# Master Thesis

February 26, 2008

# Contents

# 1  Introduction

These are our notes on DFT and pseudopotentials calculations using a finite element method. We tried to include all the necessary theory starting from the many body Schrödinger equation. There are numerous books and articles on this topic, but the problem is first that the information is scattered among many publications and second that the theory in other sources is many times very lenghty and sometimes not very clear, sometimes it suppresses several important details, which are needed in order to write a computer program. So the reader should be able to understand every single detail just from this text.

    We have added the appendix with all the necessary information one has to know about a delta function, variations and functional derivatives, spherical harmonics and the dirac notation in order to be able to use it in calculations and understand formulas.

# 2  Derivation from QED

## 2.1  Introduction

The correct theory for electrons is the quantum electrodynamics (QED). Thus all the equations which we are going to use, follow from the QED and this chapter is devoted for their derivation. The standard nonrelativistic quantum mechanics (QM) contains a lot of ad-hoc assumptions (Schrödinger equation, electromagnetic coupling, spin, etc.), that cannot be satisfactorily explained. It is true that the whole nonrelativistic QM can be derived from a small set of assumptions (mostly commutation relations and some particular form of QM operators like spin), but it is not at all clear why these assumptions look the way they look. The QM is a set of rules which are known to work quite well, but obviously, we are not interested in postulating correct forms of operators or commutation relations, we want a better explanation, if there is one. And fortunately, there is.

The correct theory so far is the Standard Model and in our case of electron structure calculations, all phenomena are explained by the QED. There are still some problems in the Standard Model, but for much higher energies than we deal with in the electron structure. Thus what we do is that we take the most complete theory so far (QED) and derive everything from it. Of course, because we are dealing with low energies and to ease calculations, we make a lot of approximations on our way, but this is the only correct way to proceed, because it will be clear that we make each approximation because of this and that. If it turns out our results are not in agreement with experiment (for example the Schrödinger equation is imprecise), we need to neglect less things.

As you will see, this chapter is not long, so there really is no excuse of doing things differently (maybe in 1940s, but certainly not now when we have the QED).

There is one argument to be made though. As you will see, we will not get any new equations by deriving everything from firt principles from QED. We sometimes get a deeper understanding, however, from the practical point of view, it is not necessary to do that. There are logical problems even in QED or the Standard Model and noone really knows what are the correct equations, because we simply don't have enough experiments. The higher the energy we go, let's say up to the QED, the more fundamental theory we get, however, also the less experiments we have and thus some logical steps are just plain hand waving, because there is a whole bunch of ways how to continue, so we chose the one which we like, or which is the mathematically simplest. Contrary, in the low energy regime, like the classic quantum mechanics, the equations are very well known and there is really no other option. On the other hand, the theory doesn't explain us "why". So deriving everything from QED answers the "why", but creates some more fundamental questions, but those needn't concern us, because for our energy scale, everything works very nice.

## 2.2   QED

The QED Lagrangian density is

$$L = \bar{\psi}(ic\gamma^{\mu}D_{\mu} - mc^2)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

where

$$\psi = (\psi_1\psi_2\psi_3\psi_4)$$

and

$$D_{\mu} = \partial_{\mu} + ieA_{\mu}$$

is the gauge covariant derivative and ($e$ is the elementary charge, which is 1 in atomic units)

$$F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$$

is the electromagnetic field tensor. It's astonishing, that this simple Lagrangian can account for all phenomena from macroscopic scales down to something like

$10^{-13}$ cm. So of course Feynman, Schwinger and Tomonaga received the 1965 Nobel Prize in Physics for such a fantastic achievement.

Plugging this Lagrangian into the Euler-Lagrange equation of motion for a field, we get:

$$(ic\gamma^\mu D_\mu - mc^2)\psi = 0$$

$$\partial_\nu F^{\nu\mu} = -ec\bar\psi\gamma^\mu\psi$$

The first equation is the Dirac equation in the electromagnetic field and the second equation is a set of Maxwell equations ($\partial_\nu F^{\nu\mu} = -ej^\mu$) with a source $j^\mu = c\bar\psi\gamma^\mu\psi$, which is a 4-current comming from the Dirac equation.

The fields $\psi$ and $A^\mu$ are quantized. The first approximation is that we take $\psi$ as a wavefunction, that is, it is a classical 4-component field. It can be shown that this corresponds to taking three orders in the perturbation theory.

The first component $A_0$ of the 4-potential is the electric potential, and because this is the potential that (as we show in a moment) is in the Schrödinger equation, we denote it by $V$:

$$A_\mu = \left(\frac{V}{ec}, A_1, A_2, A_3\right)$$

So in the non-relativistic limit, the $\frac{V}{e}$ corresponds to the electric potential. We multiply the Dirac equation by $\gamma^0$ from left to get:

$$0 = \gamma^0(ic\gamma^\mu D_\mu - mc^2)\psi = \gamma^0(ic\gamma^0(\partial_0 + i\frac{V}{c}) + ic\gamma^i(\partial_i + ieA_i) - mc^2)\psi =$$

$$= (ic\partial_0 + ic\gamma^0\gamma^i\partial_i - \gamma^0 mc^2 - V - ce\gamma^0\gamma^i A_i)\psi$$

and we make the following substitutions (it's just a formalism, nothing more): $\beta = \gamma^0$, $\alpha^i = \gamma^0\gamma^i$, $p_j = -i\partial_j$, $\partial_0 = \frac{1}{c}\frac{\partial}{\partial t}$ to get

$$(i\frac{\partial}{\partial t} - c\alpha^i p_i - \beta mc^2 - V - ce\alpha^i A_i)\psi = 0\,.$$

This, in most solid state physics texts, is usually written as

$$i\frac{\partial\psi}{\partial t} = H\psi\,,$$

where the Hamiltonian is given by

$$H = c\alpha^i(p_i + eA_i) + \beta mc^2 + V\,.$$

The right hand side of the Maxwell equations is the 4-current, so it's given by:

$$j^\mu = c\bar\psi\gamma^\mu\psi$$

Now we make the substitution $\psi = e^{-imc^2 t}\varphi$, which states, that we separate the largest oscillations of the wavefunction and we get

$$j^0 = c\bar\psi\gamma^0\psi = c\psi^\dagger\psi = c\varphi^\dagger\varphi$$

4

$$j^i = c\bar{\psi}\gamma^i\psi = c\psi^\dagger\alpha^i\psi = c\varphi^\dagger\alpha^i\varphi$$

The Dirac equation implies the Klein-Gordon equation:

$$(-ic\gamma^\mu D_\mu - mc^2)(ic\gamma^\nu D_\nu - mc^2)\psi = (c^2\gamma^\mu\gamma^\nu D_\mu D_\mu + m^2c^4)\psi =$$

$$= (c^2 D^\mu D_\mu - ic^2[\gamma^\mu, \gamma^\nu]D_\mu D_\nu + m^2c^4)\psi = 0$$

Note however, the $\psi$ in the true Klein-Gordon equation is just a scalar, but here we get a 4-component spinor. Now:

$$D_\mu D_\nu = (\partial_\mu + ieA_\mu)(\partial_\nu + ieA_\nu) = \partial_\mu\partial_\nu + ie(A_\mu\partial_\nu + A_\nu\partial_\mu + (\partial_\mu A_\nu)) - e^2 A_\mu A_\nu$$

$$[D_\mu, D_\nu] = D_\mu D_\nu - D_\nu D_\mu = ie(\partial_\mu A_\nu) - ie(\partial_\nu A_\mu)$$

We rewrite $D^\mu D_\mu$:

$$D^\mu D_\mu = g^{\mu\nu}D_\mu D_\nu = \partial^\mu\partial_\mu + ie((\partial^\mu A_\mu) + 2A^\mu\partial_\mu) - e^2 A^\mu A_\mu =$$

$$= \partial^\mu\partial_\mu + ie((\partial^0 A_0) + 2A^0\partial_0 + (\partial^i A_i) + 2A^i\partial_i) - e^2(A^0 A_0 + A^i A_i) =$$

$$= \partial^\mu\partial_\mu + i\frac{1}{c^2}\frac{\partial V}{\partial t} + 2i\frac{V}{c^2}\frac{\partial}{\partial t} + ie(\partial^i A_i) + 2ieA^i\partial_i - \frac{V^2}{c^2} - e^2 A^i A_i$$

We use the identity $\frac{\partial}{\partial t}\left(e^{-imc^2t}f(t)\right) = e^{-imc^2t}(-imc^2 + \frac{\partial}{\partial t})f(t)$ to get:

$$L = c^2\partial^\mu\psi^*\partial_\mu\psi - m^2c^4\psi^*\psi = \frac{\partial}{\partial t}\psi^*\frac{\partial}{\partial t}\psi - c^2\partial^i\psi^*\partial_i\psi - m^2c^4\psi^*\psi =$$

$$= (imc^2 + \frac{\partial}{\partial t})\varphi^*(-imc^2 + \frac{\partial}{\partial t})\varphi - c^2\partial^i\varphi^*\partial_i\varphi - m^2c^4\varphi^*\varphi =$$

$$= 2mc^2\left[\frac{1}{2}i(\varphi^*\frac{\partial\varphi}{\partial t} - \varphi\frac{\partial\varphi^*}{\partial t}) - \frac{1}{2m}\partial^i\varphi^*\partial_i\varphi + \frac{1}{2mc^2}\frac{\partial\varphi^*}{\partial t}\frac{\partial\varphi}{\partial t}\right]$$

The constant factor $2mc^2$ in front of the Lagrangian is of course irrelevant, so we drop it and then we take the limit $c \to \infty$ (neglecting the last term) and we get

$$L = \frac{1}{2}i(\varphi^*\frac{\partial\varphi}{\partial t} - \varphi\frac{\partial\varphi^*}{\partial t}) - \frac{1}{2m}\partial^i\varphi^*\partial_i\varphi$$

After integration by parts we arrive at

$$L = i\varphi^*\frac{\partial\varphi}{\partial t} - \frac{1}{2m}\partial^i\varphi^*\partial_i\varphi$$

The nonrelativistic limit can also be applied directly to the Klein-Gordon equation:

$$0 = (c^2 D^\mu D_\mu + m^2c^4)\psi =$$

$$= \left(c^2\partial^\mu\partial_\mu + i\frac{\partial V}{\partial t} + 2iV\frac{\partial}{\partial t} + iec^2(\partial^i A_i) + 2iec^2 A^i\partial_i - V^2 - e^2c^2 A^i A_i + m^2c^4\right)e^{-imc^2t}\varphi =$$

$$= \left( \frac{\partial^2}{\partial t^2} - c^2\nabla^2 + 2iV\frac{\partial}{\partial t} + i\frac{\partial V}{\partial t} + iec^2(\partial^i A_i) + 2iec^2 A^i\partial_i - V^2 - e^2c^2 A^i A_i + m^2c^4 \right) e^{-imc^2 t}\varphi =$$

$$= e^{-imc^2 t}\left( (-imc^2 + \frac{\partial}{\partial t})^2 - c^2\nabla^2 + 2iV(-imc^2 + \frac{\partial}{\partial t}) + i\frac{\partial V}{\partial t} + iec^2(\partial^i A_i) + 2iec^2 A^i\partial_i - V^2 + \right.$$
$$\left. -e^2c^2 A^i A_i + m^2c^4 \right)\varphi =$$

$$= e^{-imc^2 t}\left( -2imc^2\frac{\partial}{\partial t} + \frac{\partial^2}{\partial t^2} - c^2\nabla^2 + 2Vmc^2 + 2iV\frac{\partial}{\partial t} + i\frac{\partial V}{\partial t} + iec^2(\partial^i A_i) + 2iec^2 A^i\partial_i - V^2 + \right.$$
$$\left. -e^2c^2 A^i A_i \right)\varphi =$$

$$= -2mc^2 e^{-imc^2 t}\left( i\frac{\partial}{\partial t} + \frac{\nabla^2}{2m} - V - \frac{1}{2mc^2}\frac{\partial^2}{\partial t^2} - \frac{i}{2mc^2}\frac{\partial V}{\partial t} + \frac{V^2}{2mc^2} - \frac{iV}{mc^2}\frac{\partial}{\partial t} + \right.$$
$$\left. -\frac{ie}{2m}\partial^i A_i - \frac{ie}{m}A^i\partial_i + \frac{e^2}{2m}A^i A_i \right)\varphi$$

Taking the limit $c \to \infty$ we again recover the Schrödinger equation:

$$i\frac{\partial}{\partial t}\varphi = \left( -\frac{\nabla^2}{2m} + V + \frac{ie}{2m}\partial^i A_i + \frac{ie}{m}A^i\partial_i - \frac{e^2}{2m}A^i A_i \right)\varphi \,,$$

we rewrite the right hand side a little bit:

$$i\frac{\partial}{\partial t}\varphi = \left( \frac{1}{2m}(\partial^i\partial_i + ie\partial^i A_i + 2ieA^i\partial_i - e^2 A^i A_i) + V \right)\varphi \,,$$

$$i\frac{\partial}{\partial t}\varphi = \left( \frac{1}{2m}(\partial^i + ieA^i)(\partial_i + ieA_i) + V \right)\varphi \,,$$

And we get the usual form of the Schrödinger equation for the vector potential $\mathbf{A} = (A_1, A_2, A_3)$:

$$i\frac{\partial}{\partial t}\varphi = \left( -\frac{(\nabla + ie\mathbf{A})^2}{2m} + V \right)\varphi \,.$$

## 2.3   Radial Schrödinger and Dirac Equations

For general treatment, together with derivation of all the different forms of radial Dirac equations used in the literature, see \citebachelor-thesis. Here we just summarize the results.

## 2.4   Radial Schrödinger Equation

We have a spherically symmetric potential energy

$$V(\mathbf{x}) = V(r) \,.$$

State with a given square of an angular momentum (eigenvalue $l(l+1)$) and its $z$ component (eigenvalue $m$) is described by the wave function

$$\psi_{nlm}(\mathbf{x}) = R_{nl}(r)\, Y_{lm}\left(\frac{\mathbf{x}}{r}\right)\,, \tag{1}$$

where $R_{nl}(r)$ obeys the equation \citeformanek (eq. 2.400)

$$R_{nl}'' + \frac{2}{r}R_{nl}' + \frac{2M}{\hbar^2}(E - V)R_{nl} - \frac{l(l+1)}{r^2}R_{nl} = 0\,. \tag{2}$$

This is called the radial Schrödinger equation which we want to solve numerically.

## 2.5  Numerical integration for a given

Equation (2) is the linear ordinary differential equation of the second order, so the general solution is a linear combination of two independent solutions. Normally, the 2 constants are determined from initial and/or boundary conditions. In our case, however, we don't have any other condition besides being interested in solutions that we can integrate on the interval $(0, \infty)$ (and which are normalizable), more exactly we want $R \in L^2$ and $\int_0^\infty r^2 R^2 \, \mathrm{d}r = 1$.

It can be easily shown by a direct substitution, that there are only two asymptotic behaviors near the origin: $r^l$ and $r^{-l-1}$. We are interested in quadratic integrable solutions only, so we are left with $r^l$ and only one integration constant, which we calculate from a normalization. This determines the solution uniquely.

All the integration algorithms needs to evaluate $R''$, which is a problem at the origin, where all the terms in the equation are infinite, although their sum is finite. We thus start to integrate the equation at some small $r_0$ (for example $r_0 = 10^{-10}$ a.u.), where all the terms in the equation are finite. If we find the initial conditions $R(r_0)$ and $R'(r_0)$, the solution is then fully determined.

If $r_0$ is sufficiently small, we can set $R(r_0) = r_0^l$ and $R'(r_0) = lr_0^{l-1}$. In the case $l = 0$ we need to set $R(r_0) = 1$ and $R'(r_0) = -\frac{1}{a}$, where $a$ is the Bohr radius, see the next section for more details.

So when somebody gives us $l$ and $E$, we are now able to compute the solution up to the the multiplicative constant that is later determined from a normalization. As was already mentioned, we used the fourth-order Runge-Kutta method that proved very suitable for this problem.

## 2.6  Asymptotic behavior

The asymptotic behavior is important for the integration routine to find the correct solution for a given $E$. It is well known, that the first term of the Taylor series of the solution is $r^l$, independent of the potential \citeformanek (eq. 2.408). This is enough information to find the correct solution for $l > 0$ because the only thing we need to know is the value of the wave function and its derivative near the origin, which is effectively $r_0^l$ and $lr_0^{l-1}$ for some small $r_0$.

The problem is with $l = 0$, where the derivative cannot be calculated just from $l$ and $r_0$.

The asymptotic behavior for $l = 0$ depends on the potential $V$, so we need to take into account it's properties. We assume $V$ to be of a form:

$$V = -\frac{Z}{r} + v_0 + v_1 r + O(r^2) \,,$$

It can be shown, that the solution is then

$$R = a_0(1 - \frac{r}{a} + O(r^2)) \,,$$

where $a = \frac{\hbar^2}{ZM}$ is the Bohr radius and $a_0$ is a normalization constant. So the initial condition for the integration for $l = 0$ is $R(r_0) = 1$ and $R'(r_0) = -\frac{1}{a}$.

## 2.7 Dirac Equation

The Dirac equation for one particle is \citestrange,zabloudil:

$$H\psi = W\psi \,, \tag{3}$$

$$H = c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2 + V(r)\mathbb{1} \,,$$

where $\psi$ is a four component vector:

$$\psi = (\psi_1\psi_2\psi_3\psi_4) = (\psi_A\psi_B) \,, \qquad \psi_A = (\psi_1\psi_2), \psi_B = (\psi_3\psi_4)$$

and $\boldsymbol{\alpha}$, $\beta$ are $4 \times 4$ matrices:

$$\boldsymbol{\alpha} = (0\boldsymbol{\sigma}\boldsymbol{\sigma}0) \,,$$

$$\beta = (\mathbb{1}00 - \mathbb{1}) \,,$$

where the Pauli matrices $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ and $\mathbb{1}$ form a basis of all $2 \times 2$ Hermitian matrices. To derive a continuity equation, we multiply (3) by $\psi^*$ and subtract the conjugate transpose of (3) multiplied by $\psi$:

$$\frac{\partial}{\partial t}(\psi^*\psi) = -\nabla \cdot (c\psi^*\boldsymbol{\alpha}\psi) \,,$$

so we identify the probability and current densities as

$$\rho = \psi^*\psi = \psi_1^*\psi_1 + \psi_2^*\psi_2 + \psi_3^*\psi_3 + \psi_4^*\psi_4 \,, \qquad \mathbf{j} = c\psi^*\boldsymbol{\alpha}\psi \,.$$

The normalization of a four-component wave function is then

$$\int \rho \, \mathrm{d}^3x = \int \psi^*\psi \, \mathrm{d}^3x = \int \psi_1^*\psi_1 + \psi_2^*\psi_2 + \psi_3^*\psi_3 + \psi_4^*\psi_4 \, \mathrm{d}^3x = 1 \,. \tag{4}$$

The probability density $\rho(x, y, z)$ is the physical quantity we are interested in, and all the four-component wavefunctions and other formalism is just a way of calculating it. This $\rho$ is also the thing we should compare with the Schrödinger equation.

## 2.8   Radial Dirac equation

We and search for a basis in the form of spin angular functions:

$$\psi_A = g\chi_\kappa^{j_3} \,, \tag{5}$$

$$\psi_B = if\chi_{-\kappa}^{j_3} \,. \tag{6}$$

Substituting all of these into (3) and some more well-known manipulations one gets:

$$\hbar c\left(-\frac{\partial f}{\partial r} + \frac{\kappa-1}{r}f\frac{\partial g}{\partial r} + \frac{\kappa+1}{r}g\right) = \left((W - V - mc^2)g(W - V + mc^2)f\right)\,. \tag{7}$$

This is the radial Dirac equation. As we shall see in the next section, the equation for $g$ is (with the exception of a few relativistic corrections) identical to the radial Schrödinger equation. And $f$ vanishes in the limit $c \to \infty$. For this reason $f$ is called the small (fein, minor) component and $g$ the large (groß, major) component.

The probability density is

$$\rho = \psi^*\psi = \psi_A^*\psi_A + \psi_B^*\psi_B = f^2\chi_{-\kappa}^{j_3*}\chi_{-\kappa}^{j_3} + g^2\chi_\kappa^{j_3*}\chi_\kappa^{j_3}\,,$$

so from the normalization condition (4) we get

$$\int \rho\,\mathrm{d}^3x = \int f^2\chi_{-\kappa}^{j_3*}\chi_{-\kappa}^{j_3}+g^2\chi_\kappa^{j_3*}\chi_\kappa^{j_3}\,\mathrm{d}^3x = \int (f^2\chi_{-\kappa}^{j_3*}\chi_{-\kappa}^{j_3}+g^2\chi_\kappa^{j_3*}\chi_\kappa^{j_3})\,r^2\,\mathrm{d}r\mathrm{d}\Omega =$$

$$= \int_0^\infty f^2r^2\,\mathrm{d}r\int\chi_{-\kappa}^{j_3*}\chi_{-\kappa}^{j_3}\,\mathrm{d}\Omega+\int_0^\infty g^2r^2\,\mathrm{d}r\int\chi_\kappa^{j_3*}\chi_\kappa^{j_3}\,\mathrm{d}\Omega = \int_0^\infty r^2(f^2+g^2)\,\mathrm{d}r = 1\,,$$

where we used the normalization of spin-angular functions. Also it can be seen, that the radial probability density is

$$\rho(r) = r^2(f^2 + g^2) \tag{8}$$

(i.e., the probability to find the electron between $r_1$ and $r_2$ is $\int_{r_1}^{r_2} r^2(f^2+g^2)\,\mathrm{d}r$). The result of integrating the radial Dirac equation are the two functions $f$ and $g$, but the physically relevant quantity is the radial probability density (8). In the nonrelativistic case, the density is given by

$$\rho(r) = r^2R^2\,,$$

so the correspondence between the Schrödinger and Dirac equation is $R^2 = f^2 + g^2$.

For numerical stability and robustness, we are not solving the equations in the form (7), but a sligthly rearranged ones. Let's use Hartree atomic units ($m = \hbar = 1$) and define $E = W - mc^2 = W - c^2$, so that $E$ doesn't contain the electron rest mass energy. Let's make the substitution \citedonald:apw

$$P_\kappa = rg_\kappa\,,$$

$$Q_\kappa = rf_\kappa$$

and plug all of this into (7). After a little manipulation we get:

$$\frac{\mathrm{d}P_\kappa}{\mathrm{d}r} = -\frac{\kappa}{r}P_\kappa + \left[\frac{E-V}{c} + 2c\right]Q_k\,,$$

$$\frac{\mathrm{d}Q_\kappa}{\mathrm{d}r} = \frac{\kappa}{r}Q_\kappa - \frac{1}{c}(E-V)P_k\,, \tag{9}$$

which can be found in \citezabloudil (eq. 8.12 and 8.13), where they have one $c$ hidden in $Q_\kappa = crf_\kappa$ and use Rydberg atomic units, so they have 1 instead of 2 in the square bracket. It can be found in \citebachelet as well, they use Hartree atomic units, but have a different notation $G_\kappa \equiv P_\kappa$ and $F_\kappa \equiv Q_\kappa$, also they made a substitution $c = \frac{1}{\alpha}$.

## 2.9 Asymptotic behavior

We calculate the functions $f_\kappa$ and $g_\kappa$ in a similar way as we calculated $R$ for the Schrödinger equation, thus we need the asymptotic behavior at the origin. The potential can always be treated as $V = 1/r + \cdots$ and in this case it can be shown \citezabloudil, that the asymptotic is

$$P_\kappa = rg_\kappa = r^\beta\,,$$

$$Q_\kappa = rf_\kappa = r^{\beta-1}\frac{\beta+\kappa}{\frac{E-V}{c}+2c}\,,$$

where

$$\beta = \sqrt{\kappa^2 - \left(\frac{Z}{c}\right)^2}\,, \tag{10}$$

or, if we write it explicitly, for $j = l + \frac{1}{2}$

$$\beta^+ = \sqrt{(-l-1)^2 - \left(\frac{Z}{c}\right)^2}$$

and $j = l - \frac{1}{2}$

$$\beta^- = \sqrt{l^2 - \left(\frac{Z}{c}\right)^2}\,.$$

In the semirelativistic case (which is an approximation — we neglect the spin-orbit coupling term) we choose

$$\beta = \sqrt{\tfrac{1}{2}(|\beta^+|^2 + |\beta^-|^2)} = \sqrt{l^2 + l + \tfrac{1}{2} - \left(\frac{Z}{c}\right)^2}\,.$$

It should be noted that in the literature we can find other types of aymptotic behaviour for the semirelativistic case, its just a question of the used approximation. One can hardly say that some of them are correct and another is not

since the semirelativistic (sometimes denoted as scalar-relativstic) approximation itself is not correct, it's just an approximation.

It follows from (10) that for $j = l + \frac{1}{2}$ the radial Dirac equation completely becomes the radial Schrödinger equation in the limit $c \to \infty$ (and gives exactly the same solutions):

$$P_\kappa = r g_\kappa \to r^{l+1}\,,$$

$$Q_\kappa = r f_\kappa \to 0\,.$$

For $j = l - \frac{1}{2}$ however, we get a wrong asymptotic: we get a radial Schrödinger equation for $l$, but the asymptotic for $l - 1$.

## 2.10 Eigenproblem

In the previous sections, we learned how to calculate the solution of both the radial Schrödinger and Dirac equations for a given $E$. For most of the energies, however, the solution for $r \to \infty$ exponentially diverges to $\pm\infty$. Only for the energies equal to eigenvalues, the solution tends exponentially to zero for $r \to \infty$. The spectrum for bounded states is discrete, so we label the energies by $n$, starting from 1.

We want to find the eigenvalue and eigenfunction for a given $n$ and $l$ (and a spin in the relativistic case). The algorithm is the same for both nonrelativistic and relativistic case and is based on two facts, first that the number of nodes (ie. the number of intersections with the $x$ axis, not counting the one at the origin and in the infinity) of $R_{nl}$ and $g_\kappa$ is $n - l - 1$ and second that the solution must tend to zero at infinity.

We calculate the solution for some (random) energy $E_0$, using the procedure described above. Then we count the number of nodes (for diverging solutions, we don't count the last one) and check, if the solution is approaching the zero from top or bottom in the infinity. From the number of nodes and the direction it is approaching the zero it can be determined whether the energy $E_0$ is below or above the eigenvalue $E$ belonging to a given $n$ and $l$. The rest is simple, we find two energies, one below $E$, one above $E$ and by bisecting the interval we calculate $E$ with any precision we want.

There are a few technical numerical problems that are unimportant from the theoretical point of view, but that need to be solved if one attempts to actually implement this algorithm. One of them is that when the algorithm (described in the previous paragraph) finishes, because the energy interval is sufficiently small, it doesn't mean the solution is near zero for the biggest $r$ of our grid. Remember, the solution goes exponentially to $\pm\infty$ for every $E$ except the eigenvalues and because we never find the exact eigenvalue, the solution will (at some point) diverge from zero.

Possible solution that we have employed is as follows: when the algorithm finishes we find the last minimum (which is always near zero) and trim the solution behind it (set it to zero).

The second rather technical problem is how to choose the initial interval of energies so that the eigenvalue lies inside the interval. We use some default

values that work for atomic calculations, while allowing the user to override it if needed.

# 3 DFT

## 3.1 Introduction

Good books about DFT are \citeDFT and \citemartin, but they both contain much more topics which we don't need and some topics are missing in each of them, so this chapter gives a self-contained explanation of all one has to know about a many body quantum mechanics and the DFT in order to be able to do DFT calculations.

## 3.2 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 $\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r}_N)$ fulfilling the many-body Schrödinger equation

$$\hat{H}\ket{\Psi} = (\hat{T} + \hat{U} + \hat{V})\ket{\Psi} = E\ket{\Psi}$$

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 $R_k$ 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 $v(\mathbf{r_i}) = -\frac{Z}{|\mathbf{r_i}|}$.

$|\Psi|^2 = \Psi^*\Psi$ gives the probability density of measuring the first electron at the position $\mathbf{r_1}$, the second at $\mathbf{r_2}$, \dots and the Nth electron at the position

$\mathbf{r}_N$. The normalization is such that $\int |\Phi|^2 \mathrm{d}^3 r_1 \mathrm{d}^3 r_2 \ldots \mathrm{d}^3 r_N = 1$. The $\Psi$ is antisymmetric, i.e. $\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r}_N) = -\Psi(\mathbf{r_2}, \mathbf{r_1}, \cdots, \mathbf{r}_N) = -\Psi(\mathbf{r_1}, \mathbf{r}_N, \cdots, \mathbf{r_2})$ 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, \dots, or the $N$th) is at the position $\mathbf{r}$ is called the particle (or charge or electron) density and is therefore given by:

$$n(\mathbf{r}) = \int \Psi^*(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N) \Psi(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N) \, \mathrm{d}^3 r_2 \, \mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_N +$$

$$+ \int \Psi^*(\mathbf{r}_1, \mathbf{r}, \cdots, \mathbf{r}_N) \Psi(\mathbf{r}_1, \mathbf{r}, \cdots, \mathbf{r}_N) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_N + \cdots$$

$$+ \int \Psi^*(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}) \Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \, \mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_{N-1} =$$

$$= \int (\delta(\mathbf{r} - \mathbf{r}_1) + \delta(\mathbf{r} - \mathbf{r}_2) + \cdots + \delta(\mathbf{r} - \mathbf{r}_N))$$

$$\Psi^*(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N) \Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \, \mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_N =$$

$$= \sum_{i=1}^{N} \int \langle \Psi | \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N \rangle \, \delta(\mathbf{r} - \mathbf{r}_i) \, \langle \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N | \Psi \rangle \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \, \mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_N =$$

$$= N \int \langle \Psi | \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N \rangle \, \delta(\mathbf{r} - \mathbf{r}_1) \, \langle \mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N | \Psi \rangle \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \, \mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_N =$$

$$= N \int \Psi^*(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N) \Psi(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N) \, \mathrm{d}^3 r_2 \, \mathrm{d}^3 r_3 \cdots \mathrm{d}^3 r_N \qquad (11)$$

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

The energy of the system is given by

$$E = \langle \Psi | \hat{H} | \Psi \rangle = \langle \Psi | \hat{T} | \Psi \rangle + \langle \Psi | \hat{U} | \Psi \rangle + \langle \Psi | \hat{V} | \Psi \rangle = T + U + V \qquad (12)$$

where

$$T = \langle \Psi | \hat{T} | \Psi \rangle = \sum_i^N \int \Psi^*(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})(-\tfrac{1}{2}\nabla_i^2)\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N}) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \cdots \mathrm{d}^3 r_N$$

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

$$V = \langle \Psi | \hat{V} | \Psi \rangle = \sum_i^N \int \Psi^*(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N}) v(\mathbf{r_i}) \Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N}) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \cdots \mathrm{d}^3 r_N =$$

$$= \sum_i^N \int \Psi^*(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N}) v(\mathbf{r_1}) \Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N}) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \cdots \mathrm{d}^3 r_N =$$

$$= N \int \Psi^*(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N}) v(\mathbf{r_1}) \Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N}) \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \cdots \mathrm{d}^3 r_N =$$

$$= \int v(\mathbf{r}) n(\mathbf{r}) \mathrm{d}^3 r = V[n] \tag{13}$$

It needs to be stressed, that $E$ generally is \it 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 $n$.

### 3.3 The Hohenberg-Kohn Theorem

The SE gives the map
$$C : V \to \Psi$$

where $\Psi$ is the ground state. 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 substracting

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

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 a contradiction with 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

$$D : \Psi \to n$$

which is also bijective, because to every $\Psi$ we can compute $n$ from (11) and two different $\Psi$ and $\Psi'$ give two different $n$ and $n'$, because different $\Psi$ and $\Psi'$ give

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

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

adding these two inequalities together gives

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

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

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 $\Psi_0(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})$, in other words, $\Psi_0$ is a unique functional of $n_0$:

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

so 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

$$E_{v_0}[n] = \langle \Psi[n] | \hat{T} + \hat{U} + \hat{V}_0 | \Psi[n] \rangle = \langle \Psi[n] | \hat{T} + \hat{U} | \Psi[n] \rangle + \int v_0(\mathbf{r}) n(\mathbf{r}) \mathrm{d}^3 r \quad (14)$$

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 minimise the functional $E_{v_0}[n]$:

$$E_0 = \min_n E_{v_0}[n] \quad (15)$$

The term
$$\langle \Psi[n] | \hat{T} + \hat{U} | \Psi[n] \rangle \equiv F[n]$$
in (14) is universal in the sense that it doesn't depend on $\hat{V}_0$. It can be proven \citeDFT, that $F[n]$ is a functional of $n$ for degenerated ground states too, so (15) stays true as well.

The ground state densities in (14) and (15) are called \it pure-state v-representable\/ because they are the densities of (possible degenerate) ground state of the Hamiltonian with some local potential $v(\mathbf{r})$. One may ask a 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 to take into account in the minimisation proccess (15). Even though not every function is v-representable \citeDFT, every density defined on a grid (finite of infinite) which 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 \citeDFT. In plain words, we are fine.

The functional $E_{v_0}[n]$ in (15) depends on the particle number $N$, so in order to get $n$, we need to solve the variational formulation

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

so

$$\frac{\delta E_v[n]}{\delta n} = \mu(N) \quad (16)$$

Let the $n_N(\mathbf{r})$ be the solution of (16) 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

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

becuase

$$E_{N+\epsilon} - E_N = E_v[n_{N+\epsilon}] - E_v[n_N] = \int \frac{\delta E_v}{\delta n}(n_{N+\epsilon} - n_N)\mathrm{d}^3 r =$$

$$= \int \mu(N)(n_{N+\epsilon} - n_N)\mathrm{d}^3 r = \mu(N)(N + \epsilon - N) = \mu(N)\epsilon$$

so

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

## 3.4 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

$$|\Psi(\mathbf{r_1}, \mathbf{r_2}, \cdots, \mathbf{r_N})\rangle = |\psi_1(\mathbf{r})\rangle |\psi_2(\mathbf{r})\rangle \cdots |\psi_N(\mathbf{r})\rangle$$

and

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

where

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

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

From (16) we get

$$\mu = \frac{\delta E_s[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + \frac{\delta V_s[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + v_s(\mathbf{r}) \qquad (17)$$

Solution to this equation gives the density $n_s$.

Now we want to express the energy in (12) 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 uquivalently

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

$$V_H(\mathbf{r}) = \frac{\delta E_H}{\delta n(\mathbf{r})} = \tfrac{1}{2} \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \mathrm{d}^3r' \qquad (18)$$

So from (14) we get

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

$$= T_s[n] + E_H[n] + E_{xc}[n] + V[n] \qquad (19)$$

The rest of the energy is denoted by $E_{xc} = U - E_H + T - T_s$ and it is called is the exchange and correlation energy functional. From (16)

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

From (18) we have

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

from (13) we get

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

we define

$$\frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} = V_{xc}(\mathbf{r}) \qquad (20)$$

so we arrive at

$$\mu = \frac{\delta E[n]}{\delta n(\mathbf{r})} = \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + v(\mathbf{r}) \qquad (21)$$

Solution to this equation gives the density $n$. Comparing (21) to (17) 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

$$(-\tfrac{1}{2}\nabla^2 + v_s(\mathbf{r}))\psi_i(\mathbf{r}) \equiv (-\tfrac{1}{2}\nabla^2 + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) + v(\mathbf{r}))\psi_i(\mathbf{r}) = \epsilon_i \psi(\mathbf{r}) \quad (22)$$

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

$$n(\mathbf{r}) \equiv n_s(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2 \qquad (23)$$

The sum is taken over the lowest $N$ energies. Some of the $\psi_i$ can be degenerated, but it doesn't matter - the index $i$ counts every eigenfunction including all the degenerated. In plain words, the trick is in realising, that the ground state energy can be found by minimising the energy functional (14) and in rewriting this functional into the form (19), which shows that the interacting system can be treated as a noninteracting one with a special potential.

17

## 3.5 The XC 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

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

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

Unfortunately, no one knows $\epsilon_{xc}(\mathbf{r}'; n)$ exactly (yet). The most simple approximation is the \it local density approximation \/ (LDA), for which the xc energy density $\epsilon_{xc}$ 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:

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

The xc potential $V_{xc}$ defined by (20) is then

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

which in the LDA becomes

$$V_{xc}(\mathbf{r}; n) = \epsilon_{xc}^{LD}(n) + n\frac{\mathrm{d}\epsilon_{xc}^{LD}(n)}{\mathrm{d}n} = \frac{\mathrm{d}}{\mathrm{d}n}\left(n\epsilon_{xc}^{LD}(n)\right) = V_{xc}^{LD}(n) \qquad (24)$$

The xc energy density $\epsilon_{xc}^{LD}$ of the homogeneous gas can be computed exactly\citemartin:

$$\epsilon_{xc}^{LD}(n) = \epsilon_x^{LD}(n) + \epsilon_c^{LD}(n)$$

where the $\epsilon_x^{LD}$ is the electron gas exchange term given by\citemartin

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

the rest of $\epsilon_{xc}^{LD}$ is hidden in $\epsilon_c^{LD}(n)$ for which there doesn't exist an analytic formula, but the correlation energies are known exactly from quantum Monte Carlo (QMC) calculations by Ceperley and Alder\citepickett. The energies were fitted by Vosko, Wilkes and Nussair (VWN) with $\epsilon_c^{LD}(n)$ and they got accurate results with errors less than $0.05\,\mathrm{mRy}$ in $\epsilon_c^{LD}$, which means that $\epsilon_c^{LD}(n)$ is virtually known exactly. VWN result:

$$\epsilon_c^{LD}(n) \approx \frac{A}{2}\left\{\ln\left(\frac{y^2}{Y(y)}\right) + \frac{2b}{Q}\arctan\left(\frac{Q}{2y+b}\right) + \right.$$

$$\left. -\frac{by_0}{Y(y_0)}\left[\ln\left(\frac{(y-y_0)^2}{Y(y)}\right) + \frac{2(b+2y_0)}{Q}\arctan\left(\frac{Q}{2y+b}\right)\right]\right\}$$

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 (24):

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

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

$$V_c^{LD} = \frac{A}{2}\left\{\ln\left(\frac{y^2}{Y(y)}\right) + \frac{2b}{Q}\arctan\left(\frac{Q}{2y+b}\right) + \right.$$

$$\left. -\frac{by_0}{Y(y_0)}\left[\ln\left(\frac{(y-y_0)^2}{Y(y)}\right) + \frac{2(b+2y_0)}{Q}\arctan\left(\frac{Q}{2y+b}\right)\right]\right\} +$$

$$-\frac{A}{6}\frac{c(y-y_0) - by_0 y}{(y-y_0)Y(y)}$$

Some people also use Perdew and Zunger formulas, but they give essentially the same results. The LDA, although very simple, is suprisingly 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 a 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 $\epsilon_x^{LD}(n) \rightarrow \epsilon_x^{LD}(n)R$:

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

where

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

and

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

We also need to calculate these derivatives:

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

$$\frac{\mathrm{d}\beta}{\mathrm{d}n} = \frac{\beta}{3n}$$

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

So

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

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

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

## 3.6 Iteration to Self-consistency

The $V_H$ and $V_{xc}$ potentials in the Kohn-Sham equations (22) 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)

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

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, since 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

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

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

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

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

$$\mathbf{R}(\mathbf{x}_{M+1}) - \mathbf{R}(\mathbf{x}_M) \approx \mathbf{J} \cdot (\mathbf{x}_{M+1} - \mathbf{x}_M) \tag{25}$$

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 (25) 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 $\mathbf{J}_0, \mathbf{J}_1, \mathbf{J}_2, \dots$:

$$\mathbf{x}_{M+1} = \mathbf{x}_M - \mathbf{J}_M^{-1} \cdot \mathbf{R}(\mathbf{x}_M) \tag{26}$$

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

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

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

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, working value is somewhere around $\alpha = 0.1$ to $\alpha = 0.3$. Unfortunately, this procedure is slow and also we do not explore all the possible densities with this mixing, which means that we don't 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's very easy to implement and it works in most cases, although slowly.

Surprisingly very good method is this:

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

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 my tests it behaves almost as well as the second Broyden method.

More sophisticated approach is the Broyden update, which updates the $\mathbf{J}$ successively at every iteration. The \it first Broyden method\/ is using 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 (26) at every iteration, which is impossible in our case. The \it second Broyden method\/ updates the inverse Jacobian directly using this formula

$$\mathbf{J}_{M+1}^{-1} = \mathbf{J}_M^{-1} + \frac{(\Delta\mathbf{x}_M - \mathbf{J}_M^{-1} \cdot \Delta\mathbf{R}(\mathbf{x}_M))\Delta\mathbf{R}(\mathbf{x}_M)^T}{|\Delta\mathbf{R}(\mathbf{x}_M)|^2} \tag{27}$$

starting with the linear mixing:

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

It is impossible to store the whole dense matrix of the inverse Jacobian, but fortunately it is not necessary, realising that the (27) has a very simple structure \citesrivastava:

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

with

$$\mathbf{u} = \Delta\mathbf{x}_M - \mathbf{J}_M^{-1} \cdot \Delta\mathbf{R}(\mathbf{x}_M) \tag{28}$$

$$\mathbf{v} = \frac{\Delta\mathbf{R}(\mathbf{x}_M)}{|\Delta\mathbf{R}(\mathbf{x}_M)|^2}$$

so the whole inverse Jacobian can be written as

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

and we only need to know how to apply such a Jacobian to an arbitrary vector, which is needed in (28) and (26):

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

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 \citevanderbiltlouie suggested a \it modified Broyden method, which incorporates weights, but Eyert \citeeyert showed that if all the weights are used to tune the iteration process to its fastest convergence, they, in fact, cancel out and the result of the scheme is called by Eyert the \it generalized Broyden method, whose scheme shown by Eyert is exactly the same as for the \it Anderson mixing:

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

$$\mathbf{x}_{M+1} = \mathbf{x}_M + \beta_M \mathbf{R}(\mathbf{x}_M) - \sum_{p=M-k}^{M-1} \gamma_p (\Delta \mathbf{x}_p + \beta_M \Delta \mathbf{R}(\mathbf{x}_p))$$

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

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

$$n > 0$$

and the normalization

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

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 more elegant solution is to mix $V_h + V_{xc}$ instead of the densities.

## 3.7   Example: Pb Atom

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 (23). One option (the correct one) is to automatically try different "n" and "l" until we are sure we got the lowest 82 energies. But for Pb the combination of "n" and "l" is well-known, it is (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:

-2701.6 -466.18 -471.87 -111.45 -108.24 -92.183 -24.498 -22.086 -15.119 -5.6606 -3.9570 -2.9743 -.92718 -.33665 -.14914

# 4   Pseudopotentials

## 4.1   Introduction

Literature about pseudopotentials is unfortunately scattered among many arcticles, so this section gives a review and should save the reader from a lot of troubles.

## 4.2    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

$$\hat{V} = \sum_{lm} |lm\rangle \, \hat{V}_l \, \langle lm| \tag{29}$$

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

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

\bf Proof: Matrix elements of a general Hermitian operator $\hat{V}$ are

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

where

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

In spherical symmetry, we have

$$\langle \mathbf{r}|\hat{V}|\varphi\rangle = \langle \mathbf{r}|R^{-1}\hat{V}R|\varphi\rangle = \langle \mathbf{r}|R^\dagger\hat{V}R|\varphi\rangle = \int \langle \mathbf{r}|R^\dagger\hat{V}R|\mathbf{r}'\rangle \, \langle \mathbf{r}'|\varphi\rangle \, \mathrm{d}^3 r' =$$

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

where $R$ is the rotation operator (it's 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 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 using (51)

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

where

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

In Dirac notation:

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

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

$$V_l(\rho, \rho') = \langle lm| \, \langle \rho|\hat{V}|\rho'\rangle \, |lm\rangle \tag{30}$$

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 (30) it doesn't depend on $m$. Also according to (30) its matrix elements are

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

## 4.3   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 (29). 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

$$\hat{V} = V(\rho) \tag{31}$$

so

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

so it turned out that the kernel is local and doesn't depend on $l$ and we get

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

$$= V(\rho) \frac{1}{\rho^2 \sin \theta} \delta(\rho - \rho') \delta(\theta - \theta') \delta(\phi - \phi') = V(\rho) \delta(\mathbf{r} - \mathbf{r}')$$

and

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

so we recover (31). But we are just fooling around, there's 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

$$\hat{V} = \sum_{lm} |lm\rangle \, V_l(\rho) \, \langle lm| \tag{32}$$

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

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

$$= \sum_{lm} Y_{lm}(\theta, \phi) V_l(\rho) \int Y_{lm}^*(\theta', \phi') \varphi(\rho \hat{\mathbf{r}}') \mathrm{d}\Omega'$$

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 (32).

So, to sum it up: 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 (32).

Now, it can be shown that if we make the assumption of radial locality, we get "correct" wavefunctions and energies in the linear approximation. We generally only take a few terms in the expansion (32), usually only $V_0$, $V_1$ and $V_2$, sometimes also $V_3$ and $V_4$.

## 4.4 Separable Potentials

The pseudopotential above (Hamman, Schlüter, Chiang) has the form

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

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 doesn't cause a problem. Let's denote the second term (which is semilocal) simply by $v$:

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

26

Let's 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 $U$ is 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

$$v = \sum_i v \, |\bar\phi_i\rangle \, \frac{1}{\langle \bar\phi_i | v | \bar\phi_i \rangle} \, \langle \bar\phi_i | \, v = \sum_i v \, |\bar\phi_i\rangle \, \langle \bar\phi_i | \, v \tag{33}$$

We could take any $|\phi_i\rangle$ and orthogonalize them. But because we have $v$ in the form of (32), 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 (33) then corresponds to the KB potential. We of course take more terms and get accurate results without ghost states.

Let's 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):

FIXME FIXME

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 new orthogonal set of functions $|\bar\phi_i\rangle$ satisfying

$$\langle \bar\phi_i | v | \bar\phi_j \rangle = \delta_{ij}$$

27

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

$$|\tilde{\phi}_1\rangle \;=\; \mathbb{1}\frac{1}{\sqrt{\langle\phi_1|v|\phi_1\rangle}}\,|\phi_1\rangle\,; \qquad\qquad\qquad |\bar{\phi}_1\rangle = \frac{1}{\sqrt{\langle\tilde{\phi}_1|v|\tilde{\phi}_1\rangle}}\,|\tilde{\phi}_1\rangle$$

$$|\tilde{\phi}_2\rangle \;=\; \left(\mathbb{1}-|\bar{\phi}_1\rangle\langle\bar{\phi}_1|\,v\right)\frac{1}{\sqrt{\langle\phi_2|v|\phi_2\rangle}}\,|\phi_2\rangle\,; \qquad\qquad |\bar{\phi}_2\rangle = \frac{1}{\sqrt{\langle\tilde{\phi}_2|v|\tilde{\phi}_2\rangle}}\,|\tilde{\phi}_2\rangle$$

$$|\tilde{\phi}_3\rangle \;=\; \left(\mathbb{1}-|\bar{\phi}_1\rangle\langle\bar{\phi}_1|\,v-|\bar{\phi}_2\rangle\langle\bar{\phi}_2|\,v\right)\frac{1}{\sqrt{\langle\phi_3|v|\phi_3\rangle}}\,|\phi_3\rangle\,; \qquad |\bar{\phi}_3\rangle = \frac{1}{\sqrt{\langle\tilde{\phi}_3|v|\tilde{\phi}_3\rangle}}\,|\tilde{\phi}_3\rangle$$

$$\ldots$$

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:

$$\langle\tilde{\phi}_1|v|\tilde{\phi}_1\rangle \;=\; 1$$

$$\langle\tilde{\phi}_2|v|\tilde{\phi}_2\rangle \;=\; 1-\frac{\langle\phi_2|v|\bar{\phi}_1\rangle\langle\bar{\phi}_1|v|\phi_2\rangle}{\langle\phi_2|v|\phi_2\rangle}$$

$$\langle\tilde{\phi}_3|v|\tilde{\phi}_3\rangle \;=\; 1-\frac{\langle\phi_3|v|\bar{\phi}_1\rangle\langle\bar{\phi}_1|v|\phi_3\rangle+\langle\phi_3|v|\bar{\phi}_2\rangle\langle\bar{\phi}_2|v|\phi_3\rangle}{\langle\phi_3|v|\phi_3\rangle}$$

$$\ldots$$

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

$$|\bar{\phi}_1\rangle \;=\; \frac{|\phi_1\rangle}{\sqrt{\langle\phi_1|v|\phi_1\rangle}}$$

$$|\bar{\phi}_2\rangle \;=\; \frac{|\phi_2\rangle\langle\phi_1|v|\phi_1\rangle-|\phi_1\rangle\langle\phi_1|v|\phi_2\rangle}{\sqrt{(\langle\phi_1|v|\phi_1\rangle\langle\phi_2|v|\phi_2\rangle-\langle\phi_2|v|\phi_1\rangle\langle\phi_1|v|\phi_2\rangle)\langle\phi_1|v|\phi_1\rangle\langle\phi_2|v|\phi_2\rangle}}$$

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 G-S 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 $|R_{nl}\rangle$:

$$|\bar{R}_{10}\rangle \;=\; \frac{|R_{10}\rangle}{\sqrt{\langle R_{10}|V_0|R_{10}\rangle}}$$

$$|\bar{R}_{20}\rangle \;=\; \frac{|R_{20}\rangle - |\bar{R}_{10}\rangle\,\langle \bar{R}_{10}|V_0|R_{20}\rangle}{\sqrt{\cdots}}$$

$$|\bar{R}_{21}\rangle \;=\; \frac{|R_{21}\rangle}{\sqrt{\langle R_{21}|V_1|R_{21}\rangle}}$$

$$|\bar{R}_{30}\rangle \;=\; \frac{|R_{30}\rangle - |\bar{R}_{10}\rangle\,\langle \bar{R}_{10}|V_0|R_{30}\rangle - |\bar{R}_{20}\rangle\,\langle \bar{R}_{20}|V_0|R_{30}\rangle}{\sqrt{\cdots}}$$

$$|\bar{R}_{31}\rangle \;=\; \frac{|R_{31}\rangle - |\bar{R}_{21}\rangle\,\langle \bar{R}_{21}|V_1|R_{31}\rangle}{\sqrt{\cdots}}$$

$$|\bar{R}_{32}\rangle \;=\; \frac{|R_{32}\rangle}{\sqrt{\langle R_{32}|V_1|R_{32}\rangle}}$$

$$|\bar{R}_{40}\rangle \;=\; \frac{|R_{40}\rangle - |\bar{R}_{10}\rangle\,\langle \bar{R}_{10}|V_0|R_{40}\rangle - |\bar{R}_{20}\rangle\,\langle \bar{R}_{20}|V_0|R_{40}\rangle - |\bar{R}_{30}\rangle\,\langle \bar{R}_{30}|V_0|R_{40}\rangle}{\sqrt{\cdots}}$$

$$|\bar{R}_{41}\rangle \;=\; \frac{|R_{41}\rangle - |\bar{R}_{21}\rangle\,\langle \bar{R}_{21}|V_1|R_{41}\rangle - |\bar{R}_{31}\rangle\,\langle \bar{R}_{31}|V_1|R_{41}\rangle}{\sqrt{\cdots}}$$

$$\cdots$$

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

$$\langle \bar{R}_{nl}|V_l|\bar{R}_{n'l}\rangle = \delta_{nn'} \tag{34}$$

so for every $V_l$, we construct $|\bar{R}_{nl}\rangle$ for $n = l+1,\ l+2,\cdots$. Let's 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

$$v = \sum_i v\,|\bar{\phi}_i\rangle\,\langle\bar{\phi}_i|\,v = \sum_i V_l(\rho)\,|\bar{R}_{nl}\rangle\,|lm\rangle\,\langle lm|\,\langle\bar{R}_{nl}|\,V_l(\rho) \tag{35}$$

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

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

$$(v\psi)(\mathbf{r}) = \langle\mathbf{r}|v|\psi\rangle = \sum_i \langle\hat{\mathbf{r}}|lm\rangle\,\langle\rho|V_l(\rho)|\bar{R}_{nl}\rangle\,\langle\bar{R}_{nl}|\,V_l(\rho)\,\langle lm|\psi\rangle =$$

$$= \sum_i Y_{lm}(\hat{\mathbf{r}})\bar{R}_{nl}(\rho)V_l(\rho)\int \bar{R}_{nl}(\rho')V_l(\rho')\int Y_{lm}^*(\hat{\mathbf{r}}')\psi(\mathbf{r}')\,\mathrm{d}\Omega'\,\rho'^2\mathrm{d}\rho' =$$

$$= \sum_i Y_{lm}(\hat{\mathbf{r}})\bar{R}_{nl}(\rho)V_l(\rho)\int \bar{R}_{nl}(\rho')V_l(\rho')Y_{lm}^*(\hat{\mathbf{r}}')\psi(\mathbf{r}')\,\mathrm{d}^3r'$$

To have some insight on what we are actually doing: we are making the local potential $V_l$ nonlocal using:

$$V_l = \sum_{n=l+1}^{\infty} V_l \, |\bar{R}_{nl}\rangle \, \langle \bar{R}_{nl}| \, V_l \qquad (36)$$

where

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

or in $\mathbf{r}$ representation:

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

which is useful when computing integrals of this type

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

$$\langle i|V_l|\bar{R}_{nl}\rangle = \int \phi_i(\rho)V_l(\rho)\bar{R}_{nl}(\rho)\rho^2\mathrm{d}\rho$$

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 (36). Of course taking only finite number of terms in (36) is only an approximation to $\hat{V}_l$. In our case, we don't 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 angular parts of $V$ are nonlocal and radial part is local (so the matrix $V_{ij}$ is dense), so 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 (36) 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 (36) 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 $|\bar{R}_{l+1;l}\rangle$. For the function $|R_{l+2;l}\rangle$ all the terms except the first two are zero, because $\langle \bar{R}_{nl}|V_0|R_{l+2;l}\rangle \neq 0$ 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 (34)) For functions, which are a little different from all $|R_{nl}\rangle$ $(n > l)$, we won't genereally get precise results taking any (finite) number of terms in (36), but the higher terms should give smaller and smaller corrections.

So, to sum it up: We take all the $V_l$ in (35) as we did in (32). Theoretically we should take $\bar{R}_{nl}$ for all $n = l+1,\ l+2,\ l+3,\ldots$, but practically it suffices to only take several $\bar{R}_{nl}$ for a given $l$ from one atomic calcuraction.

Let's 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 $i = 1, 3, 4, 5, 10, 11, 12, 13, 14$ and on these 9 functions we get a precise result with only one term in the expansion (36). For the other 5 functions ($i = 2, 6, 7, 8, 9$) we will have to take into account more terms. Let's look in more detail at the case $l = 0$ (i.e. $i = 1, 2, 6$). Then

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

and for the case $i = 1$ we see that one term in (36) 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 (36)

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

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 (36), for 14 electrons actually only 3 terms. For several atoms calculation, we won't get precise results, but it should be a sufficiently good approximation.

The described method is general, the only drawback is that if we don't take functions $|R_{nl}\rangle$ which are similar to the solution, we need to take a lot of terms in (36), resulting in many matrices of the form $p_i p_j^*$, which we don't want, even though, theoretically we can get a solution with any precision we want taking more and more terms in (36).

See also \citeblochl.

# 5   FEM

## 5.1   Introduction

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

## 5.2   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

$$\int -\left(\frac{\hbar^2}{2m}\nabla^2\psi\right) v \, \mathrm{d}V = \int (E-V)\psi v \, \mathrm{d}V \,, \tag{37}$$

and using the vector identity

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

we rewrite the left hand side of (37)

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

now we apply Gauss Theorem

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

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

$$\int \frac{\hbar^2}{2m}\nabla\psi \cdot \nabla v \, \mathrm{d}V + \int vV\psi \, \mathrm{d}V = \int E\psi v \, \mathrm{d}V + \oint \frac{\hbar^2}{2m}\frac{\mathrm{d}\psi}{\mathrm{d}n}v \, \mathrm{d}S \,, \tag{38}$$

which is the weak formulation. The problem reads: find a function $\psi$ such that (38) holds for every $v$.

## 5.3   Finite Elements

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

$$\left(\int \frac{\hbar^2}{2m}\nabla\phi_j \cdot \nabla\phi_i \, \mathrm{d}V + \int \phi_i V \phi_j \, \mathrm{d}V\right) q_j = \left(\int E\phi_j \phi_i \, \mathrm{d}V\right) q_j + \oint \frac{\hbar^2}{2m}\frac{\mathrm{d}\psi}{\mathrm{d}n}\phi_i \, \mathrm{d}S \,, \tag{39}$$

which can be written in a matrix form

$$(K_{ij} + V_{ij})\, q_j = EM_{ij}q_j + F_i \,,$$

where

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

Usually we set $F_i = 0$.

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

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

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

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

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.

## 5.4 Pseudopotentials Formulation

There are no problems with other matrix elements in (39) except

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

where

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

so

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

where the complex vector $p_i$ is given by

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

and

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

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

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

$$= \int \int Y_{lm}(\hat{\mathbf{r}}) Y_{lm}^*(\hat{\mathbf{r}}') V_l(\rho) \bar{R}_{nl}(\rho) \phi_i(\mathbf{r}) \phi_j(\mathbf{r}') V_l(\rho') \bar{R}_{nl}(\rho') \mathrm{d}^3 r \mathrm{d}^3 r' =$$

and using (50) we get

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

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 cancels 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 $p_i = a_i + i b_i$ then

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

and

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

where

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

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

just don't confuse the basis function $\phi_i(\mathbf{r})$ with the spherical integration variable $\phi$.

## 5.5   Example on Si

There are only two valence electrons to take into account, thus we have only 2 $n$ in the summation $(nlm)$. The potentials in the Schrödinger equation are

$$V = V_{local} + V_{nonlocal}$$

$$V_{local} = V_H + V_{XC} + V_{loc}$$

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

There are just $V_0$, $V_1$ and $V_2$.

We probably need to calculate $\bar{R}_{10}$, $\bar{R}_{20}$, $\bar{R}_{21}$, $\bar{R}_{30}$, $\bar{R}_{31}$, $\bar{R}_{32}$, but I am not completely sure. We get maybe around 18 complex matrices, which means 36 real matrices of the form $a_i a_j$. The input for the solver is 36 real vectors $a$, $b$, $c$, $d$, $e$, \dots and sparse matrices $V^{loc}$, $M$, and $K$. The solver needs to solve

$$(K + V^{loc} + a^T a + b^T b + c^T c + d^T d + e^T e + \cdots) q = EMq$$

34

# 6 Appendix

## 6.1 Polar and Spherical Coordinates

Polar coordinates (radial, azimuth) $(r, \phi)$ are defined by

$$
\begin{aligned}
x &= r \cos \phi \\
y &= r \sin \phi
\end{aligned}
$$

Spherical coordinates (radial, zenith, azimuth) $(\rho, \theta, \phi)$:

$$
\begin{aligned}
x &= \rho \sin \theta \cos \phi \\
y &= \rho \sin \theta \sin \phi \\
z &= \rho \cos \theta
\end{aligned}
$$

Note: this meaning of $(\theta, \phi)$ is mostly used in the USA and in many books. In Europe people usualy use different symbols, like $(\phi, \theta)$, $(\vartheta, \varphi)$ and others.

## 6.2 Delta Function

Delta function $\delta(x)$ is defined such that this relation holds:

$$
\int f(x)\delta(x - t)\mathrm{d}x = f(t) \tag{40}
$$

No such function exists, but one can find many sequences "converging" to a delta function:

$$
\lim_{\alpha \to \infty} \delta_\alpha(x) = \delta(x) \tag{41}
$$

more precisely:

$$
\lim_{\alpha \to \infty} \int f(x)\delta_\alpha(x)\mathrm{d}x = \int f(x) \lim_{\alpha \to \infty} \delta_\alpha(x)\mathrm{d}x = f(t) \tag{42}
$$

one example of such a sequence is:

$$
\delta_\alpha(x) = \frac{1}{\pi x} \sin(\alpha x)
$$

It's clear that (42) holds for any well behaved function $f(x)$. Mathematicians like to say it's incorrect to use such a notation when in fact the integral (40) doesn't "exist", but they are wrong, because it is not important if something "exist" or not, but rather if it is clear what we mean by our notation: (40) is a shorthand for (42) and (41) gets a mathematically rigorous meaning when you integrate both sides and use (40) to arrive at (42). Thus one uses the relations (40), (41), (42) to derive all properties of the delta function.

Let's give an example. Let $\hat{\mathbf{r}}$ be the unit vector in 3D and we can label it using spherical coordinates $\hat{\mathbf{r}} = \hat{\mathbf{r}}(\theta, \phi)$. We can also express it in cartesian coordinates as $\hat{\mathbf{r}}(\theta, \phi) = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta)$.

$$f(\hat{\mathbf{r}}') = \int \delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') f(\hat{\mathbf{r}}) \, \mathrm{d}\hat{\mathbf{r}} \tag{43}$$

Expressing $f(\hat{\mathbf{r}}) = f(\theta, \phi)$ as a function of $\theta$ and $\phi$ we have

$$f(\theta', \phi') = \int \delta(\theta - \theta')\delta(\phi - \phi') f(\theta, \phi) \, \mathrm{d}\theta\mathrm{d}\phi \tag{44}$$

Expressing (43) in spherical coordinates we get

$$f(\theta', \phi') = \int \delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') f(\theta, \phi) \sin\theta \, \mathrm{d}\theta\mathrm{d}\phi$$

and comparing to (44) we finally get

$$\delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') = \frac{1}{\sin\theta} \delta(\theta - \theta')\delta(\phi - \phi')$$

In exactly the same manner we get

$$\delta(\mathbf{r} - \mathbf{r}') = \delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') \frac{\delta(\rho - \rho')}{\rho^2}$$

See also (45) for an example of how to deal with more complex expressions involving the delta function like $\delta^2(x)$.

## 6.3 Variations and functional derivatives

Functional derivatives are a common source of confusion and especially the notation. The reason is similar to the delta function — it's due to rigid mathematicians, who refuse to use a clear and exact notation and unfortunately they are allowed to teach the basic calculus to physics students.

Let's have $\mathbf{x} = (x_1, x_2, \ldots, x_N)$. The function $f(\mathbf{x})$ assigns a number to each $\mathbf{x}$. We define a differential of $f$ as

$$\mathrm{d}f \equiv \left. \frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x} + \varepsilon\mathbf{h}) \right|_{\varepsilon=0} = \lim_{\varepsilon \to 0} \frac{f(\mathbf{x} + \varepsilon\mathbf{h}) - f(\mathbf{x})}{\varepsilon} = \mathrm{a} \cdot \mathrm{h}$$

The last equality follows from the fact, that $\left. \frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x} + \varepsilon\mathbf{h}) \right|_{\varepsilon=0}$ is a linear function of h. We define $\frac{\partial f}{\partial x_i}$ as

$$\mathrm{a} \equiv \left( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_N} \right)$$

This also gives a formula for computing $\frac{\partial f}{\partial x_i}$: we set $h_j = \delta_{ij}h_i$ and

$$\frac{\partial f}{\partial x_i} = a_i = \mathrm{a} \cdot \mathrm{h} = \left. \frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(\mathbf{x} + \varepsilon(0, 0, \ldots, 1, \ldots, 0)) \right|_{\varepsilon=0} =$$

$$= \lim_{\varepsilon \to 0} \frac{f(x_1, x_2, \ldots, x_i + \varepsilon, \ldots, x_N) - f(x_1, x_2, \ldots, x_i, \ldots, x_N)}{\varepsilon}$$

But this is just the way the partial derivative is usually defined. Every variable can be treated as a function (very simple one):

$$x_i = g(x_1, \ldots, x_N) = \delta_{ij} x_j$$

and so we define

$$\mathrm{d}x_i \equiv \mathrm{d}g = \mathrm{d}(\delta_{ij} x_j) = h_i$$

and thus we write $h_i = \mathrm{d}x_i$ and $\mathrm{h} = \mathrm{d}\mathbf{x}$ and

$$\mathrm{d}f = \frac{\mathrm{d}f}{\mathrm{d}x_i} \mathrm{d}x_i$$

So $\mathrm{d}\mathbf{x}$ has two meanings — it's either $\mathrm{h} = \mathbf{x} - \mathbf{x}_0$ (a finite change in the independent variable $\mathbf{x}$) or a differential, depending on the context. Even mathematicians use this notation.

Functional $F[f]$ assigns a number to each function $f(x)$. The variation is defined as

$$\delta F[f] \equiv \frac{\mathrm{d}}{\mathrm{d}\varepsilon} F[f + \varepsilon h] \Big|_{\varepsilon=0} = \lim_{\epsilon \to 0} \frac{F[f + \epsilon h] - F[f]}{\epsilon} = \int a(x) h(x) \mathrm{d}x$$

We define $\frac{\delta F}{\delta f(x)}$ as

$$a(x) \equiv \frac{\delta F}{\delta f(x)}$$

This also gives a formula for computing $\frac{\delta F}{\delta f(x)}$: we set $h(y) = \delta(x - y)$ and

$$\frac{\delta F}{\delta f(x)} = a(x) = \int a(y) \delta(x - y) \mathrm{d}y = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} F[f(y) + \varepsilon \delta(x - y)] \Big|_{\varepsilon=0} =$$

$$= \lim_{\varepsilon \to 0} \frac{F[f(y) + \varepsilon \delta(x - y)] - F[f(y)]}{\varepsilon}$$

Every function can be treated as a functional (although a very simple one):

$$f(x) = G[f] = \int f(y) \delta(x - y) \mathrm{d}y$$

and so we define

$$\delta f \equiv \delta G[f] = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} G[f(x) + \varepsilon h(x)] \Big|_{\varepsilon=0} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} (f(x) + \varepsilon h(x)) \Big|_{\varepsilon=0} = h(x)$$

thus we write $h = \delta f$ and

$$\delta F[f] = \int \frac{\delta F}{\delta f(x)} \delta f(x) \mathrm{d}x$$

so $\delta f$ have two meanings — it's either $h(x) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(x) + \varepsilon h(x))\big|_{\varepsilon=0}$ (a finite change in the function $f$) or a variation of a functional, depending on the context. Mathematicians never write $\delta f$ in the meaning of $h(x)$, they always write the latter, but it's ridiculous, because it is completely analogous to $\mathrm{d}\mathbf{x}$.

The correspondence between the finite and infinite dimensional case can be summarized as:

$$
\begin{aligned}
f(x_i) &\iff F[f] \\
\mathrm{d}f = 0 &\iff \delta F = 0 \\
\frac{\partial f}{\partial x_i} = 0 &\iff \frac{\delta F}{\delta f(x)} = 0 \\
f &\iff F \\
x_i &\iff f(x) \\
x &\iff f \\
i &\iff x
\end{aligned}
$$

More generally, $\delta$-variation can by applied to any function $g$ which contains the function $f(x)$ being varied, you just need to replace $f$ by $f + \epsilon h$ and apply $\frac{\mathrm{d}}{\mathrm{d}\epsilon}$ to the whole $g$, for example (here $g = \partial_\mu \phi$ and $f = \phi$):

$$
\delta \partial_\mu \phi = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \partial_\mu (\phi + \varepsilon h)\bigg|_{\varepsilon=0} = \partial_\mu \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(\phi + \varepsilon h)\bigg|_{\varepsilon=0} = \partial_\mu \delta \phi
$$

This notation allows us a very convinient computation, as shown in the following examples. First, when computing a variation of some integral, when can interchange $\delta$ and $\int$:

$$
F[f] = \int K(x) f(x) \mathrm{d}x
$$

$$
\delta F = \delta \int K(x) f(x) \mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int K(x)(f + \varepsilon h)\mathrm{d}x\bigg|_{\varepsilon=0} = \int \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(K(x)(f + \varepsilon h))\mathrm{d}x\bigg|_{\varepsilon=0} =
$$

$$
= \int \delta(K(x) f(x)) \mathrm{d}x
$$

In the expression $\delta(K(x)f(x))$ we must understand from the context if we are treating it as a functional of $f$ or $K$. In our case it's a functional of $f$, so we have $\delta(Kf) = K\delta f$.

A few more examples:

$$
\frac{\delta}{\delta f(t)} \int \mathrm{d}t' f(t') g(t') = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \int \mathrm{d}t'(f(t') + \varepsilon \delta(t - t')) g(t')\bigg|_{\varepsilon=0} = g(t)
$$

$$
\frac{\delta f(t')}{\delta f(t)} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(t') + \varepsilon \delta(t - t'))\bigg|_{\varepsilon=0} = \delta(t - t')
$$

$$\frac{\delta f(t_1)f(t_2)}{\delta f(t)} = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}(f(t_1) + \varepsilon\delta(t-t_1))(f(t_2) + \varepsilon\delta(t-t_2))\bigg|_{\varepsilon=0} = \delta(t-t_1)f(t_2)+f(t_1)\delta(t-t_2)$$

$$\frac{\delta}{\delta f(t)}\tfrac{1}{2}\int \mathrm{d}t_1\mathrm{d}t_2 K(t_1,t_2)f(t_1)f(t_2) = \tfrac{1}{2}\int \mathrm{d}t_1\mathrm{d}t_2 K(t_1,t_2)\frac{\delta f(t_1)f(t_2)}{\delta f(t)} =$$

$$= \tfrac{1}{2}\left(\int \mathrm{d}t_1 K(t_1,t)f(t_1) + \int \mathrm{d}t_2 K(t,t_2)f(t_2)\right) = \int \mathrm{d}t_2 K(t,t_2)f(t_2)$$

The last equality follows from $K(t_1,t_2) = K(t_2,t_1)$ (any antisymmetrical part of a $K$ would not contribute to the symmetrical integration).

$$\frac{\delta}{\delta f(t)}\int f^3(x)\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int (f(x) + \varepsilon\delta(x-t))^3\mathrm{d}x\bigg|_{\varepsilon=0} =$$

$$= \int 3(f(x) + \varepsilon\delta(x-t))^2\delta(x-t)\mathrm{d}x\bigg|_{\varepsilon=0} = \int 3f^2(x)\delta(x-t)\mathrm{d}x = 3f^2(t)$$

Some mathematicians would say the above calculation is incorrect, because $\delta^2(x-t)$ is undefined. But that's not true, because in case of such problems the above notation automatically implies working with some sequence $\delta_\alpha(x) \to \delta(x)$ (for example $\delta_\alpha(x) = \frac{1}{\pi x}\sin(\alpha x)$) and taking the limit $\alpha \to \infty$:

$$\frac{\delta}{\delta f(t)}\int f^3(x)\mathrm{d}x = \lim_{\alpha\to\infty}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\int (f(x) + \varepsilon\delta_\alpha(x-t))^3\mathrm{d}x\bigg|_{\varepsilon=0} =$$

$$= \lim_{\alpha\to\infty}\int 3(f(x) + \varepsilon\delta_\alpha(x-t))^2\delta_\alpha(x-t)\mathrm{d}x\bigg|_{\varepsilon=0} = \lim_{\alpha\to\infty}\int 3f^2(x)\delta_\alpha(x-t)\mathrm{d}x =$$

$$= \int 3f^2(x)\lim_{\alpha\to\infty}\delta_\alpha(x-t)\mathrm{d}x = \int 3f^2(x)\delta(x-t)\mathrm{d}x = 3f^2(t) \qquad (45)$$

As you can see, we got the same result, with the same rigor, but using an obfuscating notation. That's why such obvious manipulations with $\delta_\alpha$ are tacitly implied.

## 6.4   Spherical Harmonics

Are defined by

$$Y_{lm}(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}\, P_l^m(\cos\theta)\, e^{im\phi}$$

where $P_l^m$ are associated Legendre polynomials defined by

$$P_l^m(x) = (-1)^m(1-x^2)^{m/2}\frac{\mathrm{d}^m}{\mathrm{d}x^m}P_l(x)$$

and $P_l$ are Legendre polynomials defined by the formula

$$P_l(x) = \frac{1}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}x^l}[(x^2-1)^l]$$

they also obey the completeness relation

$$\sum_{l=0}^{\infty} \frac{2l+1}{2} P_l(x') P_l(x) = \delta(x - x') \tag{46}$$

The spherical harmonics are ortonormal:

$$\int Y_{lm} Y_{l'm'}^* \, d\Omega = \int_0^{2\pi} \int_0^{\pi} Y_{lm}(\theta, \phi) \, Y_{l'm'}^*(\theta, \phi) \sin\theta \, d\theta \, d\phi = \delta_{mm'} \delta_{ll'} \tag{47}$$

and complete (both in the $l$-subspace and the whole space):

$$\sum_{m=-l}^{l} |Y_{lm}(\theta, \phi)|^2 = \frac{2l+1}{4\pi} \tag{48}$$

$$\sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_{lm}(\theta, \phi) Y_{lm}^*(\theta', \phi') = \frac{1}{\sin\theta} \delta(\theta - \theta') \delta(\phi - \phi') = \delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') \tag{49}$$

The relation (48) is a special case of an addition theorem for spherical harmonics

$$\sum_{m=-l}^{l} Y_{lm}(\theta, \phi) Y_{lm}^*(\theta', \phi') = \frac{4\pi}{2l+1} P_l(\cos\gamma) \tag{50}$$

where $\gamma$ is the angle between the unit vectors given by $\hat{\mathbf{r}} = (\theta, \phi)$ and $\hat{\mathbf{r}}' = (\theta', \phi')$:

$$\cos\gamma = \cos\theta \cos\theta' + \sin\theta \sin\theta' \cos(\phi - \phi') = \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'$$

## 6.5 Dirac Notation

The Dirac notation allows a very compact and powerful way of writing equations that describe a function expansion into a basis, both discrete (e.g. a fourier series expansion) and continuous (e.g. a fourier transform) and related things. The notation is designed so that it is very easy to remember and it just guides you to write the correct equation.

Let's have a function $f(x)$. We define

$$
\begin{aligned}
\langle x | f \rangle &\equiv f(x) \\
\langle x' | f \rangle &\equiv f(x') \\
\langle x' | x \rangle &\equiv \delta(x' - x) \\
\int |x\rangle \langle x| \, dx &\equiv \mathbb{1}
\end{aligned}
$$

The following equation

$$f(x') = \int \delta(x' - x) f(x) dx$$

then becomes

$$\langle x'|f\rangle = \int \langle x'|x\rangle \langle x|f\rangle \, \mathrm{d}x$$

and thus we can interpret $|f\rangle$ as a vector, $|x\rangle$ as a basis and $\langle x|f\rangle$ as the coefficients in the basis expansion:

$$|f\rangle = \mathbb{1} \, |f\rangle = \int |x\rangle \langle x| \, \mathrm{d}x \, |f\rangle = \int |x\rangle \langle x|f\rangle \, \mathrm{d}x$$

That's all there is to it. Take the above rules as the operational definition of the Dirac notation. It's like with the delta function - written alone it doesn't have any meaning, but there are clear and nonambiguous rules to convert any expression with $\delta$ to an expression which even mathematicians understand (i.e. integrating, applying test functions and using other relations to get rid of all $\delta$ symbols in the expression – but the result is usually much more complicated than the original formula). It's the same with the ket $|f\rangle$: written alone it doesn't have any meaning, but you can always use the above rules to get an expression that make sense to everyone (i.e. attaching any bra to the left and rewritting all brackets $\langle a|b\rangle$ with their equivalent expressions) – but it will be more complex and harder to remember and – that is important – less general.

Now, let's look at the spherical harmonics:

$$Y_{lm}(\hat{\mathbf{r}}) \equiv \langle \hat{\mathbf{r}}|lm\rangle$$

on the unit sphere, we have

$$\int |\hat{\mathbf{r}}\rangle \langle \hat{\mathbf{r}}| \, \mathrm{d}\hat{\mathbf{r}} = \int |\hat{\mathbf{r}}\rangle \langle \hat{\mathbf{r}}| \, \mathrm{d}\Omega = \mathbb{1}$$

$$\delta(\hat{\mathbf{r}} - \hat{\mathbf{r}}') = \langle \hat{\mathbf{r}}|\hat{\mathbf{r}}'\rangle$$

thus

$$\int_0^{2\pi} \int_0^{\pi} Y_{lm}(\theta,\phi) \, Y_{l'm'}^*(\theta,\phi) \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi = \int \langle l'm'|\hat{\mathbf{r}}\rangle \langle \hat{\mathbf{r}}|lm\rangle \, \mathrm{d}\Omega = \langle l'm'|lm\rangle$$

and from (47) we get

$$\langle l'm'|lm\rangle = \delta_{mm'}\delta_{ll'}$$

now

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

from (49) we get

$$\sum_{lm} \langle \hat{\mathbf{r}}|lm\rangle \langle lm|\hat{\mathbf{r}}'\rangle = \langle \hat{\mathbf{r}}|\hat{\mathbf{r}}'\rangle$$

so we have

$$\sum_{lm} |lm\rangle \langle lm| = \mathbb{1}$$

so $|lm\rangle$ forms an ortonormal basis. Any function defined on the sphere $f(\hat{\mathbf{r}})$ can be written using this basis:

$$f(\hat{\mathbf{r}}) = \langle \hat{\mathbf{r}}|f\rangle = \sum_{lm} \langle \hat{\mathbf{r}}|lm\rangle \langle lm|f\rangle = \sum_{lm} Y_{lm}(\hat{\mathbf{r}})f_{lm}$$

where

$$f_{lm} = \langle lm|f\rangle = \int \langle lm|\hat{\mathbf{r}}\rangle \langle \hat{\mathbf{r}}|f\rangle \, d\Omega = \int Y_{lm}^*(\hat{\mathbf{r}})f(\hat{\mathbf{r}})d\Omega$$

If we have a function $f(\mathbf{r})$ in 3D, we can write it as a function of $\rho$ and $\hat{\mathbf{r}}$ and expand only with respect to the variable $\hat{\mathbf{r}}$:

$$f(\mathbf{r}) = f(\rho\hat{\mathbf{r}}) \equiv g(\rho, \hat{\mathbf{r}}) = \sum_{lm} Y_{lm}(\hat{\mathbf{r}})g_{lm}(\rho)$$

In Dirac notation we are doing the following: we decompose the space into the angular and radial part

$$|\mathbf{r}\rangle = |\hat{\mathbf{r}}\rangle \otimes |\rho\rangle \equiv |\hat{\mathbf{r}}\rangle |\rho\rangle$$

and write

$$f(\mathbf{r}) = \langle \mathbf{r}|f\rangle = \langle \hat{\mathbf{r}}| \langle \rho|f\rangle = \sum_{lm} Y_{lm}(\hat{\mathbf{r}}) \langle lm| \langle \rho|f\rangle$$

where

$$\langle lm| \langle \rho|f\rangle = \int \langle lm|\hat{\mathbf{r}}\rangle \langle \hat{\mathbf{r}}| \langle \rho|f\rangle \, d\Omega = \int Y_{lm}^*(\hat{\mathbf{r}})f(\mathbf{r})d\Omega$$

Let's calculate $\langle \rho|\rho'\rangle$

$$\langle \mathbf{r}|\mathbf{r}'\rangle = \langle \hat{\mathbf{r}}| \langle \rho|\rho'\rangle |\hat{\mathbf{r}}'\rangle = \langle \hat{\mathbf{r}}|\hat{\mathbf{r}}'\rangle \langle \rho|\rho'\rangle$$

so

$$\langle \rho|\rho'\rangle = \frac{\langle \mathbf{r}|\mathbf{r}'\rangle}{\langle \hat{\mathbf{r}}|\hat{\mathbf{r}}'\rangle} = \frac{\delta(\rho - \rho')}{\rho^2}$$

We must stress that $|lm\rangle$ only acts in the $|\hat{\mathbf{r}}\rangle$ space (not the $|\rho\rangle$ space) which means that

$$\langle \mathbf{r}|lm\rangle = \langle \hat{\mathbf{r}}| \langle \rho|lm\rangle = \langle \hat{\mathbf{r}}|lm\rangle \langle \rho| = Y_{lm}(\hat{\mathbf{r}}) \langle \rho|$$

and $V|lm\rangle$ leaves $V|\rho\rangle$ intact. Similarly,

$$\sum_{lm} |lm\rangle \langle lm| = \mathbb{1}$$

is a unity in the $|\hat{\mathbf{r}}\rangle$ space only (i.e. on the unit sphere).

Let's rewrite the equation (50):

$$\sum_{m} \langle \hat{\mathbf{r}}|lm\rangle \langle lm|\hat{\mathbf{r}}'\rangle = \frac{4\pi}{2l+1} \langle \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'|P_l\rangle$$

Using the completeness relation (46):

$$\sum_l \frac{2l+1}{2} \langle x'|P_l\rangle \langle P_l|x\rangle = \langle x'|x\rangle$$

$$\sum_l |P_l\rangle \frac{2l+1}{2} \langle P_l| = \mathbb{1}$$

we can now derive a very important formula true for every function $f(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')$:

$$f(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') = \langle \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'|f\rangle = \sum_l \langle \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}'|P_l\rangle \frac{2l+1}{2} \langle P_l|f\rangle = \sum_{lm} \langle \hat{\mathbf{r}}|lm\rangle \langle lm|\hat{\mathbf{r}}'\rangle \frac{(2l+1)^2}{8\pi} \langle P_l|f\rangle =$$

$$= \sum_{lm} \langle \hat{\mathbf{r}}|lm\rangle f_l \langle lm|\hat{\mathbf{r}}'\rangle$$

where

$$f_l = \frac{(2l+1)^2}{8\pi} \langle P_l|f\rangle = \frac{(2l+1)^2}{8\pi} \int_{-1}^1 \langle P_l|x\rangle \langle x|f\rangle \, \mathrm{d}x = \frac{(2l+1)^2}{8\pi} \int_{-1}^1 P_l(x) f(x) \mathrm{d}x$$

or written explicitly

$$f(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_{lm}(\hat{\mathbf{r}}) f_l Y_{lm}^*(\hat{\mathbf{r}}') \tag{51}$$