CZ4275: Computational Condensed Matter Physics

Lecture given by
A/P Wang Jian-Sheng

Typeset by Yang Shiyang
June 1999

Contents

Chapter 1

Fast Fourier Transform (FFT)

1.1 Introduction to FFT

Why use FFT?

- For periodical system, wave function is best represented by Fourier series

- The convolution of the form $\int f(x)g(y-x)dx$ can be done in time of order Nlog N.

1.2 Discrete Fourier Transform (DFT)

1.2.1 General definition

The standart definition of DFT is as:

$$y_k = \sum_{j=0}^{N-1} x_j e^{-i\frac{2\pi jk}{N}} \tag{1.1}$$

Or, in the matrix form, equation 1.1 can be written as:

$$\mathbf{y} = F_n \mathbf{x} \tag{1.2}$$

where $\mathbf{x} = \begin{pmatrix} x_0 \\ x_1 \\ \dots \\ x_{n-1} \end{pmatrix}$ and $y = 7 \begin{pmatrix} y_0 \\ y_1 \\ \dots \\ y_{n-1} \end{pmatrix}$. F_n is an n⊗n matrix with $(F_n)_{pq} = \omega_n^{pq} = e^{-i2\pi \frac{pq}{n}}$, $\omega_n = e^{-i2\pi/n}$. Notice that $\omega_n^2 = \omega_{n/2}$, $\omega_n^n = 1$, $\omega_n^{n/2} = $ -1.

1.2.2 Some properties of F_n

- (1) F is symmetric.
  This is easily proven by the definition of F.

- (2) $F^H F = nI$ where $F^H = (F^T)^*$ is the Hermitian Complex Conjugate of F.
  $$(F^H F)_{pq} = \sum_{r=0}^{n-1} F_{pr}^H F_{qr} = \omega_n^{-pr+qr} = \begin{cases} 0 & p \neq q \\ n & p = q \end{cases}$$

- (3) $Q = \frac{F}{\sqrt{n}}$ is unitary.
  $Q^H Q = \frac{F^H}{\sqrt{n}} \times \frac{F}{\sqrt{n}} = \frac{nI}{n} = I$

- (4) $F^{-1} = \frac{F^*}{n}$.
  Use properties (1) and (2), $F^{-1} = \frac{F^H}{n} = \frac{F^*}{n}$

Property (4) implies $\mathbf{x} = F^{-1} \mathbf{y}$, or $x_j = \frac{1}{n} \sum_{k=0}^{n-1} y_k e^{i \frac{2\pi jk}{n}}$.

There is no need to write two seperate functions for FFT and inverse FFT. One can simply change the sign in the exponent. Also, the factor $1/n$ is done by the user. To check your program, you can do the following:

Start from $\mathbf{x}$, $F\mathbf{x} = \mathbf{y}$, $\frac{F^*}{n} \mathbf{y} = \mathbf{x}'$, then $\frac{||x - x'||}{||x||} < \epsilon$.

1.2.3 Examples of F_n

$$F_1 = [1], \quad F_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad F_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix}.$$

1.2.4 FFT idea

Start from an analogue: If we compute $b = a^n = \underbrace{a \cdot a \cdots a}_{n}$ by simple repeated multiplications, the time of running is of order O(n). However, if we use the following procedure:

```
b = a_i
for p = 1:log_2 n
{     b=b*b;     }
end
```

It will only take O(log n) time, assuming n is a power of 2. This is exactly the spirit of FFT.

1.3 Cooley-Tukey Factorization

1.3.1 Outer-product factorization

The idea is to factor F_n into products of smaller matrices.

$$F_n P_n = (I_1 \otimes B_n)\,(I_2 \otimes B_{n/2})\,\dots\,(I_{n/2} \otimes B_2)$$

P_n is an n×n permutation operator. $I_k = \begin{pmatrix} 1 & & & & 0 \\ & 1 & & & \\ & & . & . & . \\ 0 & & & & 1 \end{pmatrix}\underbrace{\qquad}_{k}$ B_k is an k×k matrix to be

specified. $A \otimes B$ is the outer product of p×q matrix A and m×n matrix B, the product is pm×qn matrix. The rule of the outer product is:

$$A \otimes B = \begin{bmatrix} A_{0,0}B & A_{0,1}B & \dots & A_{0,q-1}B \\ \dots & & & \\ A_{p-1,0}B & A_{p-1,1}B & \dots & A_{p-1,q-1}B \end{bmatrix}$$

Now consider:

$$
\begin{aligned}
y_k &= \sum_{j=0}^{n-1} x_j \omega_n^{kj} = \sum_{j\ even} x_j \omega_n^{kj} + \sum_{j\ odd} x_j \omega_n^{kj} \\
&= x_0 \omega_n^{0\cdot k} + x_2 \omega_n^{2\cdot k} + x_4 \omega_n^{4\cdot k} + \dots + x_1 \omega_n^{1\cdot k} + x_3 \omega_n^{3\cdot k} + x_5 \omega_n^{5\cdot k} + \dots \\
&= (x_0 \omega_{n/2}^{0\cdot k} + x_2 \omega_{n/2}^{1\cdot k} + x_4 \omega_{n/2}^{2\cdot k} + \dots) + \omega_n^k (x_1 \omega_{n/2}^{0\cdot k} + x_3 \omega_{n/2}^{1\cdot k} + x_5 \omega_{n/2}^{2\cdot k} + \dots) \\
&= kth\ component\ of\ the\ Fourier\ Transform\{x(0:2:n-1)\} \\
&+ \omega_n^k (\ kth\ component\ of\ the\ Fourier\ Transform\{x(1:2:n-1)\}\ ) \quad (1.3)
\end{aligned}
$$

Or, in the matrix form,

$$\mathbf{y} = \begin{bmatrix} F_{n/2} & \Omega_{n/2}F_{n/2} \\ F_{n/2} & -\Omega_{n/2}F_{n/2} \end{bmatrix} P_n \mathbf{x} \qquad (1.4)$$

where $\Omega_{n/2} = \begin{pmatrix} 1 & 0 & 0 & . & 0 \\ 0 & \omega_n & 0 & . & 0 \\ 0 & 0 & \omega_n^2 & . & 0 \\ \dots & & & & \\ 0 & 0 & 0 & . & \omega_n^{n/2-1} \end{pmatrix}.$

Proof of the equavilence of (1.3) and (1.4). Start from (1.4):

$$
\begin{aligned}
y_k &= \sum_{j=0}^{n/2-1} (F_{n/2})_{kj} x_{2j} + \sum_{j=0}^{n/2-1} (\Omega_{n/2})_{kk}(F_{n/2})_{kj} x_{2j+1} \\
&= \sum_{j=0}^{n/2-1} \omega_{n/2}^{kj} x_{2j} + \sum_{j=0}^{n/2-1} \omega_n^k \omega_{n/2}^{kj} x_{2j+1} = (1.3)
\end{aligned}
$$

1.3.2 Permutation operator P_n

The permutation operator P_n is defined as: $P_n \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \ldots \end{bmatrix} = \begin{bmatrix} x_0 \\ x_2 \\ x_4 \\ \ldots \\ x_1 \\ x_3 \\ x_5 \\ \ldots \end{bmatrix}$

Notation $x(i{:}k{:}n) = \begin{bmatrix} x_i \\ x_i + k \\ x_i + 2k \\ \ldots \end{bmatrix}$

We see that:

$$F_n = \begin{bmatrix} I_{n/2} & \Omega_{n/2} \\ I_{n/2} & -\Omega_{n/2} \end{bmatrix} [I_2 \otimes F_{n/2}] \, P_n \tag{1.5}$$

Proof of (1.5):
Start from (1.4)

$$\begin{aligned}
\mathbf{y} &= \begin{bmatrix} F_{n/2} & \Omega_{n/2}F_{n/2} \\ F_{n/2} & -\Omega_{n/2}F_{n/2} \end{bmatrix} P_n\mathbf{x} \\
&= \begin{bmatrix} I_{n/2} & \Omega_{n/2} \\ I_{n/2} & -\Omega_{n/2} \end{bmatrix} \begin{bmatrix} F_{n/2} & 0 \\ 0 & F_{n/2} \end{bmatrix} \\
&= \begin{bmatrix} I_{n/2} & \Omega_{n/2} \\ I_{n/2} & -\Omega_{n/2} \end{bmatrix} [I_2 \otimes F_{n/2}]P_n\mathbf{x} \\
\text{Compare with } (1.2) \\
F_n &= \begin{bmatrix} I_{n/2} & \Omega_{n/2} \\ I_{n/2} & -\Omega_{n/2} \end{bmatrix} [I_2 \otimes F_{n/2}] \, P_n
\end{aligned}$$

1.3.3 Example: n=16

$F_{16}\mathbf{x}$ is a combination of $F_8\mathbf{x}(0{:}2{:}15)$ and $F_8\mathbf{x}(1{:}2{:}15)$. Let $y^A = F_8\mathbf{x}(0:2:15), y^B = F_8\mathbf{x}(1:2:15)$. Then equation (1.3) gives:

$$\begin{aligned}
y_{(k<n/2)} &= \textit{kth component of the Fourier Transform}\{x(0:2:n-1)\} \\
&+ \; \omega_n^k(\textit{ kth component of the Fourier Transform}\{x(1:2:n-1)\} \,) \\
&= y_k^A + \omega_n^k y_k^B \\
y_{(k+n/2)} &= y_k^A + \omega_n^{k+n/2} y_k^B \\
&= y_k^A - \omega_n^k y_k^B \tag{1.6}
\end{aligned}$$

$$y_0 = y_0^A + y_0^B$$
$$y_1 = y_1^A + \omega_n y_1^B$$
$$y_2 = y_1^A + \omega_n^2 y_2^B$$
$$\dots$$
$$y_{n/2-1} = y_{n/2-1}^A + \omega_n^{n/2-1} y_{n/2-1}^B$$

$\left.\right\}$ First half of the vector

$$y_{n/2} = y_{n/2}^A + y_{n/2}^B$$
$$y_{n/2+2} = y_{n/2+2}^A + \omega_n^2 y_{n/2+2}^B$$
$$\dots$$
$$y_{n-1} = y_{n-1}^A + \omega_n^{n/2-1} y_{n-1}^B$$

$\left.\right\}$ Second half of the vector

The decomposition is done recursively as:

$$[0:8:15] \Longrightarrow \{ \begin{matrix} 0 \\ 8 \end{matrix}$$
$$[0:4:15] \Longrightarrow$$
$$[4:8:15] \Longrightarrow \{ \begin{matrix} 4 \\ 12 \end{matrix}$$
$$[0:2:15] \Longrightarrow$$
$$[2:8:15] \Longrightarrow \{ \begin{matrix} 2 \\ 10 \end{matrix}$$
$$[2:4:15] \Longrightarrow$$
$$[6:8:15] \Longrightarrow \{ \begin{matrix} 6 \\ 14 \end{matrix}$$
$$[0:1:15] \Longrightarrow$$
$$[1:8:15] \Longrightarrow \{ \begin{matrix} 1 \\ 9 \end{matrix}$$
$$[1:4:15] \Longrightarrow$$
$$[5:8:15] \Longrightarrow \{ \begin{matrix} 5 \\ 13 \end{matrix}$$
$$[1:2:15] \Longrightarrow$$
$$[3:8:15] \Longrightarrow \{ \begin{matrix} 3 \\ 11 \end{matrix}$$
$$[3:4:15] \Longrightarrow$$
$$[7:8:15] \Longrightarrow \{ \begin{matrix} 7 \\ 15 \end{matrix}$$

1.3.4 FFT recursive radix-2 procedure

```
function y= fft(x,n)
 if n=1
      y   <- x
 else
      m   <- n/2
      ω   <- e^{-i2π/n}
      Ω   <- diag(1, ω, ω², ..., ω^{n-1})
      Z_A <- fft( x(0:2:n-1),m )
      Z_B <- Ω· fft( x(1:2:n-1),m )
      y <-  [ I_m   I_m  ] [ Z_A ]
            [ I_m  -I_m  ] [ Z_B ]
  end
end function
```

Theorem 1.3.3 (Cooley-Tukey factorization)
If $n = 2^t$, then $F_n = A_t A_{t-1}...A_1 P_n^t$, where P_n is an index reverse permutation.
$A_q = I_r \otimes B_L, L = 2^q, r = n/L.$
$$B_L = \begin{bmatrix} I_{L^*} & \Omega_{L^*} \\ I_{L^*} & -\Omega_{L^*} \end{bmatrix}, L^* = L/2.$$
$\Omega_{L^*} = diag(1, \omega_L, \omega_L^2, ...\omega_L^{L^*-1}), \omega_L = e^{-i2\pi/L}.$

Cooley-Tukey algorithm
x <- P_n^T **x**
for q=1:t
    x <- A_q **x**
end

Bit reversal permutation

k	after permutation (reversal)			**x**
0000	0000	x_0	->	x_0
0001	1000	x_1	->	x_8
0010	0100	x_2	->	x_4
0011	1100	x_3	->	x_{12}
0100	0010	x_4	->	x_2
0101	1010	x_5	->	x_{10}
0110	0110	x_6	->	x_6
0111	1110	x_7	->	x_{14}
1000	0001	x_8	->	x_1
...				
1111	1111	x_{15}	->	x_{15}

1.3.5   Cooley-Tukey bit reversal in-place FFT algorithm (1965)

Algorithm 1.6.1
If $\mathbf{x} \in C^n, n = 2^t$, then the following algorithm overwrites **x** with $F_n\mathbf{x}$.
$\mathbf{x} \leftarrow P_n\mathbf{x}$
for q=1:t
    $L \leftarrow 2^q; r \leftarrow n/L; L^* \leftarrow L/2$
    for j=0:L^*-1
        $\omega = e^{-i2\pi j/L}$
        for k=0:r-1
            $\tau = \omega \cdot \mathbf{x}(k \cdot L + j + L^*)$
            $\mathbf{x}(k \cdot L + j + L^*) \leftarrow \mathbf{x}(k \cdot L + j) - \tau$
            $\mathbf{x}(k \cdot L + j) \leftarrow \mathbf{x}(k \cdot L + j) + \tau$
        end
    end
end

1.4   Stockham Auto-sorting algorithm

Algorithm 1.7.2
If $\mathbf{x} \in C^n, n = 2^t$, then the following algorithm overwrites $\mathbf{x}$ with $F_n\mathbf{x}$. A workspace $\mathbf{y} \in C^n$ is required.

for q=1:t
       $L \leftarrow 2^q; r \leftarrow n/L; L^* \leftarrow L/2; r^* \leftarrow n/L^*$
       $\mathbf{y} \leftarrow \mathbf{x}$
       for j=0:L^*-1
              $\omega = e^{-i2\pi j/L}$
              for k=0:r-1
                     $\tau = \omega \cdot \mathbf{y}(j \cdot r^* + k + r)$
                     $\mathbf{x}(j \cdot r + k) \leftarrow \mathbf{y}(j \cdot r^* + k) + \tau$
                     $\mathbf{x}((j + L^*) \cdot r + k) \leftarrow \mathbf{y}(j \cdot r^* + k) - \tau$
              end
       end
end

1.5   3-dimensional FFT

The definition:

$$\mathbf{y}_{\hat{i},\hat{j},\hat{k}} = \sum_{\hat{l},\hat{m},\hat{n}} \mathbf{x}_{\hat{l},\hat{m},\hat{n}} e^{-i\frac{2\pi\hat{i}\hat{l}}{n_1}} e^{-i\frac{2\pi\hat{j}\hat{m}}{n_2}} e^{-i\frac{2\pi\hat{k}\hat{n}}{n_3}}$$

We can think of 3D FFT as applying F_{n_1} to the first index $\hat{i}$, F_{n_2} to the second index $\hat{j}$, F_{n_3} to the third index $\hat{k}$ separately. Here are some issues concerning the programming efficiency.

- Uni-stride:
  Memory should be accessed sequentially, rather than jumps randomly.

- Cooley-Tukey:
  It uses minimum memory (in-place) by bit permutaiton, yet it is slow.

- Stockham:
  It does not do permutation, but uses extra memory of the size of array $\mathbf{x}$.

1.6   Lab 1

Implement a 1-dimensional and 3-dimensional FFT in C-code. It must be efficient (fast) and correct with the necessary demonstration. Time the program for 3D FFT for size $16^3, 32^3, 64^3$ in units of seconds.

1.7 Tutorial 1

Hand simulate the FFT algorithms, and compare the results with the correct answer (either by Matlab or Mathematica).

1.8 Reading Materials

- *Computational frameworks for the Fast Fourier Transform*
  There are detailed proof of all the relevant theorem in FFT.

- *Numerical recipes in C*
  It provides a quick working knowledge of FFT.

Chapter 2

Density Functional Calculation of Energy

2.1 How to solve the problem of many-body system?

2.1.1 Quantum mechanics of many-electron system

The Hamiltonian of a system of N electrons, and M nuclei each of charge Z_i is:

$$
\begin{aligned}
\mathcal{H} &= \sum_{i=1}^{N} \frac{-\hbar^2}{2m_i} \nabla_i^2 + \sum_{(i<j)=1}^{N} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i=1}^{N}\sum_{k=1}^{M} \frac{Z_k e^2}{|\mathbf{r}_i - \mathbf{R}_k|} + \sum_{k=1}^{M} \frac{-\hbar^2}{2m_k} \nabla_k^2 + \sum_{(k<l)=1}^{M} \frac{Z_k Z_l e^2}{|\mathbf{R}_k - \mathbf{R}_l|} \\
&= KE(e^-) + V(e^- - e^-) + V(e^- - n) + KE(n) + V(n-n) \quad\quad (2.1)
\end{aligned}
$$

2.1.2 Born-Oppenheimer approximation

Assume nuclei is much heavier than electron, such that nuclei don't move when we solve for the motion of electrons. We thus treat nuclei classically, while treating electrons quantum mechanically. Thus $\mathbf{R}$ is taken as a parameter instead of a variable, i.e. $\psi(\mathbf{r}_i, \mathbf{R}_k) \to \psi_{\mathbf{R}_k}(\mathbf{r}_i)$.

We can solve for the motion of electrons first, given the Hamiltonian $\mathcal{H}$ for the electrons:

$$
\mathcal{H} = \sum_{i=1}^{N} \frac{-\hbar^2}{2m_i} \nabla_i^2 + \sum_{(i<j)=1}^{N} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i=1}^{N} V(r_i) \quad\quad (2.2)
$$

where $V(\mathbf{r}_i)$ is the potential due to the nuclei.

The time-independent Schrodinger equation for the electrons is:

$$
\mathcal{H}\psi = \epsilon\psi \quad\quad (2.3)
$$

Then we solve for the motion of nuclei classically, using Newton's laws or Hamiltonian equations instead of Schrodinger equations.

2.1.3   Idea of density function

$|\psi(\mathbf{r}_1, \mathbf{r}_2, ...\mathbf{r}_N)|^2 d\mathbf{r}_1, d\mathbf{r}_2, ...d\mathbf{r}_N$ is the probability that electrons r are located at $\mathbf{r}_r$ (r=1,2,...N). Thus, the probability P_i that electron i at at $\mathbf{r}$ is:

$$P_i = \int |\psi(\mathbf{r}_1, ...\mathbf{r}_{i-1}, \mathbf{r}, \mathbf{r}_{i+1}, ...\mathbf{r}_N)|^2 d\mathbf{r}_1, ...d\mathbf{r}_{i-1}, d\mathbf{r}_{i+1}, ...d\mathbf{r}_N \qquad (2.4)$$

Since P_i is same for all permutation of ψ_j, total probability P of finding one electron at $\mathbf{r}$ is:

$$P = NP_i = N \int |\psi(\mathbf{r}_1, ...\mathbf{r}_{i-1}, \mathbf{r}, \mathbf{r}_{i+1}, ...\mathbf{r}_N)|^2 d\mathbf{r}_1, ...d\mathbf{r}_{i-1}, d\mathbf{r}_{i+1}, ...d\mathbf{r}_N \qquad (2.5)$$

Since electrons are fermions, we should take $\mathbf{r}$ as $\mathbf{x} = (\mathbf{r}, \mathbf{S})$, where spin $\mathbf{S} = \pm\frac{1}{2}$. Wave functions for fermions must be antisymmetric with respect to permutation of $\mathbf{x}$, i.e.

$$\mathbf{P}\psi(\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_N) = (-1)^P \psi(\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_N) \qquad (2.6)$$

where P is the parity of the permutation. P=1 for old permutation, and P=0 for even permutation.

2.1.4   Variational principle for ground state energy

Define

$$E[\psi] = \frac{<\psi|\mathcal{H}|\psi>}{<\psi|\psi>} \qquad (2.7)$$

E is a functional of the wave function ψ. Let E_0 be the ground state energy of $\mathcal{H}$, then:

$$E_0 \leq E[\psi] \qquad (2.8)$$

Above equation (2.8) is true for all normalized wave function ψ.

2.1.5   Theorem regarding solving the eigenvalue equation

Solving the eigenvalue equation is equivalent to solving the functional equation of E[ψ].

$$\delta E[\psi] = 0 \Leftrightarrow \mathcal{H}|\psi> = \epsilon|\psi> \qquad (2.9)$$

2.2 Hartree-Fock approximation

Assume that each electron is independent of others and each occupies the orbit $\psi(\mathbf{x})$. Then the total wave function is $\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)...\psi_N(\mathbf{x}_N)$. Since the wave function umst be antisymmetric, the actual total wave function is the linear combination of all the permutation functions for the above one, which is given by the Slater Determinant:

$$
\begin{aligned}
|\psi> \; &= \; \sum_P \mathbf{P} \; |\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)...\psi_N(\mathbf{x}_N) > \\
&= \; \frac{1}{\sqrt{N!}}
\begin{vmatrix}
\psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & ... & \psi_N(\mathbf{x}_1) \\
\psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & ... & \psi_N(\mathbf{x}_2) \\
... & & & \\
\psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & ... & \psi_N(\mathbf{x}_N)
\end{vmatrix}
\end{aligned}
\tag{2.10}
$$

More about the HF and reduced HF approximations are available in the appendix. For more reading materials, you can read the book *Density-Functional theory of atoms and molecules.*

2.3 Hohenberg-Kohn theorem

2.3.1 The first Hohenberg-Kohn theorem (1964)

Given the electron system with Hamiltonian:

$$
\mathcal{H} = \mathbf{T} + \sum_{(i<j)=1}^{N} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i=1}^{N} v(\mathbf{r}_i)
\tag{2.11}
$$

the ground state electron density $n_0(\mathbf{r})$ is determined uniquely by $\mathcal{H}$. If we fix the number of electrons N, the only thing that can change is v($\mathbf{r}$). So v($\mathbf{r}$) uniquely determines $n_0(\mathbf{r})$.

Proof:
Assume ground state ϵ_0 and ϵ_0' for two H and H '. According to the variational principle, we have:

$$
\mathcal{H}|\psi> \; = \; \epsilon_0|\psi>
$$

$$
\Rightarrow \epsilon_0 \; < \; <\psi'|\mathcal{H}|\psi'> \; = \; <\psi'|\mathcal{H}' + \mathcal{H} - \mathcal{H}'|\psi'> \; = \; \epsilon_0' + <\psi'|\sum_{i=1}^{N}(v(\mathbf{r}_i) - v(\mathbf{r}_i'))|\psi'>
$$

$$
\mathcal{H}'|\psi'> \; = \; \epsilon_0'|\psi'>
$$

$$
\Rightarrow \epsilon_0' \; < \; <\psi|\mathcal{H}'|\psi> \; = \; <\psi|\mathcal{H} + \mathcal{H}' - \mathcal{H}|\psi> \; = \; \epsilon_0 + <\psi|\sum_{i=1}^{N}(v(\mathbf{r}_i') - v(\mathbf{r}_i))|\psi>
$$

$$
\Rightarrow \epsilon_0 + \epsilon_0' \; < \; \epsilon_0 + \epsilon_0'
$$

The contradiction implies that v($\mathbf{r}$) uniquely determines ϵ_0 and uniquely determines $\mathcal{H}$ up to a constant, thus $|\psi>$ is also uniquely determined, which in turn uniquely determines the density function $n_0(\mathbf{r})$.

2.3.2   The second Hohenberg-Kohn theorem

According to first Hohenberg-Kohn theorem, we can compute:

$$E_v[n(\mathbf{r})] = <\psi|\mathcal{H}|\psi> = <\psi|\mathbf{T}|\psi> + <\psi|V_{(e^-e^-)}|\psi> + \int v \cdot n(\mathbf{r})d^3\mathbf{r} \qquad (2.12)$$

The second Hohenberg-Kohn theorem states that:

$$E_0 \leq E_v[n(\mathbf{r})] \qquad (2.13)$$

for fixed potential v($\mathbf{r}$).

2.4   Kohn-Sham energy functional and Kohn-Sham equation

Hehenberg-Kohn density functional for electrons gives total energy:

$$
\begin{aligned}
E[n(\mathbf{r})] &= T[n(\mathbf{r})] + V_{(e^-e^-)}[n(\mathbf{r})] + \int v(\mathbf{r}) \cdot n(\mathbf{r})d\mathbf{r} \\
E[n] &= T[n] + V_{(e^-e^-)}[n] + \int v \cdot n d\mathbf{r} &(2.14) \\
\mathbf{T}[n] &= <\psi|\mathbf{T}|\psi> &(2.15) \\
&= \sum_{\mathbf{S}_i} \int d\mathbf{r}_1 d\mathbf{r}_2 ... d\mathbf{r}_N \ \psi^*(\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_N) \ (\sum_{i=1}^{N} \frac{-\hbar^2}{2m_i}\nabla_i^2) \ \psi(\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_N) \\
V_{(e^-e^-)}[n] &= <\psi|V_{(e^-e^-)}|\psi> \\
&= \sum_{\mathbf{S}_i} \int d\mathbf{r}_1 d\mathbf{r}_2 ... d\mathbf{r}_N \ \psi^*(\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_N) \ (\sum_{(i<j)=1}^{N} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}\psi(\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_N) \\
&= \sum_{\mathbf{S}_i} \int d\mathbf{r}_1 d\mathbf{r}_2 ... d\mathbf{r}_N \ |\psi(\mathbf{x}_1, \mathbf{x}_2, ...\mathbf{x}_N)|^2 \ (\sum_{(i<j)=1}^{N} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} &(2.16)
\end{aligned}
$$

Kohn-Sham approach consider N non-interacting electrons in an external potential v_s:

$$
\begin{aligned}
\mathcal{H}_s &= \mathbf{T}_s + \sum_{i=1}^{N} v(\mathbf{r}_i) \\
&= \sum_{i=1}^{N} \frac{-\hbar^2}{2m_i}\nabla_i^2 + \sum_{i=1}^{N} v_s(\mathbf{r}_i) &(2.17)
\end{aligned}
$$

Apply Hohenberg-Kohn theorem to above:

$$E_s[n] \;=\; T_s[n] + \int v_s \cdot n \; d\mathbf{r} \tag{2.18}$$

Since H_s is the Hamiltonian for N-independent electrons, we have:

$$H_s = \sum_{i=1}^{N} h_s(\mathbf{r}_i) \tag{2.19}$$

$$h_s(\mathbf{x})\phi_i(\mathbf{x}) = \epsilon_i \phi_i(\mathbf{x}) \tag{2.20}$$

where ϕ is the wave function for single electron.

Total wave function for N-independent electrons is given by Slater determinant (2.10):

$$|\psi_s> \;=\; \sum_{P} \mathbf{P} \; |\phi_1(\mathbf{x}_1)\phi_2(\mathbf{x}_2)...\phi_N(\mathbf{x}_N)>$$

$$= \; \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & ... & \phi_N(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & ... & \phi_N(\mathbf{x}_2) \\ ... & & & \\ \phi_1(\mathbf{x}_N) & \phi_2(\mathbf{x}_N) & ... & \phi_N(\mathbf{x}_N) \end{vmatrix} \tag{2.21}$$

Then, the total kinetic energy is:

$$\mathbf{T}_s[n] = <\psi_s| \sum_{i=1}^{N} \frac{-\hbar^2}{2m} \nabla_i^2 |\psi_s> = \sum_{i=1}^{N} <\phi_i| \frac{-\hbar^2}{2m} \nabla^2 |\phi_i> = \sum_{i=1}^{N} K_i \tag{2.22}$$

2.4.1 Kohn-Sham energy functional

For the N interacting electrons, the exchange and correlation terms should also be taken into account, and the *Kohn-Sham energy functional* is:

$$E[n] \;=\; T_s[n] + J[n] + E_{xc}[n] + \int v \cdot n \; d\mathbf{r} \tag{2.23}$$

Proof of (2.23):
Use (2.14) and (2.18),

$$E[n] \;=\; E_s[n] + E[n] - E_s[n]$$

$$= \; T_s[n] + [\; T[n] - T_s[n] + V_{(e^-e^-)}[n] + \int v \cdot n d\mathbf{r}]$$

$$V_{(e^-e^-)}[n] \;=\; J[n](Coulomb\ potential) + K[n](exchange\ potential)$$

$$= \; \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + K[n]$$

$$E_{xc}[n] \;=\; [\; T[n] - T_s[n] \;] + K[n] = \delta\; T[n] + \delta\; V_{(e^-e^-)}[n]$$

$$\Rightarrow E[n] \;=\; T_s[n] + J[n] + E_{xc}[n] + \int v \cdot n \; d\mathbf{r}$$

2.4.2 Kohn-Sham equation

Consider the free electron system which has the same energy functional as the interacting system, the density is determined by:

$$\delta E[n] = 0$$

The famous Kohn-Sham equation is given as:

$$h_s \phi_i = \epsilon_i \phi_i \tag{2.24}$$

$$h_s = -\frac{\hbar^2}{2m}\nabla^2 + v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}' + v_{xc}(\mathbf{r}) \tag{2.25}$$

$$v_{xc} = \frac{\delta E_{xc}}{\delta n}$$

Derivation of above Kohn-Sham equation:

$$Use (2.9) \quad ,$$

$$n = \sum_{i=1}^{N} |\phi_i|^2$$

$$T_s[n] = \sum_{i=1}^{N} < \phi_i| -\frac{\hbar^2}{2m}\nabla^2|\phi_i >$$

$$\delta E[n] = \delta E[\phi_i] = 0$$

$$\Rightarrow$$

$$h_s\phi_i = \epsilon_i \phi_i$$

$$h_s = -\frac{\hbar^2}{2m}\nabla^2 + v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}' + v_{xc}(\mathbf{r})$$

$$= T_s + v_{(ion-e^-)} + v_{(e^- e^-)} + v_{xc}$$

$$= T_s + v_{eff}$$

Solving the Kohn-Sham eigen-function, the sum of Kohn-Sham equation eigenvalue is:

$$\sum_{i=1}^{N} \epsilon_i = \sum_{i=1}^{N} < \phi_i|h_s|\phi_i >$$

$$= \sum_{i=1}^{N} < \phi_i| -\frac{\hbar^2}{2m}\nabla^2|\phi_i > + \sum_{i=1}^{N} < \phi_i|v_{eff}(\mathbf{r})|\phi_i >$$

$$= T_s[n] + \int v_s(\mathbf{r})n(\mathbf{r})d\mathbf{r} = E_s[n]$$

$$= T_s[n] + \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r} + \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \int v_{xc}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \tag{2.26}$$

From above (2.26), we can solve for $T_s[n]$, together with (2.23), we get:

$$
\begin{aligned}
E &= minE[n] = T_s[n] + J[n] + E_{xc}[n] + \int v \cdot n \ d\mathbf{r} \\
&= \{ \sum_{i=1}^{N} \epsilon_i - \int v(\mathbf{r})n(\mathbf{r})d\mathbf{r} - \int \frac{n(\mathbf{r})n(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|} - \int v_{xc}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \} \\
&+ \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|} + E_{xc}[n] + \int v \cdot n \ d\mathbf{r} \\
&= \sum_{i=1}^{N} \epsilon_i - \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|} + E_{xc}[n] - \int v_{xc}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \}
\end{aligned}
\tag{2.27}
$$

2.4.3 Solving the Kohn-Sham equation

- Assume an initial wave function ψ_i to compute n($\mathbf{r}$).

- Use n($\mathbf{r}$) to compute h_s.

- Use h_s to compute the new ψ_i.

Iteration is performed until the self-consistency is reached, i.e. the iteration converges.

2.4.4 Functional derivative

Let F[f] be a functional, f(x) is a function.

$$
\delta F = F[f + \delta f] - F[f] = \int \frac{\delta F}{\delta f} \delta f(x)dx
$$

E.g.

$$
\begin{aligned}
J[n] &= \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|}d\mathbf{r}d\mathbf{r'} \\
\delta J[n] &= \frac{1}{2}\int \frac{n(\mathbf{r'})\delta n(\mathbf{r}) + n(\mathbf{r})\delta n(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|}d\mathbf{r}d\mathbf{r'} \\
&= \int \frac{n(\mathbf{r'})\delta n(\mathbf{r})}{|\mathbf{r} - \mathbf{r'}|}d\mathbf{r}d\mathbf{r'} \\
\frac{\delta J(n)}{\delta n(\mathbf{r})} &= \int \frac{n(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|}d\mathbf{r'}
\end{aligned}
$$

The result of functional derivative is a function.

2.5 Planewave representation of Kohn-Sham equation

2.5.1 Periodicity and Bloch theorem

The unit cell is introduced to represent the periodicity in the solid, which is characterized by three vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ with volume $v = |(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}| = det \begin{pmatrix} \mathbf{a}_x & \mathbf{a}_y & \mathbf{a}_z \\ \mathbf{b}_x & \mathbf{b}_y & \mathbf{b}_z \\ \mathbf{c}_x & \mathbf{c}_y & \mathbf{c}_z \end{pmatrix}$.

The periodicity is due to the fact that the physical quantities are the same at location $\mathbf{r}$ and $\mathbf{r} + \mathbf{R}$, where $\mathbf{R} = l\mathbf{a} + m\mathbf{b} + n\mathbf{c}, l, m, n \in Z$.

Bloch theorem:

- If potential v is periodic, $v(\mathbf{r}) = v(\mathbf{r} + \mathbf{R})$, then the wave function if also periodic:

$$\begin{aligned} \psi(\mathbf{r}) &= e^{i\mathbf{k} \cdot \mathbf{r}} u_k(\mathbf{r}) \\ u_k(\mathbf{r}) &= u_k(\mathbf{r} + \mathbf{R}) \end{aligned} \qquad (2.28)$$

  where $\mathbf{k}$ is determined by the Born-Von Karmen boundary condition.

- Another way to state the Bloch theorem is:

$$\psi(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k} \cdot \mathbf{R}} \psi(\mathbf{r}) \qquad (2.29)$$

Proof of the Bloch theorem:

- Schrodinger equation: $H = -\frac{\hbar^2}{2m}\nabla^2 + U(\mathbf{r})$
  Define $T_R f(\mathbf{r}) = f(\mathbf{r} + \mathbf{R})$ where T_R is the translation operator.
  $T_R(H(\mathbf{r})\psi(\mathbf{r})) = H(\mathbf{r} + \mathbf{R})\psi(\mathbf{r} + \mathbf{R}) = H(\mathbf{r})(T_R\psi(\mathbf{r})) = H(\mathbf{r})\psi(\mathbf{r} + \mathbf{R})$
  $\Rightarrow T_R H = H T_R$ (*1)
  $T_R T_{R'}\psi(\mathbf{r}) = T_{R'}T_R\psi(\mathbf{r}) = \psi(\mathbf{r} + \mathbf{R} + \mathbf{R}')$
  $\Rightarrow T_R T_{R'} = T_{R'}T_R$ (*2)
  (*1) and (*2) $\Rightarrow T_R$ and H are set of commuting operators, so the eigenstates of H are also eigenstates of T_R.
  $\Rightarrow H\psi = \epsilon\psi, T_R\psi = c(\mathbf{R})\psi$ (*3)
  $\Rightarrow \begin{cases} T_{R'}T_R\psi = c(\mathbf{R})T_{R'}\psi = c(\mathbf{R})c(\mathbf{R}')\psi \\ T_{R'}T_R\psi = T_{R+R'}\psi = c(\mathbf{R} + \mathbf{R}')\psi \end{cases}$
  $\Rightarrow c(\mathbf{R}) c(\mathbf{R}') = c(\mathbf{R} + \mathbf{R}')$ (*4)

- Now, write $c(a_i) = e^{2\pi i x_i}$ such that (*4) is satisfied, where a_i(i=1,2,3) is the primitive latice vector.
  $\mathbf{R} = n_1 a_1 + n_2 a_2 + n_3 a_3$

$$c(\mathbf{R}) = c(n_1 a_1 + n_2 a_2 + n_3 a_3) = c(a_1)^{n_1} c(a_2)^{n_2} c(a_3)^{n_3} = e^{2\pi i x_1 n_1} e^{2\pi i x_2 n_2} e^{2\pi i x_3 n_3} \quad (*5)$$

i.e. $c(\mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}}$ (*6), where $\mathbf{k} = x_1 b_1 + x_2 b_2 + x_3 b_3$, $b_1 = 2\pi \frac{a_2 \times a_3}{a_1 \cdot a_2 \times a_3}$ etc, or equivalently $b_i a_j = 2\pi \delta_{ij}$, such that $\mathbf{k}\cdot\mathbf{R} = (x_1 b_1 + x_2 b_2 + x_3 b_3)\cdot(n_1 a_1 + n_2 a_2 + n_3 a_3) = 2\pi(x_1 n_1 + x_2 n_2 + x_3 n_3)$, which agrees with (*5).

- (*3) and (*6) $\Rightarrow T_R \psi(\mathbf{r}) = \psi(\mathbf{r}+\mathbf{R}) = c(\mathbf{R})\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{R}}\psi(\mathbf{r})$ which gives the Bloch theorem (2.29).

Since u_k is periodic, we make a Fourier expansion of it to get:

$$u_k(\mathbf{r}) = \sum_{\mathbf{G}} C_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{v}}$$

$$\psi(\mathbf{r}) = \sum_{\mathbf{G}} C_{\mathbf{G}} e^{i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}} \frac{1}{\sqrt{v}}$$

When an electron is in the state $\psi(\mathbf{r})$, we say that the electron occupies the orbit $\psi(\mathbf{r})$. Due to the Pauli's Exclusion Principle, we have at most two electrons per orbit (one spin up, and one spin down). Actually, the wave function of N non-interacting electrons is constructed from single electron wave function by Slater determinant.

2.5.2   Planewave method

Kohn-Sham energy functional is expressed in the terms of the planewave coefficients $C_{\mathbf{G}}$ in a supercell. Here, the supercell is a repeating unit cell which may contain a number of primitive cell which is the smallest possible repeating unit.

Kohn-Sham energy functional (2.23) in the real space is:

$$\begin{aligned}
E[n] &= T_s[n] + J[n] + E_{xc}[n] + \int v \cdot n \; d\mathbf{r} \\
&= \sum_{i=1}^{N} < \psi_i| - \frac{hbar^2}{2m}\nabla^2 |\psi_i > + \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{xc}[n(\mathbf{r})] + \int v \cdot n(\mathbf{r}) d\mathbf{r}
\end{aligned}$$

Now, we want to rewrite the above Kohn-Sham energy functional in the Fourier space, and that will be the aim of the rest of this chapter.

2.5.3 Kinetic energy

Consider only one electron in the state $\psi(\mathbf{r}) = \sum_{\mathbf{G}} C_{\mathbf{G}} e^{i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}} \frac{1}{\sqrt{v}}$.

$$
\begin{aligned}
T_s &= \; <\psi_i| - \frac{\hbar^2}{2m}\nabla^2|\psi_i> \\
&= \int d^3\mathbf{r} \; \psi^*(\mathbf{r})(-\frac{\hbar^2}{2m}\nabla^2)\psi(\mathbf{r}) \\
&= \int d^3\mathbf{r} \sum_{\mathbf{G}'} C_{\mathbf{G}'}^* e^{-i(\mathbf{G}'+\mathbf{k})\cdot\mathbf{r}}\frac{1}{\sqrt{v}}\;(-\frac{\hbar^2}{2m}\nabla^2)\sum_{\mathbf{G}} C_{\mathbf{G}} e^{i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}}\frac{1}{\sqrt{v}} \\
&= \int d^3\mathbf{r} \sum_{\mathbf{G}'} C_{\mathbf{G}'}^* e^{-i(\mathbf{G}'+\mathbf{k})\cdot\mathbf{r}}\frac{1}{\sqrt{v}}\;\sum_{\mathbf{G}} C_{\mathbf{G}}(\frac{\hbar^2}{2m})|\mathbf{G}+\mathbf{k}|^2 e^{i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}}\frac{1}{\sqrt{v}} \\
&= \sum_{\mathbf{G}} C_{\mathbf{G}}(\frac{\hbar^2}{2m})|\mathbf{G}+\mathbf{k}|^2 \sum_{\mathbf{G}'} C_{\mathbf{G}'}^* \int d^3\mathbf{r} \; e^{-i(\mathbf{G}'-\mathbf{G})\cdot\mathbf{r}}\frac{1}{v} \\
&= \sum_{\mathbf{G}} C_{\mathbf{G}}(\frac{\hbar^2}{2m})\;|\mathbf{G}+\mathbf{k}|^2\;C_{\mathbf{G}}^* \\
&= \sum_{\mathbf{G}} \frac{\hbar^2\,|\mathbf{G}+\mathbf{k}|^2}{2m}\;|C_{\mathbf{G}}|^2
\end{aligned}
$$

The total kinetic energy of N electrons is:

$$
\begin{aligned}
T_s &= \sum_{i}^{N\;states} \sum_{\mathbf{G}} \frac{\hbar^2\,|\mathbf{G}+\mathbf{k}|^2}{2m}\;|C_{\mathbf{G}}^i|^2 \\
&= \sum_{b} n_b \sum_{\mathbf{G}} \frac{\hbar^2}{2m}\;|\mathbf{G}+\mathbf{k}|^2\;|C_{\mathbf{G}}^b|^2
\end{aligned}
\qquad (2.30)
$$

where $n_b = \begin{cases} 2 & doubly\;occupied \\ 1 & singly\;occupied \\ 0 & unoccupied \end{cases}$.

2.5.4 Electron density

In real space, electron density,

$$
n_{\mathbf{r}} = \sum_{i=1}^{N} |\psi_i|^2 = \sum_{b} n_b |\psi_b|^2
\qquad (2.31)
$$

In the Fourier space,

$$
\begin{aligned}
n_{\mathbf{r}} &= \sum_{\mathbf{G}} n_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{v}} \tag{2.32}\\
n_{\mathbf{G}} &= \int_{v} n_{\mathbf{r}} e^{-i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{v}}\, d\mathbf{r} \tag{2.33}\\
&= \sum_{b} n_{b} \int_{v} \psi_{b}^{*}\psi_{b} e^{-i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{v}}\, d\mathbf{r}\\
&= \sum_{b} n_{b} \int_{v} \Big(\sum_{\mathbf{G''}} \bar{c}_{\mathbf{G''}}^{b} e^{-i\mathbf{G''}\cdot\mathbf{r}} \frac{1}{\sqrt{v}}\Big) e^{-i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{v}}\Big(\sum_{\mathbf{G'}} c_{\mathbf{G'}}^{b} e^{i\mathbf{G'}\cdot\mathbf{r}} \frac{1}{\sqrt{v}}\Big)\Big)\\
&= \sum_{b} n_{b} \sum_{\mathbf{G''}} \sum_{\mathbf{G'}} \bar{c}_{\mathbf{G''}}^{b} c_{\mathbf{G'}}^{b} \frac{1}{\sqrt{v}} \int_{v} e^{i(\mathbf{G'}-\mathbf{G}-\mathbf{G''})\cdot\mathbf{r}} \frac{1}{v}\\
&= \sum_{b} n_{b} \sum_{\mathbf{G'}} \bar{c}_{\mathbf{G'}-\mathbf{G}}^{b} c_{\mathbf{G'}}^{b} \frac{1}{\sqrt{v}} \tag{2.34}
\end{aligned}
$$

which is the convolution of $C_{\mathbf{G}}^{b}$.

2.5.5 Hartree term

As shown, the density function in the real space and in the Fourier space are given by (2.32) and (2.33). The Coulomb term, which is normally called Hartree term, in the Fourier space is:

$$
\begin{aligned}
V_{H} &= \frac{e^2}{2} \int_{v} d\mathbf{r}' \int_{\infty} d\mathbf{r} \frac{n_{\mathbf{r}} n_{\mathbf{r}'}}{|\mathbf{r}-\mathbf{r}'|}\\
&= \frac{e^2}{2} \int_{v} d\mathbf{r}' \int_{\infty} d\mathbf{r} \Big(\sum_{\mathbf{G}} n_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{v}}\Big) \frac{n_{\mathbf{r}'}}{|\mathbf{r}-\mathbf{r}'|}\\
&= \frac{e^2}{2} \sum_{\mathbf{G}} n_{\mathbf{G}} \int_{v} d\mathbf{r}' n_{\mathbf{r}'} e^{i\mathbf{G}\cdot\mathbf{r}'} \frac{1}{\sqrt{v}} \int_{\infty} d\mathbf{r} \frac{e^{i\mathbf{G}\cdot(\mathbf{r}-\mathbf{r}')}}{|\mathbf{r}-\mathbf{r}'|}\\
&= \frac{e^2}{2} \sum_{\mathbf{G}} n_{\mathbf{G}} \int_{v} d\mathbf{r}' n_{\mathbf{r}'} e^{i\mathbf{G}\cdot\mathbf{r}'} \frac{1}{\sqrt{v}} \frac{4\pi}{\mathbf{G}^2}\\
&= \frac{e^2}{2} \sum_{\mathbf{G}} n_{\mathbf{G}}\, n_{\mathbf{G}}^{*} \frac{4\pi}{\mathbf{G}^2}\\
&= \frac{e^2}{2} \sum_{\mathbf{G}} \frac{4\pi}{\mathbf{G}^2} |n_{\mathbf{G}}|^2 \tag{2.35}
\end{aligned}
$$

2.5.6 More on the electron density

To descritize the system of size $(a \times b \times c)$ into $(L_x \times L_y \times L_z)$ grids. $x_j = j\Delta x$, j=0,1,2...L_x-1, $\Delta x = \frac{a}{L_x}$.

$$
\begin{aligned}
V &= abc \\
\Delta V &= \Delta x \cdot \Delta y \cdot \Delta z = \frac{a}{L_x} \cdot \frac{b}{L_y} \cdot \frac{c}{L_z} = \frac{V}{L_x L_y L_z} \\
\mathbf{r} &= (l\hat{x}, m\hat{y}, n\hat{z}) \\
\mathbf{G} &= (G_x, G_y, G_z) = (\frac{2\pi}{a}\hat{l}, \frac{2\pi}{b}\hat{m}, \frac{2\pi}{c}\hat{n}) \\
G_x \cdot x &= \frac{2\pi}{a}\hat{l} \cdot l\Delta x = \frac{2\pi}{a}\hat{l} \cdot l\frac{a}{L_x} = \frac{2\pi l}{L_x}\hat{l} \\
n_{\mathbf{r}} &= n_{(l,m,n)} \\
&= \sum_{\mathbf{G}} n_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{V}} \\
&= \sum_{\mathbf{G}} n_{\mathbf{G}} e^{i(G_x l\Delta x + G_y m\Delta y + G_z n\Delta z)} \frac{1}{\sqrt{V}} \\
&= \sum_{\hat{l},\hat{m},\hat{n}} n_{(\hat{l},\hat{m},\hat{n})} e^{i(\frac{2\pi l}{L_x}\hat{l} + \frac{2\pi m}{L_y}\hat{m} + \frac{2\pi n}{L_z}\hat{n})} \frac{1}{\sqrt{V}} \\
n_{\mathbf{G}} &= n_{(\hat{l},\hat{m},\hat{n})} \\
&= \int_v n_{\mathbf{r}} e^{-i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{V}} \, d\mathbf{r} \\
&= \sum_{l,m,n} n_{l,m,n} e^{-i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{V}}\Delta V \\
&= \frac{\sqrt{V}}{L_x L_y L_z} \sum_{l,m,n} n_{l,m,n} e^{-i(\frac{2\pi l}{L_x}\hat{l} + \frac{2\pi m}{L_y}\hat{m} + \frac{2\pi n}{L_z}\hat{n})}
\end{aligned}
$$

2.5.7 Ewald's sum – Ion-ion interaction

Use Born-Oppenheimer approximation, the nuclei-nuclei interaction is treated classically:

$$
V_{ii} = \frac{1}{2}\sum_{i,j}^{N} \frac{Z^2 e^2}{|R_i - R_j|}
$$

Assume only one type of atom located at $\{R_I\}$, I=1,2,...N in a unit cell. Periodicity gives the images of ion R_I at $R_I + R$. Thus, the electrostatic potential created by ions from

R_I and all its images is:

$$
\begin{aligned}
V_{ii} &= \frac{1}{2}\sum_{I,J,R,r\neq 0}^{N}\frac{Z^2e^2}{|R_I - R_J + R|}\\
&= \frac{1}{2}\sum_{I,J}Z^2e^2\{\sum_{R,r\neq 0}\frac{erfc(\eta\,|R_I - R_J + R|)}{|R_I - R_J + R|} - \frac{2\eta}{\sqrt{\pi}}\delta_{IJ}\\
&\quad + \frac{4\pi}{V}\sum_{\mathbf{G}\neq 0}\frac{1}{|\mathbf{G}|^2}e^{-\mathbf{G}^2/4\eta^2}cos[(R_I - R_J)\cdot\mathbf{G}] - \frac{\pi}{\eta^2 V}\} + \infty
\end{aligned}
\tag{2.36}
$$

which is called Ewald's sum.

Proof of Ewald's sum (2.36):

- Firstly, prove the Fourier integral of $e^{-r^2u^2}$ is

$$
f(\mathbf{G}) = \frac{1}{V}\frac{\pi^{3/2}}{u^3}e^{-G^2/u^2}
\tag{2.37}
$$

  Proof of (2.37):

$$
\begin{aligned}
f(\mathbf{G}) &= \int_v d\mathbf{r}\frac{1}{V}e^{-r^2u^2}e^{-i\mathbf{G}\cdot\mathbf{r}}\\
&= \frac{1}{V}\int_v r^2dr\,sin\theta d\theta\,d\phi e^{-r^2u^2}e^{-i\mathbf{G}\cdot\mathbf{r}}\\
&= \frac{2\pi}{V}\int_{r=0}^{\infty}r^2dr\,e^{-r^2u^2}\int_{-1}^{1}dcos\theta\,e^{-iGrcos\theta}\\
&= \frac{2\pi}{V}\int_{r=0}^{\infty}r^2dr\,e^{-r^2u^2}\frac{1}{iGr}(e^{iGr} - e^{-iGr})\\
&= \frac{2\pi}{iGV}\int_{r=0}^{\infty}dr\,re^{-r^2u^2}(e^{iGr} - e^{-iGr})
\end{aligned}
$$

  let $z = r - \frac{i}{2}\frac{G}{u^2}$, changing the variable of integral get:

$$
\begin{aligned}
f(\mathbf{G}) &= \{\frac{A}{2u^2}e^{-u^2B^2} + AB\int_{-B}^{0}e^{-u^2z^2}\,dz + \frac{\pi^{3/2}}{2Vu^3}e^{-G^2/4u^2}\}\\
&\quad - \{\frac{A}{2u^2}e^{-u^2B^2} + AB\int_{-B}^{0}e^{-u^2z^2}\,dz - \frac{\pi^{3/2}}{2Vu^3}e^{-G^2/4u^2}\}
\end{aligned}
$$

  where $A = \frac{2\pi}{iGV}e^{-G^2/4u^2}$, $B = \frac{i}{2}\frac{G}{u^2}$. Simplify to get (2.37).

- Secondly, prove

$$
\frac{1}{r} = \frac{1}{(2\pi)^3}\int d\mathbf{G}\frac{4\pi}{\mathbf{G}^2}e^{-\mathbf{G}^2/4\eta^2+i\mathbf{G}\cdot\mathbf{r}} + \frac{erfc(\eta r)}{r}
\tag{2.38}
$$

Proof of (2.38):

$$
\begin{aligned}
\frac{1}{r} &= \frac{1}{r}\frac{2}{\sqrt{\pi}}\int_0^\infty e^{-r^2u^2}d(ru) \\
&= \frac{2}{\sqrt{\pi}}\int_0^\eta e^{-r^2u^2}du + \frac{2}{\sqrt{\pi}}\int_\eta^\infty e^{-r^2u^2}du
\end{aligned}
$$

Since only the 2nd term of the above equation converges fast in real space, while the 1st term converges relatively slow in the real space, we have to perform Fourier transform to calculate the 1st term in Fourier space. The Fourier transform of $e^{-r^2u^2}$ is already proven in (2.37), do the Inverse Fourier transform to get $e^{-r^2u^2}$:

$$
\begin{aligned}
e^{-r^2u^2} &= \frac{V}{(2\pi)^3}\int_0^\infty f(\mathbf{G})e^{i\mathbf{G}\cdot\mathbf{r}}d\mathbf{G} \\
&= \frac{V}{(2\pi)^3}\int_0^\infty \frac{1}{V}\frac{\pi^{3/2}}{u^3}e^{-G^2/u^2}e^{i\mathbf{G}\cdot\mathbf{r}}d\mathbf{G} \\
&= \frac{1}{(2\pi)^3}\int_0^\infty \frac{\pi^{3/2}}{u^3}e^{-G^2/u^2+i\mathbf{G}\cdot\mathbf{r}}d\mathbf{G}
\end{aligned}
$$

Substitute into the expression of $\frac{1}{r}$:

$$
\begin{aligned}
\frac{2}{\sqrt{\pi}}\int_0^\eta e^{-r^2u^2}du &= \frac{2}{\sqrt{\pi}}\int_0^\eta du \frac{1}{(2\pi)^3}\int_0^\infty d\mathbf{G}\frac{\pi^{3/2}}{u^3}e^{-G^2/u^2+i\mathbf{G}\cdot\mathbf{r}} \\
&= \frac{1}{(2\pi)^2}\int_0^\infty d\mathbf{G}e^{i\mathbf{G}\cdot\mathbf{r}}\int_0^\eta du\frac{1}{u^3}e^{-G^2/u^2} \\
&= \frac{1}{(2\pi)^2}\int_0^\infty d\mathbf{G}e^{i\mathbf{G}\cdot\mathbf{r}}\frac{2}{G^2}e^{-G^2/4\eta^2} \\
\frac{1}{r} &= \frac{1}{(2\pi)^3}\int_0^\infty d\mathbf{G}\frac{4\pi}{G^2}e^{-G^2/4\eta^2+i\mathbf{G}\cdot\mathbf{r}} + \frac{2}{\sqrt{\pi}}\int_\eta^\infty e^{-r^2u^2}du \\
&= \frac{1}{(2\pi)^3}\int_0^\infty d\mathbf{G}\frac{4\pi}{G^2}e^{-G^2/4\eta^2+i\mathbf{G}\cdot\mathbf{r}} + \frac{erfc(\eta r)}{r}
\end{aligned}
$$

- Finally, prove (2.36)

$$
\begin{aligned}
V_{ii} &= \frac{1}{2}\sum_{I,J,R,r\neq 0}^{N}\frac{Z^2e^2}{|R_I-R_J+R|} \\
&= \frac{1}{2}\sum_{I,J,R'}Z^2e^2\left\{\frac{1}{(2\pi)^3}\int_0^\infty d\mathbf{G}\frac{4\pi}{G^2}e^{-G^2/4\eta^2+i\mathbf{G}\cdot|R_I-R_J+R|}\right. \\
&\quad + \frac{erfc(\eta|R_I-R_J+R|)}{|R_I-R_J+R|}\left.\right\} \\
&= \frac{1}{2}\sum_{I,J,R}Z^2e^2\frac{1}{(2\pi)^3}\int_0^\infty d\mathbf{G}\frac{4\pi}{G^2}e^{-G^2/4\eta^2+i\mathbf{G}\cdot|R_I-R_J+R|} \\
&\quad - \frac{1}{2}\sum_{I,J,r=0}Z^2e^2\frac{1}{(2\pi)^3}\int_0^\infty d\mathbf{G}\frac{4\pi}{G^2}e^{-G^2/4\eta^2+i\mathbf{G}\cdot|R_I-R_J+R|} \\
&\quad + \frac{1}{2}\sum_{I,J,R'}Z^2e^2\frac{erfc(\eta|R_I-R_J+R|)}{|R_I-R_J+R|}
\end{aligned}
$$

$$\begin{aligned}
V_1 &= \frac{1}{2}\sum_{I,J,R} Z^2 e^2 \frac{1}{(2\pi)^3} \int_0^\infty d\mathbf{G}\, \frac{4\pi}{G^2} e^{-G^2/4\eta^2 + i\mathbf{G}\cdot|R_I - R_J + R|} \\
&= \frac{1}{2}\sum_{I,J,R} Z^2 e^2 \frac{1}{(2\pi)^3} \sum_{\mathbf{G}} \frac{4\pi}{G^2} e^{-G^2/4\eta^2 + i\mathbf{G}\cdot|R_I - R_J + R|} \frac{(2\pi)^3}{V} \\
&= V_1(G \neq 0) + V_1(G \to 0) \\
V_1(G \to 0) &= \frac{1}{2}\sum_{I,J,R} Z^2 e^2 \frac{4\pi}{V} \frac{1}{G^2}(1 - \frac{G^2}{4\eta^2}) \\
&= \frac{1}{2}\sum_{I,J} Z^2 e^2 \frac{4\pi}{V} \frac{1}{G^2}\Big|_{G\to 0} - \frac{1}{2}\sum_{I,J} Z^2 e^2 \frac{\pi}{V\eta^2} \\
V_2 &= \frac{1}{2}\sum_{I,J,r=0} Z^2 e^2 \frac{1}{(2\pi)^3} \int_0^\infty d\mathbf{G}\, \frac{4\pi}{G^2} e^{-G^2/4\eta^2 + i\mathbf{G}\cdot|R_I - R_J + R|} \\
&= \frac{1}{2}\sum_{I,J,I=J} Z^2 e^2 \frac{1}{(2\pi)^3} \int_0^\infty dG\, 4\pi G^2\, \frac{4\pi}{G^2} e^{-G^2/4\eta^2} \\
&= \frac{1}{2}\sum_{I,J,I=J} Z^2 e^2 \frac{2\eta}{\sqrt{\pi}} \\
&= \frac{1}{2}\sum_{I,J} Z^2 e^2 \frac{2\eta}{\sqrt{\pi}} \delta_{I,J} \\
V_{ii} &= V_1 - V_2 + \frac{1}{2}\sum_{I,J,R'} Z^2 e^2 \frac{erfc(\eta|R_I - R_J + R|)}{|R_I - R_J + R|} \\
&= \frac{1}{2}\sum_{I,J} Z^2 e^2 \Big\{ \sum_{r\neq 0} \frac{erfc(\eta|R_I - R_J + R|)}{|R_I - R_J + R|} \\
&\quad - \frac{2\eta}{\sqrt{\pi}}\delta_{I,J} + \frac{4\pi}{V}\sum_{G\neq 0} \frac{1}{G^2} e^{-G^2/4\eta^2 + i\mathbf{G}\cdot|R_I - R_J + R|} - \frac{\pi}{V\eta^2} \Big\} \\
&\quad + \frac{1}{2}\sum_{I,J} Z^2 e^2 \frac{4\pi}{V} \frac{1}{G^2}\Big|_{G\to 0} \\
&= \frac{1}{2}\sum_{I,J} Z^2 e^2 \Big\{ \sum_{r\neq 0} \frac{erfc(\eta|R_I - R_J + R|)}{|R_I - R_J + R|} - \frac{2\eta}{\sqrt{\pi}}\delta_{I,J} \\
&\quad + \frac{4\pi}{V}\sum_{G\neq 0} \frac{1}{G^2} e^{-G^2/4\eta^2} \cos(\mathbf{G}\cdot|R_I - R_J + R|) - \frac{\pi}{V\eta^2} \Big\} \\
&\quad + \frac{1}{2}\sum_{I,J} Z^2 e^2 \frac{4\pi}{V} \frac{1}{G^2}\Big|_{G\to 0}
\end{aligned}$$

2.5.8 Exchange Correlation E_{xc}

As derived in the Khom-Sham energy functionaly (2.23),

$$E_{xc} = T - T_s + V_{e^- e^-} - J$$

This term is extremely difficult to compute, thus we have to use approximate method. The simplest and most popular one is the so called LDA (Local Density Approximation). In this approximation, the exchange-correlation energy of an electronic system is constructed by assuming that the exchange-correlation energy per electron at apoint $\mathbf{r}$ in the electron gas, $\epsilon_{xc}(\mathbf{r})$, is equal to the exchange-correlation energy per electron in a homogeneous electron gas that has the same density as the electron gas at point $\mathbf{r}$.

$$E_{xc}(n(\mathbf{r})) = \int_V \epsilon_{xc}(\mathbf{r}) \, n(\mathbf{r}) \, d\mathbf{r} \qquad (2.39)$$
$$\epsilon_{xc}(\mathbf{r}) = \epsilon(n(\mathbf{r}))$$

Many calculation is based on a simple formula:

$$\epsilon_{xc}(n(\mathbf{r})) = -An^{1/3}(\mathbf{r})$$
$$A = 0.4582(\frac{4\pi}{3})^{1/3}$$

A better approximation is given by Teter:

$$\epsilon_{xc}(n(\mathbf{r})) = -\frac{a_0 + a_1 r_s + a_2 r_s^2 + a_3 r_s^3}{b_1 r_s + b_2 r_s^2 + b_3 r_s^3 + b_4 r_s^4} \qquad (2.40)$$
$$r_s^{-1} = (\frac{4\pi}{3}n(\mathbf{r}))^{1/3} \qquad (2.41)$$

The ratio of two polynomials is known as padi approximation, if ϵ is know as series in r_s. Padi approximation improves the convergence of the function.

The constants in the Teter's approximation are:

$$
\begin{aligned}
a_0 &= 0.4581652932831429 \\
a_1 &= 2.217058676663745 \\
a_2 &= 0.7405551735357053 \\
a_3 &= 0.01968227878617998 \\
b_1 &= 1 \\
b_2 &= 4.504130959426697 \\
b_3 &= 1.110667363742916 \\
b_4 &= 0.02359291751427506
\end{aligned}
$$

There is another improved approximation over LDA, which is gradient expansion:

$$E_{xc}[n] = \int_V f(n, \nabla n, ...) d\mathbf{r}$$

However, the more complicated approximation is not necessary.

2.5.9   Atomic units

For elctronic structure calculation, it ismost convenient to use atomic units. In this unit, $e=\hbar=m=1$. 1 a.u. of length $= a_0 = \frac{\hbar^2}{me^2}=0.53A$. 1 a.u. of energy $= \frac{e^2}{a_0} = 27.2$ eV $= 1$ Hartree $= 2$ Rydberg.

2.5.10   Pseudo-potentials

The electrons of an atom is divided into core electrons and valence electrons. For Si, total electron number is 14, the electron shell structure orbit is $\underbrace{[1s^2 2s^2 2p^6]}_{core}\underbrace{3s^2 3p^2}_{valence}$. We treat Si as having only 4 electrons, and group the 10 core electrons into the nuclei by an effective potential.

Non-interacting electron-atom model

we first consider a non-interacting electron-atom model.

$$\psi = \phi + \sum_c a_c \phi_c$$

where ϕ_c is the wavefunction for the core electron state,and ψ is the valence electron wavefunction, ϕ is a smooth pserdo wavefunction. Require the valence wavefunction orthogonal to the core state:

$$< \phi_c|\psi >=< \phi_c|\phi > +a_c = 0$$
$$\psi = \phi - \sum_c < \phi_c|\phi > \phi_c$$

ψ must satisfy the Schrodinger equation, thus:

$$\mathcal{H}|\psi >= \epsilon|\psi >$$
$$\mathcal{H} = \frac{p^2}{2m} + V_c$$
$$\mathcal{H}|\phi > -\mathcal{H}\sum_c < \phi_c|\phi > |\phi_c >= \epsilon|\phi > -\epsilon \sum_c < \phi_c|\phi > |\phi_c >$$
$$\mathcal{H}|\phi > +\sum_c < \phi_c|\phi > \underbrace{(\epsilon - \epsilon_c)|\phi_c >}_{V_R} = \epsilon|\phi >$$
$$(\frac{p^2}{2m} + \underbrace{V_c + V_R}_{V^{PS}})|\psi >= \epsilon|\phi >$$

So, the pseudo-potential is defined as:

$$V^{PS} = V_c + \sum_c (\epsilon - \epsilon_c)|\phi_c >< \phi_c| \tag{2.42}$$

Some properties of the pseudo-potential:

- The pseudo-potential is generally nonlocal.
  If $V(\mathbf{r})$ is local, then $V(\mathbf{r})$ acting on ϕ is $V(\mathbf{r})\phi(\mathbf{r})$.
  If $V(\mathbf{r})$ is nonlocal, then $V(\mathbf{r})|\phi> = \int V(\mathbf{r},\mathbf{r}')\phi(\mathbf{r}')d\mathbf{r}'$.

- The pseudo-potential V_c and V_R tends to cancel for small r, and $V_R \to 0$ at large r.
  For r¿r_c, the pseudo-potential and wavefunction agree with the true potential and real wavefunction.

Empirical local pseudo-potential for Si

Now consider an empirical local pseudo-potential for Si (Appelbaum and Hamann).

$$V^{PS} = (V_1 + V_2 r^2)e^{-\alpha r^2} - \frac{z}{r} erfc(\sqrt{\alpha} r)$$

where V_1, V_2, α are parameters.

Some concepts of pseudo-potential:

- Soft and hard pseudo-potential.
  If the pseudo-potential is flat when r→0, it is soft.
  If the pseudo-potential is like $\frac{1}{r}$ when r→0, it is hard.

- Norm-conservation.
  The total charge inside the cut-off radius r_c for the pseudo-wavefunction is the same as that for the real wavefunction.

- Transferability.
  The pseudo-potential generated for an atom should be readily usable in a solid state environment.

Various pseudo-potentials

The most commonly used pseudo-potentials are of the form:

$$V^{PS} = V^{loc}(\mathbf{r}) + \sum V_l^{NL}(\mathbf{r})P_l$$
$$P_l = \sum_m |lm><lm| \quad (m = -l, -l+1, ...l)$$

Among various pseudo-potentials, the one by Bachelet, Hamann and Schluter is most popular. Another type called Ultrasoft pseudo-potential is developped by Vanderbilt. In this course, we use the *Ihm-Cohen potential* (J.Ihm & M.L.Cohen, Solid State Comm. v29, 711, 1979) defined as:

$$V^{PS}(q) = \frac{a_1}{q^2}[\cos(a_2 q) + a_3]e^{a_4 q^4} \tag{2.43}$$

$V^{PS}(q)$ is the Fourier transform of $V^{PS}(r)$.

$$
\begin{aligned}
a_1 &= -\frac{4\pi Z}{1 + a_3} \\
a_2 &= 0.79065 \\
a_3 &= -0.35201 \\
a_4 &= -0.01807
\end{aligned}
$$

2.5.11 Ion-electron interaction energy

Assume that the ions are located at $\mathbf{R}_I$ in a cell for I=1,2,3,...n, where n is the number of ions in the unit cell. Let $V_I(\mathbf{r} - \mathbf{R}_I)$ be the potential seen by an electron at $\mathbf{r}$, then the total potential of the electron is:

$$
\begin{aligned}
v(\mathbf{r}) &= \sum_{cell\ i} \sum_I v_I(\mathbf{r} - \mathbf{R}_I) \\
&= \sum_R \sum_I v_I(\mathbf{r} - (\mathbf{R}_I + \mathbf{R})) \tag{2.44}
\end{aligned}
$$

Total ion-electron interaction energy (per unit cell) is:

$$
\begin{aligned}
V_{i-e^-} &= \int_{cell} v(\mathbf{r})\, n(\mathbf{r})\, d\mathbf{r} \\
&= \int_V v(\mathbf{r}) \sum_\mathbf{G} n_\mathbf{G} e^{i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{V}} d\mathbf{r} \\
&= \sum_\mathbf{G} v_{(-\mathbf{G})} n_\mathbf{G} \tag{2.45}
\end{aligned}
$$

where $v_{(\mathbf{G})} = \int_V v(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{V}}\, d\mathbf{r}$ is the Fourier transform of v($\mathbf{r}$).

$$
\begin{aligned}
v_{(\mathbf{G})} &= \int_V v(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{V}}\, d\mathbf{r} \\
&= \sum_R \sum_I \int_V v_I(\mathbf{r} - (\mathbf{R}_I + \mathbf{R})) e^{-i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{V}}\, d\mathbf{r} \\
&= \sum_R \sum_I \int_V v_I(\mathbf{r}') e^{-i\mathbf{G}\cdot(\mathbf{r}'+\mathbf{R}_I+\mathbf{R})} \frac{1}{\sqrt{V}}\, d\mathbf{r} \\
&= \sum_I e^{-i\mathbf{G}\cdot\mathbf{R}_I} \frac{1}{\sqrt{V}} \int_\infty v_I(\mathbf{r}') e^{-i\mathbf{G}\cdot\mathbf{r}'}\, d\mathbf{r} \\
&= S_\mathbf{G} V_{I(\mathbf{G})}
\end{aligned}
$$

where $S_\mathbf{G} = \sum_I \frac{e^{-i\mathbf{G}\cdot\mathbf{R}_I}}{\sqrt{V}}$ is the structure factor, $V_{I(\mathbf{G})}$ is the Fourier transform of $v_I(\mathbf{r})$ that is given by (2.43). So, finally get:

$$
v_{(\mathbf{G})} = S_\mathbf{G} V^{PS}(q) \tag{2.46}
$$

Infinities in V_{i-e^-}

To handle the infinities in the ion-electron interaction energy, seperate out the $\mathbf{G} = 0$ term of the summation in (2.45):

$$
\begin{aligned}
V^{PS}(G)\Big|_{G\to 0} &= \frac{a_1}{G^2}\left(1 - \frac{1}{2}a_2^2 G^2 + a_3\right) \\
&= \frac{(1+a_3)a_1}{G^2} - \frac{1}{2}a_1 a_2^2 \\
&= -\frac{4\pi Z}{G^2} - \frac{1}{2}a_1 a_2^2 \\
&= -\frac{1}{2}a_1 a_2^2 - \infty \\
S_{-\mathbf{G}}\Big|_{G\to 0} &= \frac{N_{ion}}{\sqrt{V}} \\
n_{\mathbf{G}}\Big|_{G\to 0} &= \sum_b n_b \sum_{\mathbf{G}'} \bar{c}_{\mathbf{G}'-\mathbf{G}}^b c_{\mathbf{G}'}^b \frac{1}{\sqrt{V}} \\
&= \sum_b n_b \frac{1}{\sqrt{V}} \\
&= \frac{N_{e^-}}{\sqrt{V}}
\end{aligned}
$$

So, finally get the expression of ion-electron interaction energy:

$$
\begin{aligned}
V_{i-e^-} &= \sum_{\mathbf{G}} S_{-\mathbf{G}} V^{PS}(G) n_{\mathbf{G}} \\
&= \sum_{\mathbf{G}\neq 0} S_{-\mathbf{G}} V_{(G)}^{PS} n_{\mathbf{G}} - \frac{1}{2}a_1 a_2^2 \frac{N_{ion} N_{e^-}}{V}
\end{aligned}
\qquad (2.47)
$$

2.5.12 Total energy E_{tot}

Putting all the terms together, the total energy per unit cell is:

$$
E_{tot} = T_s + V_H + V_{ii} + E_{xc}(n(\mathbf{r})) + V_{i-e^-}
\qquad (2.48)
$$

where the five terms are given as:

$$(2.30): \quad T_s = \sum_b n_b \sum_{\mathbf{G}} \frac{\hbar^2}{2m} |\mathbf{G} + \mathbf{k}|^2 |C_{\mathbf{G}}^b|^2$$

$$(2.35): \quad V_H = \frac{e^2}{2} \sum_{\mathbf{G}} \frac{4\pi}{\mathbf{G}^2} |n_{\mathbf{G}}|^2$$

$$(2.36): \quad V_{ii} = \frac{1}{2} \sum_{I,J} Z^2 e^2 \{ \sum_{R, r \neq 0} \frac{erfc(\eta |R_I - R_J + R|)}{|R_I - R_J + R|} - \frac{2\eta}{\sqrt{\pi}} \delta_{IJ}$$

$$+ \frac{4\pi}{V} \sum_{\mathbf{G} \neq 0} \frac{1}{|\mathbf{G}|^2} e^{-\mathbf{G}^2/4\eta^2} cos[(R_I - R_J) \cdot \mathbf{G}] - \frac{\pi}{\eta^2 V} \} + \infty$$

$$(2.39): \quad E_{xc}(n(\mathbf{r})) = \int_V \epsilon_{xc}(\mathbf{r}) \, n(\mathbf{r}) \, d\mathbf{r}$$

$$(2.40): \quad \epsilon_{xc}(n(\mathbf{r})) = -\frac{a_0 + a_1 r_s + a_2 r_s^2 + a_3 r_s^3}{b_1 r_s + b_2 r_s^2 + b_3 r_s^3 + b_4 r_s^4}$$

$$(2.47): \quad V_{i-e^-} = \sum_{\mathbf{G}} v_{(-\mathbf{G})} n_{\mathbf{G}} = \sum_{\mathbf{G} \neq 0} S_{-\mathbf{G}} V_{(G)}^{PS} n_{\mathbf{G}} - \frac{1}{2} a_1 a_2^2 \frac{N_{ion} N_{e^-}}{V}$$

2.6 Lab 2

Write a program to compute the total energy given by (2.48) for the Silicon system. Here are some guidelines.

- Use a box with length a, b, c in atomic units. Use a cut-off for $\mathbf{G}$ in the Reciprocal (Fourier) space in the expressions like $\sum_{\mathbf{G}} n_{\mathbf{G}} e^{i\mathbf{G} \cdot \mathbf{r}} \frac{1}{\sqrt{V}}$. The exact solutions need to sum over G from $-\infty$ to ∞. To compute analytically, we need a cut-off of $\mathbf{G}$ at large value N, such that $G = -\frac{N}{2} + 1, -\frac{N}{2} + 2, ... \frac{N}{2}$.

- You can test the sub-function of computing Ewald's sum (ion-ion interaction) for the Sodium Chloride (NaCl) system.
  NaCl is two fcc lattices. The lattice constant a=5.64 A. The value of Ewald's sum is about -1.43×10^{-11} erg per ion pair (per NaCl).

- There will be a lot of adjustable parameters in the program, such as the number of bands, number of ions per unit cell, choice of cut-off G, and how the descretization is done. All these parameters and programming details should be handed in together with the source code.

- Si-diamond structure.
  There are 8 atoms per unit cell, located at
  $(0,0,0), (\frac{1}{2}, \frac{1}{2}, 0), (\frac{1}{2}, 0, \frac{1}{2}), (0, \frac{1}{2}, \frac{1}{2}), (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}), (\frac{3}{4}, \frac{3}{4}, \frac{1}{4}), (\frac{3}{4}, \frac{1}{4}, \frac{3}{4}), (\frac{1}{4}, \frac{3}{4}, \frac{3}{4})$.
  Each atom has 4 electrons, so there are totally 32 electrons per unit cell occupying 16 bands.

2.7 Tutorial 2

Q1: Prove variation principle (2.8).

Q2: Prove theorem (2.9).

Q3: Prove the second Hoherberg-Kohn theorem (2.13).

Q4: Prove the equation (2.22) in the Kohn-Sham approach.

Q5: Prove the two statements of Bloch theorem (2.28) and (2.29) are equivalent.

Q6: Compute $\frac{\partial T_s}{\partial \bar{c}_{\mathbf{G}}^b}$ and $\frac{\partial V_H}{\partial \bar{c}_{\mathbf{G}}^b}$.

Q7: What is the 1 a.u. of time?

Q8: Why is ϵ_{xc} proportional to $n^{1/3}$ for small n?

Q9: Calcellation of infinities for $e^- e^-, ion - ion, ion - e^-$ potential.

2.8 Reading Materials

- *Density-Functional theory of atoms and molecules*
  R.G.Parr & W.Yang

- *Reviews of Modern physics*
  M.C.Payne, vol 64, No. 4, 1045 (1992)

Chapter 3

Minimization of Total Energy

3.1 Kohn-Sham equation in Fourier space

In the tutorial 2, Q6, it is proven that:

$$\frac{\partial T_s}{\partial \bar{C}_{\mathbf{G}}^b} = \frac{1}{2} n_b |\mathbf{G} + \mathbf{k}|^2 C_{\mathbf{G}}^b$$

$$\frac{\partial V_H}{\partial \bar{C}_{\mathbf{G}}^b} = \frac{n_b}{\sqrt{V}} \sum_{\mathbf{G}'} \frac{4\pi}{\mathbf{G}'^2} \, n_{\mathbf{G}'} \, C_{\mathbf{G}-\mathbf{G}'}^b$$

The ion-electron energy term:

$$\frac{\partial V_{i-e^-}}{\partial \bar{C}_{\mathbf{G}}^b} = \sum_{\mathbf{G}'} v_{(-\mathbf{G}')} \frac{\partial n_{\mathbf{G}'}}{\partial \bar{C}_{\mathbf{G}}^b}$$

$$= \frac{n_b}{\sqrt{V}} \sum_{\mathbf{G}'} v_{(\mathbf{G}')} \, C_{\mathbf{G}-\mathbf{G}'}^b$$

The exchange-correlation term:

$$\frac{\partial E_{xc}}{\partial \bar{C}_{\mathbf{G}}^b} = \int_V \frac{\partial(\epsilon_{xc}(\mathbf{r})n(\mathbf{r}))}{\partial n(\mathbf{r})} \frac{\partial n(\mathbf{r})}{\partial \bar{C}_{\mathbf{G}}^b} \, d\mathbf{r}$$

$$= \int_V \frac{\partial(\epsilon_{xc}(\mathbf{r})n(\mathbf{r}))}{\partial n(\mathbf{r})} \sum_{\mathbf{G}'} \frac{\partial n(\mathbf{G}')}{\partial \bar{C}_{\mathbf{G}}^b} e^{i\mathbf{G}'\cdot\mathbf{r}} \frac{1}{\sqrt{V}} \, d\mathbf{r}$$

$$= \sum_{\mathbf{G}'} V_{xc(-\mathbf{G}')} \frac{\partial n(\mathbf{G}')}{\partial \bar{C}_{\mathbf{G}}^b}$$

$$= \sum_{\mathbf{G}'} V_{xc(-\mathbf{G}')} \, C_{\mathbf{G}+\mathbf{G}'}^b \frac{n_b}{\sqrt{V}}$$

$$V_{xc(-\mathbf{G})} = \int_V \frac{\partial(\epsilon_{xc}(\mathbf{r})n(\mathbf{r}))}{\partial n(\mathbf{r})} e^{-i\mathbf{G}\cdot\mathbf{r}} \frac{1}{\sqrt{V}} \, d\mathbf{r} = Fourier\ Transform\{\frac{\partial(\epsilon_{xc}(\mathbf{r})n(\mathbf{r}))}{\partial n(\mathbf{r})}\}$$

Putting all the above terms together:

$$\frac{\partial E_{tot}}{\partial \bar{C}_{\mathbf{G}}^{b}} = \frac{1}{2}n_b|\mathbf{G}+\mathbf{k}|^2 C_{\mathbf{G}}^{b} + \frac{n_b}{\sqrt{V}}\sum_{\mathbf{G'}} V_{\mathbf{G'}}^{eff}\, C_{\mathbf{G-G'}}^{b} \qquad (3.1)$$

$$V_{\mathbf{G}}^{eff} = \frac{4\pi}{\mathbf{G}^2}\, n_{\mathbf{G}} + v_{(\mathbf{G})} + V_{xc(\mathbf{G})}$$

Matrix eigenvalue problem

Now, let's compute $\frac{\partial E_{tot}}{\partial C_{\mathbf{G}}^{b}}$.

The orthonormal condition gives:

$$\int \bar{\phi}_b(\mathbf{r})\phi_{b'}(\mathbf{r})d\mathbf{r} = \delta_{bb'}$$

Substitute $\phi_b(\mathbf{r}) = \sum_{\mathbf{G}} C_{\mathbf{G}}^{b}\frac{e^{i\mathbf{G}\cdot\mathbf{r}}}{\sqrt{V}}$ into the above equation:

$$\begin{aligned}
\int \bar{\phi}_b(\mathbf{r})\phi_{b'}(\mathbf{r})d\mathbf{r} &= \sum_{\mathbf{G}} \bar{C}_{\mathbf{G}}^{b}\int \frac{e^{-i\mathbf{G}\cdot\mathbf{r}}}{\sqrt{V}}\sum_{\mathbf{G'}} C_{\mathbf{G'}}^{b'}\frac{e^{i\mathbf{G'}\cdot\mathbf{r}}}{\sqrt{V}}d\mathbf{r} \\
&= \sum_{\mathbf{G}}\sum_{\mathbf{G'}} \bar{C}_{\mathbf{G}}^{b}C_{\mathbf{G'}}^{b'}\delta_{\mathbf{G},\mathbf{G'}} \\
&= \sum_{\mathbf{G}} \bar{C}_{\mathbf{G}}^{b}C_{\mathbf{G}}^{b} = \delta_{bb'}
\end{aligned}$$

The Lagrange Multiplier method gives:

$$\lambda_{bb'}\sum_{\mathbf{G}} \bar{C}_{\mathbf{G}}^{b}C_{\mathbf{G}}^{b} = \lambda_{bb'}\delta_{bb'}$$

$$\delta E = \sum_{b,\mathbf{G}} \frac{\partial E}{\partial \bar{C}_{\mathbf{G}}^{b}}\delta\bar{C}_{\mathbf{G}}^{b} + \sum_{b,b',\mathbf{G}} \lambda_{b,b'} C_{\mathbf{G}}^{b'}\delta\bar{C}_{\mathbf{G}}^{b}$$

After diagonalizing $\lambda_{b,b'}$, we get the standard Schrodinger equation:

$$\frac{\partial E}{\partial \bar{C}_{\mathbf{G}}^{b}} = n^b \epsilon^b C_{\mathbf{G}}^{b}$$

Use equation (3.1),

$$\frac{1}{2}n_b|\mathbf{G}+\mathbf{k}|^2 C_{\mathbf{G}}^{b} + \frac{n_b}{\sqrt{V}}\sum_{\mathbf{G'}} V_{\mathbf{G'}}^{eff}\, C_{\mathbf{G-G'}}^{b} = n^b \epsilon^b C_{\mathbf{G}}^{b}$$

$$\sum_{\mathbf{G''}}\frac{1}{2}n_b|\mathbf{G}+\mathbf{k}|^2 C_{\mathbf{G''}}^{b}\delta_{\mathbf{G},\mathbf{G''}} + \frac{n_b}{\sqrt{V}}\sum_{\mathbf{G''}} V_{\mathbf{G-G''}}^{eff}\, C_{\mathbf{G''}}^{b} = n^b \epsilon^b C_{\mathbf{G}}^{b}$$

where $\mathbf{G}-\mathbf{G'} = \mathbf{G''}$.

So, finally get:

$$\sum_{\mathbf{G'}} \left[\frac{1}{2}|\mathbf{G}+\mathbf{k}|^2 \delta_{\mathbf{G},\mathbf{G'}} + \frac{1}{\sqrt{V}} V^{eff}_{\mathbf{G}-\mathbf{G'}} \right] C^b_{\mathbf{G'}} = \epsilon^b C^b_{\mathbf{G}} \tag{3.2}$$

This is equivalent to the matrix eigenvalue problem $\mathcal{H}\mathbf{c} = \epsilon\mathbf{c}$, where Hermitian matrix

$$H_{\mathbf{G}\mathbf{G'}} = \frac{1}{2}|\mathbf{G}+\mathbf{k}|^2 \delta_{\mathbf{G},\mathbf{G'}} + \frac{1}{\sqrt{V}} V^{eff}_{\mathbf{G}-\mathbf{G'}} \tag{3.3}$$

$$c = \begin{pmatrix} c_1 \\ c_2 \\ . \\ c_{\mathbf{G}} \end{pmatrix}$$

3.2   Direct minimization methods

The matrix eigenvalue method of the last section is efficient for small systems but impractical for large system. Thus the most important and practical method is direct minimization or the conjugate-gradient method.

Let's review the methods of minimization for simple one dimensional functions before going into the detail of conjugate-gradient method.

3.2.1   Golden section search

Suppose we know the function values at three points f(a), f(b), f(c), and if f(b)<f(a) and f(b)<f(c), then there must be a minimum between a and c. Choose the lengths such that $\frac{|ab|}{|bc|} = \frac{|bc|}{|ac|} = \frac{\sqrt{5}-1}{2} = 0.61803 = \gamma$. Then choose a point x at the longer segment (suppose it's bc) such that $\frac{|xb|}{|cx|} = \gamma$. If f(x)<f(b) and f(x)<f(c), then update (a,b,c) by (b,x,c); else, update (a,b,c) by (a,b,x). Repeat the whole procedure again for new point x'. The value of γ is called golden ratio.

Think over the advantages and disadvantages of Golden section search method.

3.2.2   Parabolic interpolation

If the function f(x) is sufficiently smooth, the parabola would give a good fit for the curve f(x). Successive applications would give a good approximation to the minimum.

3.2.3   Brent's method

It is a sutble combination of the above two methods.

3.2.4 Newton's iteration method

To find the minimum of f(x) is same as solving $\frac{\partial f}{\partial x} = 0$. The Newton's method gives:

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

However, this is not suitable for the present case because it may not converge and it is difficult to compute $\frac{\partial^2 E}{(\partial C_{\mathbf{G}}^b)^2}$.

3.2.5 Steepest descents

Let $F(\mathbf{x})$ be a function over an n-dimensional vector $\mathbf{x}$. The steepest descent direction is:

$$g^1 = -\frac{\partial F}{\partial \mathbf{x}}\Big|_{\mathbf{x}=\mathbf{x}^1} g^1 = -\mathbf{G}\mathbf{x}^1$$

Consider $\mathbf{x} = \mathbf{x}^1 + b^1 g^1$. Fix $\mathbf{x}^1$ and g^1, $f(b^1) = F(\mathbf{x}^1 + b^1 g^1)$. Find b^1 such that $F(\mathbf{x}^1 + b^1 g^1)$ is minimum. This is a one-dimensional problem. The minimum is obtained at $\frac{\partial f}{\partial b^1} = 0$ or $\frac{\partial \mathbf{x}}{\partial b^1}\frac{\partial F}{\partial \mathbf{x}}\Big|_{\mathbf{x}^1+b^1 g^1} = g^1 \mathbf{G}(\mathbf{x}^1 + b^1 g^1) = 0$. Thus we can find the value of b^1 at which $F(\mathbf{x}^1 + b^1 g^1)$ is minimized, then find the value of $\mathbf{x}^1 + b^1 g^1$, which is the new value x^2. Repeat the procedure to get $g^2 = -\mathbf{G}\mathbf{x}^2$ and $\mathbf{x}^3 = \mathbf{x}^2 + b^2 g^2$. The iteration is stopped when $|x^{n+1} - x^n| < \epsilon$ for some tolerance ϵ.

Example:

Consider a simple function $F(x,y) = \frac{1}{2}(x^2 + y^2)$. The gradient at point $\mathbf{x} = (x,y)$ is $\mathbf{G}\mathbf{x} = (\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}) = (x,y)$. The starting point is now chosen at $\mathbf{x}^1 = (1,1)$, then

$$g^1 = -\mathbf{G}\mathbf{x}^1 = (-1,-1)$$
$$x^2 = \mathbf{x}^1 + b^1 g^1 = (1 - b^1, 1 - b^1)$$
$$f(b^1) = F(x^2) = \frac{1}{2}((1 - b^1)^2 + (1 - b^1)^2) = (1 - b^1)^2$$
$$\frac{\partial f}{\partial b^1} = 2(b^1 - 1) = 0$$
$$b^1 = 1$$
$$x^2 = (0,0)$$
$$g^2 = -\mathbf{G}\mathbf{x}^2 = (0,0)$$
$$x^3 = \mathbf{x}^2 + b^2 g^2 = x^2$$

The minimum is $x^2 = (0,0)$ which is found in one step. In general, we might need infinite number of iterations to converge to the minimum when $\mathbf{G}\mathbf{x}_{min} = 0$.

3.2.6   Conjugate gradient method

The conjugate gradient method is one of the best method in multi-dimensional minimization. It converges much faster than steepest decent method. The basic idea is that the current search direction should be "orthogonal" to the previous search direction. More precisely, the two directions should be conjugate to each other.

Consider a function $F(\mathbf{x}) = \frac{1}{2}\mathbf{x} \cdot \mathbf{G}\mathbf{x}$, which is a matrix multiplication. $g = -\mathbf{G}\mathbf{x}$. $F(\mathbf{x})$ is minimized along direction d^1 from x^1. $f(b^1) = F(\mathbf{x}^1 + b^1 d^1)$. $\frac{\partial f}{\partial b^1} = 0$ implies

$$\frac{\partial \mathbf{x}}{\partial b^1}\frac{\partial F}{\partial \mathbf{x}}\bigg|_{\mathbf{x}^1 + b^1 g^1} = d^1 \mathbf{G}(\mathbf{x}^1 + b^1 d^1) = 0 \tag{3.4}$$

At $\mathbf{x}^2 = \mathbf{x}^1 + b^1 d^1$, $F(\mathbf{x})$ is minimized along d^2 and gives $\mathbf{x}^3 = \mathbf{x}^2 + b^2 d^2$. The condition for minimization is

$$d^2 \mathbf{G}(\mathbf{x}^2 + b^2 d^2) = 0 \tag{3.5}$$

However, the best choice of b^1 and b^2 is that $f(b^1, b^2) = F(x^3 = x^1 + b^1 d^1 + b^2 d^2)$ is minimized with respect to b^1 and b^2 separately, i.e.

$$\frac{\partial f}{\partial b^1} = d^1 \frac{\partial F}{\partial \mathbf{x}}\bigg|_{\mathbf{x}^3} = d^1 \mathbf{G}(\mathbf{x}^1 + b^1 d^1 + b^2 d^2) = 0 \tag{3.6}$$

$$\frac{\partial f}{\partial b^2} = d^2 \frac{\partial F}{\partial \mathbf{x}}\bigg|_{\mathbf{x}^3} = d^2 \mathbf{G}(\mathbf{x}^1 + b^1 d^1 + b^2 d^2) = 0 \tag{3.7}$$

In order for (3.4) and (3.5) consistent with (3.6) and (3.7), it is required

$$d^1 \mathbf{G} d^2 = 0$$
$$d^2 \mathbf{G} d^1 = 0 \tag{3.8}$$

This is the conjugate condition.

The search direction algorithm is

$$\begin{cases} d^m = g^m + \gamma^m d^{m-1} \\ \gamma^m = \frac{g^m g^m}{g^{m-1} g^{m-1}} \end{cases} \tag{3.9}$$

where $m \geq 1$, $\gamma^1 = 0$.

3.3 Minimization of $E[C_{\mathbf{G}}^b]$

This is a multi-dimensional minimization problem. Since there are many variables $C_{\mathbf{G}}^b$ for various b and $\mathbf{G}$ value, the standard way is to update one band each time.

$$
\begin{aligned}
g_{\mathbf{G}}^b &= -\frac{\partial E}{\partial C_{\mathbf{G}}^b} \\
&\qquad (3.1) gives \\
\frac{\partial E}{\partial C_{\mathbf{G}}^b} &= \frac{1}{2}n_b|\mathbf{G}+\mathbf{k}|^2 C_{\mathbf{G}}^b + \frac{n_b}{\sqrt{V}}\sum_{\mathbf{G}'} V_{\mathbf{G}'}^{eff}\, C_{\mathbf{G}-\mathbf{G}'}^b \\
&= n_b\sum_{\mathbf{G}'} \mathcal{H}_{\mathbf{G}\mathbf{G}'} C_{\mathbf{G}'}^b \\
&\qquad define\ a\ new\ direction \\
g'^b_{\mathbf{G}} &= -\frac{1}{n_b}\frac{\partial E}{\partial C_{\mathbf{G}}^b} = -\sum_{\mathbf{G}'}\mathcal{H}_{\mathbf{G}\mathbf{G}'}C_{\mathbf{G}'}^b \\
&\qquad where\ \mathcal{H}_{\mathbf{G}\mathbf{G}'}\ is\ given\ by\ (3.3) \\
g^b &= -\mathcal{H}\psi^b \qquad\qquad\qquad\qquad\qquad (3.10)
\end{aligned}
$$

The wave function is updated by $\phi^b = \psi^b + cg^b$.

3.3.1 Constraint of normalization

The new wave function is required to be normalized, which gives

$$
\begin{aligned}
\zeta^b &= -(\mathcal{H}-\lambda_b)\psi^b \\
\lambda_b &= <\psi^b|\mathcal{H}|\psi^b> = \sum_{\mathbf{G},\mathbf{G}'} \bar{C}_{\mathbf{G}}^b \mathcal{H}_{\mathbf{G}\mathbf{G}'} C_{\mathbf{G}'}^b \\
\phi^b &= \psi^b + \Delta\zeta^b \qquad\qquad\qquad\qquad\qquad (3.11)
\end{aligned}
$$

Proof that ϕ given by (3.11) is normalized:

$$
\begin{aligned}
<\phi^b|\phi^b> &= \sum_{\mathbf{G}}(<\psi^b|+\Delta<\zeta^b|)(|\psi^b>+\Delta|\zeta^b>) \\
&= <\psi^b|\psi^b> +\Delta(<\psi^b|\zeta^b>+<\zeta^b|\psi^b>)+O(\Delta^2) \\
<\psi^b|\zeta^b> &= -<\psi^b|\mathcal{H}|\psi^b>+<\psi^b|\lambda_b|\psi^b>=0 \\
<\phi^b|\phi^b> &= 1+O(\Delta^2)
\end{aligned}
$$

So, $|\phi^b>$ is approximately normalized, provided that Δ is small. A nice feature is that if $|\psi^b>$ satisfies the Schrodinger equation, then $|\zeta^b>$, i.e. the minimum is already reached.

Now the tricky problem is that the wave function must be orthogonal as well.

3.3.2   Constraint of orthogonal

To satisfy the constraint of orthogonality $< \phi^b | \phi^{b'} >= \delta_{bb'}$, where $\phi^b = \psi^b + \Delta \cdot \bar{\zeta}^b$, the new wave function is computed by:

$$\bar{\zeta}^b = \zeta^b - \sum_{b' \neq b} < \psi^{b'} | \zeta^b > \psi^{b'}$$

$$\zeta^b = -(\mathcal{H} - \lambda_b) \psi^b$$
$$\lambda_b =< \psi^b | \mathcal{H} | \psi^b >$$
$$\phi^b = \psi^b + \Delta \cdot \bar{\zeta}^b \tag{3.12}$$

Proof that ϕ given by (3.12) is orthogonal:

$$
\begin{aligned}
< \psi^b | \bar{\zeta}^{b'} > &= < \psi^b | \zeta^{b'} > - \sum_{b'' \neq b'} < \psi^{b''} | \zeta^{b'} >< \psi^b | \psi^{b''} > \\
&= < \psi^b | \zeta^{b'} > - < \psi^b | \zeta^{b'} >= 0 \\
< \psi^b | \psi^{b'} > &= \delta_{bb'} \\
< \phi^b | \phi^{b'} > &= < \psi^b + \Delta \cdot \bar{\zeta}^b | \psi^{b'} + \Delta \cdot \bar{\zeta}^{b'} > \\
&= \delta_{bb'} + O(\Delta^2)
\end{aligned}
$$

Changing wave function ψ^b imp[lies changing the electron density n($\mathbf{r}$), thus we need to compute n($\mathbf{r}$) after each iteration. $V_{\mathbf{r}}^{eff}$ needs to be updated as well. However, only one band under consideration is updated every iteration.

3.3.3   Constraint of ortho-normalization

The new wave function ϕ that is both normalized and orthogonal to each other is given by:

$$
\begin{aligned}
\phi^b &= \psi^b \cos\theta + \eta \sin\theta \\
\eta &= \eta_b = \frac{\bar{\zeta}^b}{\sqrt{< \zeta^b | \zeta^b >}} \tag{3.13}
\end{aligned}
$$

Proof that ϕ given by (3.13) is ortho-normal:

$$
\begin{aligned}
< \zeta^b | \zeta^{b' \neq b} > &= 0 \quad \Rightarrow \quad < \eta | \eta_{b' \neq b} >= 0 \\
< \eta_b | \eta_b > &= 1 \quad \Rightarrow \quad < \eta_b | \eta_{b'} >= \delta_{bb'} \\
< \psi^b | \bar{\zeta}^{b'} > &= 0 \quad \Rightarrow \quad < \psi^b | \eta_{b'} >= 0 \\
< \phi^b | \phi^{b'} > &= < \psi^b \cos\theta + \eta \sin\theta | \psi^{b'} \cos\theta + \eta_{b'} \sin\theta > \\
&= < \psi^b | \psi^{b'} > \cos^2\theta + < \eta_b | \eta_{b'} > \sin^2\theta \\
&= \delta_{bb'} \cos^2\theta + \delta_{bb'} \sin^2\theta \\
&= \delta_{bb'}
\end{aligned}
$$

3.3.4   Search for minimum E_{tot}

In each iteration, the steepest descent vector is given by η as in (3.13), and the wave function ψ is also updated according to (3.13). In each iteration, the only parameter that is varied is θ, thus E_{tot} is a function of ψ.

$$E(\theta) \;\;=\;\; E_{avg} + \sum_{n=1}^{\infty}[A_n \cos(2n\theta) + B_n \sin(2n\theta)]$$

The reason we can make a Fourier series expansion of $E(\theta)$ is that $E(\theta)$ is periodic.

It turns out that the whole expansion is dominated by the first term n=1.

$$E(\theta) \;\;=\;\; E_{avg} + A_1 \cos(2\theta) + B_1 \sin(2\theta) \tag{3.14}$$

Solve E_{avg}, A_1 and B_1

To fix E_{avg}, A_1 and B_1, we can compute three terms $\frac{\partial E}{\partial \theta}\big|_{\theta \to 0}$, $E_{(\theta=0)}$ and $E_{(\theta=\frac{\pi}{2})}$. Express every terms in $C_{\mathbf{G}}^b$, then (3.10 gives $\frac{\partial E}{\partial \phi^b} = \mathcal{H}\phi^b$.

$$\frac{\partial E}{\partial \theta}\bigg|_{\theta \to 0} \;\;=\;\; \frac{\partial \phi^b}{\partial \theta}\frac{\partial E}{\partial \phi^b} + \frac{\partial \bar{\phi}^b}{\partial \theta}\frac{\partial E}{\partial \bar{\phi}^b}$$

$$\frac{\partial \phi^b}{\partial \theta}\bigg|_{\theta \to 0} \;\;=\;\; -\psi^b \sin\theta + \eta\cos\theta\bigg|_{\theta \to 0} = \eta$$

$$\frac{\partial E}{\partial \theta}\bigg|_{\theta \to 0} \;\;=\;\; <\eta|\mathcal{H}|\phi^b> + complex\ conjugate\ term$$

$$\;\;=\;\; 2Re. <\eta|\mathcal{H}|\phi^b> \tag{3.15}$$

$$E_{(\theta=0)} \;\;=\;\; E_{(\psi^b)} \tag{3.16}$$

$$E_{(\theta=\frac{\pi}{2})} \;\;=\;\; E_{(\eta)} \tag{3.17}$$

We may also compute $E_{(\theta=0)}$, $E_{(\theta=\frac{\pi}{4})}$ and $E_{(\theta=\frac{\pi}{2})}$ to solve for E_{avg}, A_1 and B_1. The results by the above two ways are slightly different. Think over the reason of it if you are interested.

After solving the value of parameters E_{avg}, A_1 and B_1, we can find θ_{min} at which $E(\theta)$ is minimized by computing $\frac{\partial E(\theta)}{\partial \theta} = 0$.

3.4   Car-Parrinello method

The Car-Parrinello method is so famous that we can not finish the course without mentioning it. This method was discussed before the use of direct minimization.

Car and Parrinello consider the minimization by introducing an artificial molecular dynamics. The Lagrangian is defined as:

$$L = \sum_{b,\mathbf{G}} \mu |\dot{C}_{\mathbf{G}}^b|^2 - E[C_{\mathbf{G}}^b]$$

However, because of the ortho-normalization constraint in $C_{\mathbf{G}}^b$, we must add a Lagrangian multiplier Λ_{ij}. So the Lagrangian with constraint is:

$$L = \sum_{b,\mathbf{G}} \mu |\dot{C}_{\mathbf{G}}^b|^2 - E[C_{\mathbf{G}}^b] + \sum_{b,b'} \Lambda_{bb'} [\sum_{\mathbf{G}} \bar{C}_{\mathbf{G}}^b C_{\mathbf{G}}^{b'} - \delta_{bb'}]$$

The equation of motion is given by Hamiltonian mechanics

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{C}_{\mathbf{G}}^b} - \frac{\partial L}{\partial C_{\mathbf{G}}^b} = 0$$

or explicitly

$$\mu \ddot{C}_{\mathbf{G}}^b = -\sum_{\mathbf{G}'} \mathcal{H}_{\mathbf{G}\mathbf{G}'} C_{\mathbf{G}'}^b + \sum_{b'} \Lambda_{bb'} C_{\mathbf{G}}^{b'}$$

$$\mu \ddot{\psi}^b = -\mathcal{H}\psi^b + \sum_{b'} \Lambda_{bb'} \psi^{b'}$$

Since $\Lambda_{bb'}$ is unknown, we use a simplification $\Lambda_{bb'} = \lambda_b \delta_{bb'}$, $\lambda_b = <\psi^b|\mathcal{H}|\psi^b>$.

$$\mu \ddot{\psi}^b = -[\mathcal{H} - \lambda_b]\psi^b \qquad (3.18)$$

Integrate the equation of motion (3.18) using the Verlet algorithm

$$\psi^b(t + \Delta t) = 2\psi^b(t) - \psi^b(t - \Delta t) + (\Delta t)^2 \ddot{\psi}^b(t)$$

This is just a central difference approximation of the second derivative.

Damping

Since our aim is to find $C_{\mathbf{G}}^b$ such that E_{tot} is minimized, but the Lagrangian dynamics (the above molecular dynamics) conserve energy, a damping scheme is essential here. The damping can be incorperated by adding a term $-\gamma \dot{\psi}^b$ in the equation (3.18).

Another approach is renormalization of $\dot{C}_{\mathbf{G}}^b$. Set the velocity $\dot{C}_{\mathbf{G}}^b$ to zero every iteration after updating the wave funcitions.

3.5   Lab 3

Based on the program of computing E_{tot} as a function of lattice constant a for the Si crystal which is already done in the lab 2, work out a complete program or several programs to find the lattice constant a at which $E(\theta_{min})$ given by (3.14) is minimized. Here are some guidelines.

- Initialization
  There are 8 atoms per unit cell, with coordinates given in lab 2. There are alltogether 16 wave functions or 16 bands b for 32 electrons. Initialize $C_{\mathbf{G}}^b$ (b=1,2...16) with random number generator rand48(). (Notice that you may also try other random number generators like random(), yet the period of random numbers generated by rand48() is largest as you may have learned in the previous course *Computational Techniques in Theoretical Physics*.)

  The wave functions should be orthogonal, so the random numbers generated must be orthogonalized using the Gram-Schmidt scheme, which you may have learned in the previous course *Matrix Computation*, or see page 1068 of Payne. The Gram-Schmidt scheme is:

$$\phi^b = \psi^b - \sum_{b' < b} < \phi^{b'} | \psi^b > \psi^{b'}$$

$$\phi^b = \frac{\phi^b}{|\phi^b|}$$

- Run over each band b
  Compute E_{tot} and $\frac{\partial E}{\partial C_{\mathbf{G}}^b}$ and compute the steepest descent η (given by (3.11) to (3.13)) and $E(\theta_{min})$ (given by (3.14) for each band b. Repeat the computation for the total 16 bands.

- The loop structure
  Initializations;
  for each value of lattice constant a, do
  while $(|E_{n-1} - E_n| > tolerance\epsilon)$, do
  for (b=1 to 16), do

  - compute $n_{\mathbf{G}}$ from $C_{\mathbf{G}}^b$
  - compute E_{tot}
  - compute $\mathcal{H} | \psi^b >$
  - compute steepest descent vector η
  - compute θ_{min} such that $E(\theta)$ is a minimum
  - update the wave functions ψ^b

  end for-loop
  end while-loop
  end for-loop

3.6   Tutorial 3

Q1: How small should be the tolerance ϵ, s.t. $|x_k - x_{k+1}| < \epsilon$ considered to converge?

Q2: Prove that algorithm (3.9) satisfies the conjugate condition $d^n \mathbf{G} d^m = 0$ for $n \neq m$.

Q3: What do the foumulae for ζ^b and $\bar{\zeta}^b$ given by (3.11) and (3.12) mean in terms of $C_{\mathbf{G}}^b$?