German Research School
for Simulation Sciences

Calculations of atomic multiplets across the periodic table

Master's Thesis

Qian Zhang

September 2014

Supervisor
Prof. Dr. Erik Koch

Examiner
Prof. Dr. Erik Koch

Co-Examiner
Prof. Dr. Eva Pavarini

JÜLICH FORSCHUNGSZENTRUM

RWTH AACHEN UNIVERSITY

Abstract

The goal of this thesis is to develop a simulation tool for the calculation and visualization of the multiplet structure of all atoms across the periodic table. The starting point are self-consistent calculations in the spherical-potential approximation. For the resulting atomic levels we calculate the ab-initio Slater parameters that define the electron-electron repulsion term of the many-body Hamiltonian. We then construct the eigen-states of the Hamiltonian on an open shell by constructing the multiplet states using the angular momentum ladder operators and, where necessary, seniority. Finally, we include spin-orbit coupling, using the ab-initio coupling constants determined from the self-consistent radial potentials.

Contents

Chapter 1

Introduction

1.1. Many-body problem in quantum mechanics

Imagine our solar system: it consists of a heavy sun with eight planets orbiting around. In classical mechanics, the goal is to calculate the position of each celestial object as a function of time. The motion of the system is governed by Newton's second law: $\mathbf{F} = m\mathbf{a}$. While it is almost impossible to solve the classical many-body problem analytically, the numerical approach is rather straightforward. However, in quantum mechanics, the many-body problem is a completely different story.

Scaling down to 10^{-10} meters, it is even philosophically profound that we find similar structures in atomic systems as in our cosmic systems. An atom, which consists of a heavy nucleus, with electric charge Ze, is surrounded by N electrons with charge $-e$. The system here is no longer governed by Newton's second law, neither is our goal to compute the "position" of each electron as a function of time. In quantum mechanics, the position of an electron is no longer well defined. Instead, what we are looking for is the electrons' wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ which is the solution of the Schrödinger equation[1]

$$H\Psi = E\Psi \tag{1.1}$$

where H, the Hamiltonian of the system, is given by

$$H = \sum_{i=1}^{N} \left[-\frac{\hbar^2}{2m_e}\nabla_i^2 - \frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r_i} \right] + \sum_{i<j}^{N} \frac{1}{4\pi\epsilon_0}\frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1.2}$$

The term in the first sum represents the kinetic plus potential energy of the ith electron, in the electric field of the nucleus. The last term, which complicates the behavior of the system, describes the electron-electron repulsions among all N electrons (the constraint $i < j$ avoids double counting over electron pairs).

In neither classical nor quantum mechanics, can one solve many-body problems exactly. But in quantum mechanics, even a numerical solution is extremely difficult to obtain. The difficulty comes from the fact that a many-body wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ is an object with dimension $3N$. One must generate a mesh grid with $3N$-dimension to represent the wave function

[1] Strictly speaking, Eqn. (1.1) is called the time-independent Schrödinger equation and the solution $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ is the time-independent part of the wave function.

numerically and such a gigantic data is not possible to be stored on an ordinary hard disk. Consequently, different methods have been developed to simplify the problem and to treat the system approximately.

1.2. Atomic units

Since we are going to solve our problems numerically, it is useful to pay attention to the choice of units. Certainly, we can use SI units, but the scales would be inconvenient. For example, in SI units, the reduced Planck constant reads $\hbar = 1.054572 \times 10^{-34}\mathrm{J}\cdot\mathrm{s}$, which is a crazy number from a computational point of view. Hence, we employ atomic units (a.u.), namely,

$$\begin{aligned} &\text{Length:} && 1\ \mathrm{a_0} \approx 5.2918 \times 10^{-11}\ \mathrm{m} \\ &\text{Mass:} && 1\ \mathrm{m_e} \approx 9.1095 \times 10^{-31}\ \mathrm{kg} \\ &\text{Time:} && 1\ \mathrm{t_0} \approx 2.4189 \times 10^{-17}\ \mathrm{s} \\ &\text{Charge:} && 1\ \mathrm{e} \approx 1.6022 \times 10^{-19}\ \mathrm{C} \end{aligned}$$

which are deliberately chosen such that

$$\begin{aligned} \hbar &= 1\ \mathrm{a_0^2 m_e t_0^{-1}} \\ m_e &= 1\ \mathrm{m_e} \\ e &= 1\ \mathrm{e} \\ 4\pi\epsilon_0 &= 1\ \mathrm{a_0^{-3} m_e t_0^2 e^2} \end{aligned} \tag{1.3}$$

By adopting atomic units, the Hamiltonian in Eqn. (1.2) simplifies to

$$H = \sum_{i=1}^{N}\left[-\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i}\right] + \sum_{i<j}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{1.4}$$

With the choice of atomic units, distances are given in units of the Bohr radius ($\mathrm{a_0}$) and energies in Hartree ($\mathrm{a_0^2 m_e t_0^{-2}}$)

$$1\ \text{Hartree} = 2\ \text{Rydberg} \approx 27.2114\ \text{eV} \approx 4.3597 \times 10^{-18}\ \text{J} \tag{1.5}$$

From now on, we will keep using atomic units to simplify our equations and discussions.

1.3. Convention of notations

I will try to make the notations in my thesis as consistent as possible to avoid confusions and to make the discussion clear.

For a many-electron wave function, we use capital psi

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$$

For a one-electron wave function, we use lower case phi

$$\varphi(\mathbf{r})$$

For a Slater determinant, we use upper case phi

$$\Phi_{\alpha_1\alpha_2\cdots\alpha_N}(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$$

Slater determinants are constructed from one-electron wave functions. They are basis functions of anti-symmetric many-electron wave functions.

$$\Phi_{\alpha_1\cdots\alpha_N}(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_{\alpha_1}(\mathbf{r}_1) & \varphi_{\alpha_2}(\mathbf{r}_1) & \cdots & \varphi_{\alpha_N}(\mathbf{r}_1) \\ \varphi_{\alpha_1}(\mathbf{r}_2) & \varphi_{\alpha_2}(\mathbf{r}_2) & \cdots & \varphi_{\alpha_N}(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{\alpha_1}(\mathbf{r}_N) & \varphi_{\alpha_2}(\mathbf{r}_N) & \cdots & \varphi_{\alpha_N}(\mathbf{r}_N) \end{vmatrix} \tag{1.6}$$

Sometimes you hear people say, "A many-electron wave function is a Slater determinant." That is wrong. A many-electron wave function is in general a linear combination of Slater determinants, not just one (although it could be). Nevertheless, this real-space representation of Slater determinant in Eqn. (1.6) will not appear in our discussion (it only appears in Appendix B). Slater determinants will be represented in the form of second quantization[2] when later we construct our multiplet states.

[2]Second quantization is a very convenient "algebra" for handling many-body states. A brief discussion of second quantization is given in Appendix B.

Chapter 2

The shooting and matching methods

2.1. The one-electron problem

To start with, we consider the simplest case: the one-electron or hydrogen-like system. For $N = 1$, Eqn. (1.1) (in a.u.) reads

$$\left[-\frac{1}{2}\nabla^2 + V(r)\right]\varphi = E\varphi \tag{2.1}$$

where $V(r)$ is a spherically symmetric potential (for one-electron case: $V(r) = -Z/r$). In spherical coordinates, by separation of variables $\varphi(r,\theta,\phi) = R(r)Y(\theta,\phi)$, Eqn. (2.1) splits into two equations, namely,

$$\text{Radial equation:} \quad \frac{d}{dr}\left(r^2\frac{dR}{dr}\right) - 2r^2\left[V(r) - E\right]R = \quad l(l+1)R \tag{2.2}$$

$$\text{Angular equation:} \quad \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2} = -l(l+1)Y \tag{2.3}$$

where l is a non-negative integer[1] known as the orbital angular momentum quantum number. Now, a three dimensional ordinary differential equation is separated into two equations. They are called radial equation and angular equation because Eqn. (2.2) has only the dependence on the radial coordinate r and Eqn. (2.3) has only the dependence on the angular coordinate θ and ϕ. It is very important to realize that Eqn. (2.3) has no dependence on the potential $V(r)$. Hence Eqn. (2.3) is universal for all spherically symmetric potentials and its analytical solutions are the well known spherical harmonics [1]. In other words, there is no need to worry about Eqn. (2.3) since its solutions, the spherical harmonics $Y_{lm}(\theta,\phi)$ (with $l = 0,\ 1,\ \ldots$ and $m = -l,\ \ldots,\ l$), are known exactly (see Table A.1). The more difficult task for us is to solve Eqn. (2.2). In fact, for this simplest one-electron system, the radial part solutions are also known analytically. As a reference, Table 2.1 lists the first few radial wave functions for hydrogen-like atoms. A rescaled coordinate $\rho = Zr$ is used in the expressions. Nevertheless, we are going to solve the radial equation numerically so that we can further deal with the more general case, the many-electron system.

[1] In fact, from the separation of variables, there is no evidence that l should be an integer. l could be any complex number. However, the boundary condition $Y(\theta,\phi) = Y(\theta,\phi+2\pi)$ and the properties of the Legendre polynomial require that l must be a non-negative integer.

Eqn. (2.2) simplifies if we change variables $u(r) \equiv rR(r)$. The radial equation becomes

$$\boxed{-\frac{1}{2}\frac{d^2u}{dr^2} + \left[V(r) + \frac{l(l+1)}{2r^2}\right] u = Eu} \tag{2.4}$$

with boundary conditions: $u(r) \propto r^{l+1}$ as $r \to 0$ and $u(r) \propto \exp\left(-\sqrt{-2E}r\right)$ as $r \to \infty$.

Eqn. (2.4) is identical in form to the one-dimensional Schrödinger equation, except that we have an extra "centrifugal" term $[l(l+1)/2r^2]$ in addition to the external potential. Our task is to solve this equation for $u(r)$ and determine the allowed energies E.

Table 2.1.: The first few analytical radial wave functions $u_{nl}(r)$ for hydrogen-like atoms with eigen-energies $E_n = -\frac{Z^2}{2n^2}$ (Hartree). A rescaled coordinate $\rho = Zr$ is used to simplify the expressions and the radial coordinates r are in units of Bohr radius (a.u.).

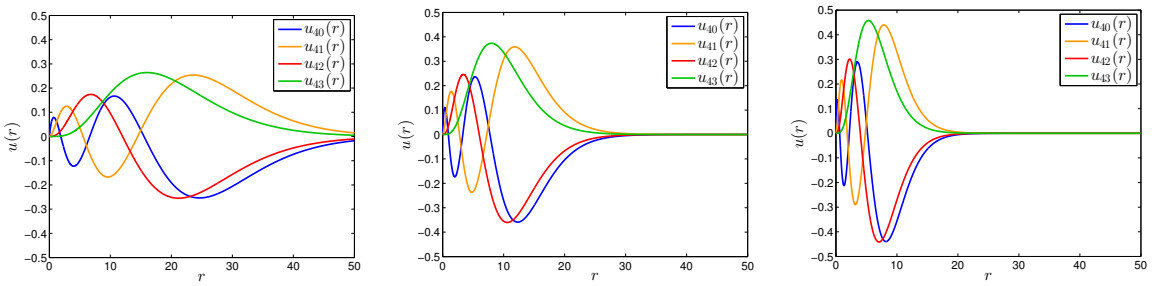
$u_{10} =$	2	$\sqrt{Z}\rho$		$\exp(-\rho)$
$u_{20} =$	$\frac{1}{\sqrt{2}}$	$\sqrt{Z}\rho$	$\left(1 - \frac{1}{2}\rho\right)$	$\exp(-\rho/2)$
$u_{21} =$	$\frac{1}{\sqrt{24}}$	$\sqrt{Z}\rho^2$		$\exp(-\rho/2)$
$u_{30} =$	$\frac{2}{\sqrt{27}}$	$\sqrt{Z}\rho$	$\left(1 - \frac{2}{3}\rho + \frac{2}{27}\rho^2\right)$	$\exp(-\rho/3)$
$u_{31} =$	$\frac{8}{27\sqrt{6}}$	$\sqrt{Z}\rho^2$	$\left(1 - \frac{1}{6}\rho\right)$	$\exp(-\rho/3)$
$u_{32} =$	$\frac{4}{81\sqrt{30}}$	$\sqrt{Z}\rho^3$		$\exp(-\rho/3)$
$u_{40} =$	$\frac{1}{4}$	$\sqrt{Z}\rho$	$\left(1 - \frac{3}{4}\rho + \frac{1}{8}\rho^2 - \frac{1}{192}\rho^3\right)$	$\exp(-\rho/4)$
$u_{41} =$	$\frac{\sqrt{5}}{16\sqrt{3}}$	$\sqrt{Z}\rho^2$	$\left(1 - \frac{1}{4}\rho + \frac{1}{80}\rho^2\right)$	$\exp(-\rho/4)$
$u_{42} =$	$\frac{1}{64\sqrt{5}}$	$\sqrt{Z}\rho^3$	$\left(1 - \frac{1}{12}\rho\right)$	$\exp(-\rho/4)$
$u_{43} =$	$\frac{1}{768\sqrt{35}}$	$\sqrt{Z}\rho^4$		$\exp(-\rho/4)$

2.2. Logarithmic grid

We are now in a position to solve the one-dimensional ordinary differential equation in (2.4) numerically. It is often convenient to solve problems on uniform grids. However, the curvature[2] of the wave function indicates that the function $u(r)$ oscillates faster if r goes smaller. This suggests us to take more points for small r but fewer for large r. For instance, Fig. 2.1 plots the exact wave functions from Table 2.1 with principal quantum number $n = 4$ for three hydrogen-like atoms. These plots demonstrate that the smaller r goes, the stronger is the oscillation of the

[2]The second derivative $u'' = -2[E - V(r) - l(l+1)/2r^2]u$.

function $u(r)$. And the wave functions are attracted more closely to the nucleus with increasing nuclear charge Z. Many grid points will be wasted to sample the non-fruitful outer regions where the wave function is almost zero (see Fig. 2.1c). Therefore, an adaptive grid with higher resolution close to the nucleus is desired.

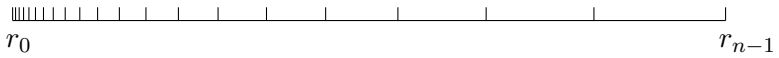


(a) $Z = 1$ (b) $Z = 2$ (c) $Z = 3$

Figure 2.1.: Plots of exact solutions for a few hydrogen-like wave functions with principal quantum number $n = 4$. (a) wave functions of $Z = 1$; (b) wave functions of $Z = 2$; (c) wave functions of $Z = 3$. The plots demonstrate the smaller r goes, the stronger is the oscillation of the function $u(r)$. And the wave functions are attracted more closely to the nucleus with increasing nuclear charge Z.

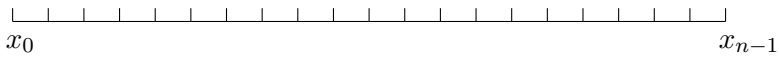
Our choice is to use a logarithmic grid[3] [2], $0 < r_0 < r_1 < \cdots < r_{n-1} < \infty$, where

$$r_i = \frac{1}{Z} e^{x_i} \quad \text{or equivalently} \quad \rho_i = e^{x_i} \tag{2.5}$$



and x is a uniformly distributed grid

$$x_i = x_0 + i\Delta x \tag{2.6}$$



The advantage of this logarithmic grid will be clear in a short while. First, we introduce a rescaled quantity[4] $\tilde{u} \equiv u/\sqrt{r}$. Recall our analytical solutions in Table 2.1, now instead of $u_{nl}(r)$, we give the relations between $\tilde{u}_{nl}$ and x in Table 2.2.

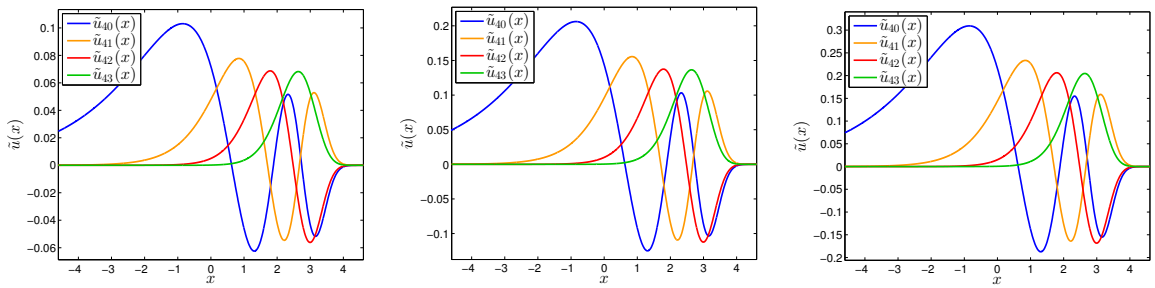
If we compare Table 2.1 and Table 2.2 carefully, we can find that in the previous formulas, the radial coordinates r are always scaled by a factor of Z, so are the plots in Fig. 2.1. But with the transformed coordinate x, we no longer have this horizontal scaling and this helps us better work with heavy nuclei. Meanwhile, the transformed coordinate x "magnifies" the inner region while shrinks the outer. As shown in Fig. 2.2, we plot the rescaled wave functions with (again) principal quantum number $n = 4$ for the three hydrogen-like atoms. In contrast to Fig. 2.1, the three plots are identical except for a vertical scaling factor Z. The inner regions are magnified clearly and the less important outer regions are reduced.

[3]The logarithmic grid is not the unique choice. One can use any adaptive grid as long as it can properly represent the wave function. But the logarithmic grid uses a change of variable technique and it is very elegant.

[4]Don't be confused with the rescalings. As a remark, the relation among R, u and $\tilde{u}$ is the following: $R(r) \equiv u(r)/r \equiv \tilde{u}(r)/\sqrt{r}$.

Table 2.2.: The first few rescaled radial wave functions $\tilde{u}_{nl}(x)$ for hydrogen-like atoms with atomic number Z. The coordinates x are the logarithmic transformations from the radial coordinate r.

$\tilde{u}_{10} =$	2	$Ze^{x/2}$		$\exp\left(-e^x\right)$
$\tilde{u}_{20} =$	$\frac{1}{\sqrt{2}}$	$Ze^{x/2}$	$\left(1-\frac{1}{2}e^x\right)$	$\exp\left(-e^x/2\right)$
$\tilde{u}_{21} =$	$\frac{1}{\sqrt{24}}$	$Ze^{3x/2}$		$\exp\left(-e^x/2\right)$
$\tilde{u}_{30} =$	$\frac{2}{\sqrt{27}}$	$Ze^{x/2}$	$\left(1-\frac{2}{3}e^x+\frac{2}{27}e^{2x}\right)$	$\exp\left(-e^x/3\right)$
$\tilde{u}_{31} =$	$\frac{8}{27\sqrt{6}}$	$Ze^{3x/2}$	$\left(1-\frac{1}{6}e^x\right)$	$\exp\left(-e^x/3\right)$
$\tilde{u}_{32} =$	$\frac{4}{81\sqrt{30}}$	$Ze^{5x/2}$		$\exp\left(-e^x/3\right)$
$\tilde{u}_{40} =$	$\frac{1}{4}$	$Ze^{x/2}$	$\left(1-\frac{3}{4}e^x+\frac{1}{8}e^{2x}-\frac{1}{192}e^{3x}\right)$	$\exp\left(-e^x/4\right)$
$\tilde{u}_{41} =$	$\frac{\sqrt{5}}{16\sqrt{3}}$	$Ze^{3x/2}$	$\left(1-\frac{1}{4}e^x+\frac{1}{80}e^{2x}\right)$	$\exp\left(-e^x/4\right)$
$\tilde{u}_{42} =$	$\frac{1}{64\sqrt{5}}$	$Ze^{5x/2}$	$\left(1-\frac{1}{12}e^x\right)$	$\exp\left(-e^x/4\right)$
$\tilde{u}_{43} =$	$\frac{1}{768\sqrt{35}}$	$Ze^{7x/2}$		$\exp\left(-e^x/4\right)$



(a) $Z = 1$ (b) $Z = 2$ (c) $Z = 3$

Figure 2.2.: The rescaled wave functions with principal quantum number $n = 4$ for three hydrogen-like atoms on the transformed grid x. (a) wave functions of $Z = 1$; (b) wave functions of $Z = 2$; (c) wave functions of $Z = 3$. The three plots are identical except for a vertical scaling factor Z. The inner regions are magnified clearly and the less important outer regions are reduced.

The original problem u on the radial coordinate r can be easily transformed to the rescaled problem $\tilde{u}$ on the transformed coordinate x by the change of variable technique. The idea is to

replace the second derivative d^2u/dr^2 in Eqn. (2.4) in terms of $d^2\tilde{u}/dx^2$:

$$\frac{d^2u}{dr^2} = -\frac{1}{4}r^{-3/2}\tilde{u} + r^{-3/2}\frac{d^2\tilde{u}}{dx^2} \tag{2.7}$$

Substituting Eqn. (2.7) into Eqn. (2.4), we obtain

$$-\frac{1}{2}\frac{d^2\tilde{u}}{dx^2} + \left[r^2V(r) + \frac{1}{2}\Big(l+\frac{1}{2}\Big)^2\right]\tilde{u} = r^2E\tilde{u} \tag{2.8}$$

This is our transformed problem. Now, instead of solving Eqn. (2.4) directly, we solve this transformed equation in (2.8). This logarithmic transformation is a key to obtaining accurate numerical solutions.

2.3. Numerov's method

The next immediate question is: how to discretize the second derivative in Eqn. (2.8) so that we can implement numerical integrations in our computer program. Perhaps the simplest way is to write this second derivative in the finite-difference form:

$$\frac{d^2\tilde{u}}{dx^2} \approx \frac{\tilde{u}_{i+1} - 2\tilde{u}_i + \tilde{u}_{i-1}}{\Delta x^2} \tag{2.9}$$

This discretization is perfectly valid and we can integrate Eqn. (2.8) numerically with 2nd order accuracy. Nevertheless, there exists a very smart trick which allows us to obtain a 4th order accurate solution with almost the same amount of computational effort. This trick is called Numerov's method.

If we take a close look at Eqn. (2.9), we get higher order terms:

$$\frac{d^2\tilde{u}}{dx^2} = \frac{\tilde{u}_{i+1} - 2\tilde{u}_i + \tilde{u}_{i-1}}{\Delta x^2} - \frac{1}{12}\tilde{u}_i^{(4)}\Delta x^2 + \mathcal{O}(\Delta x^4) \tag{2.10}$$

But the fourth derivative $\tilde{u}_i^{(4)}$ can be also written in the finite-difference form:

$$\frac{d^2\tilde{u}}{dx^2} = \frac{\tilde{u}_{i+1} - 2\tilde{u}_i + \tilde{u}_{i-1}}{\Delta x^2} - \frac{1}{12}\frac{\tilde{u}_{i+1}^{''} - 2\tilde{u}_i^{''} + \tilde{u}_{i-1}^{''}}{\Delta x^2}\Delta x^2 + \mathcal{O}(\Delta x^4) \tag{2.11}$$

Here comes the smart trick, instead of treating the second derivative $\tilde{u}_i^{''}$ in Eqn. (2.11) numerically, we can simply replace the $\tilde{u}_i^{''}$ by the relation from the original ODE in (2.8):

$$\tilde{u}_i^{''} = -2k_i^2\tilde{u}_i \tag{2.12}$$

where,

$$k_i^2 \equiv r_i^2E - r_i^2V(r_i) - \frac{1}{2}\Big(l+\frac{1}{2}\Big)^2 \tag{2.13}$$

What remains are simply substitutions. We substitute Eqn. (2.12) into Eqn. (2.11) and then substitute Eqn. (2.11) into Eqn. (2.8), we obtain the following relation:

$$\boxed{\tilde{u}_{i\pm1} = \frac{(2 - \frac{5\Delta x^2}{3}k_i^2)\tilde{u}_i - (1 + \frac{\Delta x^2}{6}k_{i\mp1}^2)\tilde{u}_{i\mp1}}{1 + \frac{\Delta x^2}{6}k_{i\pm1}^2}} \tag{2.14}$$

Eqn. (2.14) is a simple 3-point recursion: with the knowledge of $\tilde{u}_{i-1}$ and $\tilde{u}_i$, we can compute $\tilde{u}_{i+1}$ easily (or from the other direction: to compute $\tilde{u}_{i-1}$ from $\tilde{u}_i$ and $\tilde{u}_{i+1}$). We are now facing two questions:

1. How do we choose the "starting points", say, $\tilde{u}_0$ and $\tilde{u}_1$?
2. The parameter k_i^2 has a dependence on the energy E, which is so far unknown to us. How do we determine this energy?

Keep those two questions in mind and we will discuss them in detail in the following section.

2.4. The two-sided shooting and matching

First of all, let's define our grid on which we solve our numerical problem:

$$\left\{ r_{\min} = \frac{0.0001}{Z}; \quad r_{\max} = 50.0; \quad \Delta x = 0.002; \right\} \tag{2.15}$$

The minimum of the radial grid $r_{\min}$ depends on the atomic number Z, because the higher the nuclear charge, the closer the electron wave function will be attracted to the origin. We cannot take $r_{\min} = 0$ as $x_{\min} = \ln(Zr_{\min})$ will be $-\infty$. The maximum of the radial grid is a constant $r_{\max} = 50.0$. This consideration is from the fact that the size of each atom will be roughly the same after self-consistent calculations.[5] Please notice that we define Δx instead of Δr, because the grid x_i is uniformly spaced whereas the spacing Δr is not a constant.[6]

Let's start with the first question we had in our last section: the initialization of the wave functions. We initialize our wave functions according to their asymptotes (referring to the boundary conditions in Eqn. (2.4)):

$$\tilde{u}(r) \propto r^{l+1}/\sqrt{r} \quad \text{as} \quad r \to 0 \tag{2.16}$$

$$\tilde{u}(r) \propto e^{-\beta r}/\sqrt{r} \quad \text{as} \quad r \to \infty \tag{2.17}$$

where $\beta = \sqrt{-2E}$. We perform the numerical integrations from both the forward and the backward directions.[7] The corresponding initializations are summarized in Table 2.3. One might worry about the sign of the energy E as a positive energy will make the square root imaginary. But we will not work with positive energies as we are interested only in the bounded states whose energies are always below zero. A positive energy will result in a scattering state which is not of our concern here.

Be careful there is a pitfall in the backward initialization. Let's say we take $r_{n-1} = 50$ and $E = -300$. The term $e^{-\beta r_{n-1}}$ will have a value approximately 1.26×10^{-532}, which is such a small number that cannot be represented by a double precision floating point number. As a result, this initialization will return us 0. But if $\tilde{u}_{n-1}$ and $\tilde{u}_{n-2}$ are zeros, the resulting $\tilde{u}_{n-3}$, $\tilde{u}_{n-4}$, ... from the integration in Eqn. (2.14) will all become zeros. And of course this is not

[5] The self-consistent calculation will be discussed in a later chapter. By saying the size of an atom, we mean the distribution range that is covered by the significant part of the electron wave functions.

[6] As a reference, number of grid points #: $Z = 1 \to \# = 6563$; $Z = 10 \to \# = 7713$; $Z = 100 \to \# = 8865$.

[7] One could also implement a one-sided only integration. However, because of the numerical instability, the solution will diverge out quickly. A two-sided integration is perhaps the best approach to minimize the effect of numerical instability.

Table 2.3.: Wave function initializations for the forward and the backward directions.

	Initialization
Forward	$\tilde{u}_0 = r_0^{l+1}/\sqrt{r_0}$ $\tilde{u}_1 = r_1^{l+1}/\sqrt{r_1}$
Backward	$\tilde{u}_{n-1} = e^{-\beta r_{n-1}}/\sqrt{r_{n-1}}$ $\tilde{u}_{n-2} = e^{-\beta r_{n-2}}/\sqrt{r_{n-2}}$

what we wanted. To overcome this problem, we should keep doing the initialization for $\tilde{u}_{n-3}$, $\tilde{u}_{n-4}$, ... until we meet the first $\tilde{u}_i$ which is nonzero. And this $\tilde{u}_i$ and the next $\tilde{u}_{i-1}$ will be our starting points for the backward integrations. Actually the same argument applies to the forward initialization. If r_0 goes extremely small, $\tilde{u}_0$ will also have an arithmetic underflow problem. However, unlike the exponential term in the backward initialization, this value will not drop to zero that quickly. In our practical problems, we do not need to worry about the underflow problems in the forward direction and can simply assign the initial values to $\tilde{u}_0$ and $\tilde{u}_1$.

There is one important hidden message in Eqn. (2.14): If any two adjacent points (not necessarily the end ones) are given, the entire wave function can be reconstructed. This is the underlying principle of how one should match the wave functions from the forward and backward integrations. Ideally, if E is an eigen-energy, the resulting wave functions from the forward and backward integrations will be identical (up to a normalization[8] factor). On the other hand, if E is not an eigen-energy, the wave functions will not match. To check whether two wave functions matched or not, we do not compare the entire wave functions (they will never match due to numerical instability). What we should match instead are two adjacent points on the two wave functions. As we mentioned before, if the solutions agree in two points, in principle they will agree everywhere. But which two points do we compare? Apparently, those two points should not be too close to the boundaries. Because of numerical instability, the wave functions will diverge out quickly if integrating into classically forbidden regions. A good choice of these two points could be around the classical turning point. A classical turning point is the position where the energy of the electron is equal to the potential[9] $E = V(r)$. Beyond the classical turning point, the electron reaches the classically forbidden region. And inside this region, the electron wave function will decay exponentially and no further node[10] will be created. Our matching points r_{M1} and r_{M2} (they are adjacent) are taken such that $E \geq V(r_{\mathrm{M1}})$ and $E < V(r_{\mathrm{M2}})$ (see Fig. 2.3). A so called matched wave function satisfies the following criterion,

$$\left|\tilde{u}_{\mathrm{F1}}\tilde{u}_{\mathrm{B2}} - \tilde{u}_{\mathrm{F2}}\tilde{u}_{\mathrm{B1}}\right| < \varepsilon \tag{2.18}$$

where ε is a small tolerance, say, $\varepsilon = 10^{-8}$. The subscripts "F" and "B" denote whether the wave function comes form the forward integration or from the other direction. One might ask why don't we simply compare the two points such that $|\tilde{u}_{\mathrm{F1}} - \tilde{u}_{\mathrm{B1}}| < \varepsilon$ and $|\tilde{u}_{\mathrm{F2}} - \tilde{u}_{\mathrm{B2}}| < \varepsilon$. But we

[8] Normalization of wave functions will be discussed in the later section.

[9] The true classical turning point should be at $E = V_{\mathrm{eff}}(r)$ where $V_{\mathrm{eff}}(r) = V(r) + \frac{l(l+1)}{2r^2}$. But it is better to only use $V(r)$ here since $V_{\mathrm{eff}}(r)$ will introduce an additional root near the origin.

[10] A node is a point along the wave function where the wave function goes through zero. See Fig. 2.1 and Fig. 2.2 for example.

should understand that the two wave functions from the forward and backward directions need to match only up to a normalization factor. The initialization in Table 2.3 does not guarantee that the wave functions from the two directions will have the same scaling factor.

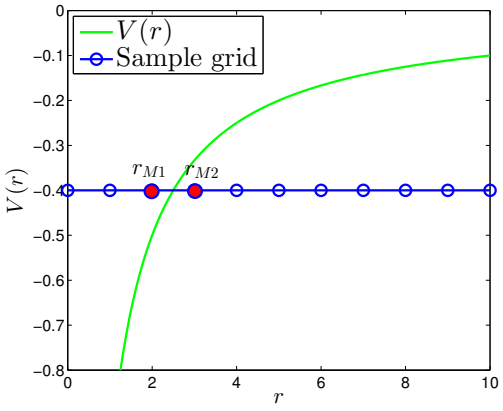
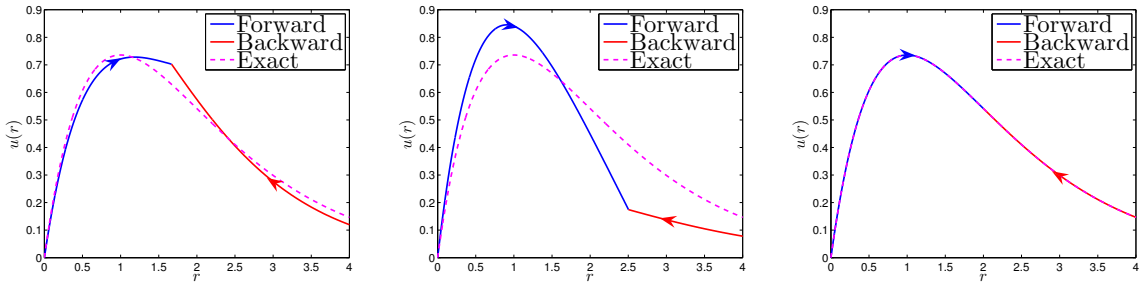


Figure 2.3.: Two matching points when $E = -0.4$ with potential $V(r) = -1/r$ on a sample grid.

As an important note, instead of comparing the two points, many people would like to compare the first derivatives of the two wave functions at a matching position. That is also very intuitive as a matched wave function should be "smooth". However, comparing the first derivatives will introduce additional numerical error unless the derivative is calculated with a formula consistent with the Numerov's integration. The comparison of two adjacent points does that automatically, which will give us a 4th order accuracy that is consistent with the Numerov's method.

So, everything is ready. What remains is to determine the electron energy E. Believe it or not, the determination of E is a trial-and-error approach. As demonstrated in Fig. 2.4, we first make a guess to the energy $E = -0.6$. Then we perform the integration (the so called shooting) and find out the wave functions at the matching point do not match and the guessed energy was too low. Next we increase our energy $E = -0.4$ and realize the energy is too high this time. Finally, by trial-and-error, we lock on the eigen-energy which in this case should be $E = -0.5$. This trial-and-error strategy is actually the spirit of the shooting and matching method.



(a) $E = -0.6$ (b) $E = -0.4$ (c) $E = -0.5$

Figure 2.4.: Forward integration and backward integration matching for the wave function u_{10} of a hydrogen atom ($Z = 1$). It is really a trial-and-error strategy to determine the eigen-energy.

2.5. The bisection method

Suppose we have our program codes ready, our next job is to modify the parameter E and run the program over and over to find out the possible eigen-energies and the corresponding eigen-functions. This is a rather tedious job. But an experienced programmer (like us) will soon notice that we could let the machine do this trial-and-error process automatically.

This automation procedural is called the bisection method. The idea is to start from two initial guesses $E_{\rm low}$ and $E_{\rm up}$, and locate the solution[11] in between. We provide the bisection algorithm for identifying the eigen-states in Algorithm 2.1.

Algorithm 2.1 Bisection method

```
1: function BISECTION(E_low, E_up)
2:     eigen_low ← SHOOT(E_low)
3:     eigen_up ← SHOOT(E_up)
4:     match_low ← eigen_low.match
5:     match_up ← eigen_up.match
6:     node_low ← eigen_low.node
7:     node_up ← eigen_up.node
8:     if node_up − node_low = 1 and match_up ∗ match_low < 0 then
9:         loop
10:            E ← (E_low + E_up)/2
11:            eigen ← SHOOT(E)
12:            match ← eigen.match
13:            if |match| < ε then                          ▷ Eigen-state found
14:                return eigen
15:            else
16:                if match ∗ match_low > 0 then
17:                    E_low ← E
18:                    match_low ← match
19:                else if match ∗ match_up > 0 then
20:                    E_up ← E
21:                    match_up ← match
22:                else
23:                    return false
24:    else
25:        return false
```

The mysterious function `Shoot()` in Algorithm 2.1 is basically the initialization and Numerov integration that we discussed in previous sections. The function `Shoot()` takes an argument E and returns an object which contains all the essential information of the resulting wave function, including the energy, the normalized wave function, number of nodes, the matching quality (as defined in Eqn. (2.18)), etc. The difference between the number of nodes from $E_{\rm low}$ and $E_{\rm up}$ determines the number of eigen-states between them. As a remark, there is a close relation

[11] It is very likely that more than one eigen-states are in between $E_{\rm low}$ and $E_{\rm up}$. But the algorithm for finding all of them will be more complicated and one has to be more careful for that situation. For simplicity, here we only provide an algorithm for locating one eigen-state.

among the number of nodes, the principal quantum number and the orbital angular momentum quantum number: $n = node + l + 1$.

2.6. Normalizing the wave function

In the previous sections, we mentioned the wave function normalization. By saying normalizing a wave function, it means to find a factor A such that

$$A \int_0^\infty dr\, |u|^2 = 1 \tag{2.19}$$

The re-assignment $u \leftarrow \sqrt{A}u$ normalizes the wave function.

It is now a matter of how to evaluate this integral numerically. First of all, it would be convenient if we transform our problem onto the uniform grid x. By change of variable $dr = (dr/dx)dx$, Eqn. (2.19) becomes

$$A \int_0^\infty dx\, r\, |u|^2 = 1 \tag{2.20}$$

There are three standard numerical methods to compute the definite integrals. The first one, which is the simplest, will be exact if the integrand $f(x)$ is a piecewise constant function. It is called the rectangle rule.

$$\int_{x_0}^{x_{n-1}} dx\, f(x) \approx \Delta x \sum_{i=0}^{n-2} f(x_i) \tag{2.21}$$

The second one, the trapezoidal rule, will be exact if the integrand is piecewise linear.

$$\int_{x_0}^{x_{n-1}} dx\, f(x) \approx \frac{\Delta x}{2} \sum_{i=0}^{n-2} [f(x_i) + f(x_{i+1})] \tag{2.22}$$

The next order is quadratic. The Simpson's rule does an interpolation on each 3-point stencil. It is exact when the integrand is piecewise quadratic. In order to concatenate every 3-point stencil together, it requires the total number of the grid points to be odd. The Simpson's rule is often more accurate than the rectangle and trapezoidal rules. Here I use (+2) in the summation to denote the index i jumps in steps of 2.

$$\int_{x_0}^{x_{n-1}} dx\, f(x) \approx \frac{\Delta x}{3} \left[f(x_0) + 4 \sum_{i=1(+2)}^{n-2} f(x_i) + 2 \sum_{i=2(+2)}^{n-3} f(x_i) + f(x_{n-1}) \right] \tag{2.23}$$

To compare those three methods, we calculate the numerical results for the following sample integration (whose analytical solution can be computed easily):

$$\int_{x_0}^{x_{n-1}} dx\, r\, \sin(x) \tag{2.24}$$

on a few grids with different spacing:

$$\left\{ r_{\min} = 1.0; \quad r_{\max} = 50.0; \quad \Delta x = 10^0, 10^{-1}, \ldots, 10^{-6}; \right\} \tag{2.25}$$

We plot the relative error against the grid size Δx on a log-log scale for the three methods, rectangle, trapezoidal and Simpson. The results are shown in Fig. 2.5. In the log-log scale plot, the slope of each curve indicates the order of accuracy of each method. Apparently, the Simpson's rule gives the highest order among the three methods. There is a "turning point" on the Simpson's curve around $\Delta x = 10^{-3}$, since the solution hits the machine precision. In our shooting method problems, we have a grid with $\Delta x = 0.002$ and the Simpson's rule is our choice.

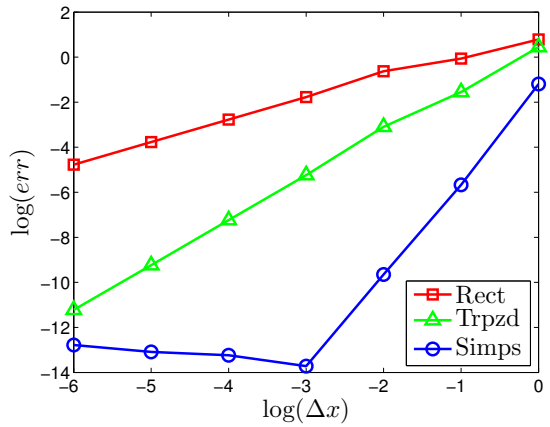


Figure 2.5.: Numerical integrations for $f(x) = r\sin(x)$ using rectangle, trapezoidal and Simpson's rules. This figure plots the relative error against the grid size Δx on a log-log scale. The slope of each curve indicates the order of the accuracy.

2.7. Numerical and exact eigen-energy comparison

To confirm the correctness of our results and to check the accuracy, we can compare our numerical solutions with the analytical eigen-energies of hydrogen-like atoms. The eigen-energies of a hydrogen-like atom is given by (in Hartree)

$$E_n = -\frac{Z^2}{2n^2} \tag{2.26}$$

We selected a few one-electron ions for comparison, namely H ($Z = 1$), C^{5+} ($Z = 6$), Fe^{25+} ($Z = 26$), Ag^{46+} ($Z = 47$) and U^{91+} ($Z = 92$). Please notice that they are not neutral atoms, but with only one electron inside (this is what we are able to deal with up to this stage). Table 2.4 lists the numerical and exact eigen-energies for the selected elements. Energies are given in units of Hartree (a.u.). The absolute error $e_{\text{abs}} = |E_{\text{num}} - E_{\text{exa}}|$ and the relative error $e_{\text{rel}} = |(E_{\text{num}} - E_{\text{exa}})/E_{\text{exa}}|$ are listed accordingly. The symbols in column 2 are the orbital names. For instance, $2p$ is an orbital with principal quantum number $n = 2$ and orbital angular momentum quantum number $l = 1$. A direct mapping between l and its "name" is given below:

l	0	1	2	3	4	5
orbital name	s	p	d	f	g	h

With this nicely tabulated result, we conclude our one-electron system computation in this section. From the next chapter, we will start to work on with more complicated, more realistic and of course more exciting cases, the many-electron systems.

Table 2.4.: The numerical and exact eigen-energies for the selected one-electron atoms. Energies are given in units of Hartree (a.u.).

Elem	Orbital	Numerical	Exact	Abs Error	Rel Error
H	$1s$	-0.500000	-0.500000	0.000000	0.000000
C	$1s$	-18.000002	-18.000000	0.000002	0.000000
	$2s$	-4.499999	-4.500000	0.000001	0.000000
	$2p$	-4.500001	-4.500000	0.000001	0.000000
Fe	$1s$	-338.000032	-338.000000	0.000032	0.000000
	$2s$	-84.499984	-84.500000	0.000016	0.000000
	$2p$	-84.500012	-84.500000	0.000012	0.000000
	$3s$	-37.555556	-37.555556	0.000000	0.000000
	$3p$	-37.555556	-37.555556	0.000000	0.000000
	$3d$	-37.555555	-37.555556	0.000001	0.000000
	$4s$	-21.125000	-21.125000	0.000000	0.000000
Ag	$1s$	-1104.500105	-1104.500000	0.000105	0.000000
	$2s$	-276.124947	-276.125000	0.000053	0.000000
	$2p$	-276.125039	-276.125000	0.000039	0.000000
	$3s$	-122.722225	-122.722222	0.000003	0.000000
	$3p$	-122.722225	-122.722222	0.000003	0.000000
	$3d$	-122.722219	-122.722222	0.000003	0.000000
	$4s$	-69.031250	-69.031250	0.000000	0.000000
	$4p$	-69.031250	-69.031250	0.000000	0.000000
	$4d$	-69.031250	-69.031250	0.000000	0.000000
	$5s$	-44.180001	-44.180000	0.000001	0.000000
U	$1s$	-4232.000404	-4232.000000	0.000404	0.000000
	$2s$	-1057.999798	-1058.000000	0.000202	0.000000
	$2p$	-1058.000151	-1058.000000	0.000151	0.000000
	$3s$	-470.222232	-470.222222	0.000010	0.000000
	$3p$	-470.222232	-470.222222	0.000010	0.000000
	$3d$	-470.222210	-470.222222	0.000012	0.000000
	$4s$	-264.500000	-264.500000	0.000000	0.000000
	$4p$	-264.500000	-264.500000	0.000000	0.000000
	$4d$	-264.500000	-264.500000	0.000000	0.000000
	$4f$	-264.500000	-264.500000	0.000000	0.000000
	$5s$	-169.280006	-169.280000	0.000006	0.000000
	$5p$	-169.280006	-169.280000	0.000006	0.000000
	$5d$	-169.280006	-169.280000	0.000006	0.000000
	$5f$	-169.280000	-169.280000	0.000000	0.000000
	$6s$	-117.555553	-117.555556	0.000003	0.000000
	$6p$	-117.555553	-117.555556	0.000003	0.000000
	$7s$	-86.367345	-86.367347	0.000002	0.000000

Chapter 3

Self-consistent field approximation

3.1. The many-electron problem

Suppose we have an Fe atom. It consists of a positively charged heavy nucleus and 26 negatively charged electrons. Unlike the one-electron systems we had before, each of those 26 electrons will not only feel the attraction from the nucleus, but will also experience the repulsions from all the other electrons. The interactions among electrons are rather complicated. But the real challenge comes from how one should represent the many-body wave function numerically. For 26 electrons, we have the wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{26})$, or in Cartesian coordinates $\Psi(x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_{26}, y_{26}, z_{26})$. This wave function has a dimension $3 \times 26 = 78$. To get an impression how crazy this dimension is, let's take a "toy" grid with 10 grid points per dimension. Then the total number of points on our toy grid will be 10^{78}. What does it imply? If the wave function on each point is a double precision floating point number, it would require us 64×10^{78} bits to store the wave function on just this toy grid. Not impressive enough? Let's say if one bit could be stored on just one atom, 64×10^{78} atoms is already beyond the amount of substances in the observable universe! So are you still thinking about storing it on your small laptop?

Nobody was able to store such a wave function on a hard disk. But that is not the end of our story. An important step to get rid of this huge dimension is to make an ansatz: the many-body solution can be written as products of one-particle wave functions

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) \approx \varphi_1(\mathbf{r}_1)\varphi_2(\mathbf{r}_2)\ldots\varphi_N(\mathbf{r}_N) \tag{3.1}$$

It is important to notice that Eqn. (3.1) is really an approximation. It assumes that electrons are distinguishable which is, however, not true. The famous Pauli exclusion principle is not included here as the wave function given by Eqn. (3.1) is not anti-symmetric.[1] Therefore, we must take extra care for the electron configurations, i.e. how one puts electrons onto different orbitals. But with Eqn. (3.1), a "$3N$ dimensional" wave function is decomposed into N "3 dimensional" wave functions. This is really a great milestone. We are not anymore embarrassed by the non-handleable gigantic wave functions. Instead, we have N lovely 3-dimensional one-electron wave

[1] An alternative ansatz that includes the anti-symmetric property of electron wave functions is the Slater determinant formulation, which leads to the Hartree-Fock method of solving many-electron problems. Our methods here are based on the density functional theory which starts from the Kohn-Sham equation and Eqn. (3.1) will be the right ansatz for us to use.

functions. But that is not good enough. Because those N electrons are coupled to each other. In order to handle those N wave functions separately, one must think about a way to decouple them. Recall our Schrödinger equation, for general N, Eqn. (1.1) (in a.u.) reads

$$\left\{\sum_{i=1}^{N}\left[-\frac{1}{2}\nabla_i^2-\frac{Z}{r_i}\right]+\sum_{i<j}^{N}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}\right\}\Psi=E\Psi \tag{3.2}$$

By analyzing this equation, we easily find out that the second term in the curly brace is causing the trouble that all electrons are coupled together. If one could decouple this electron-electron interaction, the problem will become as easy as for the one-electron case. That is where the self-consistent field approximation plays an important role. It suggests us to approximate the electron-electron repulsion term by a spherically symmetric mean-field potential

$$\sum_{i<j}^{N}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}\approx V_{\text{Hartree}}(r) \tag{3.3}$$

This $V_{\text{Hartree}}(r)$ is called the Hartree potential. Under this mean-field approximation, the N electrons are not coupled anymore. The equation

$$\left\{\sum_{i=1}^{N}\left[-\frac{1}{2}\nabla_i^2-\frac{Z}{r_i}\right]+V_{\text{Hartree}}(r)\right\}\Psi=E\Psi \tag{3.4}$$

can be decomposed into

$$\left[-\frac{1}{2}\nabla^2+V_{\text{ext}}(r)+V_{\text{Hartree}}(r)\right]\varphi_i=E_i\varphi_i\quad\text{for } i=1,2,\cdots,N \tag{3.5}$$

where $V_{\text{ext}}(r)=-Z/r$ is the external potential from the nucleus. In fact, a so called exchange-correlation potential $V_{\text{xc}}(r)$ should be added into the Hamiltonian to include the exchange-correction effect [2] from the electrons:

$$\left[-\frac{1}{2}\nabla^2+V_{\text{ext}}(r)+V_{\text{Hartree}}(r)+V_{\text{xc}}(r)\right]\varphi_i=E_i\varphi_i\quad\text{for } i=1,2,\cdots,N \tag{3.6}$$

This is the Kohn-Sham equation [3] that we are going to solve. If we compare Eqn. (3.6) with Eqn. (2.1), we immediately see that they are exactly the same except the potential here has three terms. This is how we break down a non-solvable many-electron problem into solvable one-electron problems. If you still remember how we solved the one-electron case in the previous chapter, we used a separation of variable technique to separate the 3-dimensional ODE into a radial equation and an angular equation. While the solutions for the angular equation are known already, our task is to solve the radial equation:

$$\boxed{-\frac{1}{2}\frac{d^2u_i}{dr^2}+\left[V_{\text{ext}}(r)+V_{\text{Hartree}}(r)+V_{\text{xc}}(r)+\frac{l(l+1)}{2r^2}\right]u_i=E_iu_i}\quad\text{for } i=1,2,\cdots,N \tag{3.7}$$

with the same boundary conditions: $u(r)\propto r^{l+1}$ as $r\to 0$ and $u(r)\propto\exp\left(-\sqrt{-2E}r\right)$ as $r\to\infty$.

Again, Eqn. (3.7) is identical to Eqn. (2.4) except the potential for the many-electron case is more complicated. While $V_{\text{ext}}(r)$ is given already, the two other potentials $V_{\text{Hartree}}(r)$ and $V_{\text{xc}}(r)$

remain unknown to us. But suppose we knew the exact expression for those two potentials, we would be able to solve Eqn. (3.7) as easily as we did in the previous chapter. Unfortunately, we do not have them explicitly. So one might complain: how are we supposed to solve a differential equation without knowing the differential equation? This is quite a reasonable question and the answer is: No, we cannot solve it within one step like what we did for the one-electron case. But with a self-consistent iterative scheme we would be able to achieve the solution. The idea is the following:

Since we do not know $V_{\text{Hartree}}(r)$ and $V_{\text{xc}}(r)$, we guess. We start from an initial guess of the two unknown potentials, say, $V^0_{\text{Hartree}}(r) = 0$ and $V^0_{\text{xc}}(r) = 0$.[2] From this initial guess, we obtain a solution $\{u^0_i(r)\}$.[3] Next, we use this solution to update $V^1_{\text{Hartree}}(r)$ and $V^1_{\text{xc}}(r)$ and obtain a new solution $\{u^1_i(r)\}$. This loop continues until $V^k_{\text{Hartree}}(r)$ and $V^k_{\text{xc}}(r)$ (and consequently the solution $\{u^k_i(r)\}$) converge. Fig. 3.1 illustrates the iteration scheme of the self-consistent field computation. The loop starts from an initial potential and continues updating the potential until the solution converges. That is why we call this scheme the self-consistent method: the wave functions produce the potential and the potential produces wave functions, which is self-consistent.

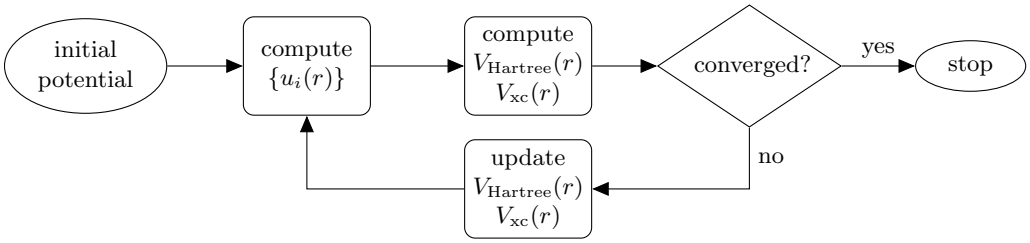


Figure 3.1.: Flow chart of self-consistent field iteration. The loop starts from an initial potential and continues updating the potential until the solution converges.

Most of the steps in Fig. 3.1 are clear to us. For example, the computation from a given potential to the wave functions has been discussed thoroughly in the previous chapter. What we need to do is simply replacing $V(r)$ in Eqn. (2.4) by $V_{\text{ext}}(r) + V_{\text{Hartree}}(r) + V_{\text{xc}}(r)$. The convergence test in the "diamond" could be checked by comparing the old total energy with the new total energy of the system.[4] What still remains unclear to us is the step from a set of given wave functions to obtaining the new potentials $V_{\text{Hartree}}(r)$ and $V_{\text{xc}}(r)$.

3.2. The Hartree potential

The computation of the Hartree potential requires not too much knowledge from quantum mechanics but almost purely electrostatics as you might have learned in your school. The keywords are: charge density, charge, electric field and electric potential. As a brief review, the relations among those quantities are summarized below.

Imagine we have a charged object with charge density $\rho(\mathbf{r})$, the total charge enclosed in volume

[2] Zero Hartree and zero exchange-correlation potentials imply the electrons are not interacting, which will result in the same solutions as for the one-electron problems.

[3] Here the notation $\{u_i(r)\}$ denotes the complete set of $u_i(r)$ for $i = 1, 2, \ldots, N$. With a superscript k, $\{u^k_i(r)\}$ indicates the solution at the k-th iteration.

[4] Computation of the total energy will be discussed in the later section.

$\mathcal{V}$ is the volume integration over the charge density

$$Q_{\text{enc}} = \int_{\mathcal{V}} d^3r\, \rho(\mathbf{r}) \tag{3.8}$$

For a spherical symmetrically distributed charge density $\rho(r)$, the total charge enclosed in a sphere with radius r can be integrated using spherical coordinate system

$$Q(r) = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta\, \sin\theta \int_0^{r} dr'\, r'^2\, \rho(r') = 4\pi \int_0^{r} dr'\, r'^2\, \rho(r') \tag{3.9}$$

The electric field created by the enclosed charge can be computed easily (recall Coulomb's law)

$$\mathbf{E}(r) = \frac{1}{4\pi\varepsilon_0}\frac{Q(r)}{r^2}\hat{\mathbf{r}} \overset{\text{in a.u.}}{=\!=} \frac{Q(r)}{r^2}\hat{\mathbf{r}} \tag{3.10}$$

Meanwhile, the electric field is the negative gradient of the electric potential $\mathbf{E}(r) = -\boldsymbol{\nabla} V(r)$ and $V(r \to \infty) \equiv 0$. Hence,

$$V(r) = -\int_{\infty}^{r} dr'\, \mathbf{E}(r') \cdot \hat{\mathbf{r}} = \int_{r}^{\infty} dr'\, E(r') \tag{3.11}$$

Now let's do a small exercise:

Question: Suppose we have a uniformly charged sphere (Fig. 3.2) of radius R with charge density $\rho(r) = \frac{3}{4\pi R^3}$. What is the electric potential it creates?

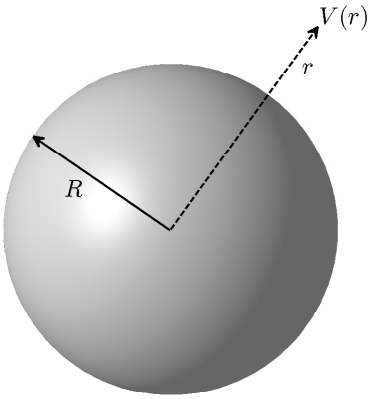


Figure 3.2.: A uniformly charged sphere of radius R with charge density $\rho(r) = \frac{3}{4\pi R^3}$.

Solution: Since there is a discontinuity on the sphere's surface, we'd better solve the problem casewise.

For $r > R$,

$$Q(r) = 4\pi \int_0^R dr'\, r'^2\, \rho(r') = 1 \tag{3.12}$$

$$E(r) = \frac{Q(r)}{r^2} = \frac{1}{r^2} \tag{3.13}$$

$$V(r) = \int_r^\infty dr'\, E(r') = \frac{1}{r} \tag{3.14}$$

For $r \leq R$,

$$Q(r) = 4\pi \int_0^r dr'\, r'^2\, \rho(r') = \frac{r^3}{R^3} \tag{3.15}$$

$$E(r) = \frac{Q(r)}{r^2} = \frac{r}{R^3} \tag{3.16}$$

$$V(r) = \int_r^\infty dr'\, E(r') = \int_r^R dr'\, \frac{r'}{R^3} + \int_R^\infty dr'\, \frac{1}{r'^2} = \frac{1}{2R} - \frac{r^2}{2R^3} + \frac{1}{R} \tag{3.17}$$

This is basically the recipe how one calculates the electric potentials. Of course, electron wave functions are not uniformly charged spheres. They usually spread out to infinity and have nodes in between. But how should a normalized wave function "look like" analogous to a uniformly charged sphere? The radial wave function should have the following piecewise constant definition (sorry for the confusion between the wave function $R(r)$ and the radius R):

$$R(r) = \begin{cases} \sqrt{\dfrac{3}{R^3}} & \text{if } r \leq R \\ 0 & \text{if } r > R \end{cases} \tag{3.18}$$

Notice that the radial wave function $R(r)$ is normalized

$$\int_0^\infty dr'\, r'^2\, |R(r')|^2 = \int_0^R dr'\, r'^2\, \frac{3}{R^3} + \int_R^\infty dr'\, r'^2\, 0 = 1 \tag{3.19}$$

Meanwhile, the angular wave function $Y(\theta, \phi)$ is always normalized

$$\int_0^{2\pi} d\phi \int_0^\pi d\theta\, \sin\theta\, |Y(\theta, \phi)|^2 = 1 \tag{3.20}$$

But if we assume the electron wave function is spherical symmetrically distributed, meaning $Y(\theta, \phi) =$ a constant $= \sqrt{\frac{1}{4\pi}}$, we get the relation between the charge density and the radial wave function:[5]

$$\rho(r) = |R(r)Y(\theta, \phi)|^2 = \frac{1}{4\pi}|R(r)|^2 \tag{3.21}$$

[5]But wait, shouldn't the charge density of an electron be negative? Yes, that is true. But as we are always working with electrons, we use a convention that charge units are negative. The same idea applies to the electric potential in the Schrödinger equation, which is the electric potential energy per negative charge.

If we have N electrons, the charge density (or electron density) is given by

$$\rho(r) = \sum_{i=1}^{N} |\varphi_i(r,\theta,\phi)|^2 = \frac{1}{4\pi} \sum_{i=1}^{N} |R_i(r)|^2 \tag{3.22}$$

Our Hartree potential (finally we come back to our main issue!) is simply the electric potential generated by this electron density, for which we have already shown the routine of calculation. Suppose we have our radial wave functions on a grid, a direct relation between $V_{\text{Hartree}}(r)$ and $\{R_i(r)\}$ is the following:

$$\boxed{\begin{aligned} \rho(r) &= \frac{1}{4\pi} \sum_{i=1}^{N} |R_i(r)|^2 \\ Q(r) &= 4\pi \int_{r_{\min}}^{r} dr'\, r'^2\, \rho(r') \\ V_{\text{Hartree}}(r) &= \int_{r}^{r_{\max}} dr'\, \frac{Q(r')}{r'^2} + \frac{N}{r_{\max}} \end{aligned}} \tag{3.23}$$

The integrals in Eqn. (3.23) can be evaluated numerically from the Simpson's rule as we discussed before. But as you might have noticed already, there is a slight difference in the integrations here. While normally an integral $\int_a^b dr\, f(r)$ returns us a "number", the integral $\int_a^r dr'\, f(r')$ returns us an "array". But that is not difficult at all. What we need to do is simply to store each intermediate integrated value while doing the `for` loop. Now, if we take the radial wave function in Eqn. (3.18) to compute the Hartree potential according to Eqn. (3.23), the numerical solution should agree with the analytical solution in Eqn. (3.17) in our exercise. Fig. 3.3 shows the solution we obtained from numerical integration versus the analytical solution. And yes, the numerical and analytical solutions happily agree with each other.

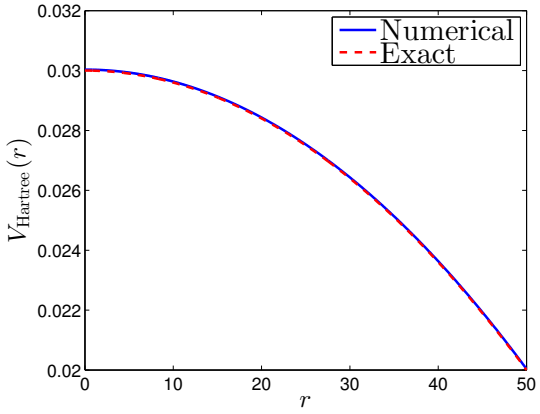


Figure 3.3.: Numerical (Simpson's rule) and analytical Hartree potentials from a normalized wave function analogous to a uniformly charged sphere of radius $R = 50$.

3.3. The exchange-correlation potential

The exchange-correlation potential is only a subtle correction[6] but definitely a key for obtaining accurate self-consistent solutions. The issue of the exchange-correlation effects comes from the framework of Kohn-Sham density functional theory (DFT) [3], which exactly maps the many-electron problem onto equivalent one-electron problems. In practical calculations, the local density approximation (LDA) is usually used to simplify the computation for the exchange-correlation potentials. LDA makes an assumption that the electron density $\rho(\mathbf{r})$ is slowly varying. As a result, the exchange-correlation energy E_{xc} can be expressed in terms of the exchange-correlation energy density ϵ_{xc},[7] as

$$E_{\mathrm{xc}}[\rho(\mathbf{r})] = \int d^3r\, \rho(\mathbf{r})\epsilon_{\mathrm{xc}}(\rho(\mathbf{r})) \tag{3.24}$$

The corresponding exchange-correlation potential is given by

$$V_{\mathrm{xc}}[\rho(\mathbf{r})] = \frac{\delta\left(\rho(\mathbf{r})\epsilon_{\mathrm{xc}}(\rho(\mathbf{r}))\right)}{\delta\rho(\mathbf{r})} \tag{3.25}$$

Don't be scared of those functional formulations. What we will do in the end is simply to input a spherically symmetric electron density $\rho(r)$ and output the exchange-correlation potential $V_{\mathrm{xc}}(r)$ (Fig. 3.4). There are a number of density functionals to approximate $V_{\mathrm{xc}}(r)$ from $\rho(r)$. A detailed discussion can be found in Reference [4]. Here we provide the Ceperley-Alder approximation with Vosko-Wilk-Nusair parameterisation (CA-VWN) which is the density functional recommended by Reference [4].

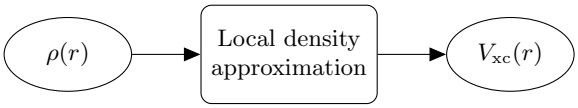


Figure 3.4.: Input a spherically symmetric electron density $\rho(r)$ and output the exchange-correlation potential $V_{\mathrm{xc}}(r)$.

First of all, we introduce two notations

$$r_s(r) = \left(\frac{3}{4\pi\rho(r)}\right)^{\frac{1}{3}} \tag{3.26}$$

$$\alpha = \left(\frac{4}{9\pi}\right)^{\frac{1}{3}} \tag{3.27}$$

As the name implies, the exchange-correlation potential consists of two terms

$$V_{\mathrm{xc}}(r) = V_{\mathrm{x}}(r) + V_{\mathrm{c}}(r) \tag{3.28}$$

The exchange term $V_{\mathrm{x}}(r)$ can be computed from the Kohn-Sham-Gaspár approximation:

$$V_x(r_s) = -\frac{1}{\pi\alpha r_s} \tag{3.29}$$

[6] The exchange-correlation potential is roughly one order smaller than the Hartree potential.

[7] The exchange-correlation energy density $\epsilon_{\mathrm{xc}}(\rho)$ is the exchange and correlation energy per electron of a uniform electron gas of density ρ. If there were no local density approximation, we would not be able to compute the exchange-correlation energy from a uniform electron gas.

The correlation term $V_{\rm c}(r)$ from CA-VWN approximation is a bit more complicated:

$$V_c(r_s) = \epsilon_c(r_s) - \frac{A}{3}\frac{1 + b_1 r_s^{1/2}}{1 + b_1 r_s^{1/2} + b_2 r_s + b_3 r_s^{3/2}} \tag{3.30}$$

where,[8]

Parameter	Paramagnetic
A	0.0310907
b	3.72744
c	12.9352
x_0	−0.10498
b_1	9.81379
b_2	2.82224
b_3	0.736412
Q	6.15199

$$\epsilon_c(r_s) = A\left[\ln\frac{r_s}{X(r_s)} + \frac{2b}{Q}\tan^{-1}\frac{Q}{2\sqrt{r_s}+b} - \frac{bx_0}{X(x_0^2)}\left(\ln\frac{(\sqrt{r_s}-x_0)^2}{X_i(r_s)} + \frac{2(b+2x_0)}{Q}\tan^{-1}\frac{Q}{2\sqrt{r_s}+b}\right)\right]$$

and $X(r_s) = r_s + b\sqrt{r_s} + c$.

3.4. Achieving self-consistency

The calculations for both the Hartree potential and the exchange-correlation potential have been discussed in the last two sections. The effective Kohn-Sham potential

$$V(r) = V_{\rm ext}(r) + V_{\rm Hartree}(r) + V_{\rm xc}(r) \tag{3.31}$$

appears much clearer to us now. Our next task is to perform the self-consistent iteration as illustrated in Fig. 3.1. To update the potential for a next iteration, naively, one would like to assign:[9]

$$V^{k+1}(r) \leftarrow V_{\rm ext}(r) + V^k_{\rm Hartree}(r) + V^k_{\rm xc}(r) \tag{3.32}$$

This is probably what most people would expect. However, this assignment normally leads to no convergence. The solutions will oscillate inside a certain region. Imagine we put electrons initially too close to each other, in the first step, due to strong repulsions, the electrons will be pushed very far away. In the next step, because the electrons are sitting too far from each other, they won't experience much repulsions and will be attracted back by the nucleus. That is the picture how the solutions oscillate. To avoid this oscillation, we use a technique called linear mixing. We assign

$$V^{k+1}(r) \leftarrow (1-\alpha)V^k(r) + \alpha\left(V_{\rm ext}(r) + V^k_{\rm Hartree}(r) + V^k_{\rm xc}(r)\right) \tag{3.33}$$

where α is a number between 0 and 1. If $\alpha = 1$, Eqn. (3.33) will be identical to Eqn. (3.32), which results in a non-converged solution. But if $\alpha = 0$, there will be no potential update at all. The assignment $V^{k+1}(r) \leftarrow V^k(r)$ makes the loop never end. So, a reasonable choice should be somewhere in between 0 and 1, and an empirical value for α could be around 0.3 to 0.5.

[8] There are parameters for both paramagnetic and ferromagnetic correlation energies. Here we use only the paramagnetic parameters since we are ignoring electron spins at this stage.

[9] Notice that the external potential $V_{\rm ext}(r) = -Z/r$ is independent of iterations.

3.5. Comparison to NIST calculations

The National Institute of Standards and Technology (NIST) provides a reference for electronic structure calculations [5], which includes exactly what we are calculating. It provides us an excellent reference for checking our solutions.

Similar to Table 2.4, we again selected the same elements for comparison, namely H (Z = 1), C (Z = 6), Fe (Z = 26), Ag (Z = 47) and U (Z = 92). In contrast to Table 2.4, this time all atoms are neutral. So we reached our goal of calculating many-electron systems. We list the orbital eigen-energies from our results versus the results from NIST in Table 3.1. Energies are given in units of Hartree (a.u.). If you watch carefully, there is a subtle difference between Table 2.4 and Table 3.1 in the second column: Previously, we didn't have superscripts over the orbital names. They are the occupation numbers, which means the number of electrons in each orbital. For example $2p^2$ means there are 2 electrons sitting in orbital $2p$. And a complete orbital information, like for C: $1s^2\ 2s^2\ 2p^2$, is called the electronic configuration, which describes the occupations of all orbitals. The electronic configuration needs to be specified by ourselves and it has to obey the Pauli exclusion principal. For example, there cannot be more than 2 electrons in an s orbital and no more than 6 electrons in a p orbital.

Table 3.1.: Comparison of eigen-energies of the occupied orbitals between my results and the results from NIST for the selected atoms. Energies are given in units of Hartree (a.u.).

Elem	Orbital	My results	NIST results	Abs Error	Rel Error
H	$1s^1$	−0.233471	−0.233471	0.000000	0.000000
C	$1s^2$	−9.947725	−9.947718	0.000007	0.000001
	$2s^2$	−0.500866	−0.500866	0.000000	0.000000
	$2p^2$	−0.199186	−0.199186	0.000000	0.000000
Fe	$1s^2$	−254.225334	−254.225505	0.000171	0.000001
	$2s^2$	−29.564863	−29.564860	0.000003	0.000000
	$2p^6$	−25.551762	−25.551766	0.000004	0.000000
	$3s^2$	−3.360622	−3.360621	0.000001	0.000000
	$3p^6$	−2.187521	−2.187523	0.000002	0.000001
	$3d^6$	−0.295047	−0.295049	0.000002	0.000007
	$4s^2$	−0.197976	−0.197978	0.000002	0.000010
Ag	$1s^2$	−900.324405	−900.324578	0.000173	0.000000
	$2s^2$	−129.859749	−129.859807	0.000058	0.000000
	$2p^6$	−120.913362	−120.913351	0.000011	0.000000
	$3s^2$	−23.678432	−23.678437	0.000005	0.000000
	$3p^6$	−20.067624	−20.067630	0.000006	0.000000
	$3d^{10}$	−13.367801	−13.367803	0.000002	0.000000
	$4s^2$	−3.223088	−3.223090	0.000002	0.000001
	$4p^6$	−2.086598	−2.086602	0.000004	0.000002
	$4d^{10}$	−0.298702	−0.298706	0.000004	0.000013
	$5s^1$	−0.157404	−0.157407	0.000003	0.000019
U	$1s^2$	−3689.356876	−3689.355141	0.001735	0.000000
	$2s^2$	−639.778647	−639.778728	0.000081	0.000000
	$2p^6$	−619.108505	−619.108550	0.000045	0.000000
	$3s^2$	−161.118060	−161.118073	0.000013	0.000000
	$3p^6$	−150.978963	−150.978980	0.000017	0.000000
	$3d^{10}$	−131.977338	−131.977358	0.000020	0.000000
	$4s^2$	−40.528086	−40.528084	0.000002	0.000000
	$4p^6$	−35.853321	−35.853321	0.000000	0.000000
	$4d^{10}$	−27.123209	−27.123212	0.000003	0.000000
	$4f^{14}$	−15.027458	−15.027460	0.000002	0.000000
	$5s^2$	−8.824083	−8.824089	0.000006	0.000001
	$5p^6$	−7.018084	−7.018092	0.000008	0.000001
	$5d^{10}$	−3.866167	−3.866175	0.000008	0.000002
	$5f^3$	−0.366535	−0.366543	0.000008	0.000022
	$6s^2$	−1.325969	−1.325976	0.000007	0.000005
	$6p^6$	−0.822530	−0.822538	0.000008	0.000010
	$6d^1$	−0.143184	−0.143190	0.000006	0.000042
	$7s^2$	−0.130943	−0.130948	0.000005	0.000038

3.6. Total energy of the system

While the Kohn-Sham orbital eigen-energies give us a nice description to the atomic systems, their physical meanings are less obvious (they are only the solutions from our ansatz (3.1)). Another very important quantity which has a key physical meaning is the total energy of the system. Naively, one might expect that the total energy of the system E_{tot} is simply the summation from all eigen-energies of the Kohn-Sham orbitals.

$$E_{\text{tot}} \stackrel{?}{=} \sum_{i=1}^{N} \varepsilon_i \tag{3.34}$$

That was a close guess, however not quite true. The discrepancy comes from the following inequalities:

$$E_{\text{Hartree}} \neq \sum_{i=1}^{N} \langle \varphi_i | \, V_{\text{Hartree}} \, | \varphi_i \rangle \tag{3.35}$$

$$E_{\text{xc}} \neq \sum_{i=1}^{N} \langle \varphi_i | \, V_{\text{xc}} \, | \varphi_i \rangle \tag{3.36}$$

where E_{Hartree} is the energy from electron-electron repulsions (Hartree energy) and E_{xc} is the exchange-correction energy. They are not simply the expectation values of their corresponding potentials. To get the exact total energy of the system, one has to be more careful:

$$E_{\text{kin}} = \sum_{i=1}^{N} \langle \varphi_i | \, \frac{1}{2}\nabla^2 \, | \varphi_i \rangle \tag{3.37}$$

$$E_{\text{ext}} = \sum_{i=1}^{N} \langle \varphi_i | \, V_{\text{ext}} \, | \varphi_i \rangle \tag{3.38}$$

$$E_{\text{Hartree}} = \frac{1}{2} \int d^3r \, \rho(\mathbf{r}) V_{\text{Hartree}}(\mathbf{r}) \tag{3.39}$$

$$E_{\text{xc}} = \int d^3r \, \rho(\mathbf{r}) \epsilon_{\text{xc}}(\mathbf{r}) \tag{3.40}$$

where E_{kin} is the kinetic energy of the electrons, E_{ext} is the energy from the external potential $V_{\text{ext}}(r) = -Z/r$, and the total energy of the system is given by the summation of those 4 energies

$$E_{\text{tot}} = E_{\text{kin}} + E_{\text{ext}} + E_{\text{Hartree}} + E_{\text{xc}} \tag{3.41}$$

Each term in Eqn. (3.41) can be evaluated explicitly. But taking the derivative in the kinetic energy (Eqn. (3.37)) introduces unnecessary numerical errors. With a small trick we can get rid of the Laplacian:

$$\begin{aligned} \sum_{i=1}^{N} \varepsilon_i &= \sum_{i=1}^{N} \langle \varphi_i | \, \frac{1}{2}\nabla^2 + V_{\text{ext}} + V_{\text{Hartree}} + V_{\text{xc}} \, | \varphi_i \rangle \qquad (3.42) \\ &= \sum_{i=1}^{N} \langle \varphi_i | \, \frac{1}{2}\nabla^2 \, | \varphi_i \rangle + \sum_{i=1}^{N} \langle \varphi_i | \, V_{\text{ext}} \, | \varphi_i \rangle + \sum_{i=1}^{N} \langle \varphi_i | \, V_{\text{Hartree}} \, | \varphi_i \rangle + \sum_{i=1}^{N} \langle \varphi_i | \, V_{\text{xc}} \, | \varphi_i \rangle \end{aligned}$$

$$\sum_{i=1}^{N} \langle \varphi_i | \, \frac{1}{2}\nabla^2 \, | \varphi_i \rangle = \sum_{i=1}^{N} \varepsilon_i - \sum_{i=1}^{N} \langle \varphi_i | \, V_{\text{ext}} \, | \varphi_i \rangle - \sum_{i=1}^{N} \langle \varphi_i | \, V_{\text{Hartree}} \, | \varphi_i \rangle - \sum_{i=1}^{N} \langle \varphi_i | \, V_{\text{xc}} \, | \varphi_i \rangle \tag{3.43}$$

Hence Eqn. (3.41) becomes

$$\boxed{E_{\text{tot}} = \sum_{i=1}^{N} \varepsilon_i - \sum_{i=1}^{N} \langle \varphi_i | \, V_{\text{Hartree}} \, | \varphi_i \rangle - \sum_{i=1}^{N} \langle \varphi_i | \, V_{\text{xc}} \, | \varphi_i \rangle + E_{\text{Hartree}} + E_{\text{xc}}} \tag{3.44}$$

This exactly agrees with the discrepancy we have mentioned in Eqns. (3.35) and (3.36). We suggest the summation of Kohn-Sham orbital energies as our total energy, then we subtract the "wrong" terms and put back the correct ones. If we expand Eqn. (3.44) and read carefully, we have to pay extra attention to the iteration numbers:

$$\begin{aligned} E_{\text{tot}}^{k} = & \sum_{i=1}^{N} \varepsilon_i^k - \int d^3r \, \rho^k(\mathbf{r}) V_{\text{Hartree}}^{k-1}(\mathbf{r}) - \int d^3r \, \rho^k(\mathbf{r}) V_{\text{xc}}^{k-1}(\mathbf{r}) \\ & + \frac{1}{2} \int d^3r \, \rho^k(\mathbf{r}) V_{\text{Hartree}}^{k}(\mathbf{r}) + \int d^3r \, \rho^k(\mathbf{r}) \epsilon_{\text{xc}}^{k}(\mathbf{r}) \end{aligned} \tag{3.45}$$

And this will be our final expression to compute the total energy of the system. To check the correctness and accuracy of our solutions, we again compare our results with the results from NIST. The comparison is listed in Table 3.2.

Table 3.2.: Comparison of self-consistent total energies between my results and the results from NIST for the selected atoms. Energies are given in units of Hartree (a.u.).

Elem	My results	NIST results	Abs Error	Rel Error
H	−0.445670	−0.445671	0.000001	0.000002
C	−37.425762	−37.425749	0.000013	0.000000
Fe	−1261.092727	−1261.093056	0.000329	0.000000
Ag	−5195.030828	−5195.031215	0.000387	0.000000
U	−25658.420765	−25658.417889	0.002876	0.000000

Chapter 4

Electron-electron interaction in second quantization

4.1. The Coulomb repulsion Hamiltonian

It was a great success that we solved the many-electron problem in the self-consistent field approximation. The solutions exhibit important physical quantities, such as the electron density distribution and the total energy of the system. If someone asks us to summarize the spirit of the methods into one word, then this word must be "mean-field". It was the mean-field potential that simplified our problem dramatically. However, by introducing a mean-field, we sacrificed detailed information how exactly individual electrons interact among themselves.

Consider a carbon atom with electronic configuration: $1s^2\ 2s^2\ 2p^2$. Both the $1s^2$ and $2s^2$ shells are fully occupied, but the $2p^2$ shell is still open. According to the Pauli exclusion principle, one p shell can contain at most 6 electrons with orbital angular projection momentum $m = 1, 0, -1$ and spin angular projection momentum $\sigma = \uparrow, \downarrow$.[1] Since the $2p^2$ shell is not fully occupied, these two electrons can take different quantum numbers among those allowed m and σ. But how exactly are these two electrons arranged, our mean-field potential does not tell a difference. In this $2p^2$ shell, energies of different m and σ are all degenerate.

It is important to understand why the mean-field potential distinguishes "n and l" but not "m or σ": Because we calculated the mean-field potential according to electron radial wave functions R_{nl}, which are distinguished by quantum numbers n and l. But for the angular part, we made an approximation that the electrons are spherical symmetrically distributed, which implies all Y_{lm} are treated equally as Y_{00}. Moreover, we didn't make a distinction between spin up and spin down. It doesn't really matter which m or σ state the electron is. However, the real physics is, due to the interaction among electrons, there will be an energy-splitting among different m and σ quantum states. Since our mean-field potential cannot resolve this energy-splitting, this might be a good time for us to revisit our "trouble maker": the very complicated Coulomb repulsion

[1] A more general convention is to write the orbital and spin projection quantum numbers as m_l and m_s, respectively. When a spin 1/2 particle (in our case electron) is considered, we can also write them as m and σ to simplify our notations.

potential

$$H_U = \sum_{i<j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{4.1}$$

Previously we replaced this electron-electron repulsion by a mean-field potential. Now, we would like to handle it directly, which is, of course, not easy. This Coulomb repulsion Hamiltonian becomes more handleable if we reformulate it into second quantization (see Appendix B)

$$H_U = \frac{1}{2} \sum_{\alpha,\beta,\gamma,\delta} U_{\alpha\beta\gamma\delta} c_\alpha^\dagger c_\beta^\dagger c_\gamma c_\delta \tag{4.2}$$

where,

$$\begin{aligned} \alpha &= \{n_1,\, l_1,\, m_1,\, \sigma_1\} \\ \beta &= \{n_2,\, l_2,\, m_2,\, \sigma_2\} \\ \gamma &= \{n_3,\, l_3,\, m_3,\, \sigma_3\} \\ \delta &= \{n_4,\, l_4,\, m_4,\, \sigma_4\} \end{aligned}$$

These four indices $(\alpha, \beta, \gamma, \delta)$ represent four sets of electron quantum numbers. Perhaps the immediate question is, "Why four indices?" It seems like we need only two indices since each electron-pair interaction in Eqn. (4.1) involves two electrons. In Eqn. (4.2), we have two electron creators and two electron annihilators. If we understand the mechanism of those creation and annihilation operators, this picture will become much clearer. Let's say initially we have a two-electron state $|\gamma, \delta\rangle$. First, the operator c_δ annihilates the electron with quantum number δ, which produces $|\gamma\rangle$. Then, c_γ annihilates the second electron and returns a vacuum state $|0\rangle$. Next, the creation operator $c_\beta^\dagger$ creates an electron from the vacuum and we get $|\beta\rangle$. Finally, $c_\alpha^\dagger$ creates another electron and we obtain our final state $|\beta, \alpha\rangle$.

$$|\gamma, \delta\rangle \xrightarrow{c_\delta} |\gamma\rangle \xrightarrow{c_\gamma} |0\rangle \xrightarrow{c_\beta^\dagger} |\beta\rangle \xrightarrow{c_\alpha^\dagger} |\beta, \alpha\rangle$$

This four-step process involves two electrons. One can think $|\gamma, \delta\rangle$ as the initial state and $|\beta, \alpha\rangle$ as the final state. That is why four indices are involved.

The next question is, "What's the range of the indices?" The answer is that the indices enumerate all possible quantum states of electrons, that is, infinitely many. In fact, Eqn. (4.2) is absolutely equivalent to Eqn. (4.1). They are just two different formulations on the same physics problem. But later we need to restrict the range of the indices, say, into the same shell, making an approximation. Otherwise the dimension of the problem will be too huge to be solvable. The very important $U_{\alpha\beta\gamma\delta}$ is called the matrix element, from which we see the connection between Eqns. (4.1) and (4.2).

$$U_{\alpha\beta\gamma\delta} = \delta_{\sigma_1\sigma_4}\delta_{\sigma_2\sigma_3} \int d^3r_1 \int d^3r_2\, \overline{\varphi_{n_1 l_1 m_1}}(\mathbf{r}_1)\overline{\varphi_{n_2 l_2 m_2}}(\mathbf{r}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \varphi_{n_3 l_3 m_3}(\mathbf{r}_2)\varphi_{n_4 l_4 m_4}(\mathbf{r}_1) \tag{4.3}$$

$U_{\alpha\beta\gamma\delta}$ is really just a "number". But the multi-dimensional integration makes the computation not trivial at all. In fact, this entire chapter is dedicated to discussing the computation of this Coulomb repulsion matrix element, which plays a crucial role in our problem. The difficulty of evaluating this integral comes from the fact that $\mathbf{r}_1$ and $\mathbf{r}_2$ are coupled together. I admit that the

term $\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}$ looks lovely. But we cannot go further if this coupling term exists. Unfortunately, we have to expand this lovely term into a monster. It is called the multipole expansion [17]

$$\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}=\sum_{k=0}^{\infty}\frac{r_<^k}{r_>^{k+1}}\frac{4\pi}{2k+1}\sum_{\mu=-k}^{k}\overline{Y_{k\mu}}(\theta_1,\phi_1)Y_{k\mu}(\theta_2,\phi_2) \tag{4.4}$$

If $r_1 \le r_2$,

$$r_< = r_1, \quad r_> = r_2 \tag{4.5}$$

If $r_1 > r_2$,

$$r_< = r_2, \quad r_> = r_1 \tag{4.6}$$

The magic of Eqn. (4.4) is that it separates $\mathbf{r}_1$ and $\mathbf{r}_2$. In other words, (r_1,θ_1,ϕ_1) and (r_2,θ_2,ϕ_2) can be integrated independently. Now, we want to substitute (4.4) into (4.3). Er... well, maybe we could do some abbreviation first to reduce our headache. The multipole expansion consists of two major parts:

The radial part,

$$\begin{aligned}R^{(k)}(n_1l_1,n_2l_2,n_3l_3,n_4l_4)&=\int_0^\infty dr_1\,r_1^2\int_0^\infty dr_2\,r_2^2\,\overline{R_{n_1l_1}}(r_1)\overline{R_{n_2l_2}}(r_2)\frac{r_<^k}{r_>^{k+1}}R_{n_3l_3}(r_2)R_{n_4l_4}(r_1)\\&=\int_0^\infty dr_1\int_0^\infty dr_2\,\overline{u_{n_1l_1}}(r_1)\overline{u_{n_2l_2}}(r_2)\frac{r_<^k}{r_>^{k+1}}u_{n_3l_3}(r_2)u_{n_4l_4}(r_1)\end{aligned} \tag{4.7}$$

The angular part,

$$\begin{aligned}A^{(k)}(l_1m_1,l_2m_2,l_3m_3,l_4m_4)=\sum_{\mu=-k}^{k}&\int_0^{2\pi}d\phi_1\int_0^{\pi}d\theta_1\,\sin\theta_1\,\overline{Y_{l_1m_1}}(\theta_1,\phi_1)\overline{Y_{k\mu}}(\theta_1,\phi_1)Y_{l_4m_4}(\theta_1,\phi_1)\\&\int_0^{2\pi}d\phi_2\int_0^{\pi}d\theta_2\,\sin\theta_2\,\overline{Y_{l_2m_2}}(\theta_2,\phi_2)Y_{k\mu}(\theta_2,\phi_2)Y_{l_3m_3}(\theta_2,\phi_2)\end{aligned} \tag{4.8}$$

Consequently, Eqn. (4.3) can be written as

$$\boxed{U_{\alpha\beta\gamma\delta}=\delta_{\sigma_1\sigma_4}\delta_{\sigma_2\sigma_3}\sum_{k=0}^{\infty}\frac{4\pi}{2k+1}R^{(k)}(n_1l_1,n_2l_2,n_3l_3,n_4l_4)A^{(k)}(l_1m_1,l_2m_2,l_3m_3,l_4m_4)} \tag{4.9}$$

Hum... Much clearer. But the difficulty remains: how to evaluate the radial part $R^{(k)}(n_1l_1,n_2l_2,n_3l_3,n_4l_4)$ and the angular part $A^{(k)}(l_1m_1,l_2m_2,l_3m_3,l_4m_4)$, respectively?

4.2. Slater-Condon parameters

Our first task is to solve the radial part:

$$R^{(k)}(n_1l_1,n_2l_2,n_3l_3,n_4l_4)=\int_0^\infty dr_1\int_0^\infty dr_2\,\overline{u_{n_1l_1}}(r_1)\overline{u_{n_2l_2}}(r_2)\frac{r_<^k}{r_>^{k+1}}u_{n_3l_3}(r_2)u_{n_4l_4}(r_1) \tag{4.10}$$

But how is it possible to evaluate such an integral with those mysterious $r_<$ and $r_>$? Recall how they are defined in Eqns. (4.5) and (4.6). We note down,

If $r_1 \leq r_2$,

$$\frac{r_<^k}{r_>^{k+1}} = \frac{r_1^k}{r_2^{k+1}} \tag{4.11}$$

If $r_1 > r_2$,

$$\frac{r_<^k}{r_>^{k+1}} = \frac{r_2^k}{r_1^{k+1}} \tag{4.12}$$

Hum... it appears not that scary now. And the integration below can be evaluated casewise:

$$\begin{aligned}
&\int_0^\infty dr_2\, \overline{u_{n_1 l_1}}(r_1)\overline{u_{n_2 l_2}}(r_2)\frac{r_<^k}{r_>^{k+1}}u_{n_3 l_3}(r_2)u_{n_4 l_4}(r_1) \\
&= \int_0^{r_1} dr_2\, \overline{u_{n_1 l_1}}(r_1)\overline{u_{n_2 l_2}}(r_2)\frac{r_2^k}{r_1^{k+1}}u_{n_3 l_3}(r_2)u_{n_4 l_4}(r_1) + \int_{r_1}^\infty dr_2\, \overline{u_{n_1 l_1}}(r_1)\overline{u_{n_2 l_2}}(r_2)\frac{r_1^k}{r_2^{k+1}}u_{n_3 l_3}(r_2)u_{n_4 l_4}(r_1) \\
&= \overline{u_{n_1 l_1}}(r_1)u_{n_4 l_4}(r_1)\left[\frac{1}{r_1^{k+1}}\int_0^{r_1} dr_2\, r_2^k \overline{u_{n_2 l_2}}(r_2)u_{n_3 l_3}(r_2) + r_1^k \int_{r_1}^\infty dr_2\, \frac{1}{r_2^{k+1}}\overline{u_{n_2 l_2}}(r_2)u_{n_3 l_3}(r_2)\right]
\end{aligned} \tag{4.13}$$

Hence, Eqn. (4.10) reads

$$\boxed{\begin{aligned}
&R^{(k)}(n_1 l_1, n_2 l_2, n_3 l_3, n_4 l_4) = \\
&\int_0^\infty dr_1 \overline{u_{n_1 l_1}}(r_1)u_{n_4 l_4}(r_1)\left[\frac{1}{r_1^{k+1}}\int_0^{r_1} dr_2\, r_2^k \overline{u_{n_2 l_2}}(r_2)u_{n_3 l_3}(r_2) + r_1^k \int_{r_1}^\infty dr_2\, \frac{1}{r_2^{k+1}}\overline{u_{n_2 l_2}}(r_2)u_{n_3 l_3}(r_2)\right]
\end{aligned}} \tag{4.14}$$

Absolutely under our control, not? We are glad to see that r_1 and r_2 are separated. This is the Slater-Condon parameter that we are going to evaluate. One should be aware that for a given index $\{n, l\}$, the radial wave function u_{nl} is not uniquely determined. It depends on the choice of the system. For instance, one can use the hydrogen-like radial wave functions, whose solutions are known analytically. But as a better estimation, we will use the radial wave functions from our self-consistent calculations. In fact, this ab-initio Slater-Condon parameter enters as a connection between our previous results and the multiplet calculations that we will work on.

It is interesting to notice that all the integrations come with a factor r^k (or r^{-k-1}). It would be ideal if we can develop a numerical integration method that takes into account those factors implicitly, so that only the wave functions are required as input. Now, we would like to extend our discussion to a general type of integration, namely, if the integral has the following form

$$I = \int_0^\infty dx\, r^k f(x) \tag{4.15}$$

where k is an arbitrary integer and $f(x)$ is a smooth and slow varying function. One can think this integral as an integration over $f(x)$ with a weight r^k (where $r = e^x/Z$).

We would like to develop a numerical scheme which is exact for any $f(x) = ax^2 + bx + c$. The recipe for constructing this scheme works as the following: First, we make an ansatz: for a

3-point stencil (uniform grid) on an interval $[x_0, x_2]$ the following relation is exact

$$\int_{x_0}^{x_2} dx\, r^k f(x) = \alpha f(x_0) + \beta f(x_1) + \gamma f(x_2) \tag{4.16}$$

Our task is to determine those magic coefficients α, β and γ. (let's first assume $k \neq 0$)

In the first step, we assume the input is a constant function $f(x) = 1$:

$$\int_{x_0}^{x_2} dx\, r^k = r_2^k \left(\frac{1}{k}\right) - r_0^k \left(\frac{1}{k}\right) = A \tag{4.17}$$

Next, for the first order, we take $f(x) = x$:

$$\int_{x_0}^{x_2} dx\, r^k x = r_2^k \left(\frac{x_2}{k} - \frac{1}{k^2}\right) - r_0^k \left(\frac{x_0}{k} - \frac{1}{k^2}\right) = B \tag{4.18}$$

Then for the second order, we take $f(x) = x^2$:

$$\int_{x_0}^{x_2} dx\, r^k x^2 = r_2^k \left(\frac{x_2^2}{k} - \frac{2x_2}{k^2} + \frac{2}{k^3}\right) - r_0^k \left(\frac{x_0^2}{k} - \frac{2x_0}{k^2} + \frac{2}{k^3}\right) = C \tag{4.19}$$

Now, our coefficients α, β and γ should be chosen such that all the three conditions are fulfilled. This is a system of equations with three unknowns. The linear system is given by

$$\begin{bmatrix} 1 & 1 & 1 \\ x_0 & x_0 + \Delta x & x_0 + 2\Delta x \\ x_0^2 & (x_0 + \Delta x)^2 & (x_0 + 2\Delta x)^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} A \\ B \\ C \end{bmatrix} \tag{4.20}$$

Thanks to `Mathematica`, a very useful tool for symbolic calculations, we obtain our solutions:

$$\alpha = -r_0^k \left(\frac{1}{k^3 \Delta x^2} + \frac{3}{2k^2 \Delta x} + \frac{1}{k}\right) + r_2^k \left(\frac{1}{k^3 \Delta x^2} - \frac{1}{2k^2 \Delta x}\right) \tag{4.21}$$

$$\beta = 2r_0^k \left(\frac{1}{k^3 \Delta x^2} + \frac{1}{k^2 \Delta x}\right) + 2r_2^k \left(-\frac{1}{k^3 \Delta x^2} + \frac{1}{k^2 \Delta x}\right) \tag{4.22}$$

$$\gamma = -r_0^k \left(\frac{1}{k^3 \Delta x^2} + \frac{1}{2k^2 \Delta x}\right) + r_2^k \left(\frac{1}{k^3 \Delta x^2} - \frac{3}{2k^2 \Delta x} + \frac{1}{k}\right) \tag{4.23}$$

Those are the magic coefficients that satisfy our ansatz. In other words, the summation $[\alpha f(x_0) + \beta f(x_1) + \gamma f(x_2)]$ will be exact for integrating any function with the form $r^k \left[ax^2 + bx + c\right]$ on an interval $[x_0, x_2]$. In our calculations we assumed that $k \neq 0$. But if $k = 0$, those three coefficients are much simpler, namely,

$$\alpha = \frac{1}{3}\Delta x, \quad \beta = \frac{4}{3}\Delta x, \quad \gamma = \frac{1}{3}\Delta x \tag{4.24}$$

They are simply the coefficients from the Simpson's rule.

Now, our task is to integrate the entire domain $[x_0, x_{n-1}]$. This is simply done by summing up each small domain. Our final expression becomes a weighted sum:

$$I \approx \sum_{i=0}^{n-1} w_i f(x_i) \tag{4.25}$$

This is the beauty of the numerical integration method. Once the weights are determined (although not trivial), we can very easily compute the integral by a weighted sum. The relation between the weights and our coefficients can be easily seen from below: (requiring n odd)

$$\begin{array}{cccccccccc}
 & \alpha_0 & \beta_0 & \gamma_0 & & & & & & \\
 & & & \alpha_2 & \beta_2 & \gamma_2 & & & & \\
 & & & & & & \ddots & & & \\
+) & & & & & & & \alpha_{n-3} & \beta_{n-3} & \gamma_{n-3} \\
\hline
 & w_0 & w_1 & w_2 & w_3 & w_4 & \cdots & w_{n-3} & w_{n-2} & w_{n-1}
\end{array}$$

In conclusion, the weights are summarized below:

$$w_i = \begin{cases} \alpha_i & \text{if } i = 0 \\ \beta_{i-1} & \text{if } i = 1, 3, \ldots, n-2 \\ \gamma_{i-2} + \alpha_i & \text{if } i = 2, 4, \ldots, n-3 \\ \gamma_{i-2} & \text{if } i = n-1 \end{cases} \tag{4.26}$$

We call this scheme a weighted Simpson's rule. If we make a two-point stencil ansatz, we could easily derive a weighted trapezoidal rule. Those weighted integration methods have the advantage that the factor r^k is taken care by the weights automatically, but the accuracy of the methods is not guaranteed to be better than a traditional numerical integration method. Now, we would like to compare the calculations of the Slater-Condon parameter (Eqn. (4.28)) with $k = 6$ for a hydrogen wave function "$4f$" using trapezoidal, weighted-trapezoidal, Simpson and weighted Simpson's rules, on a few grids with different spacing:

$$\{r_{\min} = 0.1; \quad r_{\max} = 150.0; \quad \Delta x = 10^0, 10^{-1}, \ldots, 10^{-6}; \} \tag{4.27}$$

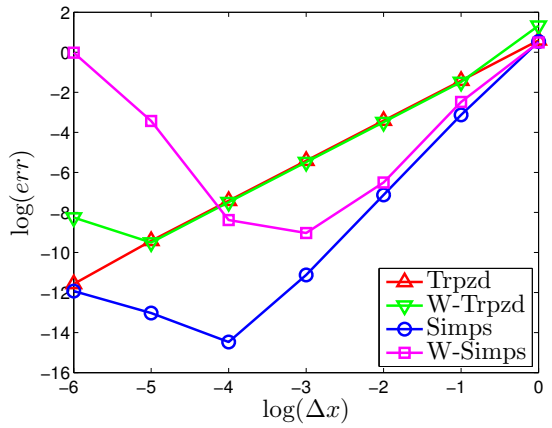


Figure 4.1.: Numerical integrations of Slater-Condon parameters with $k = 6$ for hydrogen wave function "$4f$" using trapezoidal, weighted-trapezoidal, Simpson and weighted Simpson's rules. This figure plots the relative error against the grid size Δx on a log-log scale.

However, the weighted integration methods do not give a better accuracy. On the contrary, the accuracy is worse than a traditional numerical integration scheme. The problem was caused by the "division by a very small number" numerical error. In the coefficients α, β and γ, we

had some terms like $1/(k^3\Delta x^2)$. If Δx goes very small, say, 10^{-6}, a large numerical error could be introduced due to machine accuracy. Although we have spent so much effort deriving the weighted integration method, we would still stick with the traditional Simpson's rule as our integration method. On the other hand, this is also a "good news" as the Simpson's rule is much easier to implement.

If we are only interested in the electron interactions in the same shell, i.e. $n_1l_1 = n_2l_2 = n_3l_3 = n_4l_4 = nl$, Eqn. (4.14) simplifies to

$$F^{(k)}(nl) = \int_0^\infty dr_1 |u_{nl}(r_1)|^2 \left[\frac{1}{r_1^{k+1}} \int_0^{r_1} dr_2\, r_2^k |u_{nl}(r_2)|^2 + r_1^k \int_{r_1}^\infty dr_2\, \frac{1}{r_2^{k+1}} |u_{nl}(r_2)|^2 \right] \tag{4.28}$$

To verify the accuracy of Simpson's integration on the Slater-Condon parameters, we take the first few hydrogen wave functions as inputs, which are known exactly. We compute the on-shell interaction parameters according to Eqn. (4.28). We used a grid $\{r_{\min} = 0.0001; r_{\max} = 150.0; \Delta x = 0.002;\}$.[2] The results are listed in Table 4.1. Notice that exact wave functions are used as input. In practical calculations, when a numerical wave function is provided, the results will be subject to the accuracy of the given numerical wave function.

Table 4.1.: Numerical (Simpson's rule) and exact Slater-Condon parameters for the first few exact hydrogen wave functions. Energies are given in units of Hartree (a.u.).

Orbital	k	Numerical	Exact	Abs Error	Rel Error
$1s$	0	0.625000	5/8	2.7842×10^{-12}	4.4547×10^{-12}
$2s$	0	0.150391	77/512	2.1716×10^{-13}	1.4440×10^{-12}
$2p$	0	0.181641	93/512	7.8798×10^{-14}	4.3381×10^{-13}
	2	0.087891	45/512	3.0531×10^{-14}	3.4738×10^{-13}
$3s$	0	0.066406	17/256	1.4978×10^{-13}	2.2556×10^{-12}
$3p$	0	0.071868	1987/27648	1.1495×10^{-13}	1.5995×10^{-12}
	2	0.035988	995/27648	3.9472×10^{-13}	1.0968×10^{-11}
$3d$	0	0.086046	793/9216	6.3449×10^{-14}	7.3739×10^{-13}
	2	0.045421	2093/46080	1.1700×10^{-13}	2.5760×10^{-12}
	4	0.029622	91/3072	6.3165×10^{-13}	2.1323×10^{-11}
$4s$	0	0.037271	19541/524288	1.4345×10^{-13}	3.8489×10^{-12}
$4p$	0	0.038935	20413/524288	1.2726×10^{-13}	3.2685×10^{-12}
	2	0.019922	10445/524288	5.2667×10^{-13}	2.6436×10^{-11}
$4d$	0	0.042673	22373/524288	9.9913×10^{-14}	2.3414×10^{-12}
	2	0.021573	56553/2621440	3.8654×10^{-13}	1.7917×10^{-11}
	4	0.014780	7749/524288	2.2203×10^{-13}	1.5022×10^{-11}
$4f$	0	0.050226	26333/524288	5.4602×10^{-14}	1.0871×10^{-12}
	2	0.028140	103275/3670016	1.4186×10^{-13}	5.0413×10^{-12}
	4	0.018802	69003/3670016	2.9456×10^{-13}	1.5667×10^{-11}
	6	0.013910	7293/524288	1.6736×10^{-12}	1.2031×10^{-10}

[2] We used a grid wider than (2.15) to be able to represent hydrogen wave functions with large principle quantum numbers, which spread out to the further region from the nucleus.

4.3. Gaunt coefficients

Our next task is to solve the angular part:

$$A^{(k)}(l_1m_1, l_2m_2, l_3m_3, l_4m_4) = \sum_{\mu=-k}^{k} \int_0^{2\pi} d\phi_1 \int_0^{\pi} d\theta_1 \, \sin\theta_1 \, \overline{Y_{l_1m_1}}(\theta_1, \phi_1) \overline{Y_{k\mu}}(\theta_1, \phi_1) Y_{l_4m_4}(\theta_1, \phi_1) \int_0^{2\pi} d\phi_2 \int_0^{\pi} d\theta_2 \, \sin\theta_2 \, \overline{Y_{l_2m_2}}(\theta_2, \phi_2) Y_{k\mu}(\theta_2, \phi_2) Y_{l_3m_3}(\theta_2, \phi_2) \tag{4.29}$$

Our life simplifies in Bra-Ket notation:

$$A^{(k)}(l_1m_1, l_2m_2, l_3m_3, l_4m_4) = \sum_{\mu=-k}^{k} \langle l_1m_1 | \, \overline{k\mu} \, | l_4m_4 \rangle \, \langle l_2m_2 | \, k\mu \, | l_3m_3 \rangle \tag{4.30}$$

We can get rid of the complex conjugate in the middle by using an important relation of spherical harmonics:[3]

$$\overline{Y_{k\mu}}(\theta, \phi) = (-1)^{\mu} Y_{k,-\mu}(\theta, \phi) \tag{4.31}$$

Now Eqn. (4.30) becomes,

$$A^{(k)}(l_1m_1, l_2m_2, l_3m_3, l_4m_4) = \sum_{\mu=-k}^{k} (-1)^{\mu} \, \langle l_1m_1 | \, k, -\mu \, | l_4m_4 \rangle \, \langle l_2m_2 | \, k\mu \, | l_3m_3 \rangle \tag{4.32}$$

A term like $\langle l_1m_1 | \, k\mu \, | l_2m_2 \rangle$ is called a Gaunt coefficient. It is purely integrals of three spherical harmonics. An important property of this integral is that the integration vanishes under certain combinations of indices. Non-trivial Gaunt coefficients must satisfy the sum rules:

$$\mu = m_1 - m_2 \tag{4.33}$$

$$|l_1 - l_2| \leq k \leq l_1 + l_2 \quad \text{and} \quad l_1 + l_2 + k \text{ is even} \tag{4.34}$$

While the relation in Eqn. (4.34) is difficult to derive, it is rather straightforward to show the sum rule in Eqn. (4.33). One can express the spherical harmonics in terms of the associated Legendre polynomials,[4]

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} e^{im\phi} P_l^m(\cos\theta) \tag{4.35}$$

Therefore, a Gaunt coefficient reads,

$$\langle l_1m_1 | \, k\mu \, | l_2m_2 \rangle = \sqrt{\frac{(2l_1+1)(2k+1)(2l_2+1)}{(4\pi)^3} \frac{(l_1-m_1)!(k-\mu)!(l_2-m_2)!}{(l_1+m_1)!(k+\mu)!(l_2+m_2)!}} \int_0^{2\pi} d\phi \, e^{i\phi(-m_1+\mu+m_2)} \int_0^{\pi} d\theta \, \sin\theta \, P_{l_1}^{m_1}(\cos\theta) P_k^{\mu}(\cos\theta) P_{l_2}^{m_2}(\cos\theta) \tag{4.36}$$

[3] You see, we put a comma to separate indices when there is an ambiguity. While $k\mu$ appears as two indices, the notation $k-\mu$ looks like "k minus μ". For the later case, we prefer $k, -\mu$ to make it clear.

[4] The associated Legendre polynomials are defined as $P_l^m(x) = (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m}(P_l(x))$.

It can be seen that if $-m_1+\mu+m_2 \neq 0$, the ϕ integral from 0 to 2π over a complex exponential function gives 0. Hence we had Eqn. (4.33). It is also important to notice that the Gaunt coefficients are real numbers. Now we use a notation for Gaunt coefficients

$$g^{(k)}_{m_1 m_2} = \langle l_1 m_1 |\, k\mu \,| l_2 m_2 \rangle \tag{4.37}$$

We dropped indices l_1 and l_2 in the notation $g^{(k)}_{m_1 m_2}$ because they are usually predefined. We also dropped the index μ since it is determined by m_1 and m_2 automatically. Because of this uniqueness of μ, only one term survives in the summations in Eqn. (4.32).

$$\boxed{A^{(k)}(l_1 m_1, l_2 m_2, l_3 m_3, l_4 m_4) = (-1)^{\mu} g^{(k)}_{m_1 m_4} g^{(k)}_{m_2 m_3}} \tag{4.38}$$

It is now a matter of evaluating the Gaunt coefficients. The integration of three spherical harmonics can be obtained from a recursion relation.

$$\langle l_1 m_1 |\, k\mu \,| l_2 m_2 \rangle = a \,\langle l_1+1, m_1 |\, k{-}1, \mu \,| l_2 m_2 \rangle + b \,\langle l_1 - 1, m_1 |\, k{-}1, \mu \,| l_2 m_2 \rangle + c \,\langle l_1 m_1 |\, k{-}2, \mu \,| l_2 m_2 \rangle \tag{4.39}$$

where,

$$a = \sqrt{\frac{(2k+1)(2k-1)(l_1+m_1+1)(l_1-m_1+1)}{(k+\mu)(k-\mu)(2l_1+3)(2l_1+1)}} \tag{4.40}$$

$$b = \sqrt{\frac{(2k+1)(2k-1)(l_1+m_1)(l_1-m_1)}{(k+\mu)(k-\mu)(2l_1+1)(2l_1-1)}} \tag{4.41}$$

$$c = -\sqrt{\frac{(2k+1)(k+\mu-1)(k-\mu-1)}{(k+\mu)(k-\mu)(2k-3)}} \tag{4.42}$$

with base case,

$$\langle l_1 m_1 |\, 00 \,| l_2 m_2 \rangle = \frac{1}{\sqrt{4\pi}} \delta_{l_1 l_2} \delta_{m_1 m_2} \tag{4.43}$$

Derivation: Warning: A bit long.

Given the relation between Y_{lm} and P^m_l and the recursion relation of the associated Legendre polynomials,

$$Y_{lm} = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} e^{im\phi} P^m_l \tag{4.44}$$

$$(l-m+1)P^m_{l+1} = (2l+1)xP^m_l - (l+m)P^m_{l-1} \tag{4.45}$$

We get,

$$(l-m+1)\sqrt{\frac{(2l+1)(l+m+1)}{(2l+3)(l-m+1)}} Y_{l+1,m} = (2l+1)xY_{lm} - (l+m)\sqrt{\frac{(2l+1)(l-m)}{(2l-1)(l+m)}} Y_{l-1,m} \tag{4.46}$$

Now, substitutions,[5]

$$
\begin{aligned}
&\langle Y_{l_1m_1}|\boxed{Y_{k\mu}}|Y_{l_2m_2}\rangle \\
=&\ \langle Y_{l_1m_1}|\frac{1}{k-\mu}\sqrt{\frac{(2k+1)(k-\mu)}{(2k-1)(k+\mu)}}\left[(2k-1)xY_{k-1,\mu}-(k+\mu-1)\sqrt{\frac{(2k-1)(k-\mu-1)}{(2k-3)(k+\mu-1)}}Y_{k-2,\mu}\right]|Y_{l_2m_2}\rangle \\
=&\sqrt{\frac{(2k+1)(2k-1)}{(k+\mu)(k-\mu)}}\boxed{\langle xY_{l_1m_1}|}Y_{k-1,\mu}|Y_{l_2m_2}\rangle-\sqrt{\frac{(2k+1)(k+\mu-1)(k-\mu-1)}{(k+\mu)(k-\mu)(2k-3)}}\langle Y_{l_1m_1}|Y_{k-2,\mu}|Y_{l_2m_2}\rangle \\
=&\sqrt{\frac{(2k+1)(2k-1)}{(k+\mu)(k-\mu)}}\left[\frac{l_1-m_1+1}{2l_1+1}\sqrt{\frac{(2l_1+1)(l_1+m_1+1)}{(2l_1+3)(l_1-m_1+1)}}\langle Y_{l_1+1,m_1}|Y_{k-1,\mu}|Y_{l_2m_2}\rangle\right. \\
&\left.+\frac{l_1+m_1}{2l_1+1}\sqrt{\frac{(2l_1+1)(l_1-m_1)}{(2l_1-1)(l_1+m_1)}}\langle Y_{l_1-1,m_1}|Y_{k-1,\mu}|Y_{l_2m_2}\rangle\right] \\
&-\sqrt{\frac{(2k+1)(k+\mu-1)(k-\mu-1)}{(k+\mu)(k-\mu)(2k-3)}}\langle Y_{l_1m_1}|Y_{k-2,\mu}|Y_{l_2m_2}\rangle \\
=&\sqrt{\frac{(2k+1)(2k-1)(l_1+m_1+1)(l_1-m_1+1)}{(k+\mu)(k-\mu)(2l_1+3)(2l_1+1)}}\langle Y_{l_1+1,m_1}|Y_{k-1,\mu}|Y_{l_2m_2}\rangle \\
&+\sqrt{\frac{(2k+1)(2k-1)(l_1+m_1)(l_1-m_1)}{(k+\mu)(k-\mu)(2l_1+1)(2l_1-1)}}\langle Y_{l_1-1,m_1}|Y_{k-1,\mu}|Y_{l_2m_2}\rangle \\
&-\sqrt{\frac{(2k+1)(k+\mu-1)(k-\mu-1)}{(k+\mu)(k-\mu)(2k-3)}}\langle Y_{l_1m_1}|Y_{k-2,\mu}|Y_{l_2m_2}\rangle \qquad \text{Q.E.D.}
\end{aligned}
\tag{4.47}
$$

To get a flavor of how this recursion works, we consider an element

$$\langle 1,-1|\,2,0\,|1,-1\rangle$$

Apply Eqn. (4.39), we find

$$
\begin{aligned}
&\langle 1,-1|\,2,0\,|1,-1\rangle \\
=&\frac{\sqrt{3}}{2}\langle 2,-1|\,1,0\,|1,-1\rangle+0\underbrace{\langle 0,-1|}_{\text{undefined}}1,0\,|1,-1\rangle-\frac{\sqrt{5}}{2}\underbrace{\langle 1,-1|\,0,0\,|1,-1\rangle}_{1/\sqrt{4\pi}} \\
=&\frac{\sqrt{3}}{2}\left(\sqrt{\frac{24}{35}}\underbrace{\langle 3,-1|\,0,0\,|1,-1\rangle}_{0}+\sqrt{\frac{3}{5}}\underbrace{\langle 1,-1|\,0,0\,|1,-1\rangle}_{1/\sqrt{4\pi}}+0\,\langle 2,-1|\ \underbrace{-1,0}_{\text{undefined}}\ |1,-1\rangle\right)-\frac{\sqrt{5}}{2\sqrt{4\pi}} \\
=&-\frac{1}{\sqrt{20\pi}}
\end{aligned}
\tag{4.48}
$$

Yeah, we found it!

Recursion (4.39) is universal for all Gaunt coefficients. So one would expect that we could implement a recursive function `Gaunt(l1, l2, k, m1, m2)` and apply to every Gaunt to obtain

[5] Notice that the following notations are equivalent: $\langle Y_{l_1m_1}|\,Y_{k\mu}\,|Y_{l_2m_2}\rangle=\langle l_1m_1|\,k\mu\,|l_2m_2\rangle$. Here we prefer the former one, since a term like "xY_{lm}" is less confusing than "xlm"

their values. That would be too good to be true. Unfortunately, our recursive function can only be used for the diagonal elements (strictly speaking, for $m_1 = m_2$). Now consider,

$$\langle 1,-1|\,2,-2\,|1,1\rangle$$

Apply Eqn. (4.39),

$$\begin{aligned}
&\langle 1,-1|\,2,-2\,|1,1\rangle \\
=&\sqrt{\frac{30}{0}}\,\langle 2,-1|\ \underbrace{1,-2}_{\text{undefined}}\ |1,1\rangle + \sqrt{\frac{0}{0}}\,\underbrace{\langle 0,-1|\,1,-2}_{\text{undefined}}|1,1\rangle - \sqrt{\frac{-15}{0}}\,\langle 1,-1|\ \underbrace{0,-2}_{\text{undefined}}\ |1,1\rangle \\
=&\text{NaN}
\end{aligned} \tag{4.49}$$

Not-a-Number.

This doesn't mean that $\langle 1,-1|\,2,-2\,|1,1\rangle$ is equal to infinity or zero or what. A ratio $0/0$ can give us anything, but the information is destroyed. We cannot extract this value from recursion (4.39). This "division by zeros" problem is caused by a non-zero μ, namely, $m_1 \neq m_2$. The term $(k+\mu)(k-\mu)$ in the denominator becomes zero if $|\mu| = k$. To get the off-diagonal elements, we have to ask help from the ladder operators L_+ and L_-. Recall,

$$L_\pm\,|lm\rangle = \alpha^\pm_{lm}\,|l, m\pm 1\rangle \tag{4.50}$$

where,

$$\alpha^+_{lm} = \sqrt{(l+m+1)(l-m)} \tag{4.51}$$

$$\alpha^-_{lm} = \sqrt{(l+m)(l-m+1)} \tag{4.52}$$

Now consider two nonzero elements

$$\Big(\langle l_1m_1|\,L_+\Big)k,\mu-1\,|l_2m_2\rangle \quad \text{and} \quad \Big(\langle l_1m_1|\,L_-\Big)k,\mu+1\,|l_2m_2\rangle$$

$$\begin{cases}
\Big(\langle l_1m_1|\,L_+\Big)k,\mu-1\,|l_2m_2\rangle = \alpha^-_{l_1m_1}\,\langle l_1,m_1-1|\,k,\mu-1\,|l_2m_2\rangle \\
\langle l_1m_1|\,\Big(L_+\,k,\mu-1\,|l_2m_2\rangle\Big) = \alpha^+_{k,\mu-1}\,\langle l_1m_1|\,k\mu\,|l_2m_2\rangle + \alpha^+_{l_2m_2}\,\langle l_1m_1|\,k,\mu-1\,|l_2,m_2+1\rangle
\end{cases} \tag{4.53}$$

$$\begin{cases}
\Big(\langle l_1m_1|\,L_-\Big)k,\mu+1\,|l_2m_2\rangle = \alpha^+_{l_1m_1}\,\langle l_1,m_1+1|\,k,\mu+1\,|l_2m_2\rangle \\
\langle l_1m_1|\,\Big(L_-\,k,\mu+1\,|l_2m_2\rangle\Big) = \alpha^-_{k,\mu+1}\,\langle l_1m_1|\,k\mu\,|l_2m_2\rangle + \alpha^-_{l_2m_2}\,\langle l_1m_1|\,k,\mu+1\,|l_2,m_2-1\rangle
\end{cases} \tag{4.54}$$

From Eqn. (4.53), we get the relation

$$\alpha^-_{l_1m_1}\,\langle l_1,m_1-1|\,k,\mu-1\,|l_2m_2\rangle = \alpha^+_{k,\mu-1}\,\langle l_1m_1|\,k\mu\,|l_2m_2\rangle + \alpha^+_{l_2m_2}\,\langle l_1m_1|\,k,\mu-1\,|l_2,m_2+1\rangle \tag{4.55}$$

which binds three Gaunt coefficients as

$$\begin{bmatrix}
\cdots & \cdots & \cdots & \cdots \\
\cdots & \boxed{g_{m_1-1,m_2}} & g_{m_1-1,m_2+1} & \cdots \\
\cdots & \boxed{g_{m_1,\ \ m_2}} & \boxed{g_{m_1,\ \ m_2+1}} & \cdots \\
\cdots & \cdots & \cdots & \cdots
\end{bmatrix}$$

From Eqn. (4.54), we get the relation

$$\alpha^{+}_{l_1 m_1} \langle l_1, m_1+1| k, \mu+1 |l_2 m_2\rangle = \alpha^{-}_{k,\mu+1} \langle l_1 m_1| k\mu |l_2 m_2\rangle + \alpha^{-}_{l_2 m_2} \langle l_1 m_1| k, \mu+1 |l_2, m_2-1\rangle \tag{4.56}$$

which binds three Gaunt coefficients as

$$\begin{bmatrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \boxed{g_{m_1,\ m_2-1}} & \boxed{g_{m_1,\ m_2}} & \cdots \\ \cdots & g_{m_1+1,m_2-1} & \boxed{g_{m_1+1,m_2}} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

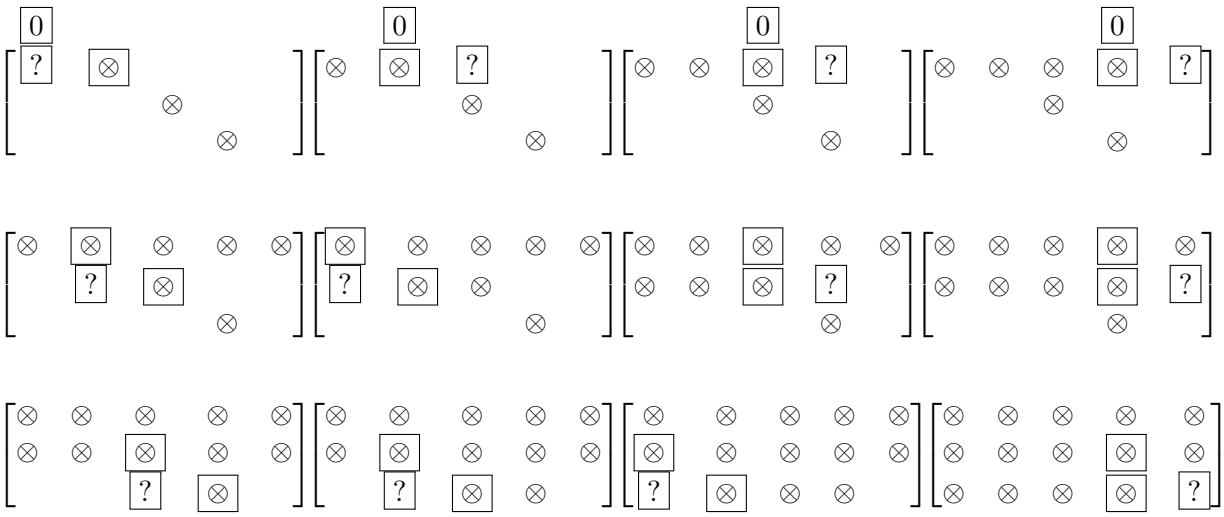
As an illustration, we consider the Gaunt coefficient matrix for $l_1 = 1$ and $l_2 = 2$. This matrix has the following rectangular shape:

$$\begin{array}{c} \\ -1 \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{ccccc} -2 & -1 & 0 & 1 & 2 \end{array} \\ \begin{bmatrix} & & & & \\ & & & & \\ & & & & \end{bmatrix} \end{array}$$

where the m_1 index traverses vertically (row number) and the m_2 index traverses horizontally (column number). A simple strategy for calculating the matrix elements is shown in the diagram below

$$\begin{array}{c} \\ -1 \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{ccccc} -2 & -1 & 0 & 1 & 2 \end{array} \\ \begin{bmatrix} \leftarrow & \otimes & \rightarrow & \rightarrow & \rightarrow \\ \leftarrow & \leftarrow & \otimes & \rightarrow & \rightarrow \\ \leftarrow & \leftarrow & \leftarrow & \otimes & \rightarrow \end{bmatrix} \end{array}$$

The elements "⊗" are the diagonal elements with $m_1 = m_2$, which can be calculated directly from the recursion relation. The arrows indicate the direction of applying the ladder operator relations. Notice that any element outside the matrix is zero, because the coefficients $\alpha^{\pm}_{lm}$ in (4.51) and (4.52) vanish for $m = l$ and $-l$, respectively. A detailed pictorial illustration is shown below: (play it as an animation, once an elements is calculated, we stamp an ⊗)



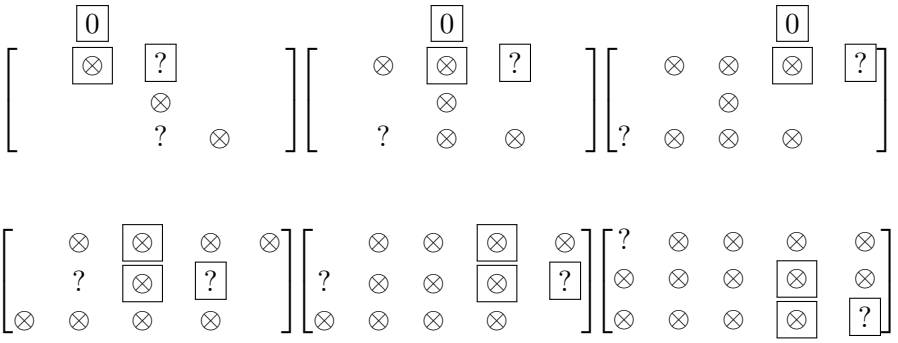
Complete! The speed of the algorithm can be doubled if we consider the symmetry of the matrix. (remember that Gaunt coefficients are real)

$$\begin{aligned}\langle l_1 m_1 |\, k\mu \,| l_2 m_2\rangle &= (-1)^{m_1+m_2+\mu} \overline{\langle l_1, -m_1 |}\; \overline{k, -\mu}\; \overline{| l_2, -m_2 \rangle} \\ &= (-1)^{2m_1} \overline{\langle l_1, -m_1 |\, k, -\mu \,| l_2, -m_2\rangle} \\ &= \langle l_1, -m_1 |\, k, -\mu \,| l_2, -m_2\rangle \end{aligned} \tag{4.57}$$

Eqn. (4.57) says, the Gaunt matrix has the following symmetry: (inverse symmetry)

$$\begin{array}{c} \\ -1 \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{ccccc} -2 & -1 & 0 & 1 & 2 \end{array} \\ \begin{bmatrix} \ominus & \times & \heartsuit & + & \diamondsuit \\ \div & \triangle & \odot & \triangle & \div \\ \diamondsuit & + & \heartsuit & \times & \ominus \end{bmatrix} \end{array}$$

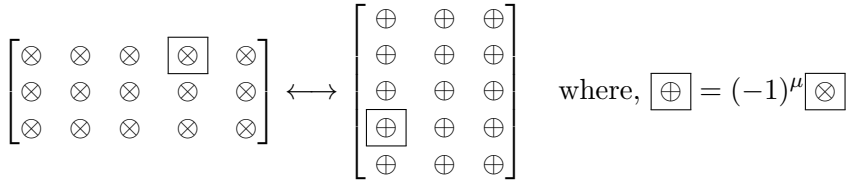
Hence, our algorithm simplifies to:



Done! Symmetry in Eqn. (4.57) allows us to compute just half of the elements thus obtaining a speed up. Yet, there is another symmetry property of Gaunt coefficients which relates two adjoint matrices. (Gaunt coefficients are real)

$$\begin{aligned}\langle l_1 m_1 |\, k\mu \,| l_2 m_2\rangle &= \overline{\langle l_1 m_1 |\, k\mu \,| l_2 m_2\rangle} \\ &= \langle l_2 m_2 |\, \overline{k\mu} \,| l_1 m_1\rangle \\ &= (-1)^{\mu} \langle l_2 m_2 |\, k, -\mu \,| l_1 m_1\rangle \end{aligned} \tag{4.58}$$

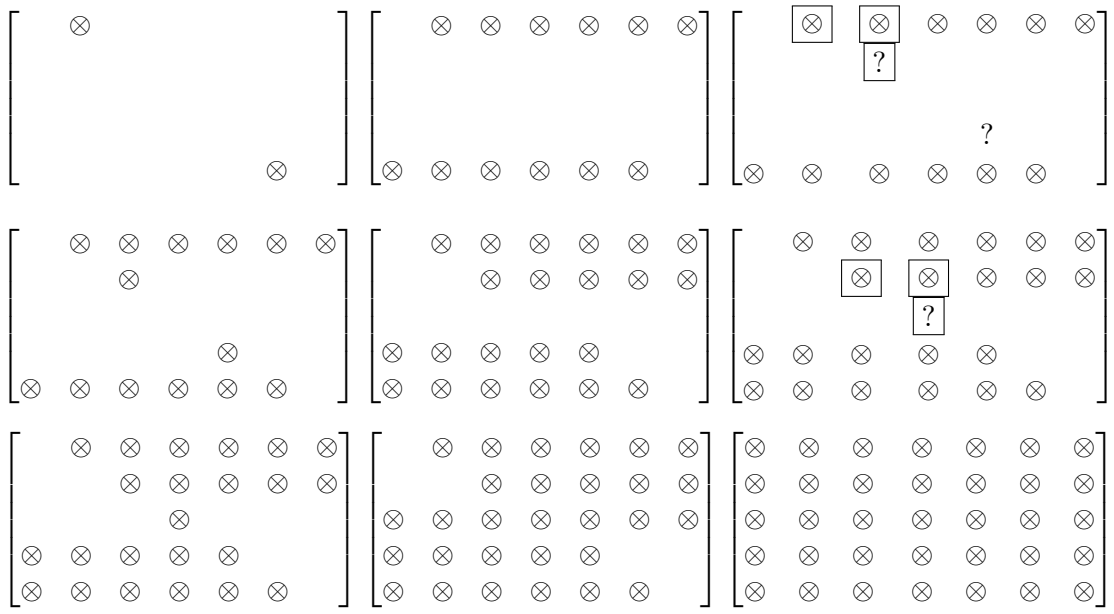
Eqn. (4.58) says,



Once the Gaunt matrix with $l_1 = 1$ and $l_2 = 2$ is determined, the Gaunt matrix with $l_1 = 2$ and $l_2 = 1$ can be read off directly without doing another calculation.

Previously, we calculated all diagonal elements from the recursion relation. That was absolutely valid, but recursive calls may cost considerable computational time especially when the matrix

gets huge. In fact, we need to calculate only one diagonal element from the recursion relation. The others can be obtained from ladder operators. For example (a 5-by-7 matrix):



This is how we implement our program to evaluate the Gaunt coefficients. It might be useful if we summarize a few Gaunt matrices with $l_1 = l_2$ for s, p, d and f shells (the on-shell interactions). The results are listed in the next pages.

s shell ($l_1 = l_2 = 0$):

$k = 0$

$$\frac{1}{\sqrt{4\pi}} \begin{bmatrix} 1 \end{bmatrix} \tag{4.59}$$

p shell ($l_1 = l_2 = 1$):

$k = 0$

$$\frac{1}{\sqrt{4\pi}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.60}$$

$k = 2$

$$\frac{1}{\sqrt{4\pi}} \frac{1}{\sqrt{5}} \begin{bmatrix} -1 & \sqrt{3} & -\sqrt{6} \\ -\sqrt{3} & 2 & -\sqrt{3} \\ -\sqrt{6} & \sqrt{3} & -1 \end{bmatrix} \tag{4.61}$$

d shell ($l_1 = l_2 = 2$):

$k = 0$

$$\frac{1}{\sqrt{4\pi}} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.62}$$

$k = 2$

$$\frac{1}{\sqrt{4\pi}} \frac{1}{7} \begin{bmatrix} -\sqrt{20} & \sqrt{30} & -\sqrt{20} & 0 & 0 \\ -\sqrt{30} & \sqrt{5} & \sqrt{5} & -\sqrt{30} & 0 \\ -\sqrt{20} & -\sqrt{5} & \sqrt{20} & -\sqrt{5} & -\sqrt{20} \\ 0 & -\sqrt{30} & \sqrt{5} & \sqrt{5} & -\sqrt{30} \\ 0 & 0 & -\sqrt{20} & \sqrt{30} & -\sqrt{20} \end{bmatrix} \tag{4.63}$$

$k = 4$

$$\frac{1}{\sqrt{4\pi}} \frac{1}{7} \begin{bmatrix} 1 & -\sqrt{5} & \sqrt{15} & -\sqrt{35} & \sqrt{70} \\ \sqrt{5} & -4 & \sqrt{30} & -\sqrt{40} & \sqrt{35} \\ \sqrt{15} & -\sqrt{30} & 6 & -\sqrt{30} & \sqrt{15} \\ \sqrt{35} & -\sqrt{40} & \sqrt{30} & -4 & \sqrt{5} \\ \sqrt{70} & -\sqrt{35} & \sqrt{15} & -\sqrt{5} & 1 \end{bmatrix} \tag{4.64}$$

f shell ($l_1 = l_2 = 3$):

$k = 0$

$$\frac{1}{\sqrt{4\pi}} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.65}$$

$k = 2$

$$\frac{1}{\sqrt{4\pi}} \frac{1}{3\sqrt{5}} \begin{bmatrix} -5 & 5 & -\sqrt{10} & 0 & 0 & 0 & 0 \\ -5 & 0 & \sqrt{15} & -\sqrt{20} & 0 & 0 & 0 \\ -\sqrt{10} & -\sqrt{15} & 3 & \sqrt{2} & -\sqrt{24} & 0 & 0 \\ 0 & -\sqrt{20} & -\sqrt{2} & 4 & -\sqrt{2} & -\sqrt{20} & 0 \\ 0 & 0 & -\sqrt{24} & \sqrt{2} & 3 & -\sqrt{15} & -\sqrt{10} \\ 0 & 0 & 0 & -\sqrt{20} & \sqrt{15} & 0 & -5 \\ 0 & 0 & 0 & 0 & -\sqrt{10} & 5 & -5 \end{bmatrix} \tag{4.66}$$

$k = 4$

$$\frac{1}{\sqrt{4\pi}} \frac{1}{11} \begin{bmatrix} 3 & -\sqrt{30} & \sqrt{54} & -\sqrt{63} & \sqrt{42} & 0 & 0 \\ \sqrt{30} & -7 & \sqrt{32} & -\sqrt{3} & -\sqrt{14} & \sqrt{70} & 0 \\ \sqrt{54} & -\sqrt{32} & 1 & \sqrt{15} & -\sqrt{40} & \sqrt{14} & \sqrt{42} \\ \sqrt{63} & -\sqrt{3} & -\sqrt{15} & 6 & -\sqrt{15} & -\sqrt{3} & \sqrt{63} \\ \sqrt{42} & \sqrt{14} & -\sqrt{40} & \sqrt{15} & 1 & -\sqrt{32} & \sqrt{54} \\ 0 & \sqrt{70} & -\sqrt{14} & -\sqrt{3} & \sqrt{32} & -7 & \sqrt{30} \\ 0 & 0 & \sqrt{42} & -\sqrt{63} & \sqrt{54} & -\sqrt{30} & 3 \end{bmatrix} \tag{4.67}$$

$k = 6$

$$\frac{1}{\sqrt{4\pi}} \frac{5}{33\sqrt{13}} \begin{bmatrix} -1 & \sqrt{7} & -\sqrt{28} & \sqrt{84} & -\sqrt{210} & \sqrt{462} & -\sqrt{924} \\ -\sqrt{7} & 6 & -\sqrt{105} & \sqrt{224} & -\sqrt{378} & \sqrt{504} & -\sqrt{462} \\ -\sqrt{28} & \sqrt{105} & -15 & \sqrt{350} & -\sqrt{420} & \sqrt{378} & -\sqrt{210} \\ -\sqrt{84} & \sqrt{224} & -\sqrt{350} & 20 & -\sqrt{350} & \sqrt{224} & -\sqrt{84} \\ -\sqrt{210} & \sqrt{378} & -\sqrt{420} & \sqrt{350} & -15 & \sqrt{105} & -\sqrt{28} \\ -\sqrt{462} & \sqrt{504} & -\sqrt{378} & \sqrt{224} & -\sqrt{105} & 6 & -\sqrt{7} \\ -\sqrt{924} & \sqrt{462} & -\sqrt{210} & \sqrt{84} & -\sqrt{28} & \sqrt{7} & -1 \end{bmatrix} \tag{4.68}$$

Chapter 5

Construction of multiplet states

5.1. Setting up the basis and Hamiltonian

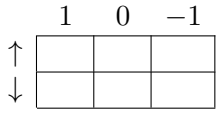
Oh, this is probably the most crucial step. All of our efforts spent so far on the Slater-Condon parameters and the Gaunt coefficients are aimed at calculating the matrix elements

$$U_{\alpha\beta\gamma\delta} = \delta_{\sigma_1\sigma_4}\delta_{\sigma_2\sigma_3} \sum_{k(\text{sum rule})} \frac{4\pi}{2k+1} R^{(k)}(n_1l_1, n_2l_2, n_3l_3, n_4l_4) A^{(k)}(l_1m_1, l_2m_2, l_3m_3, l_4m_4) \tag{5.1}$$

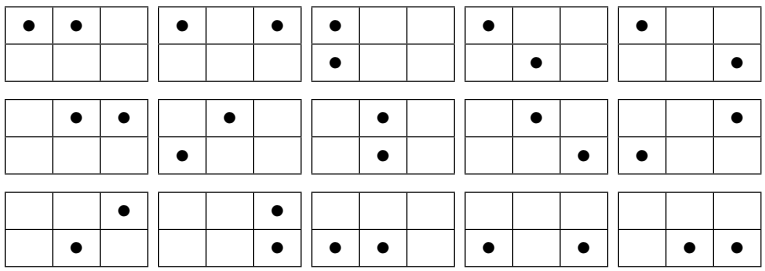
so that we can build up the Coulomb repulsion Hamiltonian.

We called $U_{\alpha\beta\gamma\delta}$ the matrix element. But as you might have noticed already, it has four indices α, β, γ and δ. It is not really a "matrix" in our common sense. Indeed, $U_{\alpha\beta\gamma\delta}$ is not the matrix that we will directly work with. The actual matrix we will be using is the matrix representation of the Hamiltonian H_U. Now, to obtain the matrix representation, we must ask ourselves what basis do we work with.

Our first step is to set up a basis in the electron configuration space. As a demonstration, we will do a case study for a p^2 orbital. According to Pauli exclusion principle, one p shell can at most contain 6 electrons with orbital and spin angular projection momenta $m = 1, 0, -1$ and $\sigma = \uparrow, \downarrow$, respectively. One can think the problem as 6 boxes where we can put (up to) 6 electrons (see the diagram below).



In our problem we have 2 electrons. How many possible ways are there to put them into our boxes? This is simply a combinatorics problem.



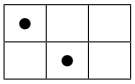
6 boxes put 2 electrons, that is "6 choose 2"

$$\binom{6}{2} = \frac{6 \times 5}{2!} = 15$$

possible ways. They are our 15 basis states (or basis vectors[1]) of electrons in the configuration space. Everyone says that, but what does it mean? Imagine you put two electrons into the shell: Of course we can arrange them in the ways as in the diagram, but electrons can do things peculiar, they can form linear combinations of those basis states.[2] Those 15 basis states (15 vectors) span a space that the electrons can live in. Recall our Cartesian coordinate, we had 3 basis vectors called $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$. Similarly, we can also assign names to those 15 basis vectors, for example, from $\hat{\mathbf{a}}$ to $\hat{\mathbf{o}}$ (if I counted correctly). But a more convenient and sophisticated naming method is the bit representation.

$$|- - - - - -\rangle$$

Here we have 6 sites, if a site is occupied by an electron, we put a "1", otherwise, we put a "0" (they are called occupation numbers). For example,

 $\equiv |100010\rangle$ (5.2)

Therefore, our 15 basis states are:

$\|110000\rangle$	$\|101000\rangle$	$\|100100\rangle$	$\|100010\rangle$	$\|100001\rangle$
$\|011000\rangle$	$\|010100\rangle$	$\|010010\rangle$	$\|010001\rangle$	$\|001100\rangle$
$\|001010\rangle$	$\|001001\rangle$	$\|000110\rangle$	$\|000101\rangle$	$\|000011\rangle$

The bit representation is convenient in the sense that they can be easily stored in the computer. For instance, a binary number 100010 is nothing but a 34 in decimal. To store $|100010\rangle$ we simply store an integer 34 in the computer. Yet, the pictorial meaning can be easily seen from its binary format.

A simple algorithm of the basis setup for a given shell `l` and a given number of electrons `Ne` is shown Algorithm 5.1. Notice that the function `countBit()` counts for the number of 1's in a given binary number.

After setting up the basis, we are now able to build the matrix representation of the Hamiltonian. To obtain one entry at row i and column j, it means to evaluate the matrix element between two basis state $|i\rangle$ and $|j\rangle$.

$$\langle i| H_U |j\rangle = \langle i| \frac{1}{2} \sum_{\alpha,\beta,\gamma,\delta} U_{\alpha\beta\gamma\delta} c_\alpha^\dagger c_\beta^\dagger c_\gamma c_\delta |j\rangle \tag{5.3}$$

It is not trivial at all to evaluate Eqn. (5.3). One needs to implement a program that loops over all six indices $(i, j, \alpha, \beta, \gamma, \delta)$. For each pair of (i, j), we sum up all the corresponding $U_{\alpha\beta\gamma\delta}$ to obtain the entry $\langle i| H_U |j\rangle$. Notice that the factor $1/2$ in front is used to compensate the double counting over electron pairs. A simplified algorithm is shown in Algorithm 5.2.

[1] By calling "states" or "vectors", we really mean the same. When we say a "state", it is more from a quantum mechanics perspective that "a state of electrons". When we use "vector", we more emphasize from the linear algebra point of view.

[2] Linear combinations of states are also valid solutions of the Schrödinger equation.

Algorithm 5.1 Set up basis

```
1: function BASIS(l, N_e)
2:     N_site ← 4 * l + 2
3:     dim ← BINOMIAL(N_site, N_e)
4:     index_i ← 0
5:     for conf_i ← 0 to 2^N_site do
6:         if COUNTBIT(conf_i) = N_e then
7:             index[conf_i] ← index_i
8:             conf[index_i] ← conf_i
9:             index_i ← index_i + 1
10:    basis.dim ← dim
11:    basis.index ← index
12:    basis.conf ← conf
13:    return basis
```

Algorithm 5.2 Set up Hamiltonian

```
1: function HAMILTONIAN(basis)
2:     dim ← basis.dim
3:     for i ← 0 to dim do
4:         conf_i ← basis.conf[i]
5:         for all α do
6:             if ISBIT(α, conf_i) then
7:                 conf_α ← CLEARBIT(α, conf_i)
8:                 for all β do
9:                     if ISBIT(β, conf_α) then
10:                        conf_β ← CLEARBIT(β, conf_α)
11:                        for all γ do
12:                            if !ISBIT(γ, conf_β) then
13:                                conf_γ ← SETBIT(γ, conf_β)
14:                                for all δ do
15:                                    if !ISBIT(δ, conf_γ) then
16:                                        conf_δ ← SETBIT(δ, conf_γ)
17:                                        conf_j ← conf_δ
18:                                        j ← basis.index[conf_j]
19:                                        H_U[i, j] ← H_U[i, j] + fsign * 0.5 * U_αβγδ
20:    return H_U
```

The function `clearBit()` clears a bit in a given configuration. For example, `clearBit(2,111111)` returns `110111` (in binary). Similarly, `setBit()` sets a bit in a given configuration. The function `isBit()` tests whether a specific site is occupied by an electron or not, from which we know if we can clear or set a bit on that site. It is (always) tricky that we have to pay extra attention to the fermi-sign `fsign`. Every time we call a function `clearBit()` or `setBit()`, it means that we applied an electron annihilation or creation operator. But when we apply an annihilation or creation operator, a fermi-sign (± 1) is involved [16].

5.2. Construction of eigen-states

Our problem is almost solved. What we want to know ultimately are the eigen-vectors and eigen-energies of the Hamiltonian. But how big is this business? Since we have the matrix, to get the eigen-vectors and eigen-energies, we can simply throw it into a matrix diagonalization solver (e.g. `Lapack`), then we are done. But, is this the end of our story?

We continue our case study for the p^2 orbital. In our previous section, we set the basis up by 15 basis vectors (let's first sort the basis vectors in an ascending order of binary numbers). Under this basis, we built up a 15-by-15 Hamiltonian. It has the following pattern:

$$H_U = \begin{bmatrix} \otimes & & & & & & & & & & & & & & \\ & \otimes & & & & & & & & & & & & & \\ & & \otimes & & & & & & & & & & & & \\ & & & \otimes & & & & & & & & & & & \\ & & & & \otimes & & \otimes & & & & & & & & \\ & & & & & \otimes & & \otimes & & \otimes & & & & & \\ & & & & \otimes & & \otimes & & & & & & & & \\ & & & & & \otimes & & \otimes & & \otimes & & & & & \\ & & & & & & \otimes & & & & \otimes & & & & \\ & & & & & & & \otimes & & & & \otimes & & & \\ & & & & & \otimes & & \otimes & & \otimes & & & & & \\ & & & & & & \otimes & & & & \otimes & & & & \\ & & & & & & & & & & & \otimes & & & \\ & & & & & & & & & & & & \otimes & & \\ & & & & & & & & & & & & & \otimes & \end{bmatrix} \begin{matrix} |000011\rangle \\ |000101\rangle \\ |000110\rangle \\ |001001\rangle \\ |001010\rangle \\ |001100\rangle \\ |010001\rangle \\ |010010\rangle \\ |010100\rangle \\ |011000\rangle \\ |100001\rangle \\ |100010\rangle \\ |100100\rangle \\ |101000\rangle \\ |110000\rangle \end{matrix}$$

The empty elements are killed by the $\delta_{\sigma_1\sigma_4}\delta_{\sigma_2\sigma_3}$ in Eqn. (5.1) (conservation of spin angular momentum) and the vanishing Gaunt coefficients. Recall Eqn. (4.38), the sum rule requires $m_1 + m_2 = m_3 + m_4$ (conservation of orbital angular momentum). At this stage, we can throw H_U into a matrix diagonalization solver, then we will obtain 15 eigen-energies plus 15 corresponding eigen-vectors. For people who are only interested in the eigen-energies (spectral lines) of the system, calling a matrix diagonalization solver will be enough to solve the problem. But if one wants to understand more physics about the system, asking, "What exactly are the quantum numbers of those eigen-states?", our solutions from the "brute-force" diagonalization cannot explain. As we shall find out later, it turns out that those 15 eigen-energies are highly degenerate. Among those 15 energies, we will obtain something like: $9 \times E_1$, $5 \times E_2$, $1 \times E_3$. Now the problem is, if the eigen-energies are degenerate, the corresponding eigen-vectors are not uniquely defined. They live in a space where any linear combinations can be eigen-vectors. Thus it is difficult to address each eigen-vector by a specific set of quantum numbers. This high degeneracy that we claimed seems to be a mystery. In fact, the reason behind is that there is a special symmetry in the Hamiltonian H_U. Based on this symmetry, we can use a special technique to diagonalize this matrix without calling a matrix diagonalization routine like `Lapack`. More importantly, with this technique, eigen-vectors can be identified uniquely with beautiful quantum numbers. The idea is to use the angular momentum ladder operators. First, we need to prove that the Hamiltonian H_U commutes with the total angular momenta $\mathbf{L}$ and $\mathbf{S}$.

Proof: By definition, the total angular momenta are $\mathbf{L} = \sum_{i=1}^{N} \boldsymbol{\ell}_i$ and $\mathbf{S} = \sum_{i=1}^{N} \mathbf{s}_i$.

$[H_U, \mathbf{S}] = 0$ is trivial, since

$$H_U = \sum_{i<j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{5.4}$$

has no dependence on spin.

$[H_U, \mathbf{L}] = 0$ is a bit more difficult. We need

$$\boldsymbol{\ell}_i = -i\mathbf{r}_i \times \boldsymbol{\nabla}_i \tag{5.5}$$

Now,

$$\begin{aligned}
\left[\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \boldsymbol{\ell}_i\right] &= \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}(-i\mathbf{r}_i \times \boldsymbol{\nabla}_i) - (-i\mathbf{r}_i \times \boldsymbol{\nabla}_i)\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\
&= i\mathbf{r}_i \times \left(\boldsymbol{\nabla}_i \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}\right) \\
&= i\mathbf{r}_i \times \left(-\frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3}\right) \\
&= i\frac{\mathbf{r}_i \times \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3} \\
&= i\mathbf{r}_j \times \frac{\mathbf{r}_j - \mathbf{r}_i}{|\mathbf{r}_j - \mathbf{r}_i|^3} \\
&= -i\mathbf{r}_j \times \left(\boldsymbol{\nabla}_j \frac{1}{|\mathbf{r}_j - \mathbf{r}_i|}\right) = -\left[\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \boldsymbol{\ell}_j\right]
\end{aligned} \tag{5.6}$$

Therefore,

$$\left[\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \boldsymbol{\ell}_i + \boldsymbol{\ell}_j\right] = 0 \tag{5.7}$$

$$\left[\sum_{i<j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \sum_{i=0}^{N} \boldsymbol{\ell}_i\right] = 0 \qquad \text{Q.E.D.} \tag{5.8}$$

Given that $[H_U, \mathbf{L}] = 0$ and $[H_U, \mathbf{S}] = 0$, we have the following relations:

$$[H_U, L^2] = 0 \qquad [H_U, L_z] = 0 \qquad [H_U, S^2] = 0 \qquad [H_U, S_z] = 0$$

As a result, an eigen-vector of H_U is simultaneously eigen-vectors of L^2, L_z, S^2, S_z. Thus, we can denote an eigen-vector as $|L, M_L, S, M_S\rangle$.[3] The commutation relations also lead to $[H_U, L_\pm] = 0$ and $[H_U, S_\pm] = 0$. Therefore, for a given eigen-vector $|L, M_L, S, M_S\rangle$ of H_U, we can find new eigen-vectors using the ladder operators, for example,

$$\begin{aligned}
L_- H_U |L, M_L, S, M_S\rangle &= L_- E |L, M_L, S, M_S\rangle \\
H_U L_- |L, M_L, S, M_S\rangle &= E L_- |L, M_L, S, M_S\rangle \\
H_U |L, M_L - 1, S, M_S\rangle &= E |L, M_L - 1, S, M_S\rangle
\end{aligned} \tag{5.9}$$

[3] At this stage, we denote an eigen-vector with the four quantum numbers $|L, M_L, S, M_S\rangle$, which is normally sufficient. However, in the next section, we will find the situation in which $|L, M_L, S, M_S\rangle$ is not able to address an eigen-vector uniquely. Then we would require an additional quantum number (called seniority number) to address the eigen-vector.

Similarly,

$$\begin{aligned} S_- H_U \left|L, M_L, S, M_S\right\rangle &= S_- E \left|L, M_L, S, M_S\right\rangle \\ H_U S_- \left|L, M_L, S, M_S\right\rangle &= E S_- \left|L, M_L, S, M_S\right\rangle \\ H_U \left|L, M_L, S, M_S - 1\right\rangle &= E \left|L, M_L, S, M_S - 1\right\rangle \end{aligned} \tag{5.10}$$

The idea is,

> **Starting from a leading vector, we can construct subsequent vectors by applying ladder operators.** (5.11)

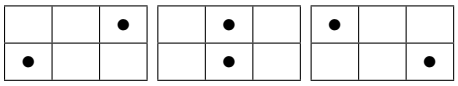
This is the technique we will use to diagonalize our Hamiltonian.

In Cartesian coordinates, we are used to arrange the basis vectors as $[\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}]$. But it really doesn't hurt if we do $[\hat{\mathbf{y}}, \hat{\mathbf{x}}, \hat{\mathbf{z}}]$. Previously we arranged the basis vectors in an ascending order of binary numbers. But we can also arrange the order such that we see clearly the pattern of the Hamiltonian.

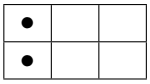
$H_U =$

	$\lvert M_L\ M_S\rangle$
$\lvert 100100\rangle$	$\lvert\ 2\ \ 0\rangle$
$\lvert 110000\rangle$	$\lvert\ 1\ \ 1\rangle$
$\lvert 010100\rangle$	$\lvert\ 1\ \ 0\rangle$
$\lvert 100010\rangle$	$\lvert\ 1\ \ 0\rangle$
$\lvert 000110\rangle$	$\lvert\ 1\ -1\rangle$
$\lvert 101000\rangle$	$\lvert\ 0\ \ 1\rangle$
$\lvert 001100\rangle$	$\lvert\ 0\ \ 0\rangle$
$\lvert 010010\rangle$	$\lvert\ 0\ \ 0\rangle$
$\lvert 100001\rangle$	$\lvert\ 0\ \ 0\rangle$
$\lvert 000101\rangle$	$\lvert\ 0\ -1\rangle$
$\lvert 011000\rangle$	$\lvert -1\ \ 1\rangle$
$\lvert 001010\rangle$	$\lvert -1\ \ 0\rangle$
$\lvert 010001\rangle$	$\lvert -1\ \ 0\rangle$
$\lvert 000011\rangle$	$\lvert -1\ -1\rangle$
$\lvert 001001\rangle$	$\lvert -2\ \ 0\rangle$

Why this pattern? Consider the configurations $|001100\rangle$, $|010010\rangle$ and $|100001\rangle$:



Those configurations have the same total $M_L = 0$ and total $M_S = 0$. Because of conservation of angular momentum, the electrons can move freely among those three configurations. This gives us the 3-by-3 block in the middle of the Hamiltonian. But a configuration like $|100100\rangle$



has no where to go. It has the maximum $M_L = 1 + 1 = 2$. Any other configuration will give a lower M_L. Consequently, there is only one element in the corresponding row and column in the

matrix. Hence, the state $|100100\rangle$ is an eigen-state of the Hamiltonian. But, as you might have noticed already, $|100100\rangle$ is not the only basis state with the unique $|M_L\, M_S\rangle$. It is extremely useful to make a table to count how many basis states are there for a given $|M_L\, M_S\rangle$.

		M_S		
		-1	0	1
	2	0	1	0
	1	1	2	1
M_L	0	1	3	1
	-1	1	2	1
	-2	0	1	0

We labeled our basis states using two quantum numbers M_L and M_S and we found out that there could be more than one basis state sharing the same M_L and M_S. An eigen-vector of the Hamiltonian is, in general, a linear combination of the basis states with the same M_L and M_S. If we are lucky, we find some basis states with unique quantum numbers. Then those states are eigen-vectors by themselves.

Here comes the technique for constructing the eigen-vectors. We start from the basis state with the largest M_S and the corresponding largest M_L (which is unique):

		M_S		
		-1	0	1
	2	0	1	0
	1	1	2	$\boxed{1}$
M_L	0	1	3	1
	-1	1	2	1
	-2	0	1	0

This gives us an eigen-vector $|1,1,1,1\rangle$. By applying L_- and S_- operators, we obtain

		M_S		
		-1	0	1
	2	0	1	0
	1	$\boxed{1}$	$\boxed{2}$	$\boxed{1}$
M_L	0	$\boxed{1}$	$\boxed{3}$	$\boxed{1}$
	-1	$\boxed{1}$	$\boxed{2}$	$\boxed{1}$
	-2	0	1	0

$|1,\ 1,1,-1\rangle$ $|1,\ 1,1,\ 0\rangle$ $|1,\ 1,1,\ 1\rangle$

$|1,\ 0,1,-1\rangle$ $|1,\ 0,1,\ 0\rangle$ $|1,\ 0,1,\ 1\rangle$

$|1,-1,1,-1\rangle$ $|1,-1,1,\ 0\rangle$ $|1,-1,1,\ 1\rangle$

Since we have extracted one vector out of each entry, we decrement each entry by 1.

		M_S		
		-1	0	1
	2	0	1	0
	1	0	1	0
M_L	0	0	2	0
	-1	0	1	0
	-2	0	1	0

Again, we start from the basis state with the largest M_S and the corresponding largest M_L:

		M_S		
		-1	0	1
	2	0	$\boxed{1}$	0
	1	0	1	0
M_L	0	0	2	0
	-1	0	1	0
	-2	0	1	0

This gives us an eigen-vector $|2,2,0,0\rangle$. Now apply both L_- and S_- operators (in this case only L_-), we obtain

		M_S			
		-1	0	1	
	2	0	$\boxed{1}$	0	$\lvert 2, \ 2,0, \ 0\rangle$
	1	0	$\boxed{1}$	0	$\lvert 2, \ 1,0, \ 0\rangle$
M_L	0	0	$\boxed{2}$	0	$\lvert 2, \ 0,0, \ 0\rangle$
	-1	0	$\boxed{1}$	0	$\lvert 2,-1,0, \ 0\rangle$
	-2	0	$\boxed{1}$	0	$\lvert 2,-2,0, \ 0\rangle$

Now we decrement each entry by 1.

		M_S		
		-1	0	1
	2	0	0	0
	1	0	0	0
M_L	0	0	$\boxed{1}$	0
	-1	0	0	0
	-2	0	0	0

$|0,0,0,0\rangle$ is the last eigen-vector in our example. It cannot be read off directly, since there are 3 basis vectors corresponding to this M_L and M_S. But two eigen-vectors have been constructed already, namely, $|1,0,1,0\rangle$ and $|2,0,0,0\rangle$. The last one, $|0,0,0,0\rangle$, should be obtained from the orthogonality relation from the previous two eigen-vectors.

Once we applied a ladder operator L_- or S_-, we changed only the "projection" quantum number M_L or M_S. The amplitude of angular momenta L and S are never changed by ladder operators. From Eqns. (5.9) and (5.10) we can conclude that eigen-vectors with the same L and S have the same eigen-energy. In other words, the eigen-vectors obtained from the ladder operators have degenerate eigen-energies. If we review the previous example, we can summarize that there are three groups of eigen-vectors:

^3P			^1D	^1S
			$\lvert 2, \ 2,0, \ 0\rangle$	
$\lvert 1, \ 1,1,-1\rangle$	$\lvert 1, \ 1,1, \ 0\rangle$	$\lvert 1, \ 1,1, \ 1\rangle$	$\lvert 2, \ 1,0, \ 0\rangle$	
$\lvert 1, \ 0,1,-1\rangle$	$\lvert 1, \ 0,1, \ 0\rangle$	$\lvert 1, \ 0,1, \ 1\rangle$	$\lvert 2, \ 0,0, \ 0\rangle$	$\lvert 0, \ 0,0, \ 0\rangle$
$\lvert 1,-1,1,-1\rangle$	$\lvert 1,-1,1, \ 0\rangle$	$\lvert 1,-1,1, \ 1\rangle$	$\lvert 2,-1,0, \ 0\rangle$	
			$\lvert 2,-2,0, \ 0\rangle$	

Each group is called a multiplet. All vectors from the same group have the same eigen-energy (this confirmed the "$9 \times E_1$, $5 \times E_2$, $1 \times E_3$" mystery we claimed in the beginning of this section). Eigen-energies are highly degenerate, that is why "multi"-plet. The convention of naming is the term symbol:

$$^{2S+1}L \tag{5.12}$$

with

$$\text{degeneracy} = (2S+1)(2L+1) \tag{5.13}$$

There is also a convention of pronouncing the multiplets. For example, ^3P, ^1D and ^1S are pronounced "triplet pee", "singlet dee" and "singlet ess". A table for mapping the superscript $2S+1$ (called multiplicity) and their names:

$2S+1$	name
1	singlet
2	doublet
3	triplet
4	quartet
5	quintet
6	sextet
7	septet
8	octet
9	nonet

The first few symbols of L are: (oh, don't ask me why there is no "J")

L	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
	S	P	D	F	G	H	I	K	L	M	N	O	Q	R	T	U	V

Of course, all eigen-vectors should be expressed in terms of our 15 basis states. It won't tell us anything if we just have a name like $|2,1,0,0\rangle$. Those expression are obtained when we apply the ladder operators. For example, (vectors should be normalized)

$$L_- \begin{array}{|c|c|c|}\hline \bullet & & \\ \hline \bullet & & \\ \hline \end{array} = \sqrt{(1+1)(1-1+1)} \begin{array}{|c|c|c|}\hline & \bullet & \\ \hline \bullet & & \\ \hline \end{array} + \sqrt{(1+1)(1-1+1)} \begin{array}{|c|c|c|}\hline \bullet & & \\ \hline & \bullet & \\ \hline \end{array}$$

$$L_- |2,2,0,0\rangle = \sqrt{2} \begin{array}{|c|c|c|}\hline & \bullet & \\ \hline \bullet & & \\ \hline \end{array} + \sqrt{2} \begin{array}{|c|c|c|}\hline \bullet & & \\ \hline & \bullet & \\ \hline \end{array}$$

$$|2,1,0,0\rangle = \frac{1}{\sqrt{2}} \begin{array}{|c|c|c|}\hline & \bullet & \\ \hline \bullet & & \\ \hline \end{array} + \frac{1}{\sqrt{2}} \begin{array}{|c|c|c|}\hline \bullet & & \\ \hline & \bullet & \\ \hline \end{array} \tag{5.14}$$

In second quantization, Eqn (5.14) can be written as

$$|2,1,0,0\rangle = \frac{1}{\sqrt{2}} \left(c^\dagger_{1\downarrow} c^\dagger_{0\uparrow} + c^\dagger_{0\downarrow} c^\dagger_{1\uparrow} \right) |0\rangle \tag{5.15}$$

We made the convention that the order of creation operators are arranged according to the configuration:

$$\begin{array}{c} \\ \uparrow \\ \downarrow \end{array}\begin{array}{c} \begin{array}{ccc} 1 & 0 & -1 \end{array} \\ \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 4 & 5 & 6 \\ \hline \end{array} \end{array} = c_6^\dagger c_5^\dagger c_4^\dagger c_3^\dagger c_2^\dagger c_1^\dagger \left|0\right\rangle \tag{5.16}$$

Hence,

$$\begin{array}{|c|c|c|} \hline \bullet & \bullet & \bullet \\ \hline \bullet & \bullet & \bullet \\ \hline \end{array} = c_{-1\downarrow}^\dagger c_{0\downarrow}^\dagger c_{1\downarrow}^\dagger c_{-1\uparrow}^\dagger c_{0\uparrow}^\dagger c_{1\uparrow}^\dagger \left|0\right\rangle \tag{5.17}$$

In summary, our 15 eigen-vectors are listed below:

${}^3P$	$\left\|1,\ 1,1,\ 1\right\rangle$	$=$	$c_{0\uparrow}^\dagger c_{1\uparrow}^\dagger \left\|0\right\rangle$
	$\left\|1,\ 1,1,\ 0\right\rangle$	$=$	$\frac{1}{\sqrt{2}}\left(-c_{1\downarrow}^\dagger c_{0\uparrow}^\dagger + c_{0\downarrow}^\dagger c_{1\uparrow}^\dagger\right)\left\|0\right\rangle$
	$\left\|1,\ 1,1,-1\right\rangle$	$=$	$c_{0\downarrow}^\dagger c_{1\downarrow}^\dagger \left\|0\right\rangle$
	$\left\|1,\ 0,1,\ 1\right\rangle$	$=$	$c_{-1\uparrow}^\dagger c_{1\uparrow}^\dagger \left\|0\right\rangle$
	$\left\|1,\ 0,1,\ 0\right\rangle$	$=$	$\frac{1}{\sqrt{2}}\left(-c_{1\downarrow}^\dagger c_{-1\uparrow}^\dagger + c_{-1\downarrow}^\dagger c_{1\uparrow}^\dagger\right)\left\|0\right\rangle$
	$\left\|1,\ 0,1,-1\right\rangle$	$=$	$c_{-1\downarrow}^\dagger c_{1\downarrow}^\dagger \left\|0\right\rangle$
	$\left\|1,-1,1,\ 1\right\rangle$	$=$	$c_{-1\uparrow}^\dagger c_{0\uparrow}^\dagger \left\|0\right\rangle$
	$\left\|1,-1,1,\ 0\right\rangle$	$=$	$\frac{1}{\sqrt{2}}\left(-c_{0\downarrow}^\dagger c_{-1\uparrow}^\dagger + c_{-1\downarrow}^\dagger c_{0\uparrow}^\dagger\right)\left\|0\right\rangle$
	$\left\|1,-1,1,-1\right\rangle$	$=$	$c_{-1\downarrow}^\dagger c_{0\downarrow}^\dagger \left\|0\right\rangle$
${}^1D$	$\left\|2,\ 2,0,\ 0\right\rangle$	$=$	$c_{1\downarrow}^\dagger c_{1\uparrow}^\dagger \left\|0\right\rangle$
	$\left\|2,\ 1,0,\ 0\right\rangle$	$=$	$\frac{1}{\sqrt{2}}\left(c_{1\downarrow}^\dagger c_{0\uparrow}^\dagger + c_{0\downarrow}^\dagger c_{1\uparrow}^\dagger\right)\left\|0\right\rangle$
	$\left\|2,\ 0,0,\ 0\right\rangle$	$=$	$\frac{1}{\sqrt{6}}\left(c_{1\downarrow}^\dagger c_{-1\uparrow}^\dagger + 2c_{0\downarrow}^\dagger c_{0\uparrow}^\dagger + c_{-1\downarrow}^\dagger c_{1\uparrow}^\dagger\right)\left\|0\right\rangle$
	$\left\|2,-1,0,\ 0\right\rangle$	$=$	$\frac{1}{\sqrt{2}}\left(c_{0\downarrow}^\dagger c_{-1\uparrow}^\dagger + c_{-1\downarrow}^\dagger c_{0\uparrow}^\dagger\right)\left\|0\right\rangle$
	$\left\|2,-2,0,\ 0\right\rangle$	$=$	$c_{-1\downarrow}^\dagger c_{-1\uparrow}^\dagger \left\|0\right\rangle$
${}^1S$	$\left\|0,\ 0,0,\ 0\right\rangle$	$=$	$\frac{1}{\sqrt{3}}\left(c_{1\downarrow}^\dagger c_{-1\uparrow}^\dagger - c_{0\downarrow}^\dagger c_{0\uparrow}^\dagger + c_{-1\downarrow}^\dagger c_{1\uparrow}^\dagger\right)\left\|0\right\rangle$

As a remark, we started from the basis state with the largest M_S and the corresponding largest M_L. Equivalently, we could also start from the largest M_L and the corresponding largest M_S, since they are both eigen-states. Actually, it is just a matter of taste that we start from either

$$\underset{S_{\max}}{\begin{array}{|c|c|c|} \hline \bullet & \bullet & \\ \hline & & \\ \hline \end{array}} \quad \text{or} \quad \underset{L_{\max}}{\begin{array}{|c|c|c|} \hline \bullet & & \\ \hline \bullet & & \\ \hline \end{array}} \quad .$$

Here, we prefer to start from the former one (maximum spin), which is naturally suggested from the famous Hund's rule. It states that the multiplet with the maximum S has the lowest energy, from which we start our multiplet construction.

5.3. Seniority

We are so far lucky enough to diagonalize the Hamiltonian using ladder operators. Recall how we constructed the multiplets for a p^2 orbital. We always start from the largest M_S and the corresponding largest M_L:

$$\begin{array}{ccccc} 0 & 1 & 0 & & 0 & \boxed{1} & 0 & & 0 & 0 & 0 \\ 1 & 2 & \boxed{1} & & 0 & 1 & 0 & & 0 & 0 & 0 \\ 1 & 3 & 1 & \rightarrow & 0 & 2 & 0 & \rightarrow & 0 & \boxed{1} & 0 \\ 1 & 2 & 1 & & 0 & 1 & 0 & & 0 & 0 & 0 \\ 0 & 1 & 0 & & 0 & 1 & 0 & & 0 & 0 & 0 \end{array}$$

We were lucky because every time we started from a "1". This single vector can be determined uniquely (by the basis vector itself or by the orthogonality relation). But are we guaranteed to always start from a "1"? No, we are not. For example, the M_L-M_S tables for the d^3 orbital:

$$\begin{array}{cccc} 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 0 \\ 1 & 4 & 4 & \boxed{1} \\ 1 & 6 & 6 & 1 \\ 2 & 8 & 8 & 2 \\ 2 & 8 & 8 & 2 \\ 2 & 8 & 8 & 2 \\ 1 & 6 & 6 & 1 \\ 1 & 4 & 4 & 1 \\ 0 & 2 & 2 & 0 \\ 0 & 1 & 1 & 0 \end{array} \rightarrow \begin{array}{cccc} 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 5 & 5 & 0 \\ 1 & 7 & 7 & \boxed{1} \\ 1 & 7 & 7 & 1 \\ 1 & 7 & 7 & 1 \\ 0 & 5 & 5 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 1 & 1 & 0 \end{array} \rightarrow \begin{array}{cccc} 0 & 1 & \boxed{1} & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 5 & 5 & 0 \\ 0 & 6 & 6 & 0 \\ 0 & 6 & 6 & 0 \\ 0 & 6 & 6 & 0 \\ 0 & 5 & 5 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 1 & 1 & 0 \end{array} \rightarrow \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 1 & \boxed{1} & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 4 & 4 & 0 \\ 0 & 5 & 5 & 0 \\ 0 & 5 & 5 & 0 \\ 0 & 5 & 5 & 0 \\ 0 & 4 & 4 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \rightarrow \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & \boxed{1} & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 4 & 4 & 0 \\ 0 & 4 & 4 & 0 \\ 0 & 4 & 4 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \rightarrow \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & \boxed{2} & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

We encounter a "2", disaster! This "2" means there are two undetermined eigen-vectors with $M_L = 2$ and $M_S = 1/2$. By the orthogonality relation, we can at most find out an eigen-space (a plane) which is spanned by the two eigen-vectors. But we cannot tell where exactly those two eigen-vectors are. That is the limitation of using ladder operators. After all, we were doing something peculiar: we diagonalized the matrix without using the matrix elements but only symmetries. At this stage, we really have to ask help from the matrix elements. With the given numerical values, we will be able to diagonalize the matrix completely. But as we mentioned already, although we cannot obtain the eigen-vectors, we can find the eigen-space spanned by these two eigen-vectors. To continue our work, we can generate two random vectors and by

applying the orthogonality relation, we get two random vectors in the eigen-space. For example,

$$|2,2,\frac{1}{2},\frac{1}{2}\rangle_1 = \Big(0.211c^\dagger_{2\downarrow}c^\dagger_{-1\uparrow}c^\dagger_{1\uparrow} - 0.259c^\dagger_{1\downarrow}c^\dagger_{0\uparrow}c^\dagger_{1\uparrow} - 0.702c^\dagger_{2\downarrow}c^\dagger_{-2\uparrow}c^\dagger_{2\uparrow} + 0.490c^\dagger_{1\downarrow}c^\dagger_{-1\uparrow}c^\dagger_{2\uparrow} - 0.279c^\dagger_{0\downarrow}c^\dagger_{0\uparrow}c^\dagger_{2\uparrow} + 0.279c^\dagger_{-1\downarrow}c^\dagger_{1\uparrow}c^\dagger_{2\uparrow}\Big)|0\rangle \tag{5.18}$$

$$|2,2,\frac{1}{2},\frac{1}{2}\rangle_2 = \Big(0.382c^\dagger_{2\downarrow}c^\dagger_{-1\uparrow}c^\dagger_{1\uparrow} - 0.468c^\dagger_{1\downarrow}c^\dagger_{0\uparrow}c^\dagger_{1\uparrow} - 0.236c^\dagger_{2\downarrow}c^\dagger_{-2\uparrow}c^\dagger_{2\uparrow} - 0.146c^\dagger_{1\downarrow}c^\dagger_{-1\uparrow}c^\dagger_{2\uparrow} + 0.528c^\dagger_{0\downarrow}c^\dagger_{0\uparrow}c^\dagger_{2\uparrow} - 0.528c^\dagger_{-1\downarrow}c^\dagger_{1\uparrow}c^\dagger_{2\uparrow}\Big)|0\rangle \tag{5.19}$$

Now apply both L_- and S_- to these two vectors, we obtain all other vectors in this (2-dimensional) multiplet:

$$\begin{array}{cccc} 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&\boxed{2}&\boxed{2}&0\\ 0&\boxed{3}&\boxed{3}&0\\ 0&\boxed{3}&\boxed{3}&0\\ 0&\boxed{3}&\boxed{3}&0\\ 0&\boxed{2}&\boxed{2}&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0 \end{array} \quad \rightarrow \quad \begin{array}{cccc} 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&1&1&0\\ 0&1&1&0\\ 0&1&1&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0 \end{array}$$

This step gives us two ${}^2D$ multiplets. We will denote it as $\overset{2\times}{{}^2D}$. Once we encounter a multiplet which appears multiple times, it implies we could only find its eigen-spaces but not eigen-vectors by using the ladder operator technique. We don't need to worry about this problem for s and p shells. But for d and higher shells, we will encounter this problem quite often. A summary of atomic multiplets for open s, p, d and f shells are tabulated in 5.1.

Recall the random vectors we generated in Eqns. (5.18) and (5.19). If we repeat the generation of random vectors, we might obtain a different set of vectors. They span the same space but the vectors will have different directions. There is a very smart method which allows us to construct those vectors uniquely (also with beautiful coefficients like $1/\sqrt{3}$), although they are still not eigen-vectors.

This new concept is called seniority [13]. The idea is that since we cannot uniquely determine the vectors in the eigen-spaces of $\overset{2\times}{{}^2D}$ in d^3 shell, we start from the ${}^2D$ in d^1 shell whose eigen-vectors can be uniquely determined.

d^1			$\boxed{{}^2D}$				
d^2	${}^1S$	${}^3P$	${}^1D$	${}^3F$		${}^1G$	
d^3		${}^2P\,{}^4P$	$\boxed{\overset{2\times}{{}^2D}}$	${}^2F\,{}^4F$		${}^2G$	${}^2H$

The d^1 and d^3 shells are different by, of course, two electrons. To go from d^1 to d^3, we need to add two more electrons. But the additional electrons should not affect the angular momenta in

^2D. Therefore, we should add electrons according to the ^1S multiplet in d^2, which has $L = 0$ and $S = 0$. This put-electron operation is called a seniority operation, one can think the ^2D in d^1 as the parent and the $\overset{2\times}{^2D}$ in d^3 as the children. The ^1S multiplet has one eigen-state

$$|0,0,0,0\rangle = \frac{1}{\sqrt{5}}\left(c^\dagger_{2\downarrow}c^\dagger_{-2\uparrow} - c^\dagger_{1\downarrow}c^\dagger_{-1\uparrow} + c^\dagger_{0\downarrow}c^\dagger_{0\uparrow} - c^\dagger_{-1\downarrow}c^\dagger_{1\uparrow} + c^\dagger_{-2\downarrow}c^\dagger_{2\uparrow}\right)|0\rangle \tag{5.20}$$

The "put-electron-operator" (seniority operator) is simply

$$T = c^\dagger_{2\downarrow}c^\dagger_{-2\uparrow} - c^\dagger_{1\downarrow}c^\dagger_{-1\uparrow} + c^\dagger_{0\downarrow}c^\dagger_{0\uparrow} - c^\dagger_{-1\downarrow}c^\dagger_{1\uparrow} + c^\dagger_{-2\downarrow}c^\dagger_{2\uparrow} \tag{5.21}$$

Apply this seniority operator to the leading vector of ^2D in d^1:

$$\begin{aligned} T\,|2,2,\frac{1}{2},\frac{1}{2}\rangle &= T\,c^\dagger_{2\uparrow}|0\rangle \\ &= \left(c^\dagger_{2\downarrow}c^\dagger_{-2\uparrow} - c^\dagger_{1\downarrow}c^\dagger_{-1\uparrow} + c^\dagger_{0\downarrow}c^\dagger_{0\uparrow} - c^\dagger_{-1\downarrow}c^\dagger_{1\uparrow} + c^\dagger_{-2\downarrow}c^\dagger_{2\uparrow}\right)c^\dagger_{2\uparrow}|0\rangle \\ &= \left(c^\dagger_{2\downarrow}c^\dagger_{-2\uparrow}c^\dagger_{2\uparrow} - c^\dagger_{1\downarrow}c^\dagger_{-1\uparrow}c^\dagger_{2\uparrow} + c^\dagger_{0\downarrow}c^\dagger_{0\uparrow}c^\dagger_{2\uparrow} - c^\dagger_{-1\downarrow}c^\dagger_{1\uparrow}c^\dagger_{2\uparrow}\right)|0\rangle \end{aligned} \tag{5.22}$$

We obtain the first leading vector of ^2D in d^3: (vectors should be normalized)

$$|2,2,\frac{1}{2},\frac{1}{2},0\rangle = \frac{1}{\sqrt{4}}\left(c^\dagger_{2\downarrow}c^\dagger_{-2\uparrow}c^\dagger_{2\uparrow} - c^\dagger_{1\downarrow}c^\dagger_{-1\uparrow}c^\dagger_{2\uparrow} + c^\dagger_{0\downarrow}c^\dagger_{0\uparrow}c^\dagger_{2\uparrow} - c^\dagger_{-1\downarrow}c^\dagger_{1\uparrow}c^\dagger_{2\uparrow}\right)|0\rangle \tag{5.23}$$

The fifth index $W = 0$ is called the seniority number. One can think it as the index of an array.[4] The remaining leading vector $|2,2,\frac{1}{2},\frac{1}{2},1\rangle$ can then be uniquely determined from the orthogonality relation. in shell d^3.

$$\begin{aligned} |2,2,\frac{1}{2},\frac{1}{2},1\rangle = \frac{1}{\sqrt{84}}\Big(&4c^\dagger_{2\downarrow}c^\dagger_{-1\uparrow}c^\dagger_{1\uparrow} - \sqrt{24}c^\dagger_{1\downarrow}c^\dagger_{0\uparrow}c^\dagger_{1\uparrow} - 5c^\dagger_{2\downarrow}c^\dagger_{-2\uparrow}c^\dagger_{2\uparrow} \\ &+c^\dagger_{1\downarrow}c^\dagger_{-1\uparrow}c^\dagger_{2\uparrow} + 3c^\dagger_{0\downarrow}c^\dagger_{0\uparrow}c^\dagger_{2\uparrow} - 3c^\dagger_{-1\downarrow}c^\dagger_{1\uparrow}c^\dagger_{2\uparrow}\Big)|0\rangle \end{aligned} \tag{5.24}$$

The vectors in Eqns. (5.23) and (5.24) span the same space as the vectors in Eqns. (5.18) and (5.19). Now those vectors can be uniquely determined and addressed by five quantum numbers $|L, M_L, S, M_S, W\rangle$. By applying ladder operators to these two leading vectors, we obtain the complete set of vectors in the multiplet $\overset{2\times}{^2D}$. One must be alerted that $|L, M_L, S, M_S, W\rangle$ is, in general, not an eigen-vector. But the complete set $|L, M_L, S, M_S, 0\rangle \ldots |L, M_L, S, M_S, n-1\rangle$ spans the complete eigen-space of multiplets $\overset{n\times}{^{2S+1}L}$.

[4] A more standard way to assign the seniority number is to name according to their parents. For example, our first vector is produced from ^2D in d^1, hence it has a seniority number 1. The second vector is produced in d^3 from the orthogonality relation, hence a seniority number 3. However, this ancestry relation gets much more complicated for an f shell. We would like to assign the seniority number just as the array index. Actually it doesn't matter how we name the seniority number since it does not really have a physical meaning. We chose it for convenience as long as it can label a vector uniquely.

Table 5.1.: Atomic multiplets for open s, p, d and f shells. All closed shells have ^1S.

Orbital	Multiplets												
s^1	^2S												
p^1 p^5		^2P											
p^2 p^4	^1S	^3P	^1D										
p^3	^4S	^2P	^2D										
d^1 d^9			^2D										
d^2 d^8	^1S	^3P	^1D	^3F	^1G								
d^3 d^7		$^2P\,^4P$	$^{2\times}{}^2D$	$^2F\,^4F$	^2G	^2H							
d^4 d^6	$^{2\times}{}^1S$	$^{2\times}{}^3P$	$^{2\times}{}^1D\,^3D\,^5D$	$^{2\times}{}^1F\,^3F$	$^{2\times}{}^1G\,^3G$	^3H	^1I						
d^5	$^2S\,^6S$	$^2P\,^4P$	$^{3\times}{}^2D\,^4D$	$^{2\times}{}^2F\,^4F$	$^{2\times}{}^2G\,^4G$	^2H	^2I						
f^1 f^{13}				^2F									
f^2 f^{12}	^1S	^3P	^1D	^3F	^1G	^3H	^1I						
f^3 f^{11}	^4S	^2P	$^{2\times}{}^2D\,^4D$	$^{2\times}{}^2F\,^4F$	$^{2\times}{}^2G\,^4G$	$^{2\times}{}^2H$	$^2I\,^4I$	^2K	^2L				
f^4 f^{10}	$^{2\times}{}^1S\,^5S$	$^{3\times}{}^3P$	$^{4\times\ 2\times}{}^1D\,^3D\,^5D$	$^{4\times}{}^1F\,^3F\,^5F$	$^{4\times\ 3\times}{}^1G\,^3G\,^5G$	$^{2\times\ 4\times}{}^1H\,^3H$	$^{3\times\ 2\times}{}^1I\,^3I\,^5I$	$^{2\times}{}^1K\,^3K$	$^{2\times}{}^1L\,^3L$	^3M	^1N		
f^5 f^9	^4S	$^{4\times\ 2\times}{}^2P\,^4P\,^6P$	$^{5\times\ 3\times}{}^2D\,^4D$	$^{7\times\ 4\times}{}^2F\,^4F\,^6F$	$^{6\times\ 4\times}{}^2G\,^4G$	$^{7\times\ 3\times}{}^2H\,^4H\,^6H$	$^{5\times\ 3\times}{}^2I\,^4I$	$^{5\times\ 2\times}{}^2K\,^4K$	$^{3\times}{}^2L\,^4L$	$^{2\times}{}^2M\,^4M$	^2N	^2O	
f^6 f^8	$^{4\times}{}^1S\,^5S$	$^{6\times}{}^1P\,^3P\,^5P$	$^{6\times\ 5\times\ 3\times}{}^1D\,^3D\,^5D$	$^{4\times\ 9\times\ 2\times}{}^1F\,^3F\,^5F\,^7F$	$^{8\times\ 7\times\ 3\times}{}^1G\,^3G\,^5G$	$^{4\times\ 9\times\ 2\times}{}^1H\,^3H\,^5H$	$^{7\times\ 6\times\ 2\times}{}^1I\,^3I\,^5I$	$^{3\times\ 6\times}{}^1K\,^3K\,^5K$	$^{4\times\ 3\times}{}^1L\,^3L\,^5L$	$^{2\times\ 3\times}{}^1M\,^3M$	$^{2\times}{}^1N\,^3N$	^3O	^1Q
f^7	$^{2\times\ 2\times}{}^2S\,^4S\,^8S$	$^{5\times\ 2\times}{}^2P\,^4P\,^6P$	$^{7\times\ 6\times}{}^2D\,^4D\,^6D$	$^{10\times\ 5\times}{}^2F\,^4F\,^6F$	$^{10\times\ 7\times}{}^2G\,^4G\,^6G$	$^{9\times\ 5\times}{}^2H\,^4H\,^6H$	$^{9\times\ 5\times}{}^2I\,^4I\,^6I$	$^{7\times\ 3\times}{}^2K\,^4K$	$^{5\times\ 3\times}{}^2L\,^4L$	$^{4\times}{}^2M\,^4M$	$^{2\times}{}^2N\,^4N$	^2O	^2Q

5.4. Eigen-energy of multiplet states

As we emphasized earlier, the eigen-energies of the Hamiltonian could be obtained by throwing the matrix into a diagonalization solver. However, from this "brute-force" approach, it cannot be seen that which eigen-energy corresponds to which multiplet. Now, since we have constructed the eigen-vectors already (up to seniority), we can easily compute the eigen-energies by matrix-vector multiplications. For a given eigen-vector $|\mathbf{v}_n\rangle$, the corresponding eigen-energy is simply

$$E_n = \langle \mathbf{v}_n | \, H_U \, | \mathbf{v}_n \rangle \tag{5.25}$$

The Hamiltonian consists of two major parts: the Slater-Condon parameters and the Gaunt coefficients. For on-shell interactions, the Slater-Condon parameters enter as pre-factors $F^{(k)}$ since they depend on only n and l (see Eqn. (4.28)). It is more convenient to take out those pre-factors and to write our Hamiltonian as,

$$H_U = \sum_{k=0(+2)}^{2l} F^{(k)} \tilde{H}_U^{(k)} \tag{5.26}$$

where $\tilde{H}_U^{(k)}$ has only a dependence on the Gaunt coefficients, which are system independent. For a given shell, the sum rule in Eqn. (4.34) requires $k = 0, 2, \ldots, 2l$. The routine for setting up $\tilde{H}_U^{(k)}$ is identical to Algorithm 5.2, except that one needs to replace $U_{\alpha\beta\gamma\delta}$ by $\delta_{\sigma_1\sigma_4}\delta_{\sigma_2\sigma_3}\frac{4\pi}{2k+1}A^{(k)}(lm_1, lm_2, lm_3, lm_4)$. The eigenvalues $\tilde{E}_n^{(k)}$ from $\tilde{H}_U^{(k)}$ are universal for all atoms. We can easily compute $\tilde{E}_n^{(k)}$ by applying our eigen-vectors to $\tilde{H}_U^{(k)}$:

$$\tilde{E}_n^{(k)} = \langle \mathbf{v}_n | \, \tilde{H}_U^{(k)} \, | \mathbf{v}_n \rangle \tag{5.27}$$

It is not trivial that the eigen-vectors of H_U are simultaneously eigen-vectors of $\tilde{H}_U^{(k)}$. This is because $\tilde{H}_U^{(k)}$ also commute with $\mathbf{L}$ and $\mathbf{S}$ (up to seniority). Now, system independently, we can write the eigen-energies as (think $F^{(k)}$ as pre-factors)

$$E_n = \sum_{k=0(+2)}^{2l} F^{(k)} \tilde{E}_n^{(k)} \tag{5.28}$$

Those pre-factors $F^{(k)}$ are subject to the input wave functions which are determined by different atoms, electronic configurations and approximation methods (e.g. self-consistent field approximation). Table 5.2 summarizes the multiplet eigen-energies for a few selected atomic shells. The terms are sorted in the descending order of multiplicity $2S + 1$. For a given multiplicity, terms are sorted in the descending order of L. The first term of each individual table is the multiplet with the lowest energy (Hund's rule).

To get a direct impression of how Table 5.2 works, we take the p^2 configuration as an example. A carbon atom with electronic configuration $1s^2\ 2s^2\ 2p^2$ would be a good candidate for this demonstration. From the self-consistent calculation, we obtained the eigen-energy of the $2p$ orbital to be -0.199186 Hartree (see Table 3.1). From the resulting wave function u_{2p}, we can easily compute the Slater-Condon parameters using Eqn. (4.28),

$$\begin{aligned} F^{(0)}(2p) &= 0.520216 \text{ (Hartree)} \\ F^{(2)}(2p) &= 0.229662 \text{ (Hartree)} \end{aligned} \tag{5.29}$$

Putting (5.29) into Table 5.2, we obtain the multiplet energies,

$$\begin{aligned} {}^3P &: 0.474284 \text{ (Hartree)} \\ {}^1D &: 0.529402 \text{ (Hartree)} \\ {}^1S &: 0.612081 \text{ (Hartree)} \end{aligned} \tag{5.30}$$

These are the three split eigen-energies of the $2p^2$ configuration.[5] It is important to understand that they are the eigen-energies from the Coulomb repulsion Hamiltonian

$$H_U = \sum_{i<j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{5.31}$$

But not the full Hamiltonian

$$H = \sum_{i=1}^{N} \left[-\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i} \right] + \sum_{i<j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{5.32}$$

The eigen-energies in (5.30) are not the absolute energies. They are correct up to a constant shift. Nevertheless, their energy splittings are normally the quantities that we are interested in.

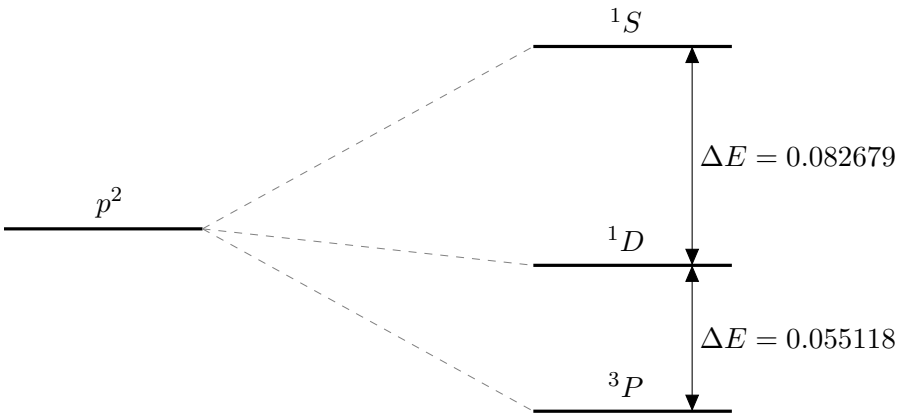


Figure 5.1.: Energy splitting of the p^2 configuration of a carbon atom. Energies are given in units of Hartree (a.u.).

We should keep in mind that we didn't obtain all eigen-vectors from the ladder operator technique. For the case where the seniority numbers are required, we can at most construct a small eigen-space that the eigen-vectors live in. From this eigen-space, namely, a set of vectors $|\mathbf{v}_i\rangle$, we compute the matrix

$$\tilde{E}_{ij}^{(k)} = \langle \mathbf{v}_i | \, \tilde{H}_U^{(k)} \, | \mathbf{v}_j \rangle \tag{5.33}$$

which contains the eigen-energies. Similar to Table 5.2, we also provide the eigen-energies of multiplets with seniority in Table 5.3. But instead of numbers being coefficients, we have small matrices as the coefficients. For a given set of $F^{(k)}$, we can sum up the small matrices and diagonalize to obtain the eigen-energies numerically.

[5] We limited our basis within the open shell. This gives reasonably good approximations since multiplet splittings are normally much smaller than the energy differences among electronic shells (first order perturbation theory).

Table 5.2.: Multiplet (without seniority cases) eigen-energies for a few selected atomic shells.

p^2	${}^3P$	$F^{(0)}$	$-$	$\frac{1}{5}F^{(2)}$		
	${}^1D$	$F^{(0)}$	$+$	$\frac{1}{25}F^{(2)}$		
	${}^1S$	$F^{(0)}$	$+$	$\frac{2}{5}F^{(2)}$		

p^3	${}^4S$	$3F^{(0)}$	$-$	$\frac{3}{5}F^{(2)}$		
	${}^2D$	$3F^{(0)}$	$-$	$\frac{6}{25}F^{(2)}$		
	${}^2P$	$3F^{(0)}$				

d^2	${}^3F$	$F^{(0)}$	$-$	$\frac{8}{49}F^{(2)}$	$-$	$\frac{1}{49}F^{(4)}$
	${}^3P$	$F^{(0)}$	$+$	$\frac{1}{7}F^{(2)}$	$-$	$\frac{4}{21}F^{(4)}$
	${}^1G$	$F^{(0)}$	$+$	$\frac{4}{49}F^{(2)}$	$+$	$\frac{1}{441}F^{(4)}$
	${}^1D$	$F^{(0)}$	$-$	$\frac{3}{49}F^{(2)}$	$+$	$\frac{4}{49}F^{(4)}$
	${}^1S$	$F^{(0)}$	$+$	$\frac{2}{7}F^{(2)}$	$+$	$\frac{2}{7}F^{(4)}$

d^3	${}^4F$	$3F^{(0)}$	$-$	$\frac{15}{49}F^{(2)}$	$-$	$\frac{8}{49}F^{(4)}$
	${}^4P$	$3F^{(0)}$			$-$	$\frac{1}{3}F^{(4)}$
	${}^2H$	$3F^{(0)}$	$-$	$\frac{6}{49}F^{(2)}$	$-$	$\frac{4}{147}F^{(4)}$
	${}^2G$	$3F^{(0)}$	$-$	$\frac{11}{49}F^{(2)}$	$+$	$\frac{13}{441}F^{(4)}$
	${}^2F$	$3F^{(0)}$	$+$	$\frac{9}{49}F^{(2)}$	$-$	$\frac{29}{147}F^{(4)}$
	${}^2P$	$3F^{(0)}$	$-$	$\frac{6}{49}F^{(2)}$	$-$	$\frac{4}{147}F^{(4)}$

d^4	${}^5D$	$6F^{(0)}$	$-$	$\frac{3}{7}F^{(2)}$	$-$	$\frac{3}{7}F^{(4)}$
	${}^3H$	$6F^{(0)}$	$-$	$\frac{17}{49}F^{(2)}$	$-$	$\frac{23}{147}F^{(4)}$
	${}^3G$	$6F^{(0)}$	$-$	$\frac{12}{49}F^{(2)}$	$-$	$\frac{94}{441}F^{(4)}$
	${}^3D$	$6F^{(0)}$	$-$	$\frac{5}{49}F^{(2)}$	$-$	$\frac{43}{147}F^{(4)}$
	${}^1I$	$6F^{(0)}$	$-$	$\frac{15}{49}F^{(2)}$	$-$	$\frac{1}{49}F^{(4)}$
	${}^1F$	$6F^{(0)}$			$-$	$\frac{4}{21}F^{(4)}$

d^5	${}^6S$	$10F^{(0)}$	$-$	$\frac{5}{7}F^{(2)}$	$-$	$\frac{5}{7}F^{(4)}$
	${}^4G$	$10F^{(0)}$	$-$	$\frac{25}{49}F^{(2)}$	$-$	$\frac{190}{441}F^{(4)}$
	${}^4F$	$10F^{(0)}$	$-$	$\frac{13}{49}F^{(2)}$	$-$	$\frac{20}{49}F^{(4)}$
	${}^4D$	$10F^{(0)}$	$-$	$\frac{18}{49}F^{(2)}$	$-$	$\frac{25}{49}F^{(4)}$
	${}^4P$	$10F^{(0)}$	$-$	$\frac{4}{7}F^{(2)}$	$-$	$\frac{5}{21}F^{(4)}$
	${}^2I$	$10F^{(0)}$	$-$	$\frac{24}{49}F^{(2)}$	$-$	$\frac{10}{49}F^{(4)}$
	${}^2H$	$10F^{(0)}$	$-$	$\frac{22}{49}F^{(2)}$	$-$	$\frac{10}{147}F^{(4)}$
	${}^2P$	$10F^{(0)}$	$+$	$\frac{20}{49}F^{(2)}$	$-$	$\frac{80}{147}F^{(4)}$
	${}^2S$	$10F^{(0)}$	$-$	$\frac{3}{49}F^{(2)}$	$-$	$\frac{65}{147}F^{(4)}$

Table 5.3.: Multiplet (seniority cases) eigen-energies for a few selected atomic shells.

d^3	${}^{2\times}_{2}D$	$\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$	$F^{(0)}$	$-$	$\begin{bmatrix} -\frac{1}{7} & \frac{3\sqrt{21}}{49} \\ \frac{3\sqrt{21}}{49} & -\frac{3}{49} \end{bmatrix}$	$F^{(2)}$	$+$	$\begin{bmatrix} \frac{1}{7} & \frac{5\sqrt{21}}{147} \\ \frac{5\sqrt{21}}{147} & -\frac{19}{147} \end{bmatrix}$	$F^{(4)}$	
d^4	${}^{2\times}_{3}F$	$\begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix}$	$F^{(0)}$	$-$	$\begin{bmatrix} \frac{2}{49} & \frac{12}{49} \\ \frac{12}{49} & \frac{8}{49} \end{bmatrix}$	$F^{(2)}$	$+$	$\begin{bmatrix} -\frac{13}{147} & \frac{20}{147} \\ \frac{20}{147} & -\frac{38}{147} \end{bmatrix}$	$F^{(4)}$	
	${}^{2\times}_{3}P$	$\begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix}$	$F^{(0)}$	$+$	$\begin{bmatrix} -\frac{1}{7} & \frac{4\sqrt{14}}{49} \\ \frac{4\sqrt{14}}{49} & -\frac{3}{49} \end{bmatrix}$	$F^{(2)}$	$-$	$\begin{bmatrix} \frac{2}{63} & \frac{20\sqrt{14}}{441} \\ \frac{20\sqrt{14}}{441} & \frac{139}{441} \end{bmatrix}$	$F^{(4)}$	
	${}^{2\times}_{1}G$	$\begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix}$	$F^{(0)}$	$-$	$\begin{bmatrix} \frac{6}{49} & \frac{4\sqrt{11}}{49} \\ \frac{4\sqrt{11}}{49} & \frac{4}{49} \end{bmatrix}$	$F^{(2)}$	$+$	$\begin{bmatrix} \frac{17}{147} & \frac{20\sqrt{11}}{441} \\ \frac{20\sqrt{11}}{441} & -\frac{64}{441} \end{bmatrix}$	$F^{(4)}$	
	${}^{2\times}_{1}D$	$\begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix}$	$F^{(0)}$	$+$	$\begin{bmatrix} \frac{15}{49} & \frac{12\sqrt{2}}{49} \\ \frac{12\sqrt{2}}{49} & \frac{3}{49} \end{bmatrix}$	$F^{(2)}$	$-$	$\begin{bmatrix} \frac{6}{49} & \frac{20\sqrt{2}}{147} \\ \frac{20\sqrt{2}}{147} & \frac{11}{49} \end{bmatrix}$	$F^{(4)}$	
	${}^{2\times}_{1}S$	$\begin{bmatrix} 6 & 0 \\ 0 & 6 \end{bmatrix}$	$F^{(0)}$	$-$	$\begin{bmatrix} -\frac{2}{7} & \frac{6\sqrt{21}}{49} \\ \frac{6\sqrt{21}}{49} & -\frac{6}{49} \end{bmatrix}$	$F^{(2)}$	$+$	$\begin{bmatrix} \frac{2}{7} & \frac{10\sqrt{21}}{147} \\ \frac{10\sqrt{21}}{147} & -\frac{38}{147} \end{bmatrix}$	$F^{(4)}$	
d^5	${}^{2\times}_{2}G$	$\begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$	$F^{(0)}$	$+$	$\begin{bmatrix} \frac{3}{49} & 0 \\ 0 & -\frac{13}{49} \end{bmatrix}$	$F^{(2)}$	$-$	$\begin{bmatrix} \frac{155}{441} & 0 \\ 0 & \frac{145}{441} \end{bmatrix}$	$F^{(4)}$	
	${}^{2\times}_{2}F$	$\begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$	$F^{(0)}$	$-$	$\begin{bmatrix} \frac{25}{49} & 0 \\ 0 & \frac{9}{49} \end{bmatrix}$	$F^{(2)}$	$-$	$\begin{bmatrix} \frac{5}{147} & 0 \\ 0 & \frac{55}{147} \end{bmatrix}$	$F^{(4)}$	
d^5	${}^{3\times}_{2}D$	$\begin{bmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \end{bmatrix}$	$F^{(0)}$	$-$	$\begin{bmatrix} 0 & 0 & \frac{6\sqrt{14}}{49} \\ 0 & \frac{4}{49} & 0 \\ \frac{6\sqrt{14}}{49} & 0 & \frac{6}{49} \end{bmatrix}$	$F^{(2)}$	$+$	$\begin{bmatrix} 0 & 0 & \frac{10\sqrt{14}}{147} \\ 0 & -\frac{40}{147} & 0 \\ \frac{10\sqrt{14}}{147} & 0 & -\frac{20}{49} \end{bmatrix}$	$F^{(4)}$	

Chapter 6

Spin-orbit coupling

6.1. The spin-orbit interaction

In the very beginning of our discussion, we claimed that our problem is to solve the Schrödinger equation with the Hamiltonian

$$H = \sum_{i=1}^{N} \left[-\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i} \right] + \sum_{i<j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{6.1}$$

We see clearly that each electron experiences a Coulomb attraction from the nucleus and repulsions from all other electrons. However, this is not quite the complete story. In fact, additionally, each electron also experiences a (weak) magnetic force. Do you see where this magnetic force comes from?

Imagine you "sit" on an electron. From your point of view, the positively charged nucleus is circling around you. This moving charge creates a current loop which generates a magnetic field $\mathbf{B}$ (can be calculated from the Biot-Savart law). On the other hand, the spinning electron has a magnetic dipole moment $\boldsymbol{\mu}_e$ which experiences the force from the magnetic field

$$\mathbf{F} = \boldsymbol{\nabla}(\boldsymbol{\mu}_e \cdot \mathbf{B}) \tag{6.2}$$

The corresponding potential energy of the electron under this magnetic field is

$$H_{\mathrm{SO}} = -\boldsymbol{\mu}_e \cdot \mathbf{B} \tag{6.3}$$

It would be more convenient if we express the magnetic dipole moment in terms of the spin angular momentum $\mathbf{S}$ and the magnetic field in terms of the orbit angular momentum $\mathbf{L}$ [6],

$$H_{\mathrm{SO}} = \xi(r)\mathbf{L} \cdot \mathbf{S} \tag{6.4}$$

where,[1]

$$\xi(r) = \frac{1}{2m_e^2 c^2} \frac{1}{r} \frac{dV}{dr} \tag{6.5}$$

[1]Some authors include an "$\hbar^2$" in the expression of $\xi(r)$. But we prefer to absorb this $\hbar^2$ into the angular momenta, since $\mathbf{L} = \mathbf{r} \times \mathbf{p} = \mathbf{r} \times (-i\hbar)\boldsymbol{\nabla}$ and $\mathbf{L} \cdot \mathbf{S}$ includes the $\hbar^2$ implicitly. Notice that in atomic units $\hbar = 1$.

and $V(r)$ is our old friend, the (spherically symmetric) electric potential. For hydrogen-like atoms,

$$V(r) = -\frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r} \quad \text{and} \quad \xi(r) = \frac{1}{2m_e^2c^2}\frac{1}{4\pi\epsilon_0}\frac{Ze^2}{r^3} \tag{6.6}$$

which agrees with the hydrogen atom spin-orbit interaction discussed in Griffiths' book [1]. Now consider an N-electron system, Eqn. (6.4) becomes

$$\boxed{H_{\text{SO}} = \sum_{i=1}^{N} \xi(r_i)\boldsymbol{\ell}_i \cdot \mathbf{s}_i} \tag{6.7}$$

We reserved the capital letters for the total angular momentum operators

$$\mathbf{L} = \sum_{i=1}^{N} \boldsymbol{\ell}_i \quad \text{and} \quad \mathbf{S} = \sum_{i=1}^{N} \mathbf{s}_i \tag{6.8}$$

This is the additional spin-orbit interaction Hamiltonian. It is rather a tiny perturbation. In Eqn. (6.5), we see a "speed of light squared" factor in the denominator. Hence, spin-orbit is in general a weak interaction. We introduced the atomic units (a.u.) in the beginning and this is the first time that we encounter the speed of light. We know that in SI units, the speed of light is

$$c = 2.99792458 \times 10^8 \text{ m/s} \tag{6.9}$$

To convert the speed of light from SI to atomic unit, we need the length and time in atomic units:

$$1 \text{ a}_0 \approx 5.2918 \times 10^{-11} \text{ m} \tag{6.10}$$

$$1 \text{ t}_0 \approx 2.4189 \times 10^{-17} \text{ s} \tag{6.11}$$

We can convert easily, in a.u.[2]

$$c \approx 137.036 \text{ a}_0/\text{t}_0 \tag{6.12}$$

Now we append H_{SO} to Hamiltonian (6.1),

$$H = \sum_{i=1}^{N}\left[-\frac{1}{2}\nabla_i^2 - \frac{Z}{r_i}\right] + \sum_{i<j}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i=1}^{N}\xi(r_i)\boldsymbol{\ell}_i \cdot \mathbf{s}_i \tag{6.13}$$

which is the full Hamiltonian with spin-orbit coupling correction.

6.2. Spin-orbit coupling within multiplet terms

If there were no spin-orbit coupling, our Hamiltonian (6.1) commuted with operators $\mathbf{L}$ and $\mathbf{S}$, which indicates that these quantities are conserved [1]. That is why we could label our eigen-vectors as

$$|L, M_L, S, M_S\rangle \tag{6.14}$$

[2]The inverse of this number is called the fine structure constant $\alpha \equiv \frac{e^2}{4\pi\epsilon_0\hbar c} \approx 1/137.036$, which is a dimensionless quantity. By the way, it is such a profound number that all good theoretical physicists put this number up on their wall and worry about it, said Mr. Feynman.

Now, in the presence of spin-orbit coupling, our Hamiltonian (6.13) no longer commutes with **L** and **S**. The spin-orbit interaction mixed them up. However, the Hamiltonian still commutes with the total angular momentum

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \tag{6.15}$$

If we only consider the spin-orbit interaction within multiplet terms (for the same L and S), we can represent our eigen-vectors by[3]

$$|L, S, J, M_J\rangle \tag{6.16}$$

which is a Clebsch-Grodan basis transformation [6] from basis (6.14). The possible values of J are $L+S, L+S-1, \ldots, |L-S|$. For instance, for multiplet ^3P with $L=1$ and $S=1$, we have $J=2,1,0$. This splits the degenerate ^3P into $^3P_2, ^3P_1, ^3P_0$. We now denote a term symbol as

$$^{2S+1}L_J \tag{6.17}$$

with

$$\text{degeneracy} = 2J+1 \tag{6.18}$$

which is less degenerate than a ^{2S+1}L term that we used previously. The sum over all $(2J+1)$ within the same multiplet is equal to $(2S+1)(2L+1)$. Now, our task is to obtain the eigen-energies of the $^{2S+1}L_J$ terms. Notice that,

$$J^2 = L^2 + 2\mathbf{L}\cdot\mathbf{S} + S^2 \tag{6.19}$$

Hence, we can evaluate the following matrix element easily,

$$\begin{aligned}\langle LSJM_J|\,\mathbf{L}\cdot\mathbf{S}\,|LSJM_J\rangle &= \frac{1}{2}\langle LSJM_J|\,J^2 - L^2 - S^2\,|LSJM_J\rangle \\ &= \frac{1}{2}[J(J+1) - L(L+1) - S(S+1)]\end{aligned} \tag{6.20}$$

But this is not really what we are looking for. The spin-orbit eigen-energies within multiplet terms should be the matrix element

$$\langle LSJM_J|\,H_{\rm SO}\,|LSJM_J\rangle = \langle LSJM_J|\sum_{i=1}^{N}\xi(r_i)\boldsymbol{\ell}_i\cdot\mathbf{s}_i\,|LSJM_J\rangle \tag{6.21}$$

which is proportional to the matrix element in Eqn. (6.20) (see Reference [6])

$$\langle LSJM_J|\sum_{i=1}^{N}\xi(r_i)\boldsymbol{\ell}_i\cdot\mathbf{s}_i\,|LSJM_J\rangle = A(nl,LS)\,\langle LSJM_J|\,\mathbf{L}\cdot\mathbf{S}\,|LSJM_J\rangle \tag{6.22}$$

The proportionality factor $A(nl, LS)$ depends on the radial wave function u_{nl} and the angular momenta L and S. The eigen-energies lifted by the spin-orbit interaction are therefore,

$$\boxed{E_{\rm SO} = A(nl,LS)\frac{1}{2}[J(J+1) - L(L+1) - S(S+1)]} \tag{6.23}$$

[3]Notice that $|L, S, J, M_J\rangle$ is in general not an eigen-vector of the full Hamiltonian since the spin-orbit interaction mixes different L and S. But if we assume the spin-orbit splitting $\ll$ the multiplet splitting, we can approximately take $|L, S, J, M_J\rangle$ as our eigen-vector (first order perturbation theory).

Our problem is to find $A(nl, LS)$. Actually, we would get the same proportionality factor if we considered the matrix elements

$$\langle LM_L SM_S| \sum_{i=1}^{N} \xi(r_i) \boldsymbol{\ell}_i \cdot \mathbf{s}_i \, |LM_L SM_S\rangle = A(nl, LS) \, \langle LM_L SM_S| \, \mathbf{L} \cdot \mathbf{S} \, |LM_L SM_S\rangle \tag{6.24}$$

from which we are able to derive the expression for $A(nl, LS)$. Let's first consider

$$\langle LM_L SM_S| \, \mathbf{L} \cdot \mathbf{S} \, |LM_L SM_S\rangle$$

To evaluate this matrix element, we expand the dot product

$$\mathbf{L} \cdot \mathbf{S} = L_x S_x + L_y S_y + L_z S_z \tag{6.25}$$

Express the x and y components in terms of ladder operators

$$L_x = \frac{L_+ + L_-}{2} \quad \text{and} \quad L_y = \frac{L_+ - L_-}{2i} \tag{6.26}$$

We obtain,

$$\begin{aligned} \langle LM_L SM_S| \, \mathbf{L} \cdot \mathbf{S} \, |LM_L SM_S\rangle &= \langle LM_L SM_S| \, \frac{1}{2} L_+ S_- + \frac{1}{2} L_- S_+ + L_z S_z \, |LM_L SM_S\rangle \\ &= \langle LM_L SM_S| \, L_z S_z \, |LM_L SM_S\rangle \\ &= M_L M_S \end{aligned} \tag{6.27}$$

Done! Keep in mind that our task is to find $A(nl, LS)$ in Eqn. (6.24). Now, we consider

$$\langle LM_L SM_S| \sum_{i=1}^{N} \xi(r_i) \boldsymbol{\ell}_i \cdot \mathbf{s}_i \, |LM_L SM_S\rangle$$

which can be split into two independent parts

$$\langle nl| \, \xi(r) \, |nl\rangle \, \langle LM_L SM_S| \sum_{i=1}^{N} \boldsymbol{\ell}_i \cdot \mathbf{s}_i \, |LM_L SM_S\rangle$$

The expectation vale $\langle nl| \, \xi(r) \, |nl\rangle$ can be calculated easily by numerical integration methods. But we encounter some difficulties with the second part, where the single electron operators act on the eigen-state which is in general a linear combination of configuration basis vectors. For example, in p^3 configuration we have,

$$|1, 1, \frac{1}{2}, \frac{1}{2}\rangle = \frac{1}{\sqrt{2}} \left(c^\dagger_{1\downarrow} c^\dagger_{-1\uparrow} c^\dagger_{1\uparrow} - c^\dagger_{0\downarrow} c^\dagger_{0\uparrow} c^\dagger_{1\uparrow} \right) |0\rangle \tag{6.28}$$

It would be convenient if we expand the eigen-vectors in terms of configuration basis vectors.

$$\begin{aligned} |LM_L SM_S\rangle &= \sum_{n=1}^{\dim} a_n \, c^\dagger_{m_N \sigma_N} \ldots c^\dagger_{m_2 \sigma_2} c^\dagger_{m_1 \sigma_1} \, |0\rangle \\ &= \sum_{n=1}^{\dim} a_n \, |L\{m_i\} S\{\sigma_i\}\rangle \end{aligned} \tag{6.29}$$

Hence,

$$\begin{aligned}
&\langle LM_L SM_S| \sum_{i=1}^{N} \boldsymbol{\ell}_i \cdot \mathbf{s}_i \, |LM_L SM_S\rangle \\
&= \left(\sum_{n=1}^{\dim} \overline{a_n} \, \langle L\{m_i\} S\{\sigma_i\}| \right) \sum_{i=1}^{N} \boldsymbol{\ell}_i \cdot \mathbf{s}_i \left(\sum_{n=1}^{\dim} a_n \, |L\{m_i\} S\{\sigma_i\}\rangle \right) \\
&= \sum_{n=1}^{\dim} \left(|a_n|^2 \, \langle L\{m_i\} S\{\sigma_i\}| \sum_{i=1}^{N} \boldsymbol{\ell}_i \cdot \mathbf{s}_i \, |L\{m_i\} S\{\sigma_i\}\rangle \right) \\
&= \sum_{n=1}^{\dim} \left(|a_n|^2 \sum_{i=1}^{N} \langle Lm_i S\sigma_i| \, \boldsymbol{\ell}_i \cdot \mathbf{s}_i \, |Lm_i S\sigma_i\rangle \right) \\
&= \sum_{n=1}^{\dim} \left(|a_n|^2 \sum_{i=1}^{N} \langle Lm_i S\sigma_i| \, \frac{1}{2}\ell_+^i s_-^i + \frac{1}{2}\ell_-^i s_+^i + \ell_z^i s_z^i \, |Lm_i S\sigma_i\rangle \right) \\
&= \sum_{n=1}^{\dim} \left(|a_n|^2 \sum_{i=1}^{N} \langle Lm_i S\sigma_i| \, \ell_z^i s_z^i \, |Lm_i S\sigma_i\rangle \right) \\
&= \sum_{n=1}^{\dim} \left(|a_n|^2 \sum_{i=1}^{N} m_i \sigma_i \right)
\end{aligned} \tag{6.30}$$

Putting everything into Eqn. (6.24), we find

$$\boxed{A(nl, LS) = \langle nl| \, \xi(r) \, |nl\rangle \, \frac{\sum_{n=1}^{\dim} \left(|a_n|^2 \sum_{i=1}^{N} m_i \sigma_i \right)}{M_L M_S}} \tag{6.31}$$

$A(nl, LS)$ has no dependence on M_L or M_S. The most convenient choice would be using the maximum values $M_L = L$ and $M_S = S$. One might worry about the denominator since L or S might be zero. However, either L or S is zero indicates that there is no spin-orbit interaction ($J = L + S = |L - S|$), hence no splitting, no worry.

I understand that the expression of $A(nl, LS)$ looks a bit complicated. It would be nice to have a few worked examples. For convenience, we break Eqn. (6.31) into two parts,

$$X(nl) = \langle nl| \, \xi(r) \, |nl\rangle \tag{6.32}$$

$$M(LS) = \frac{\sum_{n=1}^{\dim} \left(|a_n|^2 \sum_{i=1}^{N} m_i \sigma_i \right)}{M_L M_S} \tag{6.33}$$

Example 1: ^3P multiplet in p^2 configuration.

The leading vector

$$|1, 1, 1, 1\rangle = c_{0\uparrow}^{\dagger} c_{1\uparrow}^{\dagger} |0\rangle$$

$$M(^3P) = \frac{0 \times \frac{1}{2} + 1 \times \frac{1}{2}}{1 \times 1} = \frac{1}{2}$$

Hence,

$$A(np, \, ^3P) = \frac{1}{2} X(np)$$

Example 2: ^2G multiplet in d^3 configuration.

The leading vector

$$|4,4,\frac{1}{2},\frac{1}{2}\rangle = \frac{1}{\sqrt{5}}\left(\sqrt{2}c^\dagger_{2\downarrow}c^\dagger_{0\uparrow}c^\dagger_{2\uparrow} - \sqrt{3}c^\dagger_{1\downarrow}c^\dagger_{1\uparrow}c^\dagger_{2\uparrow}\right)|0\rangle$$

$$M(^2G) = \frac{\frac{2}{5}(2\times-\frac{1}{2}+0\times\frac{1}{2}+2\times\frac{1}{2})+\frac{3}{5}(1\times-\frac{1}{2}+1\times\frac{1}{2}+2\times\frac{1}{2})}{4\times\frac{1}{2}} = \frac{3}{10}$$

Hence,

$$A(nd,^2G) = \frac{3}{10}X(nd)$$

There is a symmetry property of the factor $M(LS)$. If the shell is less than half-filled, $M(LS) > 0$; if the shell is more than half-filled, $M(LS) < 0$; and if the shell is exactly half-filled, $M(LS) = 0$ (no splitting).

The remaining part $X(nl)$, is relatively more straightforward.

$$\begin{aligned} X(nl) &= \langle nl|\,\xi(r)\,|nl\rangle = \int_0^\infty dr\, r^2\,\overline{R_{nl}}(r)\xi(r)R_{nl}(r) \\ &= \int_0^\infty dr\,\overline{u_{nl}}(r)\xi(r)u_{nl}(r) = \int_0^\infty dr\,|u_{nl}(r)|^2\xi(r) \end{aligned} \tag{6.34}$$

and in atomic unit,

$$\xi(r) = \frac{1}{2\times 137.036^2}\frac{1}{r}\frac{dV}{dr} \tag{6.35}$$

In our self-consistent field approximation, the potential $V(r)$ is the mean-field electric potential that all electrons experience in. The first order derivative can be obtained approximately from a finite difference formula

$$\frac{dV}{dx} \approx \frac{V(x+\Delta x)-V(x-\Delta x)}{2\Delta x} \tag{6.36}$$

Remember the logarithmic grid transformation in Eqn (2.5), we have,

$$dr = rdx \quad \text{and} \quad \frac{dV}{dr} = \frac{1}{r}\frac{dV}{dx} \tag{6.37}$$

Hence, numerically,

$$X(nl) = \frac{1}{2\times 137.036^2}\int_{x_{\min}}^{x_{\max}} dx\,|u_{nl}(x)|^2\frac{1}{r}\frac{dV}{dx} \tag{6.38}$$

To get an intuitive understanding of this spin-orbit energy splitting, we again take the carbon atom as an example. A carbon atom has electronic configuration $1s^2\ 2s^2\ 2p^2$. From the self-consistent calculation, we obtained both the mean-field potential $V(r)$ and the wave function $u_{2p}(r)$. From Eqn. (6.38), we can calculate

$$X(2p) = 0.000218 \text{ (Hartree)} \tag{6.39}$$

Meanwhile, we have discovered that a p^2 configuration produces ^3P, ^1D and ^1S multiplets. Among these three multiplets, only ^3P splits due to spin-orbit interaction. Both ^1D and ^1S do not split since $S = 0$. In the previous examples, we found for the p^2 configuration,

$$M(^3P) = \frac{1}{2} \tag{6.40}$$

Therefore,

$$A(2p^2, {}^3P) = 0.000218 \times \frac{1}{2} = 0.000109 \text{ (Hartree)} \tag{6.41}$$

Now the spin-orbit energy splittings within ${}^3P$ are (Eqn. (6.23))

$$\begin{aligned} E_{\rm SO}({}^3P_2) &= 0.000109\frac{1}{2}[2(2+1) - 1(1+1) - 1(1+1)] = \quad 0.000109 \text{ (Hartree)} \\ E_{\rm SO}({}^3P_1) &= 0.000109\frac{1}{2}[1(1+1) - 1(1+1) - 1(1+1)] = -0.000109 \text{ (Hartree)} \\ E_{\rm SO}({}^3P_0) &= 0.000109\frac{1}{2}[0(0+1) - 1(1+1) - 1(1+1)] = -0.000218 \text{ (Hartree)} \end{aligned} \tag{6.42}$$

with this additional splitting, Fig. 5.1 gets an extra column.

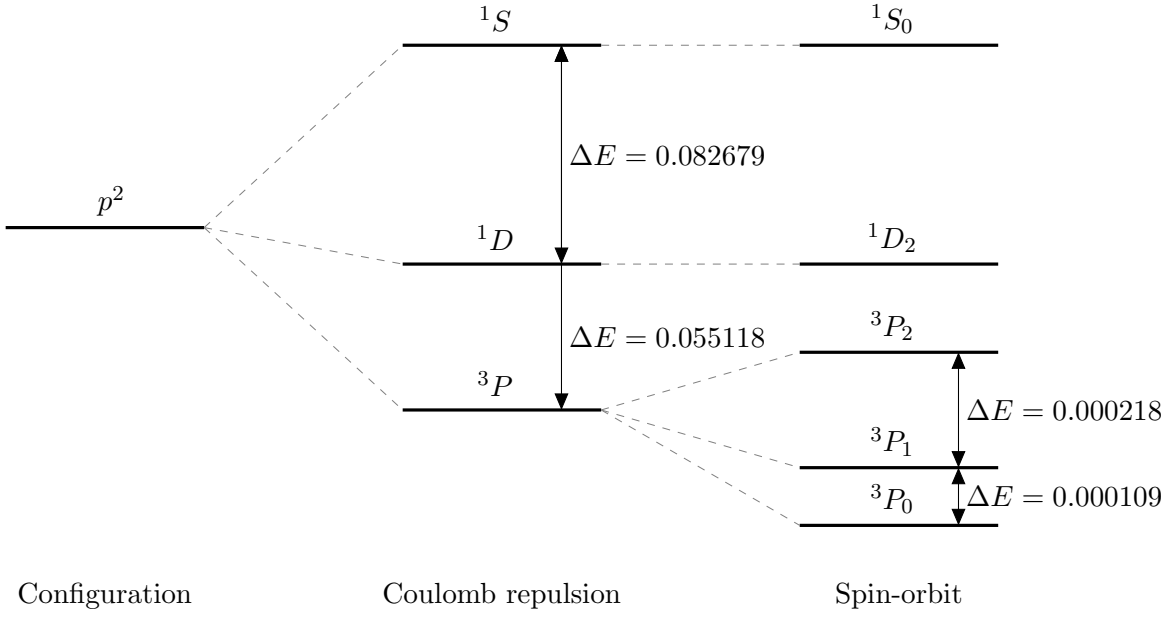


Figure 6.1.: Energy splitting (Coulomb repulsion plus spin-orbit interaction) of the p^2 configuration of a carbon atom. Energies are given in units of Hartree (a.u.). The spin-orbit splitting is magnified for a better plotting. (see Fig. 6.2 for realistic scale.)

The spin-orbit splitting in Fig. 6.1 is magnified for a better plotting. The actual splitting is rather tiny (remember the $1/c^2$ factor). But I want to emphasize that there are two issues about the spin-orbit splitting, namely, the "scale" and the "shape". This can be seen from Eqn. (6.23). The part

$$[J(J+1) - L(L+1) - S(S+1)]$$

leads to the splitting "shape", which is determined by the multiplet term ${}^{2S+1}L$. While the other part

$$A(nl, LS)$$

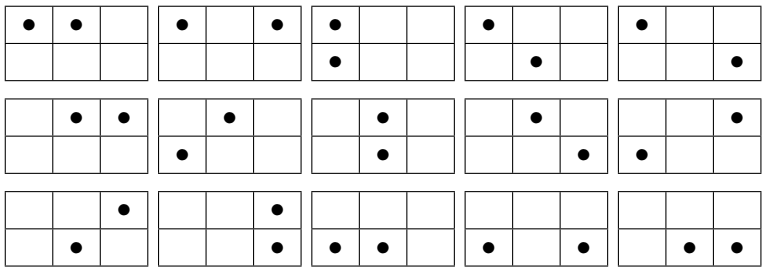
controls the "scale" of the splitting. We have seen that $A(nl, LS)$ can be separated into $X(nl)$ and $M(LS)$. For a given multiplet term, $M(LS)$ is fixed. The final term that governs the scale is

$$X(nl) = \langle nl| \, \xi(r) \, |nl\rangle$$

which is the coupling strength constant. For a not-so-heavy atom, $X(nl)$ is usually weak (see Eqn.(6.39) for a carbon atom). We can safely simplify the spin-orbit interaction within a multiplet term and obtain surprisingly accurate energies. But for a heavy atom, like uranium, this spin-orbit coupling constant becomes strong due to the deep radial potential. If the order of the spin-orbit splitting reaches the order of the multiplet energy splitting, our "spin-orbit within multiplet terms" may no longer be a good approximation. In this case, we should consider the spin-orbit interactions within the entire shell and diagonalize the complete Hamiltonian.

6.3. Spin-orbit coupling within the entire shell

By saying "within entire shell", it means to solve the problem in the complete basis of a given configuration, say, p^2,



We have previously used the same basis when solving the Coulomb repulsion problem. Since we have the basis already, the remaining task is to set up the matrix representation of the spin-orbit Hamiltonian in our basis. To construct the matrix representation, the first step is to reformulate the Hamiltonian

$$H_{\mathrm{SO}} = \sum_{i=1}^{N} \xi(r_i) \boldsymbol{\ell}_i \cdot \mathbf{s}_i \tag{6.43}$$

into second quantization,

$$H_{\mathrm{SO}} = \sum_{\alpha,\beta} V_{\alpha\beta} c_\alpha^\dagger c_\beta \tag{6.44}$$

where,

$$\alpha = \{n_1,\, l_1,\, m_1,\, \sigma_1\}$$
$$\beta = \{n_2,\, l_2,\, m_2,\, \sigma_2\}$$

enumerate all possible quantum states of electrons. Here we consider only interactions within the same shell, so we restrict $n_1 l_1 = n_2 l_2 = nl$. If you still remember, we devoted an entire chapter calculating the Coulomb repulsion matrix element $U_{\alpha\beta\gamma\delta}$ since it was extremely complicated. However, today, our spin-orbit matrix element

$$V_{\alpha\beta} = \langle \alpha |\, \xi(r) \boldsymbol{\ell} \cdot \mathbf{s} \,| \beta \rangle \tag{6.45}$$

can be calculated with zero difficulty. This spin-orbit matrix element can be split into a radial dependent part and an angular dependent part

$$V_{\alpha\beta} = \langle nl |\, \xi(r) \,| nl \rangle \, \langle m_1\sigma_1 |\, \boldsymbol{\ell} \cdot \mathbf{s} \,| m_2\sigma_2 \rangle \tag{6.46}$$

The radial part $\langle nl|\,\xi(r)\,|nl\rangle$ is identical to $X(nl)$ which we have discussed in Eqn. (6.34). And the angular part,

$$\begin{aligned}
\langle m_1\sigma_1|\,\boldsymbol{\ell}\cdot\mathbf{s}\,|m_2\sigma_2\rangle &= \langle m_1\sigma_1|\,\ell_x s_x + \ell_y s_y + \ell_z s_z\,|m_2\sigma_2\rangle \\
&= \langle m_1\sigma_1|\,\frac{1}{2}\ell_+ s_- + \frac{1}{2}\ell_- s_+ + \ell_z s_z\,|m_2\sigma_2\rangle \\
&= \frac{1}{2}\sqrt{(l+m_2+1)(l-m_2)\left(\frac{1}{2}+\sigma_2\right)\left(\frac{1}{2}-\sigma_2+1\right)}\,\langle m_1\sigma_1\,|\,m_2+1,\sigma_2-1\rangle \\
&\quad + \frac{1}{2}\sqrt{(l+m_2)(l-m_2+1)\left(\frac{1}{2}+\sigma_2+1\right)\left(\frac{1}{2}-\sigma_2\right)}\,\langle m_1\sigma_1\,|\,m_2-1,\sigma_2+1\rangle \\
&\quad + m_2\sigma_2\,\langle m_1\sigma_1\,|\,m_2\sigma_2\rangle
\end{aligned} \tag{6.47}$$

can be computed easily with the orthonormality of angular wave functions,

$$\langle m_1\sigma_1\,|\,m_2\sigma_2\rangle = \delta_{m_1m_2}\delta_{\sigma_1\sigma_2} \tag{6.48}$$

Setting up the spin-orbit Hamiltonian is simpler than setting up the Coulomb repulsion Hamiltonian since we have only α and β indices

$$\langle i|\,H_{\rm SO}\,|j\rangle = \langle i|\sum_{\alpha,\beta} V_{\alpha\beta}c^\dagger_\alpha c_\beta\,|j\rangle \tag{6.49}$$

Hence, the algorithm is also simpler with less for loops (comparing with Algorithm 5.2).

Algorithm 6.1 Set up Hamiltonian

```
1:  function HAMILTONIAN(basis)
2:      dim ← basis.dim
3:      for i ← 0 to dim do
4:          conf_i ← basis.conf[i]
5:          for all α do
6:              if ISBIT(α, conf_i) then
7:                  conf_α ← CLEARBIT(α, conf_i)
8:                  for all β do
9:                      if !ISBIT(β, conf_α) then
10:                         conf_β ← SETBIT(β, conf_α)
11:                         conf_j ← conf_β
12:                         j ← basis.index[conf_j]
13:                         H_SO[i, j] ← H_SO[i, j] + fsign * V_αβ
14:     return H_SO
```

If we diagonalize $H_{\rm SO}$ directly, we would obtain the eigen-energies of the pure spin-orbit interaction. To include both Coulomb repulsion and spin-orbit coupling, we should diagonalize (numerically) the sum $(H_U + H_{\rm SO})$. For not-so-heavy atoms, like carbon, the resulting eigen-energies are surprisingly close to the eigen-energies we obtained from the "within multiplet terms" approximation, which is pretty remarkable, since the solutions from two different approaches agree each other. However, for heavy atoms, there is a large discrepancy between those two solutions.

For heavy atoms, the deep potential leads to large values in the derivative dV/dr. Hence, the spin-orbit coupling constant $\langle nl|\,\xi(r)\,|nl\rangle$ is large. For strong spin-orbit interactions, the order of energy splitting can reach the order of the Coulomb repulsion splitting. In this case the approximation using spin-orbit coupling within multiplet terms are no longer appropriate.

This can be clearly demonstrated by a comparison between a carbon (C) and a lead (Pb) atom, which are from the same group with the same open shell configuration p^2. Table 6.1 tabulated the numerical energies of spin-orbit interactions within multiplet terms and within entire shell.

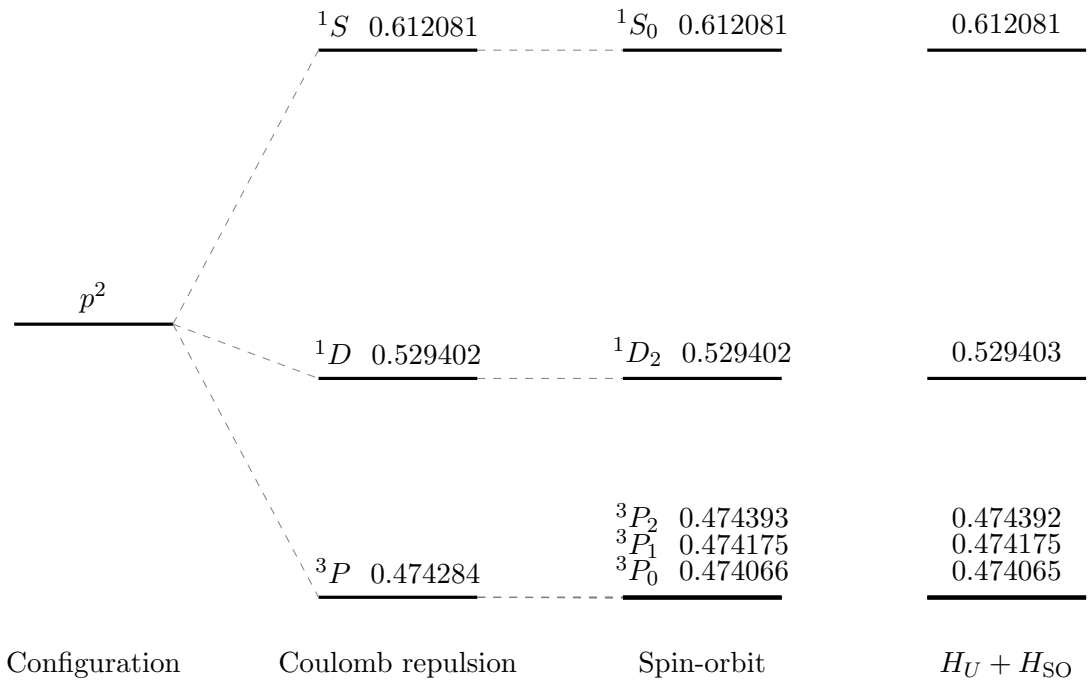
Table 6.1.: Comparison of open shell spin-orbit energies within multiplet terms and within entire shell for a carbon atom and a lead atom. Energies are given in units of Hartree (a.u.).

Elem	Orbital	Energy within multiplet terms (×degeneracy)	Energy within entire shell (×degeneracy)	Abs Error	Rel Error
C	$2p^2$	0.612081 (×1)	0.612081 (×1)	0.000000	0.000000
		0.529402 (×5)	0.529403 (×5)	0.000001	0.000002
		0.474393 (×5)	0.474392 (×5)	0.000001	0.000002
		0.474175 (×3)	0.474175 (×3)	0.000000	0.000000
		0.474066 (×1)	0.474065 (×1)	0.000001	0.000002
Pb	$6p^2$	0.323500 (×1)	0.335963 (×1)	0.012463	0.037096
		0.272826 (×5)	0.284981 (×5)	0.012155	0.042652
		0.252988 (×5)	0.240833 (×5)	0.012155	0.050471
		0.225100 (×3)	0.225100 (×3)	0.000000	0.000000
		0.211156 (×1)	0.198693 (×1)	0.012463	0.062725

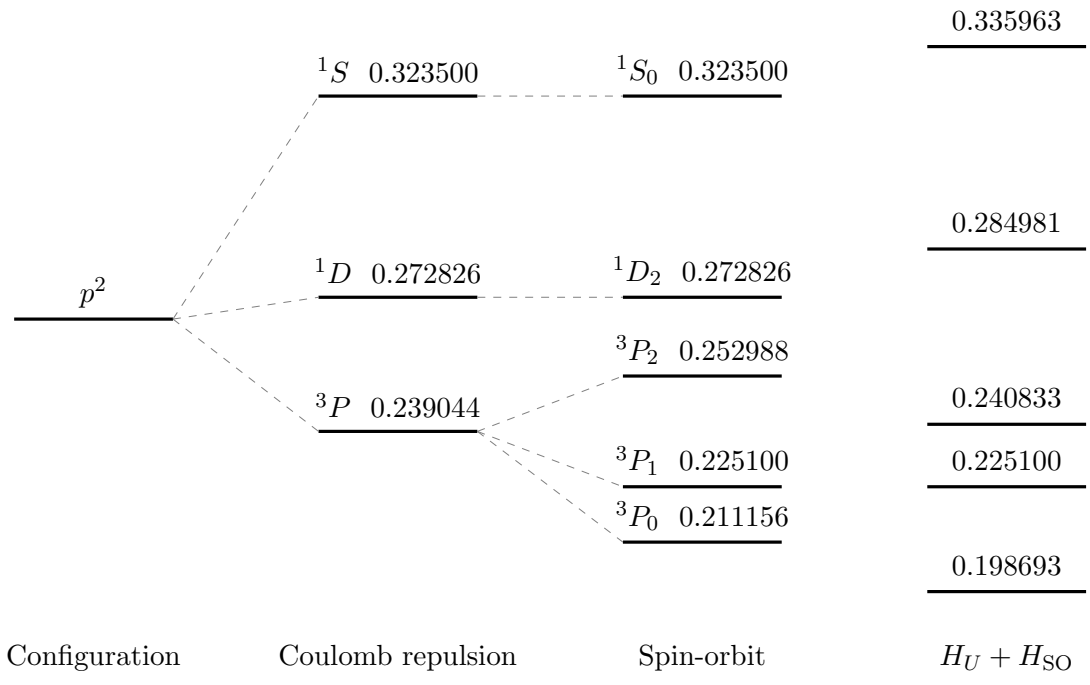
The discrepancy can be seen more easily from the spectrum plot in Fig. 6.2. The spin-orbit splitting within multiplet terms and within the entire shell are plotted in the 3rd and 4th column of the plot, respectively. It is difficult to resolve the spin-orbit splitting in the carbon atom plot, since the energy differences are so tiny. But this tiny splitting gives a good approximation when considering spin-orbit coupling within multiplet terms. On the other hand, the amplitude of spin-orbit splitting in the lead atom reaches the amplitude of Coulomb repulsion splitting. In this case, the energies from the "within multiplet terms" approximation do not match the (more accurate) full shell diagonalization.

In the 4th column in Fig. 6.2, we labeled each energy level by their numerical values. But we didn't put a label like ${}^{2S+1}L$ or ${}^{2S+1}L_J$. This is because when we consider the eigen-energies from the sum $(H_U + H_{\mathrm{SO}})$, the energy levels are mixed with contributions from different angular momenta L and S. We can no longer label each energy level as purely ${}^{2S+1}L$ or ${}^{2S+1}L_J$. Nevertheless, from the plot, we do see some strong correspondence between the multiplet terms and the energy levels from $(H_U + H_{\mathrm{SO}})$. This correspondence can be calculated from the overlap between the eigen-vectors of multiplet terms and the eigen-vectors of $(H_U + H_{\mathrm{SO}})$, which is known as the character of eigen-vectors.

For a multiplet term ${}^{2S+1}L$ (with seniority number W if necessary), we have eigen-vectors $|L, M_L, S, M_S\rangle$ with $M_L = L, \ldots, -L$ and $M_S = S, \ldots, -S$. Those vectors span a "small space" of this specific multiplet term. Now, suppose we have an eigen-vector $|v\rangle$ of $(H_U + H_{\mathrm{SO}})$ from our numerical diagonalization. To check if this eigen-vector $|v\rangle$ lives inside this "small space",



(a) Carbon (C)



(b) Lead (Pb)

Figure 6.2.: Comparison of open shell energy splitting between a "light" atom carbon (C) and a "heavy" atom lead (Pb). Energies are given in units of Hartree (a.u.). From left to right, 1st column: electronic configuration; 2nd column: Coulomb repulsion energy splitting; 3rd column: spin-orbit interaction within multiplet terms; 4th column: eigen-energies of the Hamiltonian ($H_U + H_{\mathrm{SO}}$). Spin-orbit interaction within multiplet terms are good approximations for "light" atoms but not for "heavy" atoms.

we compute the character

$$\boxed{\lambda = \sum_{M_L, M_S} \left| \langle L, M_L, S, M_S \,|\, v \rangle \right|^2} \tag{6.50}$$

If $|v\rangle$ lives completely in the space spanned by $|L, M_L, S, M_S\rangle$, we shall get $\lambda = 1$. On the contrary, if $|v\rangle$ is completely off, we will get $\lambda = 0$. However, in our problem, this $|v\rangle$ is often partly in one multiplet term and partly in the others. In this case we get $0 < \lambda < 1$. The closer to 1, the stronger is the contribution from this specific multiplet term.

Continuing with our discussion, we compute all characters for each numerical eigen-vectors of $(H_U + H_{\mathrm{SO}})$ within the three multiplet spaces ${}^1S$, ${}^1D$ and ${}^3P$. The characters are listed in Table 6.2.

Table 6.2.: Character of numerical eigen-vectors in different multiplet term spaces. We highlight the main contribution using underlines. Zero values are left as empty entries so that the structure can be seen clearly.

Elem	Orbital	Energy within entire shell (×degeneracy)	Character in ${}^1S$	Character in ${}^1D$	Character in ${}^3P$
C	$2p^2$	0.612081 (×1)	<u>0.999995</u>		0.000005
		0.529403 (×5)		<u>0.999992</u>	0.000008
		0.474392 (×5)		0.000008	<u>0.999992</u>
		0.474175 (×3)			<u>1.000000</u>
		0.474065 (×1)	0.000005		<u>0.999995</u>
Pb	$6p^2$	0.335963 (×1)	<u>0.909208</u>		0.090792
		0.284981 (×5)		<u>0.724683</u>	0.275317
		0.240833 (×5)		0.275317	<u>0.724683</u>
		0.225100 (×3)			<u>1.000000</u>
		0.198693 (×1)	0.090792		<u>0.909208</u>

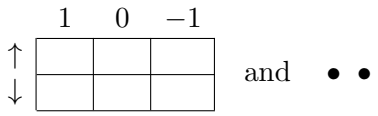
If you watch carefully and compare Table 6.2 with Fig. 6.2, you will notice that the terms are mixed if they have the same J. Maybe you also noticed the interesting "1.000000" which never mixes with the others. That is because the vectors are from the term with a unique J.

This table of characters directly indicates how strongly are the eigen-states mixed among different multiplet terms (different angular momenta). It again evidenced that with spin-orbit effect, the eigen-states in carbon (light atom) are slightly mixed with different multiplet terms, but the eigen-states in lead (heavy atom) are strongly mixed with different multiplet terms. I must point out that, the numerical diagonalization of $(H_U + H_{\mathrm{SO}})$ can always give us a better estimation of the eigen-energies (because it uses the complete basis in the open shell), which is especially important for heavy atoms. Nevertheless, our construction of multiplet states and the first order perturbation theory in spin-orbit coupling give us a very important understanding of the problem and a deep insight into the physical system.

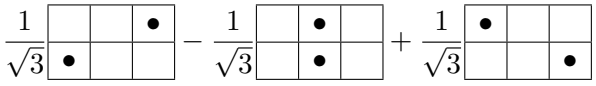
Chapter 7

Summary

If your friend asks you, "What is multiplet?" A short answer is, "Multiplets are the many-electron eigen-states in atoms." But probably he won't be satisfied since he knows the name but doesn't really understand the problem. Then you give him the following box and two electrons,



and ask, "Let's put these two electrons into this p shell. In which configuration do you think this system has the highest energy?" (don't ask for the lowest one, since it can be known easily from Hund's rule) If he complains there is no difference in the way of putting electrons, then he is speaking mean-field language, which is exactly what we assumed in our Chapter 3. Fortunately, your friend is convinced that electrons with different orbital and spin angular momenta do repel each other differently (you showed him the plots in Appendix A). But still, he won't recognize which configuration has the highest energy, because this is not at all a trivial problem! If you are also curious for the answer, I put it here directly: the state



is an eigen-state of the p^2 configuration with the highest eigen-energy, and it corresponds to the ${}^1S$ multiplet (you see, nobody could answer this easily). The complete problem is worked out step by step in Chapter 5. Finally, in addition to Coulomb repulsion, we also included spin-orbit coupling in Chapter 6, where we see the multiplet spectral lines further split into finer structures. Our work can be extended by introducing the jj-coupling, where we consider each electron's total angular momentum. We can also introduce external crystal field (our present work are in the frame of isolated atoms) to study how our multiplet states respond to different external potential fields.

In the very end I must advertise our simulation tool: all the discussions in this thesis have been implemented on a web page. Programming codes are written in JavaScript. You can run simulations of all atoms on the periodic table directly in a modern browser (no installation, no compilation, and no plug-in). This simulation tool can be accessed via the link:

`www.cond-mat.de/sims/multiplet`

Appendix A

How to draw spherical harmonics

A.1. The spherical harmonics

For the first time students encounter spherical harmonics, we are most likely scared away by the complicated expressions and bizarre geometries of the plots. Complaining,

> "I can never understand these functions and crazy plots. It's all so complicated!"

This is, however, always the case when we learn something new. Things usually look incomprehensible until we understand them and set up a good friendship.

Spherical harmonics are the solutions of the angular equation in (2.3):

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}=-l(l+1)Y \tag{A.1}$$

The derivations are nicely discussed in Griffiths' book [1]. In this short appendix, we are not going to repeat the derivations, but will emphasize another interesting perspective: how to draw the spherical harmonics. Not only for impressing your friends, but more importantly, once we understood how they are plotted, we will get a direct feeling of spherical harmonics and essentially comprehend their meanings.

The spherical harmonics $Y_{lm}(\theta,\phi)$ (with $l=0,\ 1,\ \ldots$ and $m=-l,\ \ldots,\ l$), are given by

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

The big square root in front is nothing but a normalization factor, simply a real number. The exponential term in the end is called the phase, which never contributes when we consider the modulus square $|Y_{lm}|^2$. Probably the most scaring term is P_l^m, the associated Legendre polynomials. But don't worry, they can be computed very easily. The computation routine is clearly provided by *Numerical Recipes* [9]. To give a direct impression, Table A.1 explicitly listed the first few spherical harmonics.

Table A.1.: The first few spherical harmonics $Y_{lm}(\theta, \phi)$.

$Y_{0,\ 0} =$	$\sqrt{\frac{1}{4\pi}}$		
$Y_{1,\ 0} =$	$\sqrt{\frac{3}{4\pi}}$	$\cos\theta$	
$Y_{1,\pm1} =$	$\mp\sqrt{\frac{3}{8\pi}}$	$\sin\theta$	$e^{\pm i\phi}$
$Y_{2,\ 0} =$	$\sqrt{\frac{5}{16\pi}}$	$(3\cos^2\theta - 1)$	
$Y_{2,\pm1} =$	$\mp\sqrt{\frac{15}{8\pi}}$	$\sin\theta\cos\theta$	$e^{\pm i\phi}$
$Y_{2,\pm2} =$	$\sqrt{\frac{15}{32\pi}}$	$\sin^2\theta$	$e^{\pm 2i\phi}$
$Y_{3,\ 0} =$	$\sqrt{\frac{7}{16\pi}}$	$(5\cos^3\theta - 3\cos\theta)$	
$Y_{3,\pm1} =$	$\mp\sqrt{\frac{21}{64\pi}}$	$\sin\theta(5\cos^2\theta - 1)$	$e^{\pm i\phi}$
$Y_{3,\pm2} =$	$\sqrt{\frac{105}{32\pi}}$	$\sin^2\theta\cos\theta$	$e^{\pm 2i\phi}$
$Y_{3,\pm3} =$	$\mp\sqrt{\frac{35}{64\pi}}$	$\sin^3\theta$	$e^{\pm 3i\phi}$

A.2. Plotting in spherical coordinates

Perhaps the most common and easiest way to make a plot is to plot in the Cartesian coordinate system. Just like how we plotted our radial wave functions $u_{nl}(r)$. We took the x-axis representing our spacial distance r and y-axis representing our wave functions u_{nl}

$$\begin{cases} x \leftarrow r \\ y \leftarrow u_{nl} \end{cases} \tag{A.3}$$

However, for our angular wave functions, namely the spherical harmonics $Y_{lm}(\theta, \phi)$, it is more natural to plot them in a spherical coordinate system, since it is where they are defined. Now the mapping is the following,

$$\begin{cases} r \leftarrow Y_{lm} \\ \theta \leftarrow \theta \\ \phi \leftarrow \phi \end{cases} \tag{A.4}$$

The important message is that we use the radius to represent the amplitude of spherical harmonics. The functions listed in Table A.1 are good enough for us to make a couple of beautiful

plots. For simplicity, we would like to restrict the azimuthal angle ϕ to 0, that is, we plot in the xz-plane. So now we have only one variable θ. To make the plots, we first define our grid in θ

$$\left\{\theta_{\min} = 0; \quad \theta_{\max} = 2\pi; \quad \Delta\theta = \frac{\pi}{12};\right\} \tag{A.5}$$

To get started, let's plot the simplest function Y_{00},

For $\theta = 0$, we have $r = \sqrt{\frac{1}{4\pi}}$;
For $\theta = \frac{\pi}{12}$, we have $r = \sqrt{\frac{1}{4\pi}}$;
For $\theta = \frac{2\pi}{12}$, we have $r = \sqrt{\frac{1}{4\pi}}$;
...
For $\theta = 2\pi$, we have $r = \sqrt{\frac{1}{4\pi}}$.

If we plot these data points and connect them, we get a circle! (Fig. A.1)

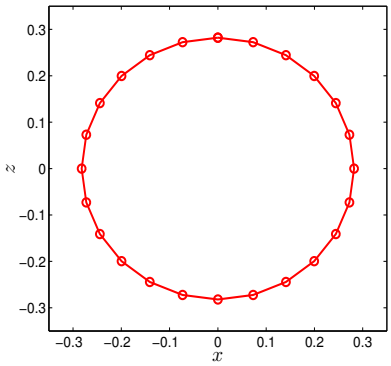


Figure A.1.: $Y_{00}(\theta, \phi)$ in the xz-plane.

Y_{00} was simple enough. Let's try a more exciting one, Y_{10},

For $\theta = 0$, we have $r \approx 0.4886$;
For $\theta = \frac{\pi}{12}$, we have $r \approx 0.4720$;
...
For $\theta = \frac{7\pi}{12}$, we have $\boxed{r \approx -0.1265}$;
...
For $\theta = \frac{23\pi}{12}$, we have $r \approx 0.4720$;
For $\theta = 2\pi$, we have $r \approx 0.4886$.

Wait! how do we plot a negative radius? Hum... we really can only plot the absolute value. To indicate the different signs, let's use two different colors. We use red color for positive Y_{10} and blue color for negative Y_{10}. Now we plot the data points and connect them. We get a red-blue colored plot in Fig. A.2

The color of the plot doesn't contribute much into the physical meaning. Never ever misunderstand them as positive or negative charges (or whatever). The physical meaning is represented by the modulus square of the wave function $|Y_{lm}|^2$, which is the probability density of finding an electron. It doesn't really matter which color ($\pm$ sign) it is.

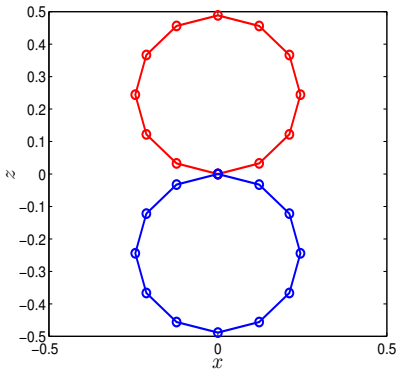


Figure A.2.: $Y_{10}(\theta, \phi)$ in the xz-plane.

Previously we restricted our azimuthal angle to 0. It wouldn't be too difficult to extend our discussion for arbitrary ϕ. When we consider a non-zero ϕ, we may have the situation that Y_{lm} is a complex number. But we simply plot the modulus of the complex number, so it won't be a problem. Nevertheless, if we want to indicate the phase of the complex number (just like for indicating the $\pm$ sign), we can do some color mapping to make a fancy plot. To get an idea how it works, we take the function $Y_{1,-1}$ as an example. This time, we restrict our inclination angle θ to $\frac{\pi}{2}$, that is, we plot in the xy-plane. Now we have freedom in ϕ.

$$\left\{\phi_{\min} = 0; \quad \phi_{\max} = 2\pi; \quad \Delta\phi = \frac{\pi}{12};\right\} \tag{A.6}$$

For $\phi = 0$, we have $r \approx 0.3455e^{\ 0.0000i}$;
For $\phi = \frac{\pi}{12}$, we have $r \approx 0.3455e^{-0.2618i}$;
For $\phi = \frac{2\pi}{12}$, we have $r \approx 0.3455e^{-0.5236i}$;
...
For $\phi = 2\pi$, we have $r \approx 0.3455e^{-6.2832i}$.

We plot the modulus of the complex numbers as the radius. And we map the colors according to the phase angles of the complex numbers, which is the angle formed by the real and imaginary parts on the complex plane. The choice of colors is arbitrary, but it is good to have some continuously interpolated colors to represent the continuous phase angles. (Fig. A.3)

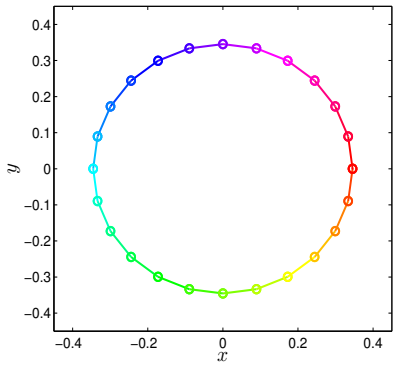


Figure A.3.: $Y_{1,-1}(\theta, \phi)$ in the xy-plane.

We have basically introduced all the essential ideas for plotting spherical harmonics. All that remains is to implement (or use) a 3D visualization program to visualize the (r, θ, ϕ) data. Writing a 3D visualization program involves some computer graphics knowledge, such as the coordinate transformations and shading programs. I have implemented a program using the WebGL technology, which can be run directly in a modern browser. I summarize some nice plots generated by WebGL in Table A.3, which are the functions listed in Table A.1.

A.3. Linear combinations of spherical harmonics

The spherical harmonics Y_{lm} (called pure harmonics) are solutions from Eqn. (A.1). Consequently, their linear combinations (with the same l) are also valid solutions. Actually, you can take any crazy combinations to make some crazy plots. But there are a handful of pre-defined linear combinations, which are typically useful for chemists. Those pre-defined combinations are called real harmonics. Because those combinations (by combining $\pm m$) eliminate the imaginary parts and result in functions which are in the real range. The first few real harmonics are listed in Table A.2. The corresponding plots are also given in Table A.4. You will see only two colors, because there are only positive and negative real numbers!

Table A.2.: The first few real harmonics.

s	$=$	$Y_{0,0}$
p_z	$=$	$Y_{1,0}$
p_x	$=$	$\sqrt{\frac{1}{2}}\,(Y_{1,-1} - Y_{1,1})$
p_y	$=$	$\sqrt{\frac{1}{2}}i(Y_{1,-1} + Y_{1,1})$
d_{3z^2-1}	$=$	$Y_{2,0}$
d_{xz}	$=$	$\sqrt{\frac{1}{2}}\,(Y_{2,-1} - Y_{2,1})$
d_{yz}	$=$	$\sqrt{\frac{1}{2}}i(Y_{2,-1} + Y_{2,1})$
$d_{x^2-y^2}$	$=$	$\sqrt{\frac{1}{2}}\,(Y_{2,-2} + Y_{2,2})$
d_{xy}	$=$	$\sqrt{\frac{1}{2}}i(Y_{2,-2} - Y_{2,2})$
$f_{z(5z^2-3)}$	$=$	$Y_{3,0}$
$f_{x(5z^2-1)}$	$=$	$\sqrt{\frac{1}{2}}\,(Y_{3,-1} - Y_{3,1})$
$f_{y(5z^2-1)}$	$=$	$\sqrt{\frac{1}{2}}i(Y_{3,-1} + Y_{3,1})$
$f_{z(x^2-y^2)}$	$=$	$\sqrt{\frac{1}{2}}\,(Y_{3,-2} + Y_{3,2})$
f_{xyz}	$=$	$\sqrt{\frac{1}{2}}i(Y_{3,-2} - Y_{3,2})$
$f_{x(x^2-3y^2)}$	$=$	$\sqrt{\frac{1}{2}}\,(Y_{3,-3} - Y_{3,3})$
$f_{y(3x^2-y^2)}$	$=$	$\sqrt{\frac{1}{2}}i(Y_{3,-3} + Y_{3,3})$

Table A.3.: The first few pure spherical harmonics.

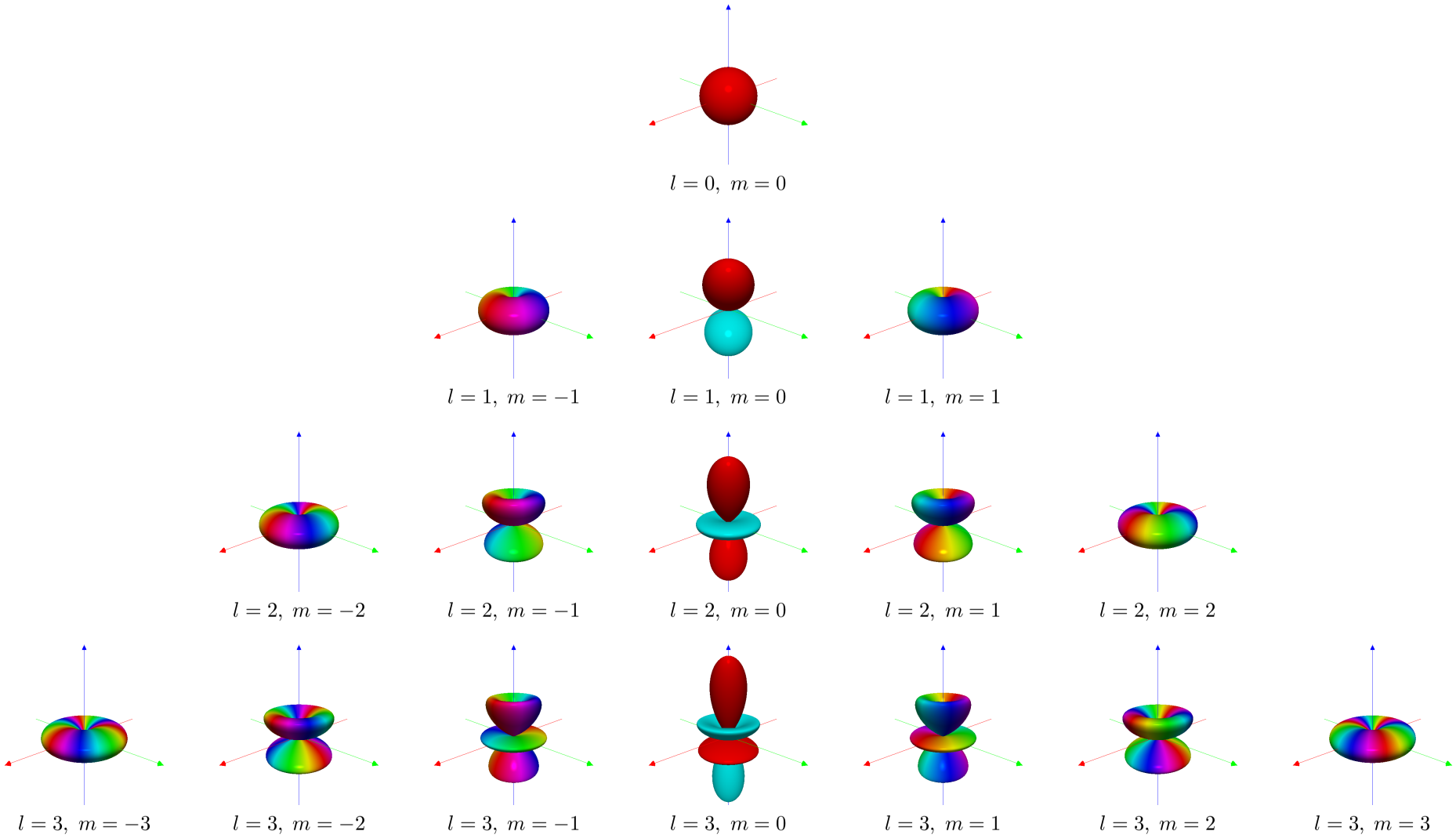
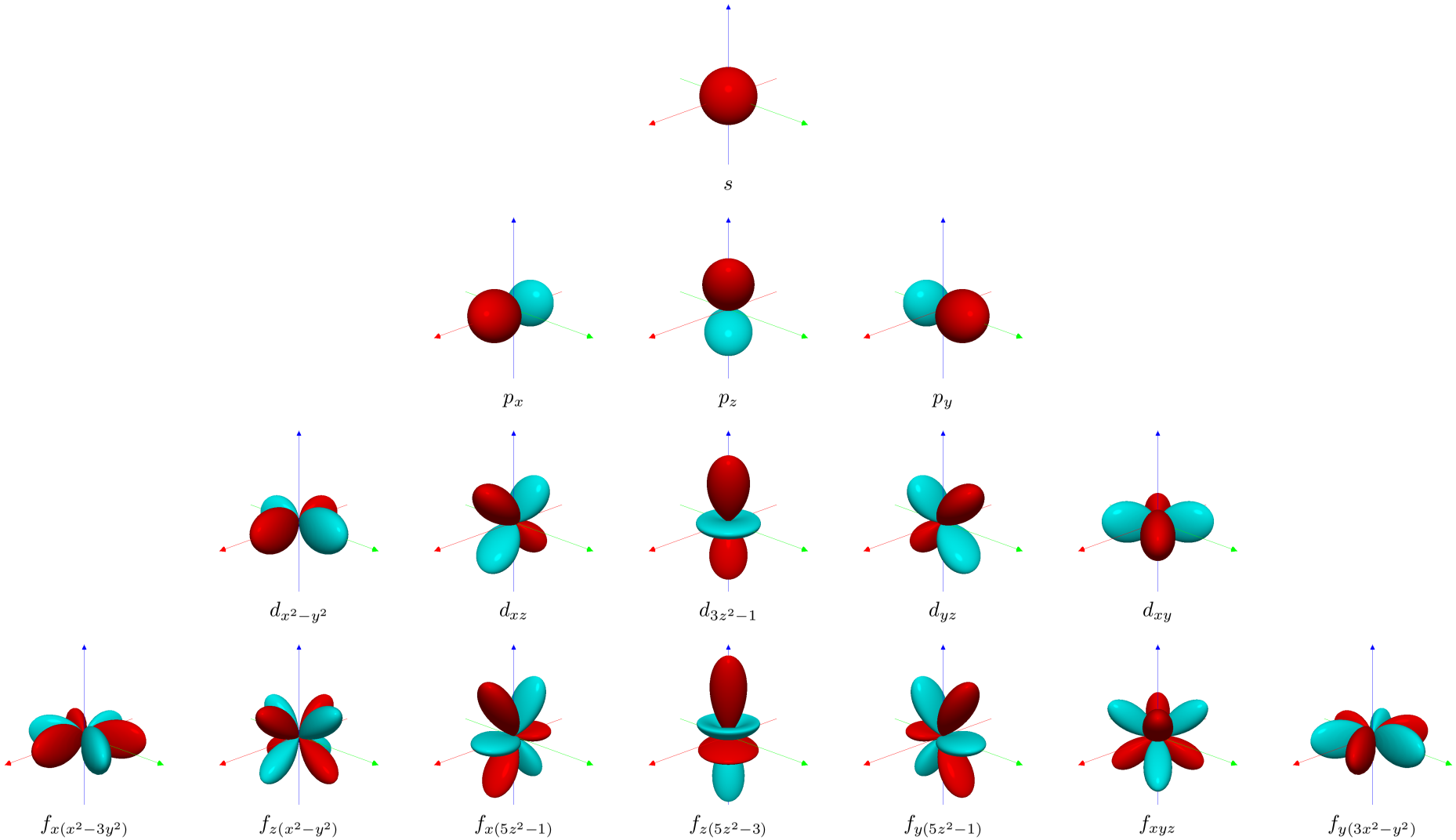


Table A.4.: The first few real spherical harmonics.



Appendix B

Second quantization

B.1. A different formalism but the same physics

In real space, electrons are described as wave functions. Because electrons are fermions, their anti-symmetric wave functions are formulated as Slater determinants [16]. (as a remark, a many-electron wave function $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ is in general a linear combination of Slater determinants. Meanwhile, we are ignoring spin to simplify our notations.)

$$\Phi_{\alpha_1 \cdots \alpha_N}(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_{\alpha_1}(\mathbf{r}_1) & \varphi_{\alpha_2}(\mathbf{r}_1) & \cdots & \varphi_{\alpha_N}(\mathbf{r}_1) \\ \varphi_{\alpha_1}(\mathbf{r}_2) & \varphi_{\alpha_2}(\mathbf{r}_2) & \cdots & \varphi_{\alpha_N}(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{\alpha_1}(\mathbf{r}_N) & \varphi_{\alpha_2}(\mathbf{r}_N) & \cdots & \varphi_{\alpha_N}(\mathbf{r}_N) \end{vmatrix} \tag{B.1}$$

For a one-electron wave function, Eqn. (B.1) is trivial:

$$\Phi_\alpha(\mathbf{r}_1) = \varphi_\alpha(\mathbf{r}_1) \tag{B.2}$$

For a two-electron wave function, Eqn. (B.1) reads:

$$\Phi_{\alpha\beta}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}(\varphi_\alpha(\mathbf{r}_1)\varphi_\beta(\mathbf{r}_2) - \varphi_\beta(\mathbf{r}_1)\varphi_\alpha(\mathbf{r}_2)) \tag{B.3}$$

For a three-electron wave function, Eqn. (B.1) becomes:

$$\begin{aligned} \Phi_{\alpha\beta\gamma}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = & \frac{1}{\sqrt{6}}\Big(\varphi_\alpha(\mathbf{r}_1)\varphi_\beta(\mathbf{r}_2)\varphi_\gamma(\mathbf{r}_3) + \varphi_\gamma(\mathbf{r}_1)\varphi_\alpha(\mathbf{r}_2)\varphi_\beta(\mathbf{r}_3) + \varphi_\beta(\mathbf{r}_1)\varphi_\gamma(\mathbf{r}_2)\varphi_\alpha(\mathbf{r}_3) \\ & -\varphi_\gamma(\mathbf{r}_1)\varphi_\beta(\mathbf{r}_2)\varphi_\alpha(\mathbf{r}_3) - \varphi_\beta(\mathbf{r}_1)\varphi_\alpha(\mathbf{r}_2)\varphi_\gamma(\mathbf{r}_3) - \varphi_\alpha(\mathbf{r}_1)\varphi_\gamma(\mathbf{r}_2)\varphi_\beta(\mathbf{r}_3)\Big) \end{aligned} \tag{B.4}$$

I wouldn't intend to write the four-electron wave function since it will be super long. This trouble is actually caused by working with real space. Since we need to label $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N$ for different degrees of freedom, we must use the Slater determinant to ensure the anti-symmetry property of the wave function, which unfortunately makes the expression very complicated. We

can get rid of this difficulty if not working with real space. To specify an electron in state α, instead of $\varphi_\alpha(\mathbf{r}_1)$, we write

$$|\alpha\rangle$$

which is known as the Dirac state [1]. Now, for a two-electron state, we write

$$|\alpha, \beta\rangle$$

But how do we ensure the anti-symmetry of this two-electron state?

$$|\alpha, \beta\rangle = -|\beta, \alpha\rangle \tag{B.5}$$

Previously, when working with real space, this was ensured by the Slater determinant. Now, this anti-symmetry will be taken care by the second quantization operators.[1]

$$|\alpha, \beta\rangle = c_\beta^\dagger c_\alpha^\dagger |0\rangle \tag{B.6}$$

These lovely operators have the property that if they change order, they produce a minus sign (the fermi-sign).

$$|\alpha, \beta\rangle = c_\beta^\dagger c_\alpha^\dagger |0\rangle = -c_\alpha^\dagger c_\beta^\dagger |0\rangle = -|\beta, \alpha\rangle \tag{B.7}$$

Surprisingly, the anti-symmetry property is automatically ensured! With second quantization, many-electron states could be written out with no pain.

For a one-electron state,

$$|\alpha\rangle = c_\alpha^\dagger |0\rangle \tag{B.8}$$

For a two-electron state,

$$|\alpha, \beta\rangle = c_\beta^\dagger c_\alpha^\dagger |0\rangle \tag{B.9}$$

For a three-electron state,

$$|\alpha, \beta, \gamma\rangle = c_\gamma^\dagger c_\beta^\dagger c_\alpha^\dagger |0\rangle \tag{B.10}$$

We do not need to worry about the anti-symmetry, because it is taken care by those operators automatically. This is the idea of second quantization. It must be pointed out, second quantization does not involve any new physics. Sometimes this name is misleading that people tend to ask, "Wait, what was the first quantization? Well, if the quantization of electron was the first, what is the second one?" No, no! Nothing is further quantized. Second quantization is just a novel "algebra" that simplifies the formalism of many-body problems.

Suppose we have a state

$$|\text{example}\rangle = \frac{1}{\sqrt{6}} \left(c_\alpha^\dagger c_\beta^\dagger c_\gamma^\dagger + 2 c_\delta^\dagger c_\epsilon^\dagger c_\zeta^\dagger + c_\eta^\dagger c_\theta^\dagger c_\iota^\dagger \right) |0\rangle \tag{B.11}$$

which is a linear combination of three different Slater determinants. It would be horrible to express it in real space (repeat Eqn. (B.4) three times). Second quantization provides us a very convenient way to handle many-body states.

[1] The order of the operators indicates we create α first and β second. Hence we write $c_\beta^\dagger c_\alpha^\dagger |0\rangle = |\alpha, \beta\rangle$.

B.2. Creation and annihilation operators

We start from the vacuum state $|0\rangle$, which is a state with no electron. Although without electron, it is defined to be normalized $\langle 0 \,|\, 0\rangle = 1$. Next, we introduce the creation operator $c_\alpha^\dagger$. If $c_\alpha^\dagger$ applies on a vacuum state, it creates one electron with state α,

$$c_\alpha^\dagger \,|0\rangle = |\alpha\rangle \tag{B.12}$$

"Hum? Create an electron out of vacuum?" No, no! We are not going to set up a lab to create electrons out of photons or whatever. This is purely an algebra. By saying "create", it is from a mathematics point of view, not a physical process. Similarly, we have an electron annihilation operator c_α. If c_α applies on a vacuum state, it returns zero

$$c_\alpha \,|0\rangle = 0 \tag{B.13}$$

Previously, we claimed that the creation operators anti-commute: $c_\alpha^\dagger c_\beta^\dagger = -c_\beta^\dagger c_\alpha^\dagger$. This is one of the definitions in second quantization. Now, the commutation relation between a creator and an annihilator is defined as

$$c_\alpha c_\beta^\dagger = \langle \alpha \,|\, \beta\rangle - c_\beta^\dagger c_\alpha \tag{B.14}$$

When working with orthonormal basis states, $\langle \alpha \,|\, \beta\rangle = \delta_{\alpha\beta}$. Let's see what happens if c_α applies on a state $|\alpha\rangle$

$$c_\alpha \,|\alpha\rangle = c_\alpha c_\alpha^\dagger \,|0\rangle = (1 - c_\alpha^\dagger c_\alpha) \,|0\rangle = |0\rangle - c_\alpha^\dagger \underbrace{c_\alpha \,|0\rangle}_{=0} = |0\rangle \tag{B.15}$$

As the name suggests, it removes one electron from $|\alpha\rangle$ and brings back the vacuum state. But this is purely an algebraic consequence, not a definition. The entire definition of second quantization algebra are summarized in Table B.1.

Table B.1.: The definition of second quantization algebra.

$\langle 0 \,\|\, 0\rangle$	$=$	1
$c_\alpha \,\|0\rangle$	$=$	0
$\{c_\alpha^\dagger, c_\beta^\dagger\}$	$=$	0
$\{c_\alpha, c_\beta\}$	$=$	0
$\{c_\alpha, c_\beta^\dagger\}$	$=$	$\langle \alpha \,\|\, \beta\rangle$

where the anti-commutator,

$$\{A, B\} \equiv AB + BA \tag{B.16}$$

Believe it or not, with simply five definitions, Table B.1 defines the complete system which formulates second quantization.

B.3. The bridge between first and second quantization

A two-electron Slater determinant in first quantization (real space),

$$\frac{1}{\sqrt{2}}(\varphi_\alpha(\mathbf{r}_1)\varphi_\beta(\mathbf{r}_2) - \varphi_\beta(\mathbf{r}_1)\varphi_\alpha(\mathbf{r}_2))$$

A two-electron Slater determinant in second quantization (configuration space),

$$c_\beta^\dagger c_\alpha^\dagger \left|0\right\rangle$$

However, these two are not the same:

$$c_\beta^\dagger c_\alpha^\dagger \left|0\right\rangle \neq \frac{1}{\sqrt{2}}(\varphi_\alpha(\mathbf{r}_1)\varphi_\beta(\mathbf{r}_2) - \varphi_\beta(\mathbf{r}_1)\varphi_\alpha(\mathbf{r}_2)) \tag{B.17}$$

A wave function is a wave function and a state is a state. They describe the same Slater determinant, but one cannot put an equal sign in between. To make the connection between real space and second quantization, we need some special electron creators and annihilators (called field operators). Although physically not possible, algebraically we can "create" an electron in such a state that it is exactly at position $\mathbf{r}$. We denote $c_\mathbf{r}^\dagger \left|0\right\rangle = \left|\mathbf{r}\right\rangle$.[2] Suppose we have an $\left|\alpha\right\rangle$ state which in real space corresponds to wave function $\varphi_\alpha(\mathbf{r})$. Considering $\varphi_\alpha(\mathbf{r})$ as an amplitude, $c_\alpha^\dagger$ and $c_\mathbf{r}^\dagger$ are (intuitively) related as

$$c_\alpha^\dagger = \int d^3r\, \varphi_\alpha(\mathbf{r}) c_\mathbf{r}^\dagger \tag{B.18}$$

Conversely, if we have a complete(!) set of single electron wave functions $\varphi_{\alpha_n}(\mathbf{r})$, we can expand the field operators in terms of the corresponding creators and annihilators

$$c_\mathbf{r}^\dagger = \sum_n \varphi_{\alpha_n}(\mathbf{r}) c_{\alpha_n}^\dagger \tag{B.19}$$

Using Eqn. (B.18), we find the anti-commutation relation

$$\{c_\mathbf{r}, c_\alpha^\dagger\} = \int d^3r'\, \varphi_\alpha(\mathbf{r}')\{c_\mathbf{r}, c_{\mathbf{r}'}^\dagger\} = \varphi_\alpha(\mathbf{r}) \tag{B.20}$$

which is such a golden relation that helps us bridge second quantization to real space. For example, a one-electron Slater determinant,

$$\left\langle \mathbf{r}_1 \,|\, \alpha \right\rangle = \left\langle 0\right| c_{\mathbf{r}_1} c_\alpha^\dagger \left|0\right\rangle = \left\langle 0\right| \varphi_\alpha(\mathbf{r}_1) - c_\alpha^\dagger c_{\mathbf{r}_1} \left|0\right\rangle = \varphi_\alpha(\mathbf{r}_1) \tag{B.21}$$

Nice! We get back our wave function in real space. Next, for a two-electron Slater determinant,

$$\begin{aligned}
\left\langle \mathbf{r}_2, \mathbf{r}_1 \,|\, \alpha, \beta \right\rangle &= \left\langle 0\right| c_{\mathbf{r}_1} c_{\mathbf{r}_2} c_\beta^\dagger c_\alpha^\dagger \left|0\right\rangle \\
&= \left\langle 0\right| c_{\mathbf{r}_1} (\varphi_\beta(\mathbf{r_2}) - c_\beta^\dagger c_{\mathbf{r}_2}) c_\alpha^\dagger \left|0\right\rangle \\
&= \left\langle 0\right| c_{\mathbf{r}_1} c_\alpha^\dagger \left|0\right\rangle \varphi_\beta(\mathbf{r}_2) - \left\langle 0\right| c_{\mathbf{r}_1} c_\beta^\dagger c_{\mathbf{r}_2} c_\alpha^\dagger \left|0\right\rangle \\
&= \varphi_\alpha(\mathbf{r}_1)\varphi_\beta(\mathbf{r}_2) - \varphi_\beta(\mathbf{r}_1)\varphi_\alpha(\mathbf{r}_2)
\end{aligned} \tag{B.22}$$

[2] Because of the importance of these special creators and annihilators, they get a name, field operators. A more standard way to write field operators are $\hat{\Psi}(\mathbf{r})$ and $\hat{\Psi}^\dagger(\mathbf{r})$ (see Reference [16]). But I would like to stick with $c_\mathbf{r}$ and $c_\mathbf{r}^\dagger$ to simplify our notations.

Impressive! Even the two-electron anti-symmetric wave function is automatically returned. A proof by induction is nicely discussed in Reference [16]. Here we quote the conclusion, for an N-electron state, its real-space Slater determinant representation is given by

$$\boxed{\Phi_{\alpha_1\cdots\alpha_N}(\mathbf{r}_1,\ldots,\mathbf{r}_N) = \frac{1}{\sqrt{N!}}\langle 0|\, c_{\mathbf{r}_1} c_{\mathbf{r}_2}\cdots c_{\mathbf{r}_N} c^\dagger_{\alpha_N}\cdots c^\dagger_{\alpha_2} c^\dagger_{\alpha_1}\,|0\rangle} \tag{B.23}$$

From our quantum mechanics lectures, we often see the relation,

$$\langle\alpha\,|\,\beta\rangle = \int d^3r\,\overline{\varphi_\alpha}(\mathbf{r})\varphi_\beta(\mathbf{r}) \tag{B.24}$$

This can also be shown using our field operators:

$$\begin{aligned}\langle\alpha\,|\,\beta\rangle &= \langle 0|\, c_\alpha c^\dagger_\beta\,|0\rangle = \langle 0|\int d^3r\,\overline{\varphi_\alpha}(\mathbf{r}) c_{\mathbf{r}}\varphi_\beta(\mathbf{r}) c^\dagger_{\mathbf{r}}\,|0\rangle \\ &= \int d^3r\,\overline{\varphi_\alpha}(\mathbf{r})\varphi_\beta(\mathbf{r})\underbrace{\langle 0|\, c_{\mathbf{r}} c^\dagger_{\mathbf{r}}\,|0\rangle}_{=1} = \int d^3r\,\overline{\varphi_\alpha}(\mathbf{r})\varphi_\beta(\mathbf{r})\end{aligned} \tag{B.25}$$

Similarly, for the two electron case,

$$\begin{aligned}\langle\alpha,\beta\,|\,\gamma,\delta\rangle &= \langle 0|\, c_\beta c_\alpha c^\dagger_\delta c^\dagger_\gamma\,|0\rangle = \langle 0|\, c_\alpha c_\beta c^\dagger_\gamma c^\dagger_\delta\,|0\rangle \\ &= \langle 0|\int d^3r_1\,\overline{\varphi_\alpha}(\mathbf{r}_1) c_{\mathbf{r}_1}\varphi_\delta(\mathbf{r}_1) c^\dagger_{\mathbf{r}_1}\int d^3r_2\,\overline{\varphi_\beta}(\mathbf{r}_2) c_{\mathbf{r}_2}\varphi_\gamma(\mathbf{r}_2) c^\dagger_{\mathbf{r}_2}\,|0\rangle \\ &= \int d^3r_1\int d^3r_2\,\overline{\varphi_\alpha}(\mathbf{r}_1)\overline{\varphi_\beta}(\mathbf{r}_2)\varphi_\gamma(\mathbf{r}_2)\varphi_\delta(\mathbf{r}_1)\underbrace{\langle 0|\, c_{\mathbf{r}_1} c^\dagger_{\mathbf{r}_1} c_{\mathbf{r}_2} c^\dagger_{\mathbf{r}_2}\,|0\rangle}_{=1}\end{aligned} \tag{B.26}$$

Those lovely operators $c_{\mathbf{r}}$ and $c^\dagger_{\mathbf{r}}$ play a role bridging first and second quantization. But they never appear explicitly in either first or second quantization!

B.4. Representation of n-body operators

In Chapter 4, we were working with the Coulomb repulsion Hamiltonian,

$$H_U = \sum_{i<j}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{B.27}$$

which is a two-body operator.

In Chapter 6, we introduced the spin-orbit coupling Hamiltonian,

$$H_{\mathrm{SO}} = \sum_{i=1}^{N}\xi(r_i)\boldsymbol{\ell}_i\cdot\mathbf{s}_i \tag{B.28}$$

which is a one-body operator.

Eqn. (B.27) and Eqn. (B.28) are in the form of the so called first quantization. They operate on real-space wave functions. A second quantization many-body state is, however, not compatible

with those operators. A beautiful discussion (you must give a look) of transforming real-space operators to second quantization operators is given in [16]. A key idea is to use the "bridge" in Eqn. (B.23). To avoid repeating the same content, I write down the results directly:

For the Coulomb repulsion Hamiltonian,

$$H_U = \frac{1}{2} \sum_{\alpha,\beta,\gamma,\delta} \langle \alpha, \beta | \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} | \gamma, \delta \rangle \, c_\alpha^\dagger c_\beta^\dagger c_\gamma c_\delta \tag{B.29}$$

which is given in Eqn. (4.2).

For the spin-orbit coupling Hamiltonian,

$$H_{\mathrm{SO}} = \sum_{\alpha,\beta} \langle \alpha | \, \xi(r) \boldsymbol{\ell} \cdot \mathbf{s} \, | \beta \rangle \, c_\alpha^\dagger c_\beta \tag{B.30}$$

which is given in Eqn. (6.44).

They are the Hamiltonians compatible with second quantization states.

B.5. Electron-hole transformation

In this section, we would like to restrict our discussion on atomic shell basis states instead of general states. In Eqn. (5.17), we made a convention that for a fully occupied shell, the electron creators are arranged in the following way:

	1	0	−1
↑	•	•	•
↓	•	•	•

$$= c_{-1\downarrow}^\dagger c_{0\downarrow}^\dagger c_{1\downarrow}^\dagger c_{-1\uparrow}^\dagger c_{0\uparrow}^\dagger c_{1\uparrow}^\dagger |0\rangle \tag{B.31}$$

Now we understand why it is important to make such a convention: the convention is arbitrary, but once it is decided, it must remain unchanged through the entire discussion, since changing the order of electron creators involves fermi-signs (± 1).

To motivate the topic of this section, let's consider an almost-full shell, say, a d^8:

	2	1	0	−1	−2
↑	•	•	•	•	
↓	•	•		•	•

We would write it in terms of electron creation operators as

$$c_{-2\downarrow}^\dagger c_{-1\downarrow}^\dagger c_{1\downarrow}^\dagger c_{2\downarrow}^\dagger c_{-1\uparrow}^\dagger c_{0\uparrow}^\dagger c_{1\uparrow}^\dagger c_{2\uparrow}^\dagger |0\rangle$$

This becomes a bit cumbersome and not very readable (but of course much simpler than its real-space form). We noticed that if we express the same state in terms of the unoccupied sites, the expression will become much shorter. What we need to do is to transform our "electron algebra" into a "hole algebra". Let's start from the fully occupied d shell

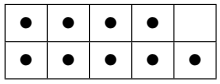
$$|\mathrm{full}\rangle = c_{-2\downarrow}^\dagger c_{-1\downarrow}^\dagger c_{0\downarrow}^\dagger c_{1\downarrow}^\dagger c_{2\downarrow}^\dagger c_{-2\uparrow}^\dagger c_{-1\uparrow}^\dagger c_{0\uparrow}^\dagger c_{1\uparrow}^\dagger c_{2\uparrow}^\dagger |0\rangle \tag{B.32}$$

Notice that a $|\mathrm{full}\rangle$ state also has $M_L = 0$ and $M_S = 0$. From the hole's point of view, the $|\mathrm{full}\rangle$ state behaves like a "vacuum" state. Creating a hole at site $(-m, -\sigma)$ on a $|\mathrm{full}\rangle$ state leaves

the system with momentum $M_L = m$ and $M_S = \sigma$. Hence we could define our hole creation operator as

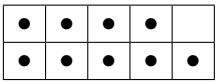
$$h^\dagger_{m\sigma} = c_{-m,-\sigma} \tag{B.33}$$

But this is not very convenient. Because what we really want is, for example,



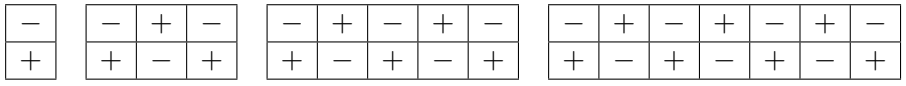
$$= h^\dagger_{2\downarrow} |\text{full}\rangle \tag{B.34}$$

However,



$$\begin{aligned}
&= c^\dagger_{-2\downarrow} c^\dagger_{-1\downarrow} c^\dagger_{0\downarrow} c^\dagger_{1\downarrow} c^\dagger_{2\downarrow} c^\dagger_{-1\uparrow} c^\dagger_{0\uparrow} c^\dagger_{1\uparrow} c^\dagger_{2\uparrow} |0\rangle \\
&= c^\dagger_{-2\downarrow} c^\dagger_{-1\downarrow} c^\dagger_{0\downarrow} c^\dagger_{1\downarrow} c^\dagger_{2\downarrow} \boxed{c_{-2\uparrow}} c^\dagger_{-2\uparrow} c^\dagger_{-1\uparrow} c^\dagger_{0\uparrow} c^\dagger_{1\uparrow} c^\dagger_{2\uparrow} |0\rangle \\
&= (-1)^5 \boxed{c_{-2\uparrow}} \underbrace{c^\dagger_{-2\downarrow} c^\dagger_{-1\downarrow} c^\dagger_{0\downarrow} c^\dagger_{1\downarrow} c^\dagger_{2\downarrow} c^\dagger_{-2\uparrow} c^\dagger_{-1\uparrow} c^\dagger_{0\uparrow} c^\dagger_{1\uparrow} c^\dagger_{2\uparrow} |0\rangle}_{|\text{full}\rangle} \\
&= -c_{-2\uparrow} |\text{full}\rangle = -h^\dagger_{2\downarrow} |\text{full}\rangle
\end{aligned} \tag{B.35}$$

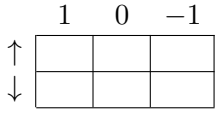
If we really want to write as the way in Eqn. (B.34), we must absorb the fermi-sign into definition (B.33). According to our full shell definition, this fermi-sign has the following pattern



Therefore, we define, (the definition is subject to how a $|\text{full}\rangle$ is defined)

$$\boxed{h^\dagger_{m\sigma} = (-1)^{l+m+\sigma-\frac{1}{2}} c_{-m,-\sigma}} \tag{B.36}$$

The next question is how to arrange these hole operators. Previously we made a convention of ordering electron creators. Now we no longer have this freedom to define new convention of ordering hole creators. As a consequence from previous convention, the hole creators should be ordered in the following way: (notice that it is the same order of putting electrons)

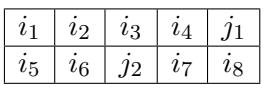


$$= h^\dagger_{1\uparrow} h^\dagger_{0\uparrow} h^\dagger_{-1\uparrow} h^\dagger_{1\downarrow} h^\dagger_{0\downarrow} h^\dagger_{-1\downarrow} |\text{full}\rangle \tag{B.37}$$

In general, an N-electron basis vector and a $(2(2l-1)-N)$-hole basis vector are equivalent by the relation,

$$\prod_{i=1}^{N} c^\dagger_{m_i\sigma_i} |0\rangle = \prod_{j=1}^{2(2l-1)-N} h^\dagger_{m_j\sigma_j} |\text{full}\rangle$$

The indices $\{j\}$ run over the complement part of indices $\{i\}$, for example,



(B.38)

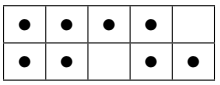
To show their equivalence,

$$\prod_{j=1}^{2(2l-1)-N} h^\dagger_{m_j\sigma_j} |\text{full}\rangle = \prod_{j=1}^{2(2l-1)-N} (-1)^{l+m_j+\sigma_j-\frac{1}{2}} c_{-m_j,-\sigma_j} \prod_{i=1}^{2(2l+1)} c^\dagger_{m_i\sigma_i} |0\rangle \tag{B.39}$$

But anti-commuting $c_{-m_j,-\sigma_j}$ into the full shell always cancels the $(-1)^{l+m_j+\sigma_j-\frac{1}{2}}$ factor (that is how this factor is designed for). After all anti-commutations, what left is,

$$\prod_{j=1}^{2(2l-1)-N} h^\dagger_{m_j\sigma_j} |\text{full}\rangle = \prod_{i=1}^{N} c^\dagger_{m_i\sigma_i} |0\rangle \tag{B.40}$$

We can verify that our example state



$$\begin{aligned}
&= h^\dagger_{0\uparrow} h^\dagger_{2\downarrow} |\text{full}\rangle = -c_{0\downarrow} c_{-2\uparrow} |\text{full}\rangle \\
&= -c_{0\downarrow} \boxed{c_{-2\uparrow}} c^\dagger_{-2\downarrow} c^\dagger_{-1\downarrow} c^\dagger_{0\downarrow} c^\dagger_{1\downarrow} c^\dagger_{2\downarrow} c^\dagger_{-2\uparrow} c^\dagger_{-1\uparrow} c^\dagger_{0\uparrow} c^\dagger_{1\uparrow} c^\dagger_{2\uparrow} |0\rangle \\
&= (-1)^6 c_{0\downarrow} c^\dagger_{-2\downarrow} c^\dagger_{-1\downarrow} c^\dagger_{0\downarrow} c^\dagger_{1\downarrow} c^\dagger_{2\downarrow} \boxed{c_{-2\uparrow}} c^\dagger_{-2\uparrow} c^\dagger_{-1\uparrow} c^\dagger_{0\uparrow} c^\dagger_{1\uparrow} c^\dagger_{2\uparrow} |0\rangle \\
&= (-1)^6 \boxed{c_{0\downarrow}} c^\dagger_{-2\downarrow} c^\dagger_{-1\downarrow} c^\dagger_{0\downarrow} c^\dagger_{1\downarrow} c^\dagger_{2\downarrow} c^\dagger_{-1\uparrow} c^\dagger_{0\uparrow} c^\dagger_{1\uparrow} c^\dagger_{2\uparrow} |0\rangle \\
&= (-1)^8 c^\dagger_{-2\downarrow} c^\dagger_{-1\downarrow} \boxed{c_{0\downarrow}} c^\dagger_{0\downarrow} c^\dagger_{1\downarrow} c^\dagger_{2\downarrow} c^\dagger_{-1\uparrow} c^\dagger_{0\uparrow} c^\dagger_{1\uparrow} c^\dagger_{2\uparrow} |0\rangle \\
&= c^\dagger_{-2\downarrow} c^\dagger_{-1\downarrow} c^\dagger_{1\downarrow} c^\dagger_{2\downarrow} c^\dagger_{-1\uparrow} c^\dagger_{0\uparrow} c^\dagger_{1\uparrow} c^\dagger_{2\uparrow} |0\rangle
\end{aligned} \tag{B.41}$$

is indeed equivalent to the expression in terms of electron creators.

Eqn. (B.36) is essentially the portal connecting the electron world with the hole world. With this definition, we can investigate the properties of hole operators. First of all, the $|\text{full}\rangle$ state is normalized, $\langle\text{full}\,|\,\text{full}\rangle = 1$. If we apply a hole annihilation operator on it, according to Pauli exclusion principle, $h_{m\sigma}|\text{full}\rangle = 0$. Now, we test the anti-commutation relation between hole operators.

$$\{h^\dagger_{m_1\sigma_1}, h^\dagger_{m_2\sigma_2}\} = (-1)^{2l+m_1+m_2+\sigma_1+\sigma_2-1}\{c_{-m_1,-\sigma_1}, c_{-m_2,-\sigma_2}\} = 0 \tag{B.42}$$

$$\{h_{m_1\sigma_1}, h_{m_2\sigma_2}\} = (-1)^{2l+m_1+m_2+\sigma_1+\sigma_2-1}\{c^\dagger_{-m_1,-\sigma_1}, c^\dagger_{-m_2,-\sigma_2}\} = 0 \tag{B.43}$$

$$\begin{aligned}
\{h_{m_1\sigma_1}, h^\dagger_{m_2\sigma_2}\} &= (-1)^{2l+m_1+m_2+\sigma_1+\sigma_2-1}\{c^\dagger_{-m_1,-\sigma_1}, c_{-m_2,-\sigma_2}\} \\
&= (-1)^{2l+m_1+m_2+\sigma_1+\sigma_2-1}\delta_{m_1m_2}\delta_{\sigma_1\sigma_2} = \delta_{m_1m_2}\delta_{\sigma_1\sigma_2}
\end{aligned} \tag{B.44}$$

It turns out (not by definition), our hole operators behave exactly as electron operators.

Table B.2.: Second quantization algebra from the hole's perspective.

$\langle\text{full}\,\|\,\text{full}\rangle$	$=$	1
$h_{m\sigma}\|\text{full}\rangle$	$=$	0
$\{h^\dagger_{m_1\sigma_1}, h^\dagger_{m_2\sigma_2}\}$	$=$	0
$\{h_{m_1\sigma_1}, h_{m_2\sigma_2}\}$	$=$	0
$\{h_{m_1\sigma_1}, h^\dagger_{m_2\sigma_2}\}$	$=$	$\delta_{m_1m_2}\delta_{\sigma_1\sigma_2}$

Acknowledgements

I am heartily thankful to my supervisor, Prof. Dr. Erik Koch, for his guidance, encouragement and support throughout my entire thesis. He was always so patient to answer my frequent questions and guided me to think and solve problems systematically. I also would like to thank Dr. Hermann Ulm who promised me that I can ask him any question. He helped me a lot in understanding many detailed problems and gave me very strong encouragement through my thesis. I am grateful to German Research School for Simulation Sciences for providing me this great opportunity of doing my master thesis. I am also thankful to Dr. Hunter Sims, Khaldoon Ghanem and Michael Baumgärtel for their many different helps during my thesis period. In the end, I want to express my many thanks to Cica Gustiani, Rodrigo Canales and Thu Hoai for our best memories studying and working together.

Bibliography

[1] D. J. Griffiths: *Introduction to Quantum Mechanics*, Pearson Prentice Hall (2005).

[2] R. M. Martin: *Electronic Structure*, Cambridge University Press (2004).

[3] W. Kohn and L. J. Sham: *Self-Consistent Equations Including Exchange and Correlation Effects*, Physical Review A **140**, 1133 (1965).

[4] J. M. MacLaren, D. P. Clougherty, M. E. McHenry and M. M. Donovan: *Parameterised local spin density exchange-correlation energies and potentials for electronic structure calculations*, Computer Physics Communications **66**, 383–391 (1991).

[5] S. Kotochigova, Z. H. Levine, E. L. Shirley, M. D. Stiles and C. W. Clark: *Atomic Reference Data for Electronic Structure Calculations*, Physical Measurement Laboratory, NIST. Link: `http://www.nist.gov/pml/data/dftdata/`

[6] M. Weissbluth: *Atoms and Molecules*, Academic Press (1978).

[7] D. J. Griffiths: *Introduction to Electrodynamics*, Prentice Hall (1999).

[8] N. J. Giordano and H. Nakanishi: *Computational Physics*, Pearson Prentice Hall (2006).

[9] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery: *Numerical Recipes in C*, Cambridge University Press (1992).

[10] S. E. Koonin and D. C. Meredith: *Compututational Physics*, Westview Press (1990).

[11] G. Racah: *Theory of Complex Spectra. I*, Physical Review **61**, 186 (1942).

[12] G. Racah: *Theory of Complex Spectra. II*, Physical Review **62**, 438 (1942).

[13] G. Racah: *Theory of Complex Spectra. III*, Physical Review **63**, 367 (1943).

[14] G. Racah: *Theory of Complex Spectra. IV*, Physical Review **76**, 1352 (1949).

[15] E. Koch: *Exchange Mechanisms*, in [18].

[16] E. Koch: *Many-Electron States*, in [19].

[17] R. Eder: *Multiplets in Transition Metal Ions*, in [18].

[18] E. Pavarini, E. Koch, F. Anders and M. Jarrell (eds.): *Correlated Electrons: From Models to Materials*, Reihe Modeling and Simulation, Vol. 2, Forschungszentrum Jülich (2012).

[19] E. Pavarini, E. Koch and U. Schollwöck (eds.): *Emergent Phenomena in Correlated Matter*, Reihe Modeling and Simulation, Vol. 3, Forschungszentrum Jülich (2013).

Declaration of authorship

I, Qian Zhang, hereby declare that I have created this work completely on my own and used no other sources or tools than the ones listed, and that I have marked any citations accordingly.

Aachen, September 2014

Qian Zhang