



## Density functional theory calculations using the finite element method

Ondřej Čertík<sup>a\*</sup>, Jiří Vackář<sup>a</sup>, and Jiří Plešek<sup>b</sup>

<sup>a</sup> Institute of Physics, Academy of Sciences of the Czech Republic, Cukrovarnická 10, 162 53 Praha 6, Czech Republic; [vackar@fzu.cz](mailto:vackar@fzu.cz)

<sup>b</sup> Institute of Thermomechanics, Academy of Sciences of the Czech Republic, Dolejškova 1402/5, 182 00 Praha 8, Czech Republic; [plesek@it.cas.cz](mailto:plesek@it.cas.cz)

Received 5 February 2008, in revised form 20 May 2008

**Abstract.** We propose a method to solve Kohn–Sham equations and to calculate electronic states, total energy, and material properties of non-crystalline, non-periodic structures with  $l$ -dependent fully non-local real-space *ab initio* pseudopotentials using finite elements. Contrary to the variety of well established  $k$ -space methods, which are based on Bloch’s theorem and applicable to periodic structures, we do not assume periodicity in any respect. Precise *ab initio* environment-reflecting pseudopotentials that have been applied in the  $k$ -space, plane wave approach so far, are connected with real space finite-element basis in this work. The main expected asset of the present approach is the combination of efficiency and high precision of *ab initio* pseudopotentials with applicability not restricted to periodic environment.

**Key words:** DFT, pseudopotentials, electronic structure, finite elements.

### 1. INTRODUCTION

In this paper we propose a method to solve Kohn–Sham equations and to calculate electronic states, total energy, and material properties of non-crystalline, non-periodic structures with  $l$ -dependent fully non-local real-space *ab initio* pseudopotentials using finite elements.

Contrary to the variety of well-established  $k$ -space methods that are based on Bloch’s theorem and applicable to periodic structures, we do not assume periodicity in any respect. Precise *ab initio* environment-reflecting pseudopotentials, which have been applied in the  $k$ -space, plane wave approach so far, are connected with real space finite-element basis in our work. The main expected asset of the present approach is the combination of efficiency and high precision of *ab initio* pseudopotentials with applicability not restricted to periodic environment.

In Section 2 we discuss a Density Functional Theory (DFT) and how it is applied to electronic structure calculations. We try to derive all equations that are needed for our work and it should serve as a tutorial. In Section 3 we show some preliminary results of our code for one atomic calculations. In the next section we show incorporation of pseudopotentials together with equations and some other technical difficulties that have to be overcome in order to solve the problem. Our own contribution is to sort all the ideas out and present them in a fashion applicable to our problem. In the last section we show how to solve the Kohn–Sham equations using a finite element method – again pieces of the information could be found elsewhere in the

\* Corresponding author, [certik@fzu.cz](mailto:certik@fzu.cz)

literature, but our contribution is showing how the density functional theory problem with pseudopotentials could be solved using the finite element method.

## 2. DENSITY FUNCTIONAL THEORY

### 2.1. Introduction

This section serves as an introduction to DFT and we try to derive all equations that are used in our program. We tried not to skip any important steps; however, we suggest that the literature be consulted if something is not clear enough. Excellent references are [2], [5], and [6].

### 2.2. Many-body Schrödinger equation

We use the Born–Oppenheimer approximation, which says that the nuclei of the treated atoms are seen as fixed. A stationary electronic state (for  $N$  electrons) is then described by a wave function  $\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$  fulfilling the many-body Schrödinger equation

$$\hat{H}|\Psi\rangle = (\hat{T} + \hat{U} + \hat{V})|\Psi\rangle = E|\Psi\rangle,$$

where

$$\hat{T} = \sum_i^N -\frac{1}{2}\nabla_i^2$$

is the kinetic term,

$$\hat{U} = \sum_{i<j} U(\mathbf{r}_i, \mathbf{r}_j) = \frac{1}{2} \sum_{i,j} U(\mathbf{r}_i, \mathbf{r}_j),$$

$$U(\mathbf{r}_i, \mathbf{r}_j) = U(\mathbf{r}_j, \mathbf{r}_i) = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

is the electron–electron interaction term, and

$$\hat{V} = \sum_i^N v(\mathbf{r}_i),$$

$$v(\mathbf{r}_i) = \sum_k -\frac{Z_k}{|\mathbf{r}_i - \mathbf{R}_k|}$$

is the interaction term between electrons and nuclei, where  $\mathbf{R}_k$  are positions of nuclei and  $Z_k$  the number of nucleons in each nucleus (we are using atomic units). So for one atomic calculation with the atom nucleus in the origin, we have just  $v(\mathbf{r}_i) = -\frac{Z}{|\mathbf{r}_i|}$ .

The formula  $|\Psi|^2 = \Psi^*\Psi$  gives the probability density of measuring the first electron at the position  $\mathbf{r}_1$ , the second at  $\mathbf{r}_2$ , ..., and the  $N$ th electron at the position  $\mathbf{r}_N$ . The normalization is such that  $\int |\Phi|^2 d^3r_1 d^3r_2 \dots d^3r_N = 1$ . The  $\Psi$  is antisymmetric, i.e.  $\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = -\Psi(\mathbf{r}_2, \mathbf{r}_1, \dots, \mathbf{r}_N) = -\Psi(\mathbf{r}_1, \mathbf{r}_N, \dots, \mathbf{r}_2)$ , etc.

Integrating  $|\Psi|^2$  over the first  $N - 1$  electrons is the probability density that the  $N$ th electron is at the position  $\mathbf{r}_N$ . Thus the probability density  $n(\mathbf{r})$  that any of the  $N$  electrons (i.e the first, or the second, or the

third, ..., or the  $N$ th) is at the position  $\mathbf{r}$  is called the particle (or charge or electron) density and is therefore given by

$$\begin{aligned}
n(\mathbf{r}) &= \int \Psi^*(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) d^3 r_2 d^3 r_3 \dots d^3 r_N \\
&\quad + \int \Psi^*(\mathbf{r}_1, \mathbf{r}, \dots, \mathbf{r}_N) \Psi(\mathbf{r}_1, \mathbf{r}, \dots, \mathbf{r}_N) d^3 r_1 d^3 r_3 \dots d^3 r_N + \dots \\
&\quad + \int \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}) \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}) d^3 r_1 d^3 r_2 d^3 r_3 \dots d^3 r_{N-1} \\
&= \int (\delta(\mathbf{r} - \mathbf{r}_1) + \delta(\mathbf{r} - \mathbf{r}_2) + \dots + \delta(\mathbf{r} - \mathbf{r}_N)) \\
&\quad \times \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) d^3 r_1 d^3 r_2 d^3 r_3 \dots d^3 r_N \\
&= \sum_{i=1}^N \int \langle \Psi | \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N \rangle \delta(\mathbf{r} - \mathbf{r}_i) \langle \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N | \Psi \rangle d^3 r_1 d^3 r_2 d^3 r_3 \dots d^3 r_N \\
&= N \int \langle \Psi | \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N \rangle \delta(\mathbf{r} - \mathbf{r}_1) \langle \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N | \Psi \rangle d^3 r_1 d^3 r_2 d^3 r_3 \dots d^3 r_N \\
&= N \int \Psi^*(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) \Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) d^3 r_2 d^3 r_3 \dots d^3 r_N. \tag{1}
\end{aligned}$$

Thus  $\int_{\Omega} n(\mathbf{r}) d^3 r$  gives the number of particles (and also the amount of charge) in the region of integration  $\Omega$ . Obviously  $\int n(\mathbf{r}) d^3 r = N$ .

The energy of the system is given by

$$E = \langle \Psi | \hat{H} | \Psi \rangle = \langle \Psi | \hat{T} | \Psi \rangle + \langle \Psi | \hat{U} | \Psi \rangle + \langle \Psi | \hat{V} | \Psi \rangle = T + U + V, \tag{2}$$

where

$$T = \langle \Psi | \hat{T} | \Psi \rangle = \sum_i^N \int \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \left(-\frac{1}{2} \nabla_i^2\right) \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) d^3 r_1 d^3 r_2 \dots d^3 r_N,$$

$$U = \langle \Psi | \hat{U} | \Psi \rangle,$$

$$\begin{aligned}
V &= \langle \Psi | \hat{V} | \Psi \rangle = \sum_i^N \int \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) v(\mathbf{r}_i) \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) d^3 r_1 d^3 r_2 \dots d^3 r_N \\
&= \sum_i^N \int \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) v(\mathbf{r}_1) \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) d^3 r_1 d^3 r_2 \dots d^3 r_N \\
&= N \int \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) v(\mathbf{r}_1) \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) d^3 r_1 d^3 r_2 \dots d^3 r_N \\
&= \int v(\mathbf{r}) n(\mathbf{r}) d^3 r = V[n]. \tag{3}
\end{aligned}$$

It needs to be stressed that generally  $E$  is *not* a functional of  $n$  alone, only the  $V[n]$  is. In the next section we show, however, that if the  $|\Psi\rangle$  is a ground state (of any system), then  $E$  becomes a functional of  $n$ .

### 2.3. The Hohenberg–Kohn theorem

The SE gives the map

$$C : V \rightarrow \Psi,$$

where  $\Psi$  is the ground state and  $C$  is bijective (one-to-one correspondence), because to every  $V$  we can compute the corresponding  $\Psi$  from SE and two different  $V$  and  $V'$  (differing by more than a constant) give two different  $\Psi$ , because if  $V$  and  $V'$  gave the same  $\Psi$ , then by subtracting

$$\hat{H}|\Psi\rangle = E_{gs}|\Psi\rangle$$

from

$$\hat{H}'|\Psi\rangle = (\hat{H} - \hat{V} + \hat{V}')|\Psi\rangle = E'_{gs}|\Psi\rangle$$

we would get  $V - V' = E - E'$ , which is in contradiction to the assumption that  $V$  and  $V'$  differ by more than a constant.

Similarly, from the ground state wavefunction  $\Psi$  we can compute the charge density  $n$  giving rise to the map

$$D : \Psi \rightarrow n,$$

which is also bijective, because to every  $\Psi$  we can compute  $n$  from (1) and two different  $\Psi$  and  $\Psi'$  give two different  $n$  and  $n'$ , because different  $\Psi$  and  $\Psi'$  give

$$E_{gs} = \langle \Psi | \hat{H} | \Psi \rangle < \langle \Psi' | \hat{H} | \Psi' \rangle = \langle \Psi' | \hat{H}' + \hat{V} - \hat{V}' | \Psi' \rangle = E'_{gs} + \int n'(\mathbf{r})(v(\mathbf{r}) - v'(\mathbf{r})) d^3r,$$

$$E'_{gs} = \langle \Psi' | \hat{H}' | \Psi' \rangle < \langle \Psi | \hat{H}' | \Psi \rangle = \langle \Psi | \hat{H} + \hat{V}' - \hat{V} | \Psi \rangle = E_{gs} + \int n(\mathbf{r})(v'(\mathbf{r}) - v(\mathbf{r})) d^3r.$$

Adding these two inequalities together gives

$$0 < \int n'(\mathbf{r})(v(\mathbf{r}) - v'(\mathbf{r})) d^3r + \int n(\mathbf{r})(v'(\mathbf{r}) - v(\mathbf{r})) d^3r = \int (n(\mathbf{r}) - n'(\mathbf{r}))(v'(\mathbf{r}) - v(\mathbf{r})) d^3r,$$

which for  $n = n'$  gives  $0 < 0$ , which is nonsense, therefore  $n \neq n'$ .

So we have proved that for a given ground state density  $n_0(\mathbf{r})$  (generated by a potential  $\hat{V}_0$ ) it is possible to calculate the corresponding ground state wavefunction  $\Psi_0(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ , in other words,  $\Psi_0$  is a unique functional of  $n_0$ :

$$\Psi_0 = \Psi_0[n_0].$$

The ground state energy  $E_0$  is also a functional of  $n_0$

$$E_0 = \langle \Psi_0[n_0] | \hat{T} + \hat{U} + \hat{V}_0 | \Psi_0[n_0] \rangle = E[n_0].$$

We define an energy functional

$$E_{v_0}[n] = \langle \Psi[n] | \hat{T} + \hat{U} + \hat{V}_0 | \Psi[n] \rangle = \langle \Psi[n] | \hat{T} + \hat{U} | \Psi[n] \rangle + \int v_0(\mathbf{r})n(\mathbf{r})d^3r, \quad (4)$$

where  $|\Psi[n]\rangle$  is any ground state wavefunction (generated by an arbitrary potential), that is,  $n$  is a ground state density belonging to an arbitrary system.  $E_0$ , which is generated by the potential  $V_0$ , can then be expressed as

$$E_0 = E_{v_0}[n_0]$$

and for  $n \neq n_0$  we have (from the Ritz principle)

$$E_0 < E_{v_0}[n]$$

and one has to minimize the functional  $E_{v_0}[n]$ :

$$E_0 = \min_n E_{v_0}[n]. \quad (5)$$

The term

$$\langle \Psi[n] | \hat{T} + \hat{U} | \Psi[n] \rangle \equiv F[n]$$

in (4) is universal in the sense that it does not depend on  $\hat{V}_0$ . It can be proven [2] that  $F[n]$  is a functional of  $n$  for degenerated ground states too, so (5) stays true as well.

The ground state densities in (4) and (5) are called *pure-state V-representable* because they are the densities of the (possible degenerate) ground state of the Hamiltonian with some local potential  $v(\mathbf{r})$ . One may ask the question if all possible functions are V-representable (this is called the V-representability problem). The question is relevant, because we need to know which functions should be taken into account in the minimization process (5). Even though not every function is V-representable [2], every density defined on a grid (finite or infinite) that is strictly positive, normalized, and consistent with the Pauli principle is ensemble V-representable. Ensemble V-representation is just a simple generalization of the above, for details see [2]. In plain words, we are fine.

The functional  $E_{v_0}[n]$  in (5) depends on the particle number  $N$ , so in order to get  $n$ , we need to solve the variational formulation

$$\frac{\delta}{\delta n} \left( E_v[n] - \mu(N) \int n(\mathbf{r}) d^3 r \right) = 0,$$

so

$$\frac{\delta E_v[n]}{\delta n} = \mu(N). \quad (6)$$

Let the  $n_N(\mathbf{r})$  be the solution of (6) with a particle number  $N$  and the energy  $E_N$ :

$$E_N = E_v[n_N].$$

The Lagrangian multiplier  $\mu$  is the exact chemical potential of the system

$$\mu(N) = \frac{\partial E_N}{\partial N}$$

because

$$\begin{aligned} E_{N+\varepsilon} - E_N &= E_v[n_{N+\varepsilon}] - E_v[n_N] = \int \frac{\delta E_v}{\delta n} (n_{N+\varepsilon} - n_N) d^3 r \\ &= \int \mu(N) (n_{N+\varepsilon} - n_N) d^3 r = \mu(N) (N + \varepsilon - N) = \mu(N) \varepsilon \end{aligned}$$

so

$$\mu(N) = \frac{E_{N+\varepsilon} - E_N}{\varepsilon} \longrightarrow \frac{\partial E_N}{\partial N}.$$

## 2.4. The Kohn–Sham equations

Consider an auxiliary system of  $N$  noninteracting electrons (noninteracting gas):

$$\hat{H}_s = \hat{T} + \hat{V}_s.$$

Then the many-body ground state wavefunction can be decomposed into single particle orbitals

$$|\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)\rangle = |\psi_1(\mathbf{r})\rangle |\psi_2(\mathbf{r})\rangle \cdots |\psi_N(\mathbf{r})\rangle$$

and

$$E_s[n] = T_s[\{\psi_i[n]\}] + V_s[n],$$

where

$$T_s[n] = \langle \Psi[n] | \hat{T} | \Psi[n] \rangle = \sum_i \langle \psi_i | -\frac{1}{2} \nabla^2 | \psi_i \rangle,$$

$$V_s[n] = \langle \Psi[n] | \hat{V} | \Psi[n] \rangle = \int v_s(\mathbf{r}) n(\mathbf{r}) d^3 r.$$

From (6) we get

$$\mu = \frac{\delta E_s[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + \frac{\delta V_s[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + v_s(\mathbf{r}). \quad (7)$$

The solution to this equation gives the density  $n_s$ .

Now we want to express the energy in (2) using  $T_s$  and  $E_H$  for convenience, where  $E_H$  is the classical electrostatic interaction energy of the charge distribution  $n(\mathbf{r})$ :

$$\nabla^2 V_H = n(\mathbf{r})$$

or equivalently

$$E_H[n] = \frac{1}{2} \int \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r d^3 r',$$

$$V_H(\mathbf{r}) = \frac{\delta E_H}{\delta n(\mathbf{r})} = \frac{1}{2} \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r'. \quad (8)$$

So from (4) we get

$$E[n] = (T + U)[n] + V[n] = T_s[n] + E_H[n] + (T - T_s + U - E_H)[n] + V[n]$$

$$= T_s[n] + E_H[n] + E_{xc}[n] + V[n]. \quad (9)$$

The rest of the energy is denoted by  $E_{xc} = U - E_H + T - T_s$  and it is called the exchange and correlation energy functional. From (6)

$$\mu = \frac{\delta E[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + \frac{\delta E_H[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} + \frac{\delta V[n]}{\delta n(\mathbf{r})}.$$

From (8) we have

$$\frac{\delta E_H}{\delta n(\mathbf{r})} = V_H(\mathbf{r}),$$

from (3) we get

$$\frac{\delta V[n]}{\delta n(\mathbf{r})} = v(\mathbf{r}),$$

we define

$$\frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} = V_{xc}(\mathbf{r}), \quad (10)$$

so we arrive at

$$\mu = \frac{\delta E[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + v(\mathbf{r}). \quad (11)$$

The solution to this equation gives the density  $n$ . Comparing (11) to (7) we see that if we choose

$$v_s \equiv V_H + V_{xc} + v,$$

then  $n_s(\mathbf{r}) \equiv n(\mathbf{r})$ . So we solve the Kohn–Sham equations of this auxiliary non-interacting system

$$\left(-\frac{1}{2}\nabla^2 + v_s(\mathbf{r})\right)\psi_i(\mathbf{r}) \equiv \left(-\frac{1}{2}\nabla^2 + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + v(\mathbf{r})\right)\psi_i(\mathbf{r}) = \varepsilon_i\psi(\mathbf{r}), \quad (12)$$

which yield the orbitals  $\psi_i$  that reproduce the density  $n(\mathbf{r})$  of the original interacting system

$$n(\mathbf{r}) \equiv n_s(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2. \quad (13)$$

The sum is taken over the lowest  $N$  energies. Some of the  $\psi_i$  can be degenerated, but it does not matter – the index  $i$  counts every eigenfunction including all the degenerated ones. In plain words, the trick is in realizing that the ground state energy can be found by minimizing the energy functional (4) and in rewriting this functional into the form (9), which shows that the interacting system can be treated as a noninteracting one with a special potential.

## 2.5. The exchange and correlation term

The exchange and correlation functional

$$E_{xc}[n] = (T + U)[n] - E_H[n] - T_S[n]$$

can always be written in the form

$$E_{xc}[n] = \int n(\mathbf{r}')\varepsilon_{xc}(\mathbf{r}';n)d^3r',$$

where the  $\varepsilon_{xc}(\mathbf{r}';n)$  is called the  $xc$  energy density.

Unfortunately, no one knows  $\varepsilon_{xc}(\mathbf{r}';n)$  exactly (yet). The most simple approximation is the *local density approximation* (LDA), for which the  $xc$  energy density  $\varepsilon_{xc}$  at  $\mathbf{r}$  is taken as that of a homogeneous electron gas (the nuclei are replaced by a uniform positively charged background, density  $n = \text{const}$ ) with the same local density:

$$\varepsilon_{xc}(\mathbf{r};n) \approx \varepsilon_{xc}^{LD}(n(\mathbf{r})).$$

The  $xc$  potential  $V_{xc}$  defined by (10) is then

$$V_{xc}(\mathbf{r};n) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} = \varepsilon_{xc}(\mathbf{r}';n) + \int n(\mathbf{r}')\frac{\delta \varepsilon_{xc}(\mathbf{r}';n)}{\delta n(\mathbf{r})}d^3r',$$

which in the LDA becomes

$$V_{xc}(\mathbf{r};n) = \varepsilon_{xc}^{LD}(n) + n\frac{d\varepsilon_{xc}^{LD}(n)}{dn} = \frac{d}{dn}(n\varepsilon_{xc}^{LD}(n)) = V_{xc}^{LD}(n). \quad (14)$$

The  $xc$  energy density  $\varepsilon_{xc}^{LD}$  of the homogeneous gas can be computed exactly [5]:

$$\varepsilon_{xc}^{LD}(n) = \varepsilon_x^{LD}(n) + \varepsilon_c^{LD}(n),$$

where the  $\varepsilon_x^{LD}$  is the electron gas exchange term given by [5]

$$\varepsilon_x^{LD}(n) = -\frac{3}{4\pi}(3\pi^2n)^{\frac{1}{3}}.$$

The rest of  $\varepsilon_{xc}^{LD}$  is hidden in  $\varepsilon_c^{LD}(n)$  for which there does not exist any analytic formula, but the correlation energies are known exactly from quantum Monte Carlo (QMC) calculations by Ceperley and Alder, see for example the review article [6] for more information. The energies were fitted by Vosko, Wilkes, and Nussair (VWN) with  $\varepsilon_c^{LD}(n)$  and they got accurate results with errors less than 0.05 mRy in  $\varepsilon_c^{LD}$ , which means that  $\varepsilon_c^{LD}(n)$  is virtually known exactly. The VWN result:

$$\varepsilon_c^{LD}(n) \approx \frac{A}{2} \left\{ \ln \left( \frac{y^2}{Y(y)} \right) + \frac{2b}{Q} \arctan \left( \frac{Q}{2y+b} \right) - \frac{by_0}{Y(y_0)} \left[ \ln \left( \frac{(y-y_0)^2}{Y(y)} \right) + \frac{2(b+2y_0)}{Q} \arctan \left( \frac{Q}{2y+b} \right) \right] \right\},$$

where  $y = \sqrt{r_s}$ ,  $Y(y) = y^2 + by + c$ ,  $Q = \sqrt{4c - b^2}$ ,  $y_0 = -0.10498$ ,  $b = 3.72744$ ,  $c = 12.93532$ ,  $A = 0.0621814$ , and  $r_s$  is the electron gas parameter, which gives the mean distance between electrons (in atomic units):

$$r_s = \left( \frac{3}{4\pi n} \right)^{\frac{1}{3}}.$$

The  $xc$  potential is then computed from (14):

$$V_{xc}^{LD} = V_x^{LD} + V_c^{LD},$$

$$V_x^{LD} = -\frac{1}{\pi} (3\pi^2 n)^{\frac{1}{3}},$$

$$V_c^{LD} = \frac{A}{2} \left\{ \ln \left( \frac{y^2}{Y(y)} \right) + \frac{2b}{Q} \arctan \left( \frac{Q}{2y+b} \right) - \frac{by_0}{Y(y_0)} \left[ \ln \left( \frac{(y-y_0)^2}{Y(y)} \right) + \frac{2(b+2y_0)}{Q} \arctan \left( \frac{Q}{2y+b} \right) \right] \right\} - \frac{A}{6} \frac{c(y-y_0) - by_0y}{(y-y_0)Y(y)}.$$

The LDA, although very simple, is surprisingly successful. More sophisticated approximations exist, for example the generalized gradient approximation (GGA), which sometimes gives better results than the LDA, but is not perfect either. Other options include orbital-dependent (implicit) density functionals or linear response type functionals, but this topic is still evolving. The conclusion is that the LDA is a good approximation to start with, and only when we are not satisfied, we will have to try some more accurate and modern approximation.

RLDA: Relativistic corrections to the energy-density functional were proposed by MacDonald and Vosko and basically are just a change in  $\varepsilon_x^{LD}(n) \rightarrow \varepsilon_x^{LD}(n)R$ :

$$R = \left[ 1 - \frac{3}{2} \left( \frac{\beta\mu - \ln(\beta + \mu)}{\beta^2} \right)^2 \right],$$

where

$$\mu = \sqrt{1 + \beta^2}$$

and

$$\beta = \frac{(3\pi^2 n)^{\frac{1}{3}}}{c}.$$

We also need to calculate these derivatives:

$$\frac{dR}{d\beta} = -6 \frac{\beta\mu - \ln(\beta + \mu)}{\beta^2} \left( \frac{1}{\mu} - \frac{\beta\mu - \ln(\beta + \mu)}{\beta^3} \right),$$



$$\frac{d\beta}{dn} = \frac{\beta}{3n},$$

$$\frac{d\epsilon_x^{LD}}{dn} = \frac{\epsilon_x^{LD}}{3n}.$$

So

$$V_x^{RLD} = \epsilon_x^{LD} R + n \frac{d\epsilon_x^{LD} R}{dn} = \frac{4}{3} \epsilon_x^{LD} R + \frac{1}{3} \epsilon_x^{LD} \frac{dR}{d\beta} \beta.$$

For  $c \rightarrow \infty$  we get  $\beta \rightarrow 0$ ,  $R \rightarrow 1$  and  $V_x^{RLD} \rightarrow \frac{4}{3} \epsilon_x^{LD} = V_x^{LD}$  as expected, because

$$\lim_{\beta \rightarrow 0} \frac{\beta \sqrt{1 + \beta^2} - \ln(\beta + \sqrt{1 + \beta^2})}{\beta^2} = 0.$$

## 2.6. Iteration to self-consistency

The  $V_H$  and  $V_{xc}$  potentials in the Kohn–Sham equations (12) depend on the solution  $n$ , thus the KS equations need to be iterated to obtain a self-consistent density. One can regard the KS procedure as a nonlinear operator  $\hat{F}$ , which satisfies (at the  $M$ th iteration)

$$n_M^{\text{out}} = \hat{F} n_M$$

and the problem is to find the self-consistent density, which satisfies

$$n = \hat{F} n.$$

Due to the long-range nature of the Coulomb interaction, a small change in the input density  $n_M$  can lead to a relatively large change in the output density  $\hat{F} n_M$ , thus it is not possible to use the output density itself as the input density for the next iteration, because large unstable charge oscillations arise. Rather it is essential to mix input and output densities in an appropriate manner to obtain a new input density.

The  $n(\mathbf{r})$  is in practice defined on some grid, or using coefficients of plane waves, local orbitals, or the like, which means that the precise relation

$$\mathbb{1} = \int |\mathbf{r}\rangle \langle \mathbf{r}| d^3 r$$

is changed for

$$\mathbb{1} \approx \sum_i |\mathbf{r}_i\rangle \langle \mathbf{r}_i|$$

in the case of a grid (or some other basis like plane waves can be used instead of  $|\mathbf{r}_i\rangle$ ) and  $n(\mathbf{r}) = \langle \mathbf{r}|n\rangle$  is approximated by  $n(\mathbf{r}_i) = \langle \mathbf{r}_i|n\rangle$ . Let

$$\mathbf{x} = (x_1, x_2, x_3, \dots), \quad x_i \equiv n(\mathbf{r}_i) = \langle \mathbf{r}_i|n\rangle$$

and

$$\mathbf{F}(\mathbf{x}_M) \equiv \hat{F} n_M, \quad F_i = (\hat{F} n_M)(\mathbf{r}_i)$$

the self-consistency is reached when  $\mathbf{F}(\mathbf{x}) = \mathbf{x}$ .

So the problem is in solving the equation

$$\mathbf{F}(\mathbf{x}) = \mathbf{x},$$

where  $\mathbf{x}$  denotes a vector in many dimensions (the number of points in the grid). It can also be expressed in the form of the residual  $\mathbf{R}(\mathbf{x}) = \mathbf{F}(\mathbf{x}) - \mathbf{x}$  as

$$\mathbf{R}(\mathbf{x}) = 0.$$

Almost all of the methods start with approximating

$$\mathbf{R}(\mathbf{x}_{M+1}) - \mathbf{R}(\mathbf{x}_M) \approx \mathbf{J} \cdot (\mathbf{x}_{M+1} - \mathbf{x}_M), \quad (15)$$

where the Jacobian

$$J_{ij} = \frac{\partial R_i}{\partial x_j}.$$

We want  $\mathbf{R}(\mathbf{x}_{M+1}) = 0$ , so substituting that into (15) we get

$$\mathbf{x}_{M+1} \approx \mathbf{x}_M + \mathbf{J}^{-1} \cdot (\mathbf{R}(\mathbf{x}_{M+1}) - \mathbf{R}(\mathbf{x}_M)) = \mathbf{x}_M - \mathbf{J}^{-1} \cdot \mathbf{R}(\mathbf{x}_M).$$

If we knew the Jacobian exactly, this would be the multidimensional Newton–Raphson method, but we can only make approximations to  $\mathbf{J}$  using a sequence of  $\mathbf{J}_0, \mathbf{J}_1, \mathbf{J}_2, \dots$ :

$$\mathbf{x}_{M+1} = \mathbf{x}_M - \mathbf{J}_M^{-1} \cdot \mathbf{R}(\mathbf{x}_M) \quad (16)$$

and the rate of convergence is determined by the quality of the Jacobian. Methods of this type are called quasi-Newton–Raphson methods.

The simplest approach is to use the *linear mixing* scheme for which

$$\mathbf{J}_M^{-1} = -\alpha \mathbf{1},$$

so

$$\mathbf{x}_{M+1} = \mathbf{x}_M + \alpha \mathbf{R}(\mathbf{x}_M) = \mathbf{x}_M + \alpha (\mathbf{F}(\mathbf{x}_M) - \mathbf{x}_M),$$

where  $0 < \alpha \leq 1$  is the mixing parameter, the working value is somewhere around  $\alpha = 0.1$  to  $\alpha = 0.3$ . Unfortunately, this procedure is slow, besides we do not explore all the possible densities with this mixing, which means that we do not get the correct density with any accuracy, because we get stuck at a ‘stiff’ situation for which continued iteration does not improve the distance  $|\mathbf{R}(\mathbf{x}_M)|$  between input and output densities. On the other hand, it is very easy to implement and it works in most cases, although slowly.

Surprisingly, a very good method is this:

$$\mathbf{J}_M^{-1} = -\text{diag}(\beta_1, \beta_2, \beta_3, \dots),$$

start with  $\beta_1 = \beta_2 = \beta_3 = \dots = \alpha$  and at every iteration adjust the parameters  $\beta_i$  according to this very simple algorithm: if  $R_i(\mathbf{x}_{M-1})R_i(\mathbf{x}_M) > 0$  then increase  $\beta_i$  by  $\alpha$  (if  $\beta_i > \alpha_{max}$ , set  $\beta_i = \alpha_{max}$ ) otherwise set  $\beta_i = \alpha$ . In our tests it behaved almost as well as the second Broyden method.

A more sophisticated approach is the Broyden update, which updates the  $\mathbf{J}$  successively at every iteration. The *first Broyden method* uses this formula:

$$\mathbf{J}_{M+1} = \mathbf{J}_M - \frac{(\Delta \mathbf{R}(\mathbf{x}_M) + \mathbf{J}_M \cdot \Delta \mathbf{x}_M) \Delta \mathbf{x}_M^T}{|\Delta \mathbf{x}_M|^2},$$

which has the disadvantage that we need to compute the inverse Jacobian in (16) at every iteration, which is impossible in our case. The *second Broyden method* updates the inverse Jacobian directly using this formula

$$\mathbf{J}_{M+1}^{-1} = \mathbf{J}_M^{-1} + \frac{(\Delta \mathbf{x}_M - \mathbf{J}_M^{-1} \cdot \Delta \mathbf{R}(\mathbf{x}_M)) \Delta \mathbf{R}(\mathbf{x}_M)^T}{|\Delta \mathbf{R}(\mathbf{x}_M)|^2} \quad (17)$$

starting with the linear mixing:

$$\mathbf{J}_0^{-1} = -\alpha \mathbf{1}.$$

It is impossible to store the whole dense matrix of the inverse Jacobian, but fortunately it is not necessary, realizing that formula (17) has a very simple structure [7]:

$$\mathbf{J}_{M+1}^{-1} = \mathbf{J}_M^{-1} + \mathbf{u} \mathbf{v}^T$$

with

$$\begin{aligned} \mathbf{u} &= \Delta \mathbf{x}_M - \mathbf{J}_M^{-1} \cdot \Delta \mathbf{R}(\mathbf{x}_M), \\ \mathbf{v} &= \frac{\Delta \mathbf{R}(\mathbf{x}_M)}{|\Delta \mathbf{R}(\mathbf{x}_M)|^2}, \end{aligned} \quad (18)$$

so the whole inverse Jacobian can be written as

$$\mathbf{J}_M^{-1} = -\alpha \mathbf{1} + \mathbf{u}_1 \mathbf{v}_1^T + \mathbf{u}_2 \mathbf{v}_2^T + \mathbf{u}_3 \mathbf{v}_3^T + \dots,$$

and we only need to know how to apply such a Jacobian to an arbitrary vector, which is needed in (18) and (16):

$$\mathbf{J}_M^{-1} \cdot \mathbf{y} = -\alpha \mathbf{y} + \mathbf{u}_1 (\mathbf{v}_1^T \mathbf{y}) + \mathbf{u}_2 (\mathbf{v}_2^T \mathbf{y}) + \mathbf{u}_3 (\mathbf{v}_3^T \mathbf{y}) + \dots$$

Thus instead of the whole dense matrix, we only need to save the vectors  $\mathbf{u}$  and  $\mathbf{v}$  from every iteration.

Vanderbilt and Louie [8] suggested a *modified Broyden method*, which incorporates weights, but Eyert [3] showed that if all the weights are used to tune the iteration process to its fastest convergence, they, in fact, cancel out. The result of the scheme is called by Eyert the *generalized Broyden method*, whose scheme shown by Eyert is exactly the same as for the *Anderson mixing*:

$$\begin{aligned} \sum_{p=M-k}^{M-1} (1 + \omega_0^2 \delta_{pn}) \Delta \mathbf{R}(\mathbf{x}_n)^T \Delta \mathbf{R}(\mathbf{x}_p) \gamma_p &= \Delta \mathbf{R}(\mathbf{x}_n)^T \mathbf{R}(\mathbf{x}_M), \\ \mathbf{x}_{M+1} &= \mathbf{x}_M + \beta_M \mathbf{R}(\mathbf{x}_M) - \sum_{p=M-k}^{M-1} \gamma_p (\Delta \mathbf{x}_p + \beta_M \Delta \mathbf{R}(\mathbf{x}_p)), \end{aligned}$$

which according to Eyert should converge even faster than the second Broyden method, but it does not in our own implementation. Addition of  $\omega_0$  is just for numerical stability; a good value is  $\omega = 0.01$ , but it can also be switched off by  $\omega_0 = 0$ . The  $p$  is the number of last iterations to use, a good value according to Eyert is  $p = 5$ ,  $\beta_M$  should not influence the convergence much for  $p = 5$ .

The problem with  $n$  is that there are two conditions that need to be satisfied

$$n > 0$$

and the normalization

$$\int n(\mathbf{r}) d^3 r = Z.$$

The Newton method converges to the correct norm, but slowly. The condition  $n > 0$  however causes great instability. One option could be to use a logistic function like

$$n(r) = \frac{C}{1 + e^{-x(r)}}$$

for sufficiently large  $C$  and solve for  $x$ , which can be both positive and negative. But a more elegant solution is to mix  $V_h + V_{xc}$  instead of the densities.

### 3. EXAMPLE OF Pb ATOM CALCULATION

To illustrate the explained theory, we will show how to calculate the Pb atom. We have  $N = 82$  and

$$v(\mathbf{r}_i) = -\frac{82}{|\mathbf{r}_i|}$$

and we need to sum over the lowest 82 eigenvalues in (13). The way to do that is to calculate the eigenfunction for the lowest  $n$  and  $l$  by iterating  $n = 0, 1, 2, \dots$  and taking  $l = 1, 2, \dots, n - 1$  for each  $n$ , while summing the number of states  $s = 2(2l + 1)$  for each pair  $n, l$ , until we exceed  $Z$  and then taking another  $n$  just to be sure not to miss any state belonging to the lowest  $Z$  states. It turns out that the combination  $n, l$  does not change for an atom and it is equal to the well-known configuration of electrons that can be found in any chemistry textbook, as an example, for Pb we get (first number is  $n$ , the letter is  $l$  and the number in superscript gives the number of times the particular eigenvalue needs to be taken into account in the sum):  $1S^2, 2S^2, 2P^6, 3S^2, 3P^6, 3D^{10}, 4S^2, 4P^6, 4D^{10}, 4F^{14}, 5S^2, 5P^6, 5D^{10}, 6S^2, 6P^2$  (notice the 5F and 5G are missing). Together it is 82 eigenvalues. The KS energies for these eigenvalues are:  $-2901.078061, -488.8433352, -470.8777849, -116.526852, -107.950391, -91.88992429, -25.75333021, -21.99056413, -15.03002657, -5.592531664, -4.206797624, -2.941656967, -0.9023926829, -0.3571868295, -0.1418313263$ .

The corresponding radial charge density is plotted in Fig. 1. Our code works for any atom (see e.g. a result for boron in Fig. 2).

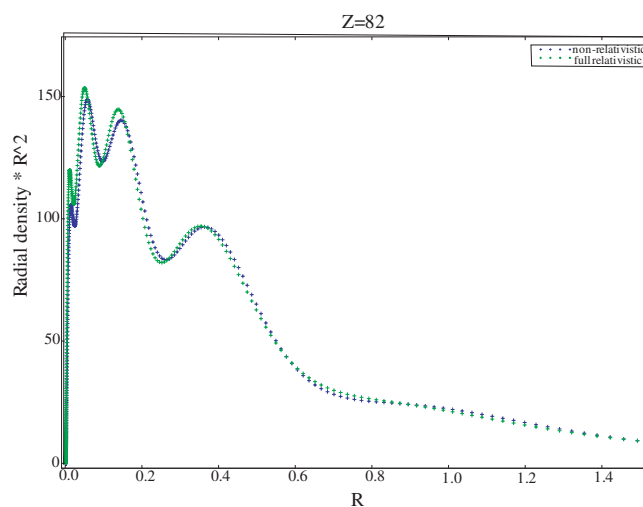


Fig. 1. Lead ( $Z = 82$ ).

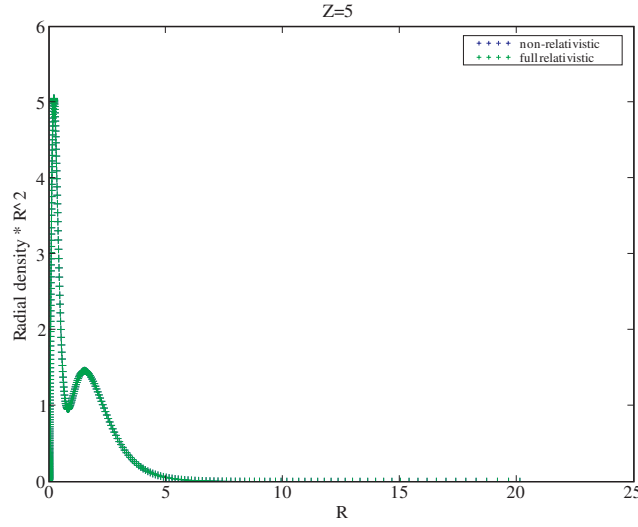


Fig. 2. Boron ( $Z = 5$ ).

## 4. PSEUDOPOTENTIALS

### 4.1. Hermitian operators in spherical symmetry

We show that every Hermitian operator  $\hat{V}$  in the spherical symmetric problem ( $\hat{V} = R^{-1}\hat{V}R$ ) can be written in the form

$$\hat{V} = \sum_{lm} |lm\rangle \hat{V}_l \langle lm|, \quad (19)$$

where the operator  $\hat{V}_l = \langle lm|\hat{V}|lm\rangle$  has matrix elements

$$\langle \rho|\hat{V}_l|\rho'\rangle = \langle lm|\langle \rho|\hat{V}|\rho'\rangle|lm\rangle = V_l(\rho, \rho').$$

*Proof.* Matrix elements of a general Hermitian operator  $\hat{V}$  are

$$\langle \mathbf{r}|\hat{V}|\varphi\rangle = \int \langle \mathbf{r}|\hat{V}|\mathbf{r}'\rangle \langle \mathbf{r}'|\varphi\rangle d^3r' = \int V(\mathbf{r}, \mathbf{r}')\varphi(\mathbf{r}')d^3r',$$

where

$$V(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r}|\hat{V}|\mathbf{r}'\rangle.$$

In spherical symmetry, we have

$$\begin{aligned} \langle \mathbf{r}|\hat{V}|\varphi\rangle &= \langle \mathbf{r}|R^{-1}\hat{V}R|\varphi\rangle = \langle \mathbf{r}|R^\dagger\hat{V}R|\varphi\rangle = \int \langle \mathbf{r}|R^\dagger\hat{V}R|\mathbf{r}'\rangle \langle \mathbf{r}'|\varphi\rangle d^3r' \\ &= \int \langle R\mathbf{r}|\hat{V}|R\mathbf{r}'\rangle \langle \mathbf{r}'|\varphi\rangle d^3r' = \int V(R\mathbf{r}, R\mathbf{r}')\varphi(\mathbf{r}')d^3r', \end{aligned}$$

where  $R$  is the rotation operator (it is unitary). We have thus derived  $V(R\mathbf{r}, R\mathbf{r}') = V(\mathbf{r}, \mathbf{r}')$  true for any  $R$ , which means that the kernel only depends on  $\rho$ ,  $\rho'$ , and  $\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'$ , where  $\mathbf{r} = \rho\hat{\mathbf{r}}$  and  $\mathbf{r}' = \rho'\hat{\mathbf{r}}'$ . So we obtain

$$V(\mathbf{r}, \mathbf{r}') = V(\rho, \rho', \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') = \sum_{lm} Y_{lm}(\hat{\mathbf{r}})V_l(\rho, \rho')Y_{lm}^*(\hat{\mathbf{r}}'),$$

where

$$V_l(\rho, \rho') = \frac{(2l+1)^2}{8\pi} \int_{-1}^1 P_l(x) V_l(\rho, \rho', x) dx.$$

In Dirac notation:

$$V(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r} | \hat{V} | \mathbf{r}' \rangle = \langle \hat{\mathbf{r}} | \langle \rho | \hat{V} | \rho' \rangle | \hat{\mathbf{r}}' \rangle = \sum_{lm'l'm'} \langle \hat{\mathbf{r}} | lm \rangle \langle lm | \langle \rho | \hat{V} | \rho' \rangle | l'm' \rangle \langle l'm' | \hat{\mathbf{r}}' \rangle.$$

From the above derivation we see that we must have

$$\langle lm | \langle \rho | \hat{V} | \rho' \rangle | l'm' \rangle = V_l(\rho, \rho') \delta_{ll'} \delta_{mm'},$$

in other words,

$$V_l(\rho, \rho') = \langle lm | \langle \rho | \hat{V} | \rho' \rangle | lm \rangle. \quad (20)$$

So we get

$$\langle \mathbf{r} | \hat{V} | \mathbf{r}' \rangle = \sum_{lm} \langle \hat{\mathbf{r}} | lm \rangle V_l(\rho, \rho') \langle lm | \hat{\mathbf{r}}' \rangle = \sum_{lm} Y_{lm}(\theta, \phi) V_l(\rho, \rho') Y_{lm}^*(\theta', \phi')$$

and

$$\hat{V} = \sum_{lm} |lm\rangle \langle lm| \hat{V} |lm\rangle \langle lm| = \sum_{lm} |lm\rangle \hat{V}_l \langle lm|,$$

where the operator  $\hat{V}_l = \langle lm | \hat{V} | lm \rangle$  only acts on the radial part of the wavefunction and according to (20) it does not depend on  $m$ . Also according to (20) its matrix elements are

$$\langle \rho | \hat{V}_l | \rho' \rangle = \langle lm | \langle \rho | \hat{V} | \rho' \rangle | lm \rangle = V_l(\rho, \rho').$$

## 4.2. Nonlocal pseudopotentials

A nonlocal pseudopotential  $\hat{V}$  is just a general Hermitian operator. We only want to construct pseudopotentials in the spherical problem, so every pseudopotential can be written in the form (19). In practice we only use either \it local\ (the operator  $\hat{V}$  is local) or \it semilocal\ (the operator  $\hat{V}$  is radially local, but angularly nonlocal) pseudopotential.

Local potential (radially and angularly local) is defined by:

$$\langle \mathbf{r} | \hat{V} | \mathbf{r}' \rangle = V(\rho) \langle \mathbf{r} | \mathbf{r}' \rangle.$$

So we can simply write

$$\hat{V} = V(\rho), \quad (21)$$

so

$$V_l(\rho, \rho') = \langle lm | \langle \rho | \hat{V} | \rho' \rangle | lm \rangle = V(\rho) \langle \rho | \rho' \rangle = V(\rho) \frac{\delta(\rho - \rho')}{\rho^2}.$$

Thus it turned out that the kernel is local and does not depend on  $l$  and we get

$$\begin{aligned} V(\mathbf{r}(\rho, \theta, \phi), \mathbf{r}'(\rho', \theta', \phi')) &= \sum_{lm} Y_{lm}(\theta, \phi) V(\rho) \frac{\delta(\rho - \rho')}{\rho^2} Y_{lm}^*(\theta', \phi') \\ &= V(\rho) \frac{1}{\rho^2 \sin \theta} \delta(\rho - \rho') \delta(\theta - \theta') \delta(\phi - \phi') = V(\rho) \delta(\mathbf{r} - \mathbf{r}') \end{aligned}$$

and

$$\langle \mathbf{r} | \hat{V} | \varphi \rangle = \int V(\rho) \delta(\mathbf{r} - \mathbf{r}') \varphi(\mathbf{r}') d^3 r' = V(\rho) \varphi(\mathbf{r}).$$

So we recover (21). But we are just fooling around, there is nothing new in these formulas.

For a semilocal potential (radially local, but angularly nonlocal), the kernel cannot depend on  $m$  and is radially local, so:

$$\langle \rho | \hat{V}_l | \rho' \rangle = V_l(\rho, \rho') = \langle lm | \langle \rho | \hat{V} | \rho' \rangle | lm \rangle = V_l(\rho) \langle \rho | \rho' \rangle = V_l(\rho) \frac{\delta(\rho - \rho')}{\rho^2},$$

so the kernel is local and does depend on  $l$  and we simply write

$$\hat{V}_l = V_l(\rho)$$

and

$$\hat{V} = \sum_{lm} |lm\rangle V_l(\rho) \langle lm|. \quad (22)$$

We can also calculate the same result explicitly in the  $\mathbf{r}$  representation:

$$V(\mathbf{r}(\rho, \theta, \phi), \mathbf{r}'(\rho', \theta', \phi')) = \sum_{lm} Y_{lm}(\theta, \phi) V_l(\rho) \frac{\delta(\rho - \rho')}{\rho^2} Y_{lm}^*(\theta', \phi')$$

and

$$\begin{aligned} \langle \mathbf{r} | \hat{V} | \varphi \rangle &= \int \sum_{lm} Y_{lm}(\theta, \phi) V_l(\rho) \frac{\delta(\rho - \rho')}{\rho^2} Y_{lm}^*(\theta', \phi') \varphi(\mathbf{r}') d^3 r' \\ &= \sum_{lm} Y_{lm}(\theta, \phi) V_l(\rho) \int Y_{lm}^*(\theta', \phi') \varphi(\rho \hat{\mathbf{r}}') d\Omega' \end{aligned}$$

or in Dirac notation

$$\langle \mathbf{r} | \hat{V} | \varphi \rangle = \sum_{lm} \langle \hat{\mathbf{r}} | lm \rangle V_l(\rho) \langle lm | \langle \rho | \varphi \rangle$$

and we recover (22).

So, to sum it up: a semilocal pseudopotential is a general Hermitian operator in the spherically symmetric problem (i.e.  $\hat{V} = R^{-1} \hat{V} R$ ) and radially local. All such operators can be written in the form (22).

Now, it can be shown that if we make the assumption of radial locality, we will get ‘correct’ wavefunctions and energies in the linear approximation. We generally only take a few terms in the expansion (22), usually only  $V_0$ ,  $V_1$ , and  $V_2$ , sometimes also  $V_3$  and  $V_4$ .

### 4.3. Separable potentials

The pseudopotential above has the form

$$\hat{V} = \sum_{lm} |lm\rangle V_l(\rho) \langle lm| = V_{loc}(\rho) + \sum_{lm} |lm\rangle [V_l(\rho) - V_{loc}(\rho)] \langle lm|,$$

which was first introduced by Hamman, Schlüter and Chiang [4]. Or, equivalently, in the  $\mathbf{r}$  representation:

$$V(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r} | \hat{V} | \mathbf{r}' \rangle = V_{loc}(\rho) \delta(\mathbf{r} - \mathbf{r}') + \frac{\delta(\rho - \rho')}{\rho^2} \sum_{lm} Y_{lm}(\hat{\mathbf{r}}) [V_l(\rho) - V_{loc}(\rho)] Y_{lm}^*(\hat{\mathbf{r}}').$$

The first term does not cause a problem. Let us denote the second term (which is semilocal) simply by  $v$ :

$$v = \sum_{lm} |lm\rangle [V_l(\rho) - V_{loc}(\rho)] \langle lm|.$$

Let us choose a complete but otherwise arbitrary set of functions  $|\phi_i\rangle$  (they contain both a radial and an angular dependence) and define a matrix  $U$  by the equation

$$\sum_j U_{ij} \langle \phi_j | v | \phi_k \rangle = \delta_{ik},$$

then  $(|\psi\rangle = |\phi_k\rangle \alpha_k)$ :

$$v|\psi\rangle = \sum_{ik} v|\phi_i\rangle \delta_{ik} \alpha_k = \sum_{ijk} v|\phi_i\rangle U_{ij} \langle \phi_j | v | \phi_k \rangle \alpha_k = \sum_{ij} v|\phi_i\rangle U_{ij} \langle \phi_j | v | \psi \rangle.$$

So any Hermitian operator (including  $v$ ) can be transformed exactly into the following form

$$v = \sum_{ij} v|\phi_i\rangle U_{ij} \langle \phi_j | v.$$

We diagonalize the matrix  $U$  by choosing such functions  $|\bar{\phi}_i\rangle$  for which the matrix  $\langle \bar{\phi}_j | v | \bar{\phi}_k \rangle$  (and hence the corresponding matrix  $U$ ) is equal to \one. We can find such functions for example using the Gram–Schmidt orthogonalization procedure on  $|\phi_i\rangle$  with a norm  $\langle f | v | g \rangle$  (for functions  $f$  and  $g$ ); more on that later. Then

$$v = \sum_i v|\bar{\phi}_i\rangle \frac{1}{\langle \bar{\phi}_i | v | \bar{\phi}_i \rangle} \langle \bar{\phi}_i | v = \sum_i v|\bar{\phi}_i\rangle \langle \bar{\phi}_i | v. \quad (23)$$

We could take any  $|\phi_i\rangle$  and orthogonalize them. However, as we have  $v$  in the form of (22), we will be using  $|\phi_i\rangle$  in the form  $|\phi_i\rangle = |R_{nl}\rangle |lm\rangle$ , because it turns out we will only need to orthogonalize the radial parts. The first term in (23) then corresponds to the KB potential. We of course take more terms and get accurate results without ghost states.

Let us look at the orthogonalization. We start with the wavefunctions:

$$|\phi_i\rangle = |R_{nl}\rangle |lm\rangle,$$

where  $R_{nl}(\rho) = \langle \rho | R_{nl} \rangle$  and  $i$  goes over all possible triplets  $(nlm)$ , for example in this order (but any order is fine):

$i$	$n$	$l$	$m$
1	1	0	0
2	2	0	0
3	2	1	-1
4	2	1	0
5	2	1	1
6	3	0	0
7	3	1	-1
8	3	1	0
9	3	1	1
10	3	2	-2
11	3	2	-1



We can also relate the  $i$  and  $n, l, m$  using this formula

$$i_{nlm} = \sum_{k=1}^{n-1} k^2 + \left( \sum_{k=0}^{l-1} (2k+1) \right) + (l+m+1) = \frac{(n-1)n(2n-1)}{6} + l(l+1) + m + 1.$$

The operator  $v$  acts on these  $|\phi_i\rangle$  like this

$$\langle \mathbf{r} | v | \phi_i \rangle = \langle \mathbf{r} | v | R_{nl} \rangle |lm\rangle = \langle \hat{\mathbf{r}} | \langle \rho | V_l(\rho) | R_{nl} \rangle |lm\rangle = V_l(\rho) R_{nl}(\rho) Y_{lm}(\hat{\mathbf{r}}).$$

Now we need to construct a new orthogonal set of functions  $|\bar{\phi}_i\rangle$  satisfying

$$\langle \bar{\phi}_i | v | \bar{\phi}_j \rangle = \delta_{ij}.$$

This can be done using several methods. We chose the Gram–Schmidt orthogonalization procedure, which works according to the following scheme:

$$\begin{aligned} |\tilde{\phi}_1\rangle &= \frac{1}{\sqrt{\langle \phi_1 | v | \phi_1 \rangle}} |\phi_1\rangle, & |\bar{\phi}_1\rangle &= \frac{1}{\sqrt{\langle \tilde{\phi}_1 | v | \tilde{\phi}_1 \rangle}} |\tilde{\phi}_1\rangle; \\ |\tilde{\phi}_2\rangle &= (\mathbb{1} - |\bar{\phi}_1\rangle \langle \bar{\phi}_1 | v) \frac{1}{\sqrt{\langle \phi_2 | v | \phi_2 \rangle}} |\phi_2\rangle, & |\bar{\phi}_2\rangle &= \frac{1}{\sqrt{\langle \tilde{\phi}_2 | v | \tilde{\phi}_2 \rangle}} |\tilde{\phi}_2\rangle; \\ |\tilde{\phi}_3\rangle &= (\mathbb{1} - |\bar{\phi}_1\rangle \langle \bar{\phi}_1 | v - |\bar{\phi}_2\rangle \langle \bar{\phi}_2 | v) \frac{1}{\sqrt{\langle \phi_3 | v | \phi_3 \rangle}} |\phi_3\rangle, & |\bar{\phi}_3\rangle &= \frac{1}{\sqrt{\langle \tilde{\phi}_3 | v | \tilde{\phi}_3 \rangle}} |\tilde{\phi}_3\rangle. \\ &\dots & & \end{aligned}$$

We can verify by a direct calculation that this procedure ensures

$$\langle \bar{\phi}_i | v | \bar{\phi}_j \rangle = \delta_{ij}.$$

It may be useful to compute the normalization factors explicitly:

$$\begin{aligned} \langle \tilde{\phi}_1 | v | \tilde{\phi}_1 \rangle &= 1, \\ \langle \tilde{\phi}_2 | v | \tilde{\phi}_2 \rangle &= 1 - \frac{\langle \phi_2 | v | \bar{\phi}_1 \rangle \langle \bar{\phi}_1 | v | \phi_2 \rangle}{\langle \phi_2 | v | \phi_2 \rangle}, \\ \langle \tilde{\phi}_3 | v | \tilde{\phi}_3 \rangle &= 1 - \frac{\langle \phi_3 | v | \bar{\phi}_1 \rangle \langle \bar{\phi}_1 | v | \phi_3 \rangle + \langle \phi_3 | v | \bar{\phi}_2 \rangle \langle \bar{\phi}_2 | v | \phi_3 \rangle}{\langle \phi_3 | v | \phi_3 \rangle}, \\ &\dots \end{aligned}$$

We can also write down a first few orthogonal vectors explicitly:

$$\begin{aligned} |\bar{\phi}_1\rangle &= \frac{|\phi_1\rangle}{\sqrt{\langle \phi_1 | v | \phi_1 \rangle}}, \\ |\bar{\phi}_2\rangle &= \frac{|\phi_2\rangle \langle \phi_1 | v | \phi_1 \rangle - |\phi_1\rangle \langle \phi_1 | v | \phi_2 \rangle}{\sqrt{(\langle \phi_1 | v | \phi_1 \rangle \langle \phi_2 | v | \phi_2 \rangle - \langle \phi_2 | v | \phi_1 \rangle \langle \phi_1 | v | \phi_2 \rangle) \langle \phi_1 | v | \phi_1 \rangle \langle \phi_2 | v | \phi_2 \rangle}}. \end{aligned}$$

Now the crucial observation is

$$\langle lm | \langle R_{nl} | v | R_{n'l'} \rangle | l'm' \rangle = \langle R_{nl} | V_l(\rho) | R_{n'l} \rangle \delta_{ll'} \delta_{mm'},$$

which means that  $\langle \phi_i | v | \phi_j \rangle = 0$  if  $|\phi_i\rangle$  and  $|\phi_j\rangle$  have different  $l$  or  $m$ . In other words  $|\phi_i\rangle$  and  $|\phi_j\rangle$  for different  $|lm\rangle$  are already orthogonal. Thus the Gram–Schmidt orthogonalization procedure only makes the  $R_{nl}$  orthogonal for the same  $|lm\rangle$ . To get explicit expressions for  $|\bar{\phi}_i\rangle$ , we simply use the formulas above and get:

$$|\phi_i\rangle = |R_{nl}\rangle |lm\rangle \quad \rightarrow \quad |\bar{\phi}_i\rangle = |\bar{R}_{nl}\rangle |lm\rangle,$$

where we have constructed new  $|\bar{R}_{nl}\rangle$  from original  $|R_{nl}\rangle$ :

$$\begin{aligned} |\bar{R}_{10}\rangle &= \frac{|R_{10}\rangle}{\sqrt{\langle R_{10} | V_0 | R_{10} \rangle}}, \\ |\bar{R}_{20}\rangle &= \frac{|R_{20}\rangle - |\bar{R}_{10}\rangle \langle \bar{R}_{10} | V_0 | R_{20} \rangle}{\sqrt{\dots}}, \\ |\bar{R}_{21}\rangle &= \frac{|R_{21}\rangle}{\sqrt{\langle R_{21} | V_1 | R_{21} \rangle}}, \\ |\bar{R}_{30}\rangle &= \frac{|R_{30}\rangle - |\bar{R}_{10}\rangle \langle \bar{R}_{10} | V_0 | R_{30} \rangle - |\bar{R}_{20}\rangle \langle \bar{R}_{20} | V_0 | R_{30} \rangle}{\sqrt{\dots}}, \\ |\bar{R}_{31}\rangle &= \frac{|R_{31}\rangle - |\bar{R}_{21}\rangle \langle \bar{R}_{21} | V_1 | R_{31} \rangle}{\sqrt{\dots}}, \\ |\bar{R}_{32}\rangle &= \frac{|R_{32}\rangle}{\sqrt{\langle R_{32} | V_1 | R_{32} \rangle}}, \\ |\bar{R}_{40}\rangle &= \frac{|R_{40}\rangle - |\bar{R}_{10}\rangle \langle \bar{R}_{10} | V_0 | R_{40} \rangle - |\bar{R}_{20}\rangle \langle \bar{R}_{20} | V_0 | R_{40} \rangle - |\bar{R}_{30}\rangle \langle \bar{R}_{30} | V_0 | R_{40} \rangle}{\sqrt{\dots}}, \\ |\bar{R}_{41}\rangle &= \frac{|R_{41}\rangle - |\bar{R}_{21}\rangle \langle \bar{R}_{21} | V_1 | R_{41} \rangle - |\bar{R}_{31}\rangle \langle \bar{R}_{31} | V_1 | R_{41} \rangle}{\sqrt{\dots}}, \\ &\dots \end{aligned}$$

OK, so we have constructed new  $|\bar{R}_{nl}\rangle$  from  $|R_{nl}\rangle$ , which obey

$$\langle \bar{R}_{nl} | V_l | \bar{R}_{n'l} \rangle = \delta_{nn'}. \quad (24)$$

For every  $V_l$ , we construct  $|\bar{R}_{nl}\rangle$  for  $n = l + 1, l + 2, \dots$ . Let us continue:

$$v |\bar{\phi}_i\rangle = V_l(\rho) |\bar{R}_{nl}\rangle |lm\rangle$$

and finally we arrive at the separable form of the  $l$  dependent pseudopotential

$$v = \sum_i v |\bar{\phi}_i\rangle \langle \bar{\phi}_i | v = \sum_i V_l(\rho) |\bar{R}_{nl}\rangle |lm\rangle \langle lm | \langle \bar{R}_{nl} | V_l(\rho). \quad (25)$$

*Note:* the  $V_l$  is actually  $V_l - V_{loc}$ , but this is just a detail.

To have some explicit formula, let us write how the separable potential acts on a wavefunction:

$$\begin{aligned}
(v\psi)(\mathbf{r}) &= \langle \mathbf{r} | v | \psi \rangle = \sum_i \langle \hat{\mathbf{r}} | lm \rangle \langle \rho | V_l(\rho) | \bar{R}_{nl} \rangle \langle \bar{R}_{nl} | V_l(\rho) \rangle \langle lm | \psi \rangle \\
&= \sum_i Y_{lm}(\hat{\mathbf{r}}) \bar{R}_{nl}(\rho) V_l(\rho) \int \bar{R}_{nl}(\rho') V_l(\rho') \int Y_{lm}^*(\hat{\mathbf{r}}') \psi(\mathbf{r}') d\Omega' \rho'^2 d\rho' \\
&= \sum_i Y_{lm}(\hat{\mathbf{r}}) \bar{R}_{nl}(\rho) V_l(\rho) \int \bar{R}_{nl}(\rho') V_l(\rho') Y_{lm}^*(\hat{\mathbf{r}}') \psi(\mathbf{r}') d^3 r'.
\end{aligned}$$

What we are actually doing is making the local potential  $V_l$  nonlocal using:

$$V_l = \sum_{n=l+1}^{\infty} V_l | \bar{R}_{nl} \rangle \langle \bar{R}_{nl} | V_l, \quad (26)$$

where

$$\langle \bar{R}_{nl} | V_l | \bar{R}_{n'l} \rangle = \delta_{nn'},$$

or in  $\mathbf{r}$  representation:

$$V_l(\rho) \psi(\rho \hat{\mathbf{r}}) = \sum_n V_l(\rho) \bar{R}_{nl}(\rho) \int \bar{R}_{nl}(\rho') V_l(\rho') \psi(\rho' \hat{\mathbf{r}}) \rho'^2 d\rho',$$

which is useful when computing integrals of this type

$$\begin{aligned}
V_{ij} &= \int \phi_i(\rho) V_l \phi_j(\rho) \rho^2 d^3 \rho = \langle i | V_l | j \rangle = \sum_n \langle i | V_l | \bar{R}_{nl} \rangle \langle \bar{R}_{nl} | V_l | j \rangle, \\
\langle i | V_l | \bar{R}_{nl} \rangle &= \int \phi_i(\rho) V_l(\rho) \bar{R}_{nl}(\rho) \rho^2 d\rho,
\end{aligned}$$

because the integral on the left hand side actually represents  $N^2$  integrals, where  $N$  is the number of basis vectors  $|\phi_i\rangle$ . The sum on the right hand side however only represents  $K \cdot N$  integrals, where  $K$  is the number of terms taken into account in (26). Of course, taking only a finite number of terms in (26) is just an approximation to  $\hat{V}_l$ . In our case, we do not need these 1D integrals (which can be easily computed directly, because  $V_l$  is local and the basis functions  $\phi_i$  are nonzero only around a node in the mesh, which means that the matrix  $V_{ij}$  is sparse), but 3D integrals, where the angular parts of  $V$  are nonlocal and the radial part is local (so the matrix  $V_{ij}$  is dense). Therefore, the above procedure is the only way how to proceed, because we decompose the matrix  $V_{ij}$  into the sum of matrices in the form  $p_i p_j^*$ , which can easily be handled and solved.

The scheme for the separation described above works for any functions  $R_{nl}(\rho)$ . Because of the form of the expansion (26) however, we will use  $R_{nl}$  from one atomic calculation. We need to approximate  $V_l$  by as few terms as possible. So imagine how the  $V_l(\rho)$  acts on the lowest radial function in the  $l$  subspace, which is  $|R_{l+1;l}\rangle$ , and we see that all the terms in (26) except the first one  $V_l | \bar{R}_{l+1;l} \rangle \langle \bar{R}_{l+1;l} | V_l$  give zero, because they are orthogonal to  $|R_{l+1;l}\rangle$ . For the function  $|R_{l+2;l}\rangle$  all the terms except the first two are zero, because  $\langle \bar{R}_{nl} | V_l | R_{l+2;l} \rangle \neq 0$  only for  $n = l + 1$  or  $n = l + 2$  (because the vectors  $|R_{l+1;l}\rangle$  and  $|R_{l+2;l}\rangle$  span the same subspace as  $|\bar{R}_{l+1;l}\rangle$  and  $|\bar{R}_{l+2;l}\rangle$  and using (24)). For functions that are a little different from all  $|R_{nl}\rangle$  ( $n > l$ ), we will generally not get precise results taking any (finite) number of terms in (26), but the higher terms should give smaller and smaller corrections.

To sum it up: we take all the  $V_l$  in (25) as we did in (22). Theoretically we should take  $\bar{R}_{nl}$  for all  $n = l + 1, l + 2, l + 3, \dots$ , but practically it suffices to only take several  $\bar{R}_{nl}$  for a given  $l$  from one atomic calculation.

Let us give an example: we are calculating 14 electrons, so we will only take into account the lowest 14 eigenvalues in the Kohn–Sham equations, which are  $|\phi_1\rangle$  up to  $|\phi_{14}\rangle$ . The lowest radial functions in each  $l$  subspace are  $|\phi_i\rangle$  for  $i = 1, 3, 4, 5, 10, 11, 12, 13, 14$ . On these 9 functions we get a precise result with only one term in the expansion (26). For the other 5 functions ( $i = 2, 6, 7, 8, 9$ ) we will have to take into account more terms. Let us look in more detail at the case  $l = 0$  (i.e.  $i = 1, 2, 6$ ). Then

$$V_0 = V_0 |\bar{R}_{10}\rangle \langle \bar{R}_{10}| V_0 + V_0 |\bar{R}_{20}\rangle \langle \bar{R}_{20}| V_0 + V_0 |\bar{R}_{30}\rangle \langle \bar{R}_{30}| V_0 + \dots$$

and for the case  $i = 1$  we see that one term in (26) is enough:

$$v|\phi_1\rangle = v|R_{10}\rangle|00\rangle = V_0|R_{10}\rangle|00\rangle = V_0|\bar{R}_{10}\rangle \langle \bar{R}_{10}| V_0|R_{10}\rangle|00\rangle$$

because  $\langle \bar{R}_{n0}|V_0|R_{10}\rangle = 0$  for  $n > 1$ . For the case  $i = 2$  we get the correct result with 2 terms in (26)

$$v|\phi_2\rangle = v|R_{20}\rangle|00\rangle = V_0|R_{20}\rangle|00\rangle = (V_0|\bar{R}_{10}\rangle \langle \bar{R}_{10}| V_0|R_{20}\rangle + V_0|\bar{R}_{20}\rangle \langle \bar{R}_{20}| V_0|R_{20}\rangle)|00\rangle$$

because  $\langle \bar{R}_{n0}|V_0|R_{20}\rangle = 0$  for  $n > 2$ . For the case  $i = 6$  we need to take into account 3 terms etc. We can see from this example that taking  $|R_{nl}\rangle$  from one atomic calculation, we get precise results (with the same atom) only taking into account a finite number of terms in (26), for 14 electrons actually only 3 terms. For several atoms calculation, we will not get precise results, but it should be a sufficiently good approximation.

The described method is general, the only drawback is that if we do not take functions  $|R_{nl}\rangle$  that are similar to the solution, we need to take a large number of terms in (26), resulting in many matrices of the form  $p_i p_j^*$ , which we do not want, even though, theoretically we can get a solution with any precision we want taking more and more terms in (26).

See also [1].

## 5. FINITE ELEMENT METHOD

This chapter explains FEM and gives concrete formulas which are needed in the calculation.

### 5.1. Weak formulation of the Schrödinger equation

One particle Schrödinger equation is

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\psi = E\psi.$$

We multiply both sides by a test function  $v$

$$-\left(\frac{\hbar^2}{2m}\nabla^2\psi\right)v = (E - V)\psi v,$$

and integrate over the whole volume we are interested in

$$\int -\left(\frac{\hbar^2}{2m}\nabla^2\psi\right)v dV = \int (E - V)\psi v dV, \quad (27)$$

and using the vector identity

$$-(\nabla^2\psi)v = \nabla\psi \cdot \nabla v - \nabla \cdot ((\nabla\psi)v),$$

we rewrite the left hand side of (27)

$$\int \frac{\hbar^2}{2m} \nabla \psi \cdot \nabla v \, dV = \int (E - V) \psi v \, dV + \int \frac{\hbar^2}{2m} \nabla \cdot ((\nabla \psi) v) \, dV,$$

now we apply the Gauss theorem

$$\int \frac{\hbar^2}{2m} \nabla \psi \cdot \nabla v \, dV = \int (E - V) \psi v \, dV + \oint \frac{\hbar^2}{2m} (\nabla \psi) v \cdot \mathbf{n} \, dS,$$

and rewriting  $\nabla \psi \cdot \mathbf{n} \equiv \frac{d\psi}{dn}$

$$\int \frac{\hbar^2}{2m} \nabla \psi \cdot \nabla v \, dV + \int v V \psi \, dV = \int E \psi v \, dV + \oint \frac{\hbar^2}{2m} \frac{d\psi}{dn} v \, dS, \quad (28)$$

which is the weak formulation. The problem reads: find a function  $\psi$  such that (28) holds for every  $v$ .

The boundary condition of all functions is that they disappear at the boundary (infinity) and have zero normal derivative there. That follows from the physical requirement of normalizability of the wavefunction.

## 5.2. Finite elements

We choose a basis  $\phi_i$  and substitute  $\phi_i$  for  $v$  and expand  $\psi = \sum q_j \phi_j$

$$\left( \int \frac{\hbar^2}{2m} \nabla \phi_j \cdot \nabla \phi_i \, dV + \int \phi_i V \phi_j \, dV \right) q_j = \left( \int E \phi_j \phi_i \, dV \right) q_j + \oint \frac{\hbar^2}{2m} \frac{d\psi}{dn} \phi_i \, dS, \quad (29)$$

which can be written in a matrix form

$$(K_{ij} + V_{ij}) q_j = E M_{ij} q_j + F_i,$$

where

$$\begin{aligned} V_{ij} &= \int \phi_i V \phi_j \, dV, \\ M_{ij} &= \int \phi_i \phi_j \, dV, \\ K_{ij} &= \frac{\hbar^2}{2m} \int \nabla \phi_i \cdot \nabla \phi_j \, dV, \\ F_i &= \frac{\hbar^2}{2m} \oint \frac{d\psi}{dn} \phi_i \, dS. \end{aligned}$$

Usually we set  $F_i = 0$ .

We decompose the domain into elements and compute the integrals as the sum over elements. For example:

$$K_{ij} = \sum_{E \in \text{elements}} K_{ij}^E,$$

where  $K_{ij}^E$  is the integral over one element only

$$K_{ij}^E = \int \frac{\hbar^2}{2m} \nabla \phi_j \cdot \nabla \phi_i dV^E \approx \sum_{q=0}^{N_q-1} \frac{\hbar^2}{2m} \nabla \phi_i(x_q) \cdot \nabla \phi_j(x_q) w_q |\det J(\hat{x}_q)|.$$

The integral is computed numerically using a Gauss integration:  $x_q$  are Gauss points (there are  $N_q$  of them),  $w_q$  is the weight of each point, and the Jacobian  $|\det J(\hat{x}_q)|$  is there because we are actually computing the integral on the reference element instead in the real space.

The surface integrals are computed similarly, but in our case they are all zero, as the normal derivative of the wavefunction is zero at the boundary.

The actual assembly is a little more complex (you need to loop over all elements, calculate the integrals above, and put the numbers to the correct place in the global matrix), but the above sketch should make the main idea clear.

As to the concrete form (shape) of the basis (shapefunctions), it is well known that quadratic elements are generally more precise than linear, and that cubes are generally more precise compared to tetrahedra. However, in practice that needs to be tried on a case by case basis, also depending on the mesh. Our code can work with both linear and quadratic elements and with tetrahedra, cubes, prisms, and pyramids.

### 5.3. Pseudopotentials formulation

There are no problems with other matrix elements in (29) except

$$V_{ij} = \int \phi_i(\mathbf{r}) V \phi_j(\mathbf{r}) d^3 r = \int \langle i | \mathbf{r} \rangle \langle \mathbf{r} | \hat{V} | j \rangle d^3 r = \langle i | \hat{V} | j \rangle,$$

where

$$\hat{V} = V_{loc}(\rho) + \sum_{nlm} V_l(\rho) |\bar{R}_{nl}\rangle |lm\rangle \langle lm| \langle \bar{R}_{nl}| V_l(\rho).$$

Thus

$$V_{ij} = \langle i | V_{loc}(\rho) | j \rangle + \langle i | \sum_{nlm} V_l(\rho) |\bar{R}_{nl}\rangle |lm\rangle \langle lm| \langle \bar{R}_{nl}| V_l(\rho) | j \rangle = V_{ij}^{loc} + \sum_{nlm} p_i p_j^*,$$

where the complex vector  $p_i$  is given by

$$p_i^{(nlm)} = \langle i | lm \rangle V_l(\rho) |\bar{R}_{nl}\rangle = \int \langle i | \mathbf{r} \rangle \langle \hat{\mathbf{r}} | lm \rangle \langle \rho | V_l(\rho) |\bar{R}_{nl}\rangle d^3 r = \int \phi_i(\mathbf{r}) Y_{lm}(\hat{\mathbf{r}}) V_l(\rho) \bar{R}_{nl}(\rho) d^3 r$$

and

$$V_{ij}^{loc} = \int \phi_i(\mathbf{r}) V_{loc}(\rho) \phi_j(\mathbf{r}) d^3 r,$$

and  $Y_{lm}(\hat{\mathbf{r}})$ ,  $\bar{R}_{nl}(\rho)$  and  $V_l(\rho)$  are given functions. Noticing that

$$\begin{aligned} \sum_m p_i p_j^* &= \int \phi_i(\mathbf{r}) Y_{lm}(\hat{\mathbf{r}}) V_l(\rho) \bar{R}_{nl}(\rho) d^3 r \int \phi_j(\mathbf{r}') Y_{lm}^*(\hat{\mathbf{r}}') V_l(\rho') \bar{R}_{nl}(\rho') d^3 r' \\ &= \int \int Y_{lm}(\hat{\mathbf{r}}) Y_{lm}^*(\hat{\mathbf{r}}') V_l(\rho) \bar{R}_{nl}(\rho) \phi_i(\mathbf{r}) \phi_j(\mathbf{r}') V_l(\rho') \bar{R}_{nl}(\rho') d^3 r d^3 r' \end{aligned}$$

we get

$$\sum_m p_i p_j^* = \int \int \frac{4\pi}{2l+1} P_l(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') V_l(\rho) \bar{R}_{nl}(\rho) \phi_i(\mathbf{r}) \phi_j(\mathbf{r}') V_l(\rho') \bar{R}_{nl}(\rho') d^3 r d^3 r',$$

which is a real number, thus  $\sum_{nlm} p_i p_j^*$  is also a real number, which means that we can calculate with only the real parts of the matrix  $p_i p_j^*$ , because the imaginary parts cancel out in the result:

$$\sum_{nlm} p_i p_j^* = \Re \left( \sum_{nlm} p_i p_j^* \right) = \sum_{nlm} \Re(p_i p_j^*).$$

Let  $p_i = a_i + ib_i$  then

$$\Re(p_i p_j^*) = \Re((a_i + ib_i)(a_j - ib_j)) = a_i a_j + b_i b_j$$

and

$$V_{ij} = V_{ij}^{loc} + \sum_{nlm} (a_i a_j + b_i b_j),$$

where

$$a_i = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} \int \phi_i(\mathbf{r}) P_l^m(\cos \theta) \cos(m\phi) V_l(\rho) \bar{R}_{nl}(\rho) d^3 r,$$

$$b_i = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} \int \phi_i(\mathbf{r}) P_l^m(\cos \theta) \sin(m\phi) V_l(\rho) \bar{R}_{nl}(\rho) d^3 r.$$

Just do not confuse the basis function  $\phi_i(\mathbf{r})$  with the spherical integration variable  $\phi$ .

## 6. CONCLUSION

We proposed a method to solve Kohn–Sham equations and to calculate electronic states and other properties of non-crystalline, non-periodic structures with  $l$ -dependent fully non-local real-space ab initio pseudopotentials using finite elements, together with some preliminary results of our program. We believe this new approach could yield some new interesting results in electronic structure calculations.

## ACKNOWLEDGEMENTS

We acknowledge support provided by the Grant Agency of the Academy of Sciences of the Czech Republic, under Grant No. IAA100100637. This research was also partly supported by the LC06040 research centre project.

## REFERENCES

1. Blochl, P. E. Generalized separable potentials for electronic-structure calculations. *Phys. Rev. B*, 1990, **41**(5414).
2. Dreizler, R. M. and Gross, E. K. U. *Density Functional Theory*. Springer-Verlag, 1990.
3. Eyert, V. A comparative study on methods for convergence acceleration of iterative vector sequences. *J. Comp. Phys.*, 1996, **124**, 271–285.
4. Hamman, D. R., Schlüter, M. and Chiang, C. Norm-conserving pseudopotentials. *Phys. Rev. Lett.*, 1979, **43**, 1494.
5. Martin, R. M. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2005.
6. Pickett, W. E. Pseudopotential methods in condensed matter applications. *Comp. Phys. Rep.*, 1989, **9**, 115–198.
7. Srivastava, G. P. Broyden's method for self-consistent field convergence acceleration. *J. Phys. A*, 1984, **17**, L317–L321.
8. Vanderbilt, D. and Louie, S. G. Total energies of diamond (111) surface reconstructions by a linear combination of atomic orbitals method. *Phys. Rev. B*, 1984, **30**(6118).

## Tihedusfunktsionaalide teooria arvutused lõplike elementidega

Ondřej Čertík, Jiří Vackář ja Jiří Plešek

On esitatud meetod Kohni–Shami võrrandi lahendamiseks ja mittekristalsete mitteperioodiliste  $l$ -sõltuvuslike, täielikult mittelokaalsete, reaalaruumis *ab initio* pseudopotsiaaliga struktuuride elektrooniliste olekute, kogenergia ning materjali omaduste arvutamiseks lõplike elementide meetodiga. Vastupidiselt hästi välja töötatud  $k$ -ruumi meetoditele, mis põhinevad Blochi teoreemil ja on rakendatavad perioodilistele struktuuridele, ei ole siin eeldatud perioodilisust üheski vormis. Käesolevas käsitluses on ühendatud seni  $k$ -ruumis tasandlainete lähenduses kasutatud täpsed *ab initio* keskkonnapeegelduse pseudopotsiaalid reaalaruumi lõplike elementide baasiga. Antud lähenduse oodatav põhiline eelis on *ab initio* pseudopotsiaalide tõhususe ja suure täpsuse ühendamine rakendustes, mis pole kitsendatud perioodilise keskkonnaga.