi_i^\sigma|^2 + \int dr V_{ext}(r)n(r) + E_{Hartree}[n] + E_{II} + E_{xc}[n].$$

- Constraint – required
- Exclusion principle for independent particles

$$\frac{\delta E_{KS}}{\delta \psi_i^{\sigma*}(r)} = 0, \quad \langle \psi_i^\sigma | \psi_j^{\sigma'} \rangle = \delta_{i,j} \delta_{\sigma,\sigma'}.$$  \hspace{1cm} (1)

Eigenvalue is a Lagrange multiplier needed for the constraint

$$(-\frac{1}{2} \nabla^2 + V_{KS}^\sigma(r), -\varepsilon_i^\sigma)\psi_i^\sigma(r) = 0 \quad (3)$$

$$V_{KS}^\sigma(r) = V_{ext}(r) + \frac{\delta E_{Hartree}}{\delta n(r, \sigma)} + \frac{\delta E_{xc}}{\delta n(r, \sigma)}$$

$$= V_{ext}(r) + V_{Hartree}(r) + V_{xc}^\sigma(r). \quad (4)$$
Another view of the Kohn-Sham problem

- Consider the Kohn-Sham energy to be a functional of the set of occupied wavefunctions and the positions of the nuclei

\[ E_{KS}[\{\psi_i\}, \{R_I\}] \]

\[
E[\{\psi_i\}, \{R_I\}] = 2 \sum_{i=1}^{N} \int \psi_i^*(r) \left( -\frac{1}{2} \nabla^2 \right) \psi_i(r) \, dr + U[n] + E_{II}[\{R_I\}], \quad (18.3)
\]

\[
U[n] = \int dr V_{ext}(r)n(r) + \frac{1}{2} \int \int drdr' \frac{n(r)n(r')}{|r - r'|} + E_{xc}[n], \quad (18.4)
\]

\[
n(r) = 2 \sum_{i=1}^{N} |\psi_i(r)|^2, \quad (18.5)
\]

\[
F_I = -\frac{\partial E}{\partial R_I}, \quad (18.6)
\]
Another view of the Kohn-Sham problem

- A problem in minimization - can apply well-known numerical methods to minimize the Kohn-Sham energy with respect to the occupied wavefunctions $E_{KS}[^{\psi_i}, \{R_i\}]$

$$E_{KS} = \frac{1}{2} \sum_\sigma \sum_{i=1} |\nabla \psi_i^\sigma|^2 + \int dr V_{ext}(r)n(r) + E_{Hartree}[n] + E_{II} + E_{xc}[n].$$

Constraint – required

Exclusion principle for independent particles

Gradient gives the “force”
That pushes the states $\psi_i$
Toward minimum energy

There is no eigenvalue equation! This is done differently.

Note that one needs only the occupied $\psi_i$ to determine density, energy, . . .

R. Martin - Iterative methods - AIMS - Capetown - 2008
Another view of the Kohn-Sham problem

- We have the ingredients for minimization methods:
  - Energy \( E_{KS}[\{\psi_i\}, \{R_i\}] \)
  - Gradient \( \frac{\delta E_{KS}}{\delta \psi_i^*(r)} \)
  - Constraint
    \[ \langle \psi_i^\sigma | \psi_j^\sigma' \rangle = \delta_{i,j} \delta_{\sigma,\sigma'} \]

- What is \( \frac{\delta E}{\delta \psi_i} \)?
- We have done this already:
  \[ -\frac{\delta E}{\delta \psi_i^*(r)} = -H \psi_i(r, t) \]
  \[ \hat{H}_{\text{eff}} \psi_i^\sigma(r) = \left[ -\frac{\hbar^2}{2m_e} \nabla^2 + V_{\text{eff}}^\sigma(r) \right] \psi_i^\sigma(r) \]

- The force on the electrons is \(-H\psi\)

- This leads to iterative methods in which the energy is minimized by changing the wavefunctions along the direction \(-H\psi\)
The operation $\mathbf{H}\psi$

- In general $\mathbf{H}\psi$ is a (matrix x vector) operation
  - $\mathbf{H}$ is ($\#$ plane waves) x ($\#$ plane waves)
  - There is one $\psi$ for each band
    (one needs only one band per electron for the density, energy, . . .)
  - Each $\psi$ is a vector of size ($\#$ plane waves)

\[
\psi_{i,k}(\mathbf{r}) = \sum_{m} c_{i,m}(\mathbf{k}) \times \frac{1}{\sqrt{\Omega}} \exp(i(\mathbf{k} + \mathbf{G}_m) \cdot \mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r}) \frac{1}{\sqrt{N_{\text{cell}}}} u_{i,k}(\mathbf{r}), \quad (12.11)
\]

where $\Omega = N_{\text{cell}} \Omega_{\text{cell}}$ and

\[
u_{i,k}(\mathbf{r}) = \frac{1}{\sqrt{\Omega_{\text{cell}}}} \sum_{m} c_{i,m}(\mathbf{k}) \exp(i\mathbf{G}_m \cdot \mathbf{r}), \quad (12.12)
\]

The wavefunction is defined by the coefficients $c_{i,m}$ for each $k$

Thus the problem is optimization (minimization) of a functional of ($\#$ bands) x ($\#$ plane waves) parameters $c_{i,m}$ for each $k$
Efficient calculation of $H\psi$

- If $H$ is sparse, then many terms are zero and one can reduce the complexity.
- How can this be done with plane waves?
- $H = -\frac{1}{2} \nabla^2 + V$
- The kinetic energy is sparse in Fourier space
  \[ \nabla^2 \exp(\imath \mathbf{q} \cdot \mathbf{r}) = -q^2 \exp(\imath \mathbf{q} \cdot \mathbf{r}) \]
- The potential energy is sparse in real space
  \[ V \exp(\imath \mathbf{q} \cdot \mathbf{r}) = V(r) \exp(\imath \mathbf{q} \cdot \mathbf{r}) \]
- Each is a vector operation

Solution:
Use Fourier space for KE, Real space for $V$
Fast Fourier Transform between the spaces!
Fourier Transforms and calculation of $H\psi$

- Operate with $H$ in the two spaces and sum the result to find $H\psi$ in either basis.

- Each FFT takes time $\sim N_{PW} \ln(N_{PW})$.

Much faster than $N_{PW}^2$ for large systems.
Minimization methods using the gradient

• Equation of motion
  – The original Car-Parrinello method
  – Fictitious dynamics

• Well known numerical methods
  – Lanczos, residual minimization, . . .
  – Steepest descent
  – Conjugate gradient
Car-Parrinello fictitious dynamics

- The Car-Parrinello simultaneous equations for the nuclei and electrons -- fictitious dynamics for the electrons

\[ \mu \ddot{\psi}_i(r, t) = -\frac{\delta E}{\delta \psi_i^*(r)} + \sum_k \Lambda_{ik} \psi_k(r, t) \]

\[ = -H \psi_i(r, t) + \sum_k \Lambda_{ik} \psi_k(r, t), \]

(18.8)

\[ M_I \ddot{R}_I = F_I = -\frac{\partial E}{\partial R_I}. \]

(18.9)

Holonomic force of constraint to maintain orthogonality
Damped Equation of motion to quench to the minimum energy

- Quenching electrons in Si to the ground state (from the original Car-Parrinello paper)

![Graphs showing total energy, lattice constant, and eigenvectors over time](image)

**Figure 18.1.** Eigenvalues at \( k = 0 \) for crystalline Si calculated by quenching the “fictitious kinetic energy” in the lagrangian to reach the steady state [156].
Iterative solution for eigenstates

We first consider the problem of solving the Schrödinger equation for a fixed hamiltonian

\[(H - \varepsilon)|\psi\rangle = 0.\]  \hspace{1cm} (M.1)

**M.2 Simple relaxation algorithms**

The algorithm [939] proposed by Jacobi in 1848 is in many ways the grandfather of iterative eigenvalue methods. The basic idea is to iterate a from of the equation

\[(H - \varepsilon^n)|\psi^n\rangle = |R[\psi^n]\rangle,\]  \hspace{1cm} (M.2)

where \(n\) is the iteration step, \(|\psi^n\rangle\) and \(\varepsilon^n\) are approximate eigenvectors and eigenvalues, and \(|R[\psi^n]\rangle\) is a “residual” vector. The iterations continue with a particular choice of the improved eigenvector \(|\psi^{n+1}\rangle\) and eigenvalue \(\varepsilon^{n+1}\) until the eigenvalue is converged or the norm of the residual vanishes to within some tolerance [930].

**The general idea in all the methods is to reduce the residual R until one reaches the solution when R =0**

**Methods:**  Jacobi, RMM – DIIS (Pulay), Newton, Davidson, . . .
Example - iterative relaxation method

If the matrix is diagonally dominant (as is the case for the hamiltonian expressed in the bases most commonly chosen in electronic structure calculations) then we can rewrite the eigenvalue problem, Eq. (M.1), as

$$|\psi\rangle = D^{-1}(H - \varepsilon)|\psi\rangle + |\psi\rangle,$$  \hspace{1cm} (M.3)

where $D$ is a non-singular matrix. This form suggests many variations and the choice of $D$ can be viewed as a “preconditioning” of the hamiltonian operator, as discussed below. If we define the iteration sequence [930]

$$\varepsilon^n = \frac{\langle\psi^n|H|\psi^n\rangle}{\langle\psi^n|\psi^n\rangle},$$

$$\delta \psi^{n+1} = D^{-1}(H - \varepsilon^n)\psi^n,$$

$$\psi^{n+1} = \psi^n + \delta \psi^{n+1},$$  \hspace{1cm} (M.4)

then the middle equation of (M.4) is just the linear set of equations

$$D\delta \psi^{n+1} = R^n \quad \text{or} \quad \psi^{n+1} = D^{-1}R^n \equiv KR^n,$$  \hspace{1cm} (M.5)

where $R^n$ is the residual at step $n$ and $K \equiv D^{-1}$. 

R. Martin - Iterative methods - AIMS - Capetown - 2008 21
Krylov Subspaces

The general approach of many methods can be described as generating a sequence of vectors $\psi^n$

An interactive subspace or Krylov space $[\psi^0, \psi^1, \psi^2, \ldots]$.

Different methods take into account different numbers of vectors:
- Use only two vectors in the iteration $[\psi^{n-1}, \psi^n]$:
  - Simple relation, Steepest descent, Conjugate Gradient
- A larger number $N$ states—diagonalizing a $N \times N$ matrix to get the best solution
  - RMM-DIIS

Many methods
Advantages of residual minimization methods

• One can set up the problem to get any eigenvector
  • Example – get eigenvector for state with energy closest to a given energy -- without explicitly requiring it to be orthogonal to the others
  • Orthogonalization is automatically satisfied if the states are eigenstates

• What are advantages? - Examples
  • In the self consistent calculation one needs all the states
  • On a parallel computer each state can be calculated independently on separate processors!
  • In many cases one cares only about some states – e.g., the states near the Fermi energy
  • One can find the band gap knowing only the two states!
    (After one has done the full self consistent calculation)

The basic algorithm used in VASP
Steepest Descent

- Problem: minimization of a function $F(x_i)$ in space of variables $x_i, i = 1,N$
- In the absence of any other information the best choice for a direction of displacement from a point $x_i^0$ to reach the minimum is the steepest descent (SD) direction:
  \[ g_i^0 = -\frac{\partial F}{\partial x_i} \bigg|_{x_i=x_i^0} \]
- The lowest energy along this direction can be found by “line minimization” in the one-dimensional space, i.e., the minimum of $F$ as a function of $\alpha^i$ where
  \[ x_i^1 = x_i^0 + \alpha^i g_i^0 \]
- the function $F$ decreases at each step, but the steps do not move directly to the minimum.
Conjugate Gradient

- Faster way to reach the minimum than to always follow the “downhill” steepest descent direction.
- After the first step, one has not only on the gradient of $F$ at the present point, but also the value and gradient at previous points. The additional information can be used to choose a more optimal direction along which the line minimization will lead to a lower energy.
- It can be demonstrated that for a quadratic functional in N dimensions, the conjugate gradient (CG) method reaches the minimum in N steps.
- Consider the quadratic functional

$$F(\{x_i\}) \equiv F(x) = \frac{1}{2} x \cdot H \cdot x$$

with gradients

$$g = -\frac{\partial F}{\partial x} = -H \cdot x$$
Conjugate Gradient

- The first step is the same as steepest-descent, i.e., minimization of F along a line $x^1 = x^0 + \alpha^1 d^0 d_0 = g_0$.
- For this and for all steps the minimum occurs for $d^n \cdot g(x^{n+1}) = 0$
- For the $n+1$ step the best choice is to move in a direction where the gradient along the previous direction $d^n$ remains zero. Since the change in gradient as we move in the new direction $d^{n+1}$ is $\Delta g = \alpha^{n+1} H \cdot d^{n+1}$ the desired condition is satisfied if

$$d^n \cdot H \cdot d^{n+1} = 0$$

that defines the “conjugate direction”
Conjugate Gradient

- The key point is that (unlike steepest-descent) each line minimization preserves the minimization done all the previous steps and only adds independent (i.e., conjugate) variations.
- This is manifested in the fact that (unlike the steepest-descent method that never reaches the minimum) for a quadratic functional the conjugate gradient method reaches the minimum exactly in $N$ steps where $N$ is the dimension of the space $x_i$, $i = 1, N$. 
Preconditioning

- The basic idea behind “preconditioning” is to modify the functional dependence upon the variables to be more “isotropic”, i.e., to make the curvature more similar for the different variables of the N-dimensional problem (basically, it makes all possible effort so that all the variables reach their minimum at the same “time”, or with the same “speed”)
- The energy minimization of the K-S problem is usually very badly conditioned, but it is also the perfect example of how one can precondition the problem in terms of the physical nature of the problem itself.
- In the energy minimization, the variables of the problems are the coefficient of the plane wave expansion of the wavefunctions, $c_{i,m}^n(k)$
- Because high Fourier components $|k + G_m|$ have high kinetic energy, the total energy varies much more rapidly as a function of the coefficient $c_{i,m}^n(k)$ with large $|G_m|$ than for coefficients with small $|G_m|$
- Preconditioning can be used to modify the gradients and cancel this effect (see for example: Teter, Payne and Allan, *Phys. Rev. B* 40, 12255 (1989)
The eigenstate for band \( n \) is found by minimizing energy subject to constraint that eigenfunction is orthogonal to all states \( m < n \).
Example of convergence with various methods

FIG. 2. The total energy error in eV/(unit cell) of an 8-atom silicon cube with a 16-hartree kinetic-energy cutoff vs iteration number for various methods. The scales are different for the different methods. The number of iterations of the original Williams-Soler method has been divided by 5 to allow comparison at the same level of computational effort.

FIG. 3. The total-energy error in eV of a row of 12 silicon unit cells with a 4-hartree kinetic-energy cutoff vs iteration number for various methods. The methods are scaled as in Fig. 2.


The basic algorithm used in ABINIT and CASTEP

R. Martin - Iterative methods - AIMS - Capetown - 2008
“Band by Band” minimization

• One issue is updating the electron density

• A well-known problem is how to best choose a new density in a self consistent loop – the “old” density at the last step, the “new” density at the current step, or a combination.
• Usually the “new” density “overshoots” and it’s better to use a mixture

• The band by band method provides an interesting variation

• One can update the density after changing only one band – a small change that avoids the overshoot

• There are many variations –
   See discussion in book “Electronic Structure”
Other ideas for efficiency

• Kleinman-Bylander separable form for pseudopotential
  – Much faster in iterative methods since it is a vector inner product operation

• Not particularly useful in matrix diagonalization methods – Why?

• Why is it more efficient in iterative codes?

• Matrix diagonalization methods
  – Setting up square matrix $H_{m,m}$ scales as $(\#\text{PWs})^2$
  – Diagonalizing scales as $(\#\text{PWs})^3$
  – Setting up matrix is faster with separable form
  – But diagonalizing dominates
Nonlocal pseudopotential in the operation $H\psi$

From before

- $H = -\frac{1}{2} \nabla^2 + V$
- The kinetic energy is sparse in Fourier space
  \[ \nabla^2 \exp(i \mathbf{q} \cdot \mathbf{r}) = -q^2 \exp(i \mathbf{q} \cdot \mathbf{r}) \]
- The potential energy is sparse in real space
  \[ V \exp(i \mathbf{q} \cdot \mathbf{r}) = V(r) \exp(i \mathbf{q} \cdot \mathbf{r}) \]
- Each is a vector operation

Note: This assumes a local potential $V(r)$
For a nonlocal potential one needs a matrix operation

\[ V_{nl} \exp(i \mathbf{q} \cdot \mathbf{r}) = \sum \mathbf{r'}, V(r, r') \exp(i \mathbf{q} \cdot \mathbf{r'}) \]

Faster to use a separable form – product of a few vector operations

\[ V_{nl, sep} \exp(i \mathbf{q} \cdot \mathbf{r}) = \sum_i F_i(r) \sum_r G_i(r') \exp(i \mathbf{q} \cdot \mathbf{r'}) \]
Other points not covered

• **Self-consistency – efficient iterative methods**
  – Formally similar to iterations for wavefunction
  – **But very different since there are many fewer variables** (one density instead of many wavefunctions)
  – One can use more expensive approaches for the smaller problem
  – Broyden scheme to generate the Hessian (second derivatives)

• **Moving the atoms**
  – Similar to above – many fewer degrees of freedom for the atoms than for the electrons

• **The original Car-Parrinello method for combined nucleus – electron problem**
Conclusions

• Plane waves solution for electronic states in crystals
  – Hamiltonian matrix given simply in terms involving Fourier components of
    the potential

• Efficient iterative methods
  – Well developed standard methods
  – Difference in this case from many applications – need many eigenvectors that
    are orthonormal

• Clever ideas!
  – New approaches, algorithms, . . .

• Well developed in plane waves because of the simplicity of the operations
  – Uniform grid operations

• Key - FFTs are fast!
  – FFTs are widely used – algorithms are highly developed -

• Most of the time is spent in FFTs!
  – One can take advantage of improvements in FFTs optimized for any given
    computer