

1 Introduction to quantum mechanics

The theory of quantum mechanics explores the behavior of matter and energy at atomic scales. At these scales, the classical rules of physics we are familiar with no longer apply. This introduction is designed for graduate students in mathematics with little to no background in physics. We will use the famous double-slit experiment as a gateway to uncover key ideas in quantum mechanics, culminating in the electronic Schrödinger equation, a fundamental tool for studying particles at the atomic level. We will come to the conclusion that electrons behave like both particles and waves which is known as the particle-wave duality. As most readers will be – at least at the intuitive level – familiar with classical mechanics, we will begin our exposition with classical physics and then transition into the quantum theory of particles.

Point particles vs classical waves

In classical physics, waves and point particles are two very different models for physical systems, each with an exceptionally large range of applications.

Point particles which are idealized particles with no spatial extension are very commonly used in physics to describe the behavior of (uniform) rigid body dynamics, such as the “projectile motion” or “billiard mechanics”. The notion of (point) particles is governed by Newtonian mechanics going back to the 17th century, which is manifested in three Newtonian Law’s of motion

1. A body remains at rest, or in motion at a constant speed in a straight line, except insofar as it is acted upon by a force.
2. At any instant of time, the net force on a body is equal to the rate at which the body’s momentum is changing with time.

$$F = \frac{dp}{dt} = m \frac{dv}{dt} = ma \quad (1)$$

3. If two bodies exert forces on each other, these forces have the same magnitude but opposite directions.

The describing quantities of (point) particles are (1) their mass, (2) their position, and (3) their forces.

Classical waves are propagating dynamic disturbances, i.e., changes from an equilibrium state. Periodic waves oscillate repeatedly about an equilibrium (resting) value at some frequency. When the entire waveform moves in one direction, it is said to be a traveling wave; by contrast, a pair of superimposed periodic waves traveling in opposite directions makes a standing wave. In a standing wave, the amplitude of vibration has nulls at some positions where the wave amplitude appears smaller or even zero. Commonly, a wave is described by a scalar field

$$u : \mathbb{R}^3 \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R} ; (\mathbf{r}, t) \mapsto u(\mathbf{r}, t) \quad (2)$$

assigning at any point in time $t \in \mathbb{R}$ and any point in space $\mathbf{r} \in \mathbb{R}^3$ a magnitude $u(\mathbf{r}, t)$. Classical waves are governed by the wave equation, i.e.,

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta_{\mathbf{r}} u, \quad (3)$$

where $c \in \mathbb{R}_{>0}$. Unlike point particles, waves have continuous values at many points in space that vary with time. The spatial extent of waves can vary with time due to diffraction or wave interferences, see Fig. 1. These are characteristic observations of waves.

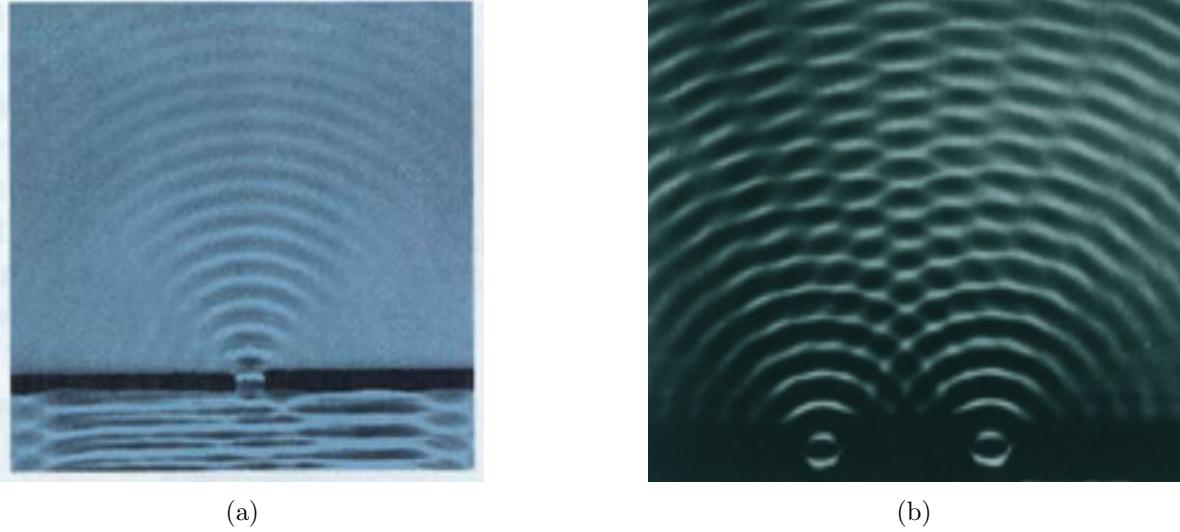


Figure 1: (a) Left panel shows diffraction of water wave on a single slit. (b) Right panel shows interference pattern of concentric waves.

Unlike waves, point particles do not exhibit diffraction or interference.

The Double-Slit Experiment

We will transition to atomic scales; imagine an apparatus where particles at the atomic scale, such as electrons, are fired onto a detector, see Fig. 2

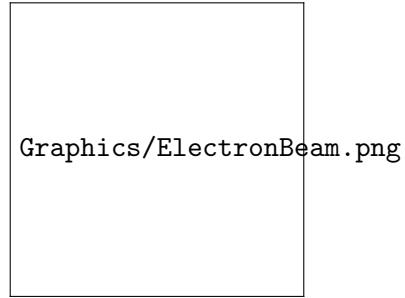


Figure 2: electron beam

The particles can be individually detected and appear to behave like point particles in classical physics. Now, imagine a setup where the particles are sent toward a barrier with two small openings, called slits, see 3. Beyond the slits is a screen that detects where the particles land.

We now consider two scenarios:

1. **Blocking one slit:** If we close one slit, the particles form a simple pattern on the screen directly behind the open slit, see ???. This result again aligns with our expectation from the point particle picture.

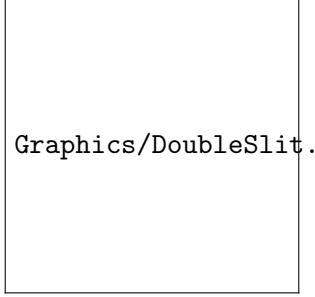


Figure 3: electron beam

2. **Opening both slits:** With both slits open, we expect the result to be the sum of the patterns from each slit. However, what we observe is far more surprising; a series of bright and dark bands, called an interference pattern, similar to what we see if waves (not particles) were passing through the slits, see 3.

This experiment shows that particles such as electrons behave like waves under certain conditions. More surprisingly, if we place detectors at the slits to observe which slit the individual electron passes through, see ??, the interference pattern disappears, and the electrons behave like particles again. This puzzling behavior is at the heart of quantum mechanics.

But, how do we interpret this result and what does it mean for our understanding of physics? The double-slit experiment suggests that particles have a dual nature: they can behave like both, particles and waves. This is called the *wave-particle duality* challenging our classical picture of physics in which these concepts are fundamentally different. The quantum mechanical framework allows us to describe the interference pattern arising in the dual slit experiment. Quantum mechanically, particles are characterized by a *wave function*, which determines the probability of finding the particle in a particular location. As the name suggests, the propagation of the particle's wave function is governed by “some type of wave equation”. In particular, when no measurement is made, the particle's wave function spreads out like a wave producing the wave-like interference patterns observed in the double-slit experiment. However, when we measure which slit the particle goes through, the wave function “collapses” into a specific location, and the interference pattern vanishes – the electron now behaves like a classical particle.

Mathematically, the wave function is a function in space and time, denoted by

$$\psi : \mathbb{R}^m \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{C} ; (\mathbf{x}, t) \mapsto \psi(\mathbf{x}, t) \quad (4)$$

where $m \in \mathbb{N}$. The wave function is the central object in quantum mechanics that encapsulates all the information about the system's state. Its squared modulus is the probability density of finding the particles at positions $\mathbf{x} \in \mathbb{R}^m$ at time $t \in \mathbb{R}_{\geq 0}$, hence, $\psi \in L^2(\mathbb{R}^m)$ for all $t \in \mathbb{R}_{\geq 0}$. The governing equation for the wave function is the Schrödinger equation, proposed by Erwin Schrödinger in 1925 [?]. In its time-dependent form, the Schrödinger equation reads

$$i \frac{\partial \psi(\mathbf{x}, t)}{\partial t} = H\psi(\mathbf{x}, t), \quad (5)$$

where i is the imaginary unit and H is the Hamiltonian operator, representing the total energy

of the system of interest. For a particle moving in a potential $V(\mathbf{x})$, the Hamiltonian is

$$H = -\frac{1}{2}\nabla^2 + V(\mathbf{x}), \quad (6)$$

where $-\frac{1}{2}\nabla^2$ describes the kinetic energy, and $V(\mathbf{x})$ describes the potential energy. In many cases, however, we are interested in systems that do not change in time, such as electrons in a stable atom or molecule. Here, the Schrödinger equation simplifies to

$$H\psi(x) = E\psi(x), \quad (7)$$

where E describes the system's energy. Narrowing down on the application to atoms and molecules, we can specify the potential further. Consider a molecule with N_{nuc} nuclei and N electrons, the wave function depends on both nuclear positions $\mathbf{R} = (\mathbf{R}_1, \dots, \mathbf{R}_{N_{\text{nuc}}}) \in \mathbb{R}^{3N_{\text{nuc}}}$ with $\mathbf{R}_i \in \mathbb{R}^3$ and the electronic positions $\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{3N}$ with $\mathbf{r}_i \in \mathbb{R}^3$

$$H\Psi(\mathbf{r}, \mathbf{R}) = E\Psi(\mathbf{r}, \mathbf{R}) \quad (8)$$

where

$$\begin{aligned} H = & - \sum_{i \in [N]} \frac{1}{2} \Delta_{\mathbf{r}_i} - \sum_{i \in [N]} \sum_{j \in [N_{\text{nuc}}]} \frac{Z_j}{\|\mathbf{r}_i - \mathbf{R}_j\|} + \frac{1}{2} \sum_{i \in [N]} \sum_{i \neq j \in [N]} \frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|} \\ & - \sum_{i \in [N_{\text{nuc}}]} \frac{1}{2M_i} \Delta_{\mathbf{R}_i} + \frac{1}{2} \sum_{i \in [N_{\text{nuc}}]} \sum_{i \neq j \in [N_{\text{nuc}}]} \frac{Z_i Z_j}{\|\mathbf{R}_i - \mathbf{R}_j\|} \end{aligned} \quad (9)$$

where M_i is the nuclear mass relative to the electronic mass, and Z_i are the atomic numbers found on the periodic table. A common simplification is the Born-Oppenheimer approximation, which assumes that nuclei move much more slowly than electrons, meaning that the kinetic potential energy coming exclusively from nuclei simply enter as a constant. This reduces the problem to the electronic Schrödinger equation:

$$H(\mathbf{R})\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \quad (10)$$

where

$$H(\mathbf{R}) = - \sum_{i \in [N]} \frac{1}{2} \Delta_{\mathbf{r}_i} - \sum_{i \in [N]} \sum_{j \in [N_{\text{nuc}}]} \frac{Z_j}{\|\mathbf{r}_i - \mathbf{R}_j\|} + \frac{1}{2} \sum_{i \in [N]} \sum_{i \neq j \in [N]} \frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|}. \quad (11)$$

Subject of this topic course will be to solve this innocent looking eigenvalue problem.

2 The hydrogen atom

For the hydrogen atom, the (stationary) Schrödinger equation takes the simple form

$$\left(-\frac{1}{2}\Delta - \frac{1}{\|\mathbf{R} - \mathbf{r}\|}\right)\Psi(\mathbf{r}) = E\Psi(\mathbf{r}), \quad (12)$$

where $\mathbf{R} \in \mathbb{R}^3$ denotes the position of the hydrogen nucleus. Without loss of generality, we can define our coordinate system such that the nucleus defines the coordinate origin simplifying the equation to

$$\left(-\frac{1}{2}\Delta - \frac{1}{\|\mathbf{r}\|}\right)\Psi(\mathbf{r}) = E\Psi(\mathbf{r}). \quad (13)$$

We note that this is a spherically symmetric problem. This can be seen as follows: Consider the rotation matrix R , then

$$\mathbf{r} \mapsto R\mathbf{r} = \mathbf{r}' \quad (14)$$

describes a rotation of the frame. Since R is unitary and the 2-norm is unitarily invariant, we have

$$\|\mathbf{r}'\| = \|R\mathbf{r}\| = \|\mathbf{r}\|. \quad (15)$$

Moreover, in the rotated frame the partial derivative is given by

$$\partial_{x'_i} = \sum_j R_{i,j} \partial_{x_j} \quad \Rightarrow \quad \partial_{x'_i}^2 = \sum_{k,j} R_{i,k} R_{j,k} \partial_{x_j} \partial_{x_k}. \quad (16)$$

Then

$$\Delta' f = \sum_i \sum_{k,j} R_{i,k} R_{j,k} \partial_{x_j} \partial_{x_k} f = \sum_{k,j} \left(\sum_i R_{i,k} R_{j,k} \right) \partial_{x_j} \partial_{x_k} f = \sum_k \partial_{x_k}^2 f = \Delta f. \quad (17)$$

This suggests that we investigate this problem in spherical coordinates. Recall the spherical coordinate transformation

$$s : [0, \pi] \times [0, 2\pi] \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^3 ; \quad \begin{bmatrix} \theta \\ \phi \\ r \end{bmatrix} \mapsto \begin{bmatrix} r \sin(\theta) \cos(\phi) \\ r \sin(\theta) \sin(\phi) \\ r \cos(\theta) \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (18)$$

and

$$s^{-1} : \mathbb{R}^3 \rightarrow [0, \pi] \times [0, 2\pi] \times \mathbb{R}_{\geq 0} ; \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} \phi \\ \theta \\ r \end{bmatrix} = \begin{bmatrix} \phi \\ \theta \\ r \end{bmatrix} \quad (19)$$

The Laplacian in spherical coordinates is given by

$$\begin{aligned} \Delta &= \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \\ &=: \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{1}{r^2} \mathcal{L}^2 \end{aligned} \quad (20)$$

where we defined

$$\mathcal{L}^2 = -\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) - \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2}. \quad (21)$$

The separation of the radial and angular part in Eq. (20) suggests to use the separation of variables ansatz, i.e.,

$$\Psi(r, \theta, \phi) = R(r)Y(\theta, \phi). \quad (22)$$

Inserting this ansatz into the Schrödinger equation (12), this yields

$$\begin{aligned} & \left(-\frac{1}{2} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{2} \frac{1}{r^2} \mathcal{L}^2 - \frac{1}{|r|} \right) R(r)Y(\theta, \phi) = ER(r)Y(\theta, \phi) \\ \Leftrightarrow & \left(-\frac{1}{2} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{1}{|r|} - E \right) R(r) + \frac{R(r)}{2r^2} \frac{\mathcal{L}^2 Y(\theta, \phi)}{Y(\theta, \phi)} = 0 \end{aligned} \quad (23)$$

We see that solving this equation is subject to characterizing the spectrum of \mathcal{L}^2 .

2.0.1 Spectrum of \mathcal{L}^2

We seek to solve the eigenvalue problem

$$\mathcal{L}^2 Y(\theta, \phi) = k Y(\theta, \phi). \quad (24)$$

Given the structure of the operator \mathcal{L}^2 , see Eq. (21), we again propose the separation of variables ansatz

$$Y(\theta, \phi) = \Theta(\theta)\Phi(\phi), \quad (25)$$

which yields

$$\begin{aligned} & \left(-\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) - \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) \Theta(\theta)\Phi(\phi) = k\Theta(\theta)\Phi(\phi) \\ \Leftrightarrow & \frac{\sin^2 \theta}{\Theta(\theta)} \left(-\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) - k \right) \Theta(\theta) - \frac{1}{\Phi(\phi)} \frac{\partial^2}{\partial \phi^2} \Phi(\phi) = 0 \end{aligned} \quad (26)$$

We first separate $\Phi(\phi)$ which yields

$$\begin{cases} -\frac{\partial^2}{\partial \phi^2} \Phi(\phi) = m^2 \Phi(\phi) \\ \Phi(0) = \Phi(2\pi) \end{cases} \quad (27)$$

with separation constant m^2 , and periodic boundary condition because of the spherical symmetry of the problem. The solution to Eq. (27) is given by

$$\Phi(\phi) = A e^{im\phi}, \quad (28)$$

with $m \in \mathbb{Z}$ to fulfill the azimuthal periodic boundary condition.

For $\Theta(\theta)$ we then obtain

$$\left(-\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{m^2}{\sin^2 \theta} \right) \Theta(\theta) = k \Theta(\theta) \quad (29)$$

Performing the coordinate transformation

$$\zeta := \cos(\theta) \quad \text{and} \quad \xi(\cos(\theta)) := \Theta(\theta) \quad (30)$$

yields

$$\frac{d}{d\zeta} \left((1 - \zeta)^2 \frac{d\xi}{d\zeta} \right) + \left(k - \frac{m^2}{1 - \zeta^2} \right) \xi = 0 \quad (31)$$

with $|\xi(1)|, |\xi(-1)| < \infty$. This differential equation is a (well-)known differential equation, namely, the generalized Legendre differential equation [?]. Recall the conventional Legendre equation, i.e.,

$$(1 - x^2) \frac{d^2y}{dx^2} - 2x \frac{dy}{dx} + \ell(\ell + 1)y = 0, \quad (32)$$

where ℓ is the degree. This equation describes solutions to Laplace's equation for azimuthally symmetric boundary conditions, i.e., when the solution has no dependence on the azimuthal angle ϕ . However, to account for non-zero azimuthal dependence, particularly in gravitational and electromagnetic potentials, the equation was extended to the generalized Legendre differential equation, i.e.,

$$(1 - x^2) \frac{d^2y}{dx^2} - 2x \frac{dy}{dx} + \left[\ell(\ell + 1) - \frac{m^2}{1 - x^2} \right] y = 0, \quad (33)$$

introducing the order m , which corresponds to the projection of angular momentum along the z -axis. The solutions to this equation are the associated Legendre functions $P_\ell^m(x)$, which are derived from the Legendre polynomials $P_\ell(x)$ using the relation

$$P_\ell^m(x) = (1 - x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_\ell(x). \quad (34)$$

These functions generalize $P_\ell(x)$ by incorporating azimuthal dependence and are crucial in describing solutions to Laplace's equation in spherical coordinates for problems with full rotational symmetry. This formalism, developed to handle problems involving angular momentum and wave-like behavior, became essential in physics, geophysics, and engineering.

The Eq. (31) is a Legendre differential equation on $[-1, 1]$, which yields

$$k = \ell(\ell + 1), \quad \ell \in \mathbb{Z}_{\geq 0} \quad (35)$$

Each eigenvalue $\ell(\ell + 1)$ corresponds to $\ell + 1$ degenerate, orthogonal eigenfunctions, denoted

$$\xi(\zeta) = P_\ell^m(\zeta) := (-1)^m (1 - \zeta^2)^{\frac{m}{2}} \frac{d^m}{d\zeta^m} P_\ell(\zeta), \quad m \in \{0, \dots, \ell\} \quad (36)$$

where P_ℓ^m are the associated Legendre polynomials and

$$P_\ell(\zeta) = \frac{(-1)^\ell}{2^\ell \ell!} \frac{d^\ell}{d\zeta^\ell} (1 - \zeta^2)^\ell \quad (37)$$

are the Legendre polynomials [?]. We also define the *negative associated Legendre polynomials* as

$$P_\ell^{-m}(\zeta) = (-1)^m \frac{(\ell - m)!}{(\ell + m)!} P_\ell^m(\zeta) \quad (38)$$

Example 2.0.1. The first six Legendre polynomials are given by

$$\begin{aligned} P_0(\zeta) &= 1 & P_1(\zeta) &= \zeta & P_2(\zeta) &= \frac{1}{2}(3\zeta^2 - 1) \\ P_3(\zeta) &= \frac{1}{2}(5\zeta^3 - 3\zeta) & P_4(\zeta) &= \frac{1}{8}(35\zeta^4 - 30\zeta^2 + 3) & P_5(\zeta) &= \frac{1}{8}(63\zeta^5 - 70\zeta^3 + 15\zeta) \end{aligned}$$

The first few associated Legendre polynomials are give by

$$\begin{aligned} P_0^0(x) &= 1 \\ P_1^{-1}(x) &= \frac{(1-x^2)^{1/2}}{2} & P_1^0(x) &= x & P_1^1(x) &= -(1-x^2)^{1/2} \\ P_2^{-2}(x) &= \frac{3(1-x^2)}{24} & P_2^{-1}(x) &= \frac{3x(1-x^2)^{1/2}}{6} & P_2^0(x) &= \frac{(3x^2-1)}{2} & P_2^1(x) &= -3x(1-x^2)^{1/2} & P_2^2(x) &= 3(1-x^2) \end{aligned}$$

Changing the variables back to θ and $\Theta(\theta)$, this yields that

$$\Theta(\theta) = P_\ell^m(\cos(\theta)) \quad (39)$$

Hence

$$Y_{\ell,m}(\theta, \phi) = C_{\ell,m} P_\ell^m(\cos(\theta)) e^{im\phi} \quad (40)$$

where $C_{\ell,m}$ a normalization constant which is chosen such that

$$\begin{aligned} &\int_{[0,\phi] \times [0,2\pi)} Y_{\ell,m}(\theta, \phi) Y_{\ell',m'}(\theta, \phi) \sin^2(\theta) d\theta d\phi = \delta_{\ell,\ell'} \delta_{m,m'} \\ \Leftrightarrow \quad C_{\ell,m} &= (-1)^m \sqrt{\frac{(2\ell+1)}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}} \end{aligned} \quad (41)$$

The functions $Y_{\ell,m}(\theta, \phi)$ are the *spherical harmonics* fulfilling

$$\mathcal{L}^2 Y_{\ell,m} = \ell(\ell+1) Y_{\ell,m} \quad (42)$$

where the eigenvalue had geometric multiplicity $2\ell+1$.

Example 2.0.2. Let's consider $\ell = 1$ and $m = 0$, then

$$Y_{1,0}(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \cos(\theta). \quad (43)$$

We see that $Y_{1,0}$ only depends on the polar angle θ and not on the azimuthal angle ϕ . A possible way to visualize this function is via the map

$$f_\ell^m : (\theta, \phi) \mapsto P_\ell^m(\cos(\theta)) e^{im\phi} \begin{bmatrix} \sin(\theta) \sin(\phi) \\ \sin(\theta) \cos(\phi) \\ \cos(\theta) \end{bmatrix} \quad (44)$$

Using this procedure, we can visualize $Y_{1,0}$ as shown in Fig. 4.

This procedure can similarly be applied to the first few spherical harmonics in Fig. 5.

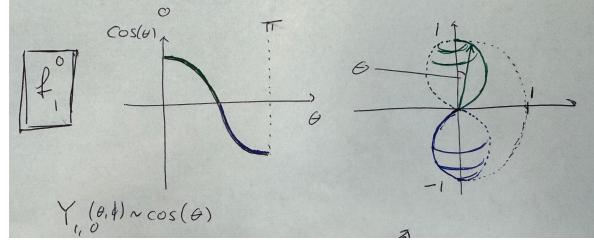


Figure 4: Construction of a visualization of $Y_{1,0}$.

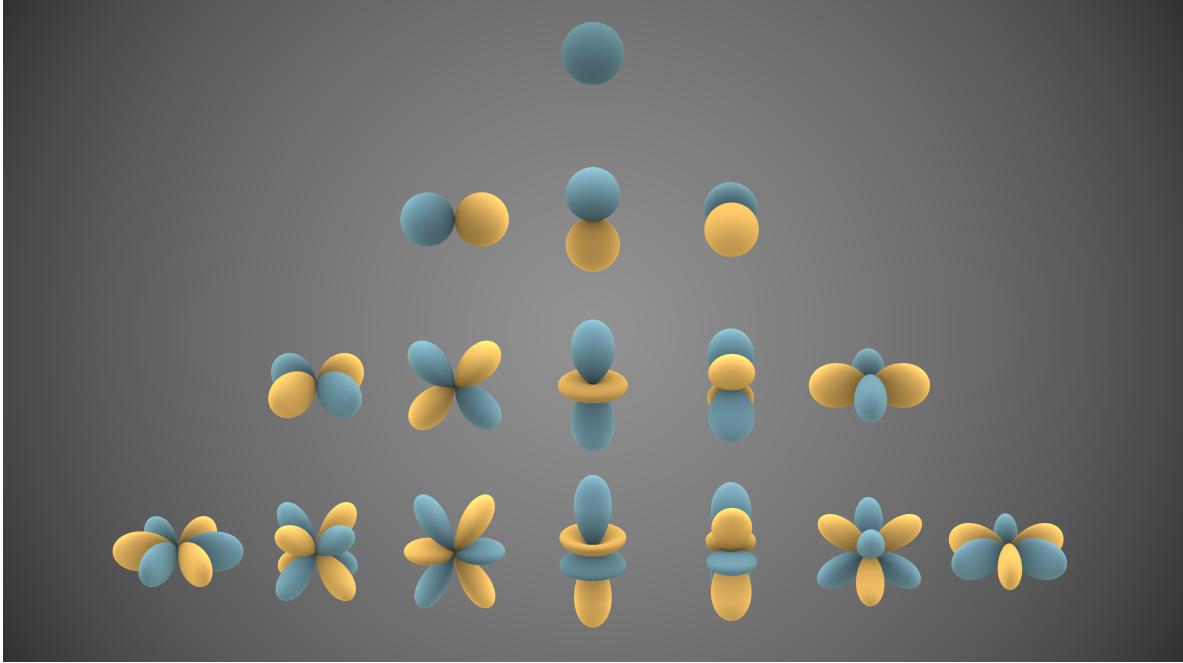


Figure 5: Graphical representation of the first few spherical harmonics, using the function f_ℓ^m in Eq. (44). Blue portions represent regions where the function is positive, and yellow portions represent where it is negative. The rows correspond to $\ell = 0$ (s), $\ell = 1$ (p), $\ell = 2$ (d), $\ell = 3$ (f); the columns correspond to $m \in \{-\ell, \dots, \ell\}$.

We shall now return to Eq. (23). Having found the spectrum of \mathcal{L}^2 , Eq. (23) simplifies to

$$-\frac{1}{2} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R(r)}{\partial r} \right) - \frac{1}{r} R(r) + \frac{R(r)}{2r^2} \ell(\ell + 1) = ER(r), \quad r > 0, \quad (45)$$

which is the radial equation. Note that

$$\begin{aligned} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R(r)}{\partial r} \right) &= \frac{1}{r^2} \left(2r \frac{\partial R(r)}{\partial r} + r^2 \frac{\partial^2 R(r)}{\partial r^2} \right) = \left(\frac{2}{r} \frac{\partial R(r)}{\partial r} + \frac{\partial^2 R(r)}{\partial r^2} \right) \\ &= \frac{1}{r} \frac{\partial}{\partial r} \left(R(r) + r \frac{\partial R(r)}{\partial r} \right) = \frac{1}{r} \frac{\partial^2}{\partial r^2} r R(r) \end{aligned} \quad (46)$$

Substituting

$$u(r) = rR(r) \quad (47)$$

yields

$$\begin{aligned} & -\frac{1}{2} \frac{\partial^2}{\partial r^2} r R(r) - R(r) + \frac{R(r)}{2r} \ell(\ell+1) = Er R(r) \\ \Leftrightarrow \quad & -\frac{1}{2} \frac{\partial^2}{\partial r^2} u(r) + \tilde{V}(r) u(r) = Eu(r) \end{aligned} \quad (48)$$

where

$$\tilde{V}(r) = \frac{\ell(\ell+1)}{2r^2} - \frac{1}{r}. \quad (49)$$

We will now investigate the solution to Eq. (48) in two limiting cases. First, in the far field. We note that

$$\tilde{V}(r) \xrightarrow{r \rightarrow \infty} 0 \quad (50)$$

which simplifies Eq. (48) to

$$-\frac{1}{2} \frac{\partial^2}{\partial r^2} u(r) = Eu(r). \quad (51)$$

For $E > 0$ the solution to Eq. (51) yields

$$u(r) = c_1 e^{i\sqrt{E}2r} + c_2 e^{-i\sqrt{E}2r}, \quad (52)$$

which cannot be square integrable, hence, any value $E > 0$ cannot be an isolated eigenvalue. For $E \leq 0$ the solution is given by

$$u(r) = c_1 e^{\sqrt{|E|}2r} + c_2 e^{-\sqrt{|E|}2r}. \quad (53)$$

Since this solution must remain finite as $r \rightarrow \infty$, c_1 must be zero, hence,

$$u(r) = ce^{-\sqrt{|E|}2r}. \quad (54)$$

Second, in the near field. We note that

$$\tilde{V}(r) \xrightarrow{r \rightarrow 0} -\frac{\ell(\ell+1)}{2r^2} \quad \text{and} \quad u(r) \xrightarrow{r \rightarrow 0} 0 \quad (55)$$

which yields

$$-\frac{1}{2} \frac{\partial^2}{\partial r^2} u(r) - \frac{\ell(\ell+1)}{r^2} u(r) = 0. \quad (56)$$

Equation (56) has the solution

$$u(r) = c_1 r^{\ell+1} + c_2 r^{-\ell} \quad (57)$$

but since $r^{-\ell} \rightarrow \infty$ for $r \rightarrow 0$, c_2 must be zero, hence,

$$u(r) = cr^{\ell+1}. \quad (58)$$

Combining the near field and far field solution then yields the general solution

$$u(r) = Cr^{\ell+1} e^{-\sqrt{\frac{|E|}{2}}r} G(r) \quad (59)$$

where

$$G(r) = A_0 + A_1 r + A_2 r^2 + \dots \quad (60)$$

“stitches” the near field and far field solution together. Inserting this ansatz into Eq. (48) yields the recursion

$$A_k = -2A_{k-1} \frac{1 - (\ell + k)\sqrt{-E2}}{(\ell + k)(\ell + k + 1) - \ell(\ell + 1)}. \quad (61)$$

For this series to terminate we require

$$0 = 1 - (\ell + k)\sqrt{-2E} \quad \Leftrightarrow \quad E_{k,\ell} = -\frac{1}{2(\ell + k)^2} \quad (62)$$

3 Spin

We now want to transition to multi-electron systems. To that end, we need to address a fundamental quantity that is – surprisingly – not explicitly incorporated into the electronic Schrödinger equation, namely, the spin. Understanding the spin of an electron is a cornerstone of quantum mechanics, yet it can be an abstract concept for learners. The Stern-Gerlach experiment provides an intuitive and visualizable way to grasp the key ideas surrounding spin. In the early 20th century, physicists were intrigued by the idea that electrons and other particles possess intrinsic angular momentum beyond what classical physics could explain. This intrinsic property, later named *spin*, was suspected but not clearly understood. The Stern-Gerlach experiment, conducted in 1922 by Otto Stern and Walther Gerlach, aimed to probe the nature of this intrinsic property by observing the deflection of particles in a magnetic field.

The core idea of the experiment is deceptively simple. Here's a breakdown of the essential components, see Fig. 6:

- Source of Silver Atoms: The experiment initially used silver atoms – which contain a single valence electron in an unpaired state (Ag has 47 electrons) – as a convenient proxy for electrons.
- Collimated Beam: The silver atoms are heated and directed into a narrow beam.
- Non-Uniform Magnetic Field: The atoms pass through a region of non-uniform magnetic field, which exerts a different force depending on the magnetic moment of each atom.
- Detection Screen: After passing through the magnetic field, the atoms strike a detection screen, revealing their distribution.

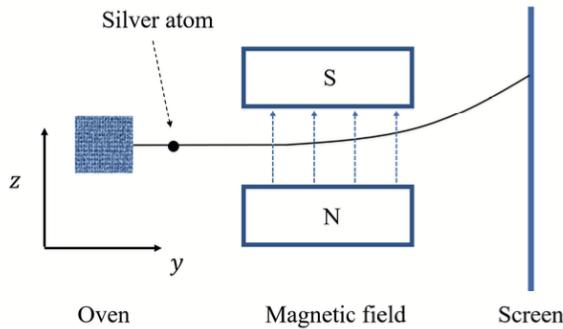


Figure 6: Setup of the Stern-Gerlach experiment

Hot silver atoms emerge from an oven with an initial velocity directed along the y -axis. These atoms then pass through a region containing an inhomogeneous magnetic field, created by combining a uniform magnetic field oriented along the z -axis with a small perturbation that introduces a gradient. As the silver atoms traverse this magnetic field, their final positions along the z -axis are recorded on a detection screen placed to the right. Inside the oven, individual silver atoms are produced carrying one unpaired valence electron that imparts a net magnetic moment. This magnetic moment, referred to as the electron's spin, is represented by a vector $\mu \in \mathbb{R}^3$. The direction of this vector determines whether the atom's trajectory bends upwards or downwards within the magnetic field gradient.

Before the experiment was conducted, classical physics predicted that the magnetic moments of the atoms would orient randomly. As a result, a continuous smear of detection points was expected on the screen, see Fig. 7.

However, quantum mechanics made a strikingly different prediction: due to the quantized nature of spin, the beam should split into discrete parts, see Fig. 7.

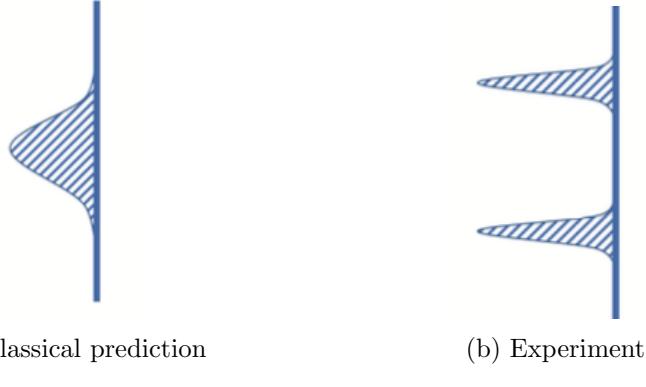


Figure 7: (a) The prediction of the result of the Stern–Gerlach experiment from classical theory and (b) the experimental result.

By blocking one particle beam, we may use the Stern–Gerlach experiment to generate filters, see Fig. 8. Embracing the Dirac notation, we denote the two states that pass through a Stern–Gerlach apparatus with magnetic field in z -direction (SG_z apparatus) by $|+_z\rangle$ and $|-_z\rangle$, respectively, see Fig. 8.



Figure 8: (a) The Stern–Gerlach apparatus along the z -direction and (b) a filtering apparatus.

Note that the spatial orientation of the magnetic field does not have any fundamental effect on the outcome. More precisely, we may rotate the apparatus in a way that the magnetic field is oriented along the x -axis instead, and the outcome would be the same as above, however, rotated around the y -axis. The resulting apparatus would be a Stern–Gerlach apparatus with a magnetic field in x -direction (SG_x apparatus) and the out-coming states would be $|+_x\rangle$ and $|-_x\rangle$, respectively.

Having the ability to combine different filters may allow us to perform a number of experiments. First, we would combine two SG_z apparatuses – or equivalently two SG_x apparatuses. Once a spin is filtered out, we do not observe a beam split if we filter twice or n times in fact, see Fig. 9. On the other hand, filtering one spin out of the beam does not affect the other spin, i.e., once an x -spin is filtered out, we still observe a beam split when transitioning through a SG_z apparatus.

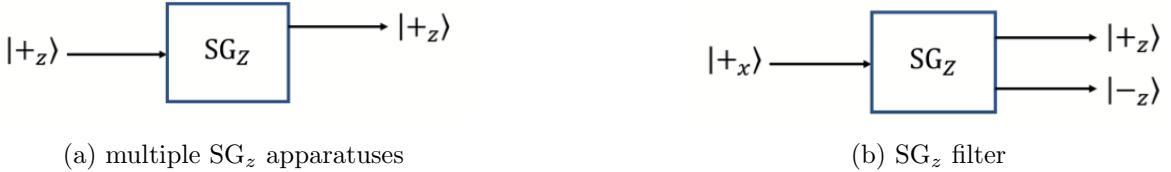


Figure 9: (a) The Stern–Gerlach apparatus along the z -direction and (b) a filtering apparatus.

An interesting observation happens when concatenating three SG filters, $\text{SZ}_z\text{--}\text{SZ}_x\text{--}\text{SZ}_z$. Intuitively, the first filter would filter out a spin component in z -direction, the second filter would filter out a spin component in x -direction, and we should no observe a beam split when the beam transitions the third SZ filter. However, the experiment shows a different outcome, see Fig. 10.

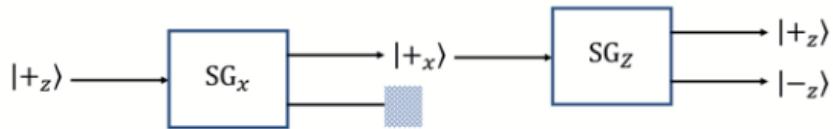


Figure 10: Sequence of three SG filters $\text{SZ}_z\text{--}\text{SZ}_x\text{--}\text{SZ}_z$.

The solution to this mysterious outcome can be obtained using only 2×2 matrices and will lead to the Heisenberg uncertainty principle. The theory of a single spin- $\frac{1}{2}$ particle can be described using a two-dimensional vector space $\mathcal{H} \simeq \mathbb{C}^2$ (isometric isomorphic) called the state vector space. The states $|\pm_z\rangle$ form a basis of \mathcal{H} , i.e., for any $|\psi\rangle \in \mathcal{H}$ there exist $c_1, c_2 \in \mathbb{C}$, s.t.

$$|\psi\rangle = c_1|+z\rangle + c_2|-z\rangle. \quad (63)$$

In particular, $|\pm_x\rangle$ are also states in \mathcal{H} and can be extended using $|\pm_z\rangle$. Moreover, \mathcal{H} is a Hilbert space, and $|\pm_z\rangle$ are orthonormal w.r.t. the inner product structure. In Dirac notation, this means

$$\langle +z|+z\rangle = \langle -z|-z\rangle = 1 \quad \text{and} \quad \langle +z|-z\rangle = 0. \quad (64)$$

Now, since the orientation of the Stern-Gerlach experiment is somewhat arbitrary, we may change $|\pm_z\rangle$ to $|\pm_x\rangle$ in the above discussion. This in turn means that $|\pm_x\rangle$ are linearly independent and therewith form a basis of \mathcal{H} .

In quantum mechanics, a measurement corresponds to a linear operator that is self-adjoint on \mathcal{H} , and the possible outcomes correspond to the operator's spectrum. The Stern-Gerlach apparatus SG_z corresponds to a measurement that measures the spin in z -direction. According to the above postulate, this means that it corresponds to a linear and self-adjoint operator on \mathcal{H} , we shall denote it S_z , and we know the eigenstates of S_z are $|\pm_z\rangle$. Similarly, we may define S_x and S_y together with their spectra $|\pm_x\rangle$ and $|\pm_y\rangle$, respectively. Since $|\pm_x\rangle \in \mathcal{H}$ and $|\pm_y\rangle \in \mathcal{H}$ we may wonder about their expansion using the basis $|\pm_z\rangle$. The Stern-Gerlach experiment provides some insight into this expansion. We first note that when $|+x\rangle$ passes through SG_z the outcome is $|\pm_x\rangle$ with equal probability, i.e., the outcome suggests a bimodal symmetric distribution. Since we moreover have a phase that can be introduced, we find

$$|+x\rangle = \frac{1}{\sqrt{2}} (|+z\rangle + e^{i\alpha} |-z\rangle) \quad (65)$$

and we find

$$|-x\rangle = \frac{1}{\sqrt{2}} (|+z\rangle - e^{i\alpha} |-z\rangle) \quad (66)$$

Similarly we find

$$|+y\rangle = \frac{1}{\sqrt{2}} (|+z\rangle + e^{i\beta} |-z\rangle) \quad (67)$$

and

$$|-y\rangle = \frac{1}{\sqrt{2}} (|+z\rangle - e^{i\beta} |-z\rangle) \quad (68)$$

In order to relate α and β we consider passing $|\pm_x\rangle$ through SG_y ; which yields a bimodal distribution! Hence

$$|\langle \pm_y | \pm_x \rangle|^2 = \frac{1}{2} \quad (69)$$

Since

$$\begin{aligned} \langle \pm_y | \pm_x \rangle &= \frac{1}{2} (\langle +z | \pm e^{-i\beta} \langle -z |) (|+z\rangle \pm e^{i\alpha} |-z\rangle) \\ &= \frac{1}{2} (\langle +z | +z \rangle \pm e^{-i\beta} \langle -z | +z \rangle \pm e^{i\alpha} \langle +z | -z \rangle + e^{-i(\beta-\alpha)} \langle -z | -z \rangle) \\ &= \frac{1}{2} (1 + e^{-i(\beta-\alpha)}) \end{aligned} \quad (70)$$

this yields

$$\frac{1}{2} = |\langle \pm_y | \pm_x \rangle|^2 = \frac{1}{2} (1 + \cos(\beta - \alpha)) \quad (71)$$

and therewith

$$\alpha - \beta = \pm \frac{\pi}{2} + 2\pi n, \quad n \in \mathbb{Z}. \quad (72)$$

We introduce the convention that $\alpha = 0$ and $\beta = \pi/2$ which yields that

$$|\pm_x\rangle = \frac{1}{\sqrt{2}} (|+z\rangle \pm |-z\rangle) \quad \text{and} \quad |\pm_y\rangle = \frac{1}{\sqrt{2}} (|+z\rangle \pm i|-z\rangle) \quad (73)$$

Since $\mathcal{H} \simeq \mathbb{C}^2$ we may perform the identification

$$|+z\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad |-z\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (74)$$

which yields

$$|+x\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad |-x\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (75)$$

and

$$|+_y\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{and} \quad |-_y\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} \quad (76)$$

This choice of functions allows us to formulate S_x , S_z , and S_z in matrix form. To that end, we use the spectral theorem which yields

$$S_z = \frac{1}{2} (|+_z\rangle\langle +_z| - |-_z\rangle\langle -_z|) = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (77)$$

as $|\pm_z\rangle$ are eigenvectors corresponding to eigenvalues $\pm\frac{1}{2}$. The choice of these eigenvalues. Similarly we find

$$\begin{aligned} S_x &= \frac{1}{2} (|+_x\rangle\langle +_x| - |-_x\rangle\langle -_x|) = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ S_y &= \frac{1}{2} (|+_y\rangle\langle +_y| - |-_y\rangle\langle -_y|) = \frac{1}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \end{aligned} \quad (78)$$

Following the convention that $S = (S_x, S_y, S_z) = \frac{1}{2}(\sigma_x, \sigma_y, \sigma_z)$ we define the Pauli matrices as

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (79)$$

One immediate observation is that linear operators for different observables do not necessarily commute, e.g.,

$$\begin{aligned} S_x S_z &= \frac{1}{4} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\ S_z S_x &= \frac{1}{4} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \end{aligned} \quad (80)$$

which yields

$$[S_z, S_x] = S_z S_x - S_x S_z = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = i S_y \quad (81)$$

This means that S_x and S_z are not simultaneously diagonalizable and we refer to them as incompatible in a quantum mechanical context. Similarly, we find that

$$[S_x, S_y] = i S_z \quad \text{and} \quad [S_y, S_z] = i S_x \quad (82)$$

Hence, S_x , S_y and S_z are mutually incompatible. On the other hand

$$S^2 = S_x^2 + S_y^2 + S_z^2 \quad (83)$$

is compatible with all spin operators along any individual direction.

Compatibility has an important physical consequence. Recall that the quantum mechanics postulation that the final state from any measurement leads to an eigenstate of the operator corresponding to a physical observable. Then if two operators can be simultaneously diagonalized using the same set of eigenstates, it means that one can simultaneously measure the corresponding quantities, e.g., the spin magnitude S^2 and the individual spin operators. The compatibility condition is sufficient and necessary. In other words, if the two operators are incompatible, then one cannot always simultaneously measure the values of the two physical observables. This is the core concept behind the *uncertainty principle*, which can be formulated in terms of an inequality for the fluctuation of the measurements for two operators A and B . For a given operator A and quantum state Ψ , we define an operator

$$\Delta A = A - \langle A \rangle I := A - \langle \Psi | A | \Psi \rangle I \quad (84)$$

Thus, the expectation value of ΔA is given by

$$\langle \Delta A \rangle = \langle A \rangle - \langle A \rangle = 0 \quad (85)$$

and its variance is

$$\langle \Delta A^2 \rangle = \langle (A - \langle \Psi | A | \Psi \rangle I)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2 \quad (86)$$

Now, if two operators A and B are compatible, and Ψ is one of their common eigenvectors, then

$$\langle \Delta A^2 \rangle = \langle \Delta B^2 \rangle = 0. \quad (87)$$

This means that there is no uncertainty in measuring the values of both A and B simultaneously. Since A is hermitian, ΔA is also hermitian; together with the Cauchy–Schwarz inequality yields

$$\begin{aligned} |\langle \Delta A \Delta B \rangle|^2 &= |(\langle \psi | \Delta A)(\Delta B | \psi \rangle)|^2 \\ &\leq \langle \psi | \Delta A (\Delta A)^* | \psi \rangle \langle \psi | (\Delta B)^* \Delta B | \psi \rangle \\ &= \langle \Delta A^2 \rangle \langle \Delta B^2 \rangle \end{aligned} \quad (88)$$

and note that

$$\Delta A \Delta B = \frac{1}{2} (\Delta A \Delta B + \Delta B \Delta A) + \frac{1}{2} (\Delta A \Delta B - \Delta B \Delta A) = \frac{1}{2} \{\Delta A, \Delta B\} + \frac{1}{2} [\Delta A, \Delta B] \quad (89)$$

Since $\{\Delta A, \Delta B\}$ is hermitian, $\langle \{\Delta A, \Delta B\} \rangle \in \mathbb{R}$. Similarly, we note that $i[\Delta A, \Delta B]$ is hermitian, hence, $\langle [\Delta A, \Delta B] \rangle$ is purely imaginary. This yields

$$\begin{aligned} |\langle \Delta A \Delta B \rangle|^2 &= \frac{1}{4} |\langle \{\Delta A, \Delta B\} \rangle + \langle [\Delta A, \Delta B] \rangle|^2 \\ &\geq \frac{1}{4} |\langle [\Delta A, \Delta B] \rangle|^2 = \frac{1}{4} |\langle [A, B] \rangle|^2 \end{aligned} \quad (90)$$

where we used that

$$\begin{aligned}
[\Delta A, \Delta B] &= [A - \langle A \rangle I, B - \langle B \rangle I] \\
&= (A - \langle A \rangle I)(B - \langle B \rangle I) - (B - \langle B \rangle I)(A - \langle A \rangle I) \\
&= AB - \langle B \rangle A - \langle A \rangle B + \langle A \rangle \langle B \rangle - BA + \langle A \rangle B + \langle B \rangle A - \langle A \rangle \langle B \rangle \\
&= AB - BA \\
&= [A, B].
\end{aligned} \tag{91}$$

Hence

$$\langle \Delta A^2 \rangle \langle \Delta B^2 \rangle \geq \frac{1}{4} | \langle [A, B] \rangle |^2 \tag{92}$$

which is known as the uncertainty principle. It states that there is a lower bound for the product of the uncertainty of two operators $\langle \Delta A^2 \rangle$ and $\langle \Delta B^2 \rangle$ given by the expectation value of the commutator. Due to the uncertainty principle, one cannot obtain simultaneously precise measurements of, e.g., S_x and S_z .

3.1 Heisenberg Uncertainty Principle

The Heisenberg uncertainty principle is defined as the uncertainty of the position operator and momentum operator. First introduced in 1927 by German physicist Werner Heisenberg, the formal inequality relating the standard deviation of position and the standard deviation of momentum was derived by Earle Hesse Kennard later that year and by Hermann Weyl in 1928.

The explanation for “why” the position and momentum operators are of the presented form requires more physical background and motivation which we will not provide here.

In real space the position operator is defined as

$$x|\psi\rangle = |x\psi\rangle \quad \Leftrightarrow \quad (x\psi)(x) = x\psi(x) \tag{93}$$

for any $|\psi\rangle \in \mathcal{H}$. Note that for an eigenstate of the position operator, i.e.,

$$x|\psi\rangle = x_0|\psi\rangle, \tag{94}$$

this means that $\Psi(x) = 0$ if $x \neq x_0$. This however contradicts that $|\psi\rangle$ is normalizable. In fact the position operator does not have an eigenstate that is square integrable. Instead the eigen decomposition of the position operator is given by the Dirac δ -distribution loosely defined as

$$\delta(x - x_0) = \begin{cases} \infty, & \text{if } x = x_0 \\ 0, & \text{else} \end{cases} \tag{95}$$

with

$$\int \delta(x - x_0) dx = 1. \tag{96}$$

Note that to fully grasp this, we require the theory of distributions. The momentum operator determines the momentum of a quantum mechanical particle and is defined as

$$p = -i \frac{d}{dx} \tag{97}$$

hence, when applied to a wavefunction $\psi(x)\psi(x)$ this yields

$$p\psi(x) = -i\psi'(x). \quad (98)$$

We may now compute the commutator of the position operator and the momentum operator

$$[x, p]\psi(x) = (xp - px)\psi(x) = x(-i\psi'(x)) + i\psi(x) + ix\psi'(x) = i\psi(x) \quad (99)$$

hence

$$[x, p] = i \quad (100)$$

which is known as the canonical commutation relation. In particular, position and momentum cannot be simultaneously determined. A more quantitative version of this statement is given by the uncertainty principle

$$\sqrt{\langle \Delta x^2 \rangle} \sqrt{\langle \Delta p^2 \rangle} = \frac{1}{2} \quad (101)$$

3.2 Quantum numbers

Recall from the hydrogen atom computation that we found the states $\psi_{n,\ell,m}$, where $n \geq 1$, $0 \leq \ell \leq n-1$, and $-\ell \leq m_\ell \leq \ell$. We have now learned that an electron also has a spin value $m_s = \{\pm \frac{1}{2}\}$. This means that every energetic state of the electron in the hydrogen is described by these four numbers:

principal quantum number	$n \geq 1$
azimuthal quantum number	$0 \leq \ell \leq n-1$
magnetic quantum number	$-\ell \leq m_\ell \leq \ell$
spin quantum number	$m_s = \{\pm \frac{1}{2}\}$

In fact, this characterization of electronic states in an atom generalizes to multi-electron systems atoms. However, there is a catch:

Pauli-exclusion principle [?]:

“In a poly-electron atom it is impossible for any two electrons to have the same two values of all four of their quantum numbers, which are: n , the principal quantum number; ℓ , the azimuthal quantum number; m_ℓ , the magnetic quantum number; and m_s , the spin quantum number.”

Consequently, in multi-electron atoms the electrons occupy different electronic states. The “filling” of electronic states follows the Aufbau principle see Fig. 11.

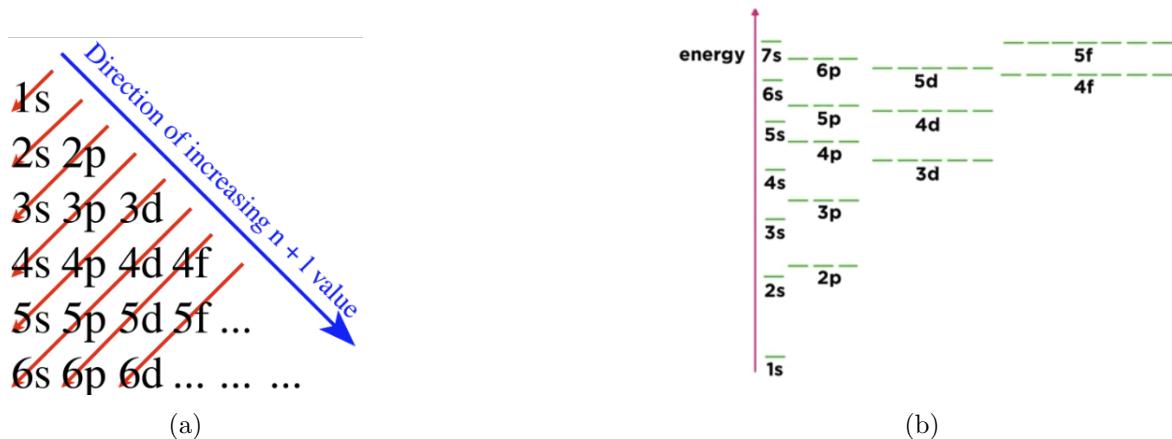


Figure 11: (a) Visualization of the Aufbau principle. and (b) the energies of the different states.

The Aufbau principle relates to the first three quantum numbers, which are sufficient for the hydrogen atom. In multi-electron atoms the spin quantum number needs to be carefully considered. Here Hund’s rule applies, which says that the orbitals of the subshell are each occupied singly with electrons of parallel spin before double occupation occurs.

Example 3.0.1. Consider the molecule chlorine which has 17 electrons, following the Aufbau principle and Hund's rule we find the following configuration, see Table. 1.

	s	p	s	p
$n = 1$	$\{\psi_{1,0,0,\frac{1}{2}}, \psi_{1,0,0,-\frac{1}{2}}\}$			
$n = 2$	$\{\psi_{2,0,0,\frac{1}{2}}, \psi_{2,0,0,-\frac{1}{2}}\}$	$\{\psi_{2,1,-1,\frac{1}{2}}, \psi_{2,1,-1,-\frac{1}{2}}, \psi_{2,1,0,\frac{1}{2}}, \psi_{2,1,0,-\frac{1}{2}}, \psi_{2,1,1,\frac{1}{2}}, \psi_{2,1,1,-\frac{1}{2}}\}$		
$n = 3$	$\{\psi_{3,0,0,\frac{1}{2}}, \psi_{3,0,0,-\frac{1}{2}}\}$	$\{\psi_{3,1,-1,\frac{1}{2}}, \psi_{3,1,-1,-\frac{1}{2}}, \psi_{3,1,0,\frac{1}{2}}, \psi_{3,1,0,-\frac{1}{2}}, \psi_{3,1,1,\frac{1}{2}}\}$		

Table 1: Orbital occupation for chlorine,

Therefore the electronic configuration is

$$1s^2 2s^2 2p^6 3s^2 3p^5 = [\text{Ne}]3s^2 3p^5 \quad (102)$$

Remark 3.0.1. The electronic configuration of an atom can be determined quicker by looking at the periodic table, see Fig. 12.

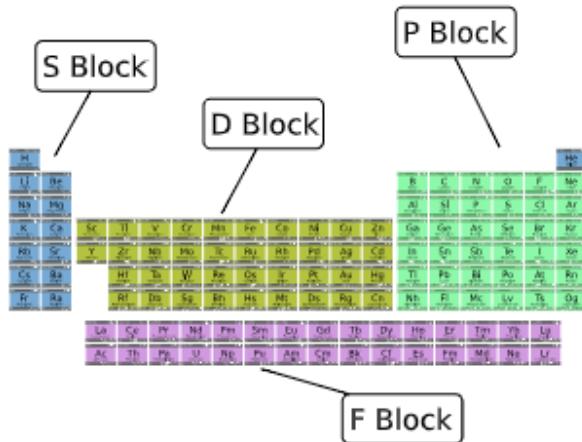


Figure 12: Orbital blocks in periodic table

We then “walk from left to right and top to bottom” to determine the electronic configuration.

4 Combining two spin- $\frac{1}{2}$ particles

When we investigated a single spin particle, we could use the $|\pm_z\rangle$ as a basis. Now what is the correct basis for two spin- $\frac{1}{2}$ particles? A natural extension to the $|\pm_z\rangle$ basis is the tensor product basis, i.e.,

$$\begin{aligned} |+z\rangle \otimes |+z\rangle &= |+z,+z\rangle & |+z\rangle \otimes |-z\rangle &= |+z,-z\rangle \\ |-z\rangle \otimes |+z\rangle &= |-z,+z\rangle & |-z\rangle \otimes |-z\rangle &= |-z,-z\rangle \end{aligned} \quad (103)$$

in which the first element describes the state of particle one and the second element describes the state of particle two. More formally, we may consider two Hilbert spaces \mathcal{H}_A and \mathcal{H}_B of dimension N_A and N_B describing the individual particles, respectively, and denote their bases $\{|\varphi_i^A\rangle\}_{i=1}^{N_A}$ and $\{|\varphi_j^B\rangle\}_{j=1}^{N_B}$. The tensor product space is then defined as

$$\mathcal{H}_A \otimes \mathcal{H}_B = \text{Span} \{ |\varphi_i^A, \varphi_j^B\rangle \mid i \in [\![N_A]\!], j \in [\![N_B]\!] \}, \quad (104)$$

where

Here, $\{|\varphi_i^A, \varphi_j^B\rangle\}$ form a new basis. Recall the tensor product of vectors

$$|+z\rangle \otimes |-z\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (105)$$

Since \mathcal{H}_A and \mathcal{H}_B are Hilbert spaces, their tensor product space inherits an inner product structure

$$\langle \varphi_i^A, \varphi_j^B | \varphi_{i'}^A, \varphi_{j'}^B \rangle_{A \otimes B} = \langle \varphi_i^A | \varphi_{i'}^A \rangle_A \langle \varphi_j^B | \varphi_{j'}^B \rangle_B \quad (106)$$

Therefore, if $\{|\varphi_i^A\rangle\}_{i=1}^{N_A}$ and $\{|\varphi_i^B\rangle\}_{i=1}^{N_B}$ are orthonormal, the tensor product basis inherits the orthonormality, i.e.,

$$\langle \varphi_i^A, \varphi_j^B | \varphi_{i'}^A, \varphi_{j'}^B \rangle = \langle \varphi_i^A | \varphi_{i'}^A \rangle \langle \varphi_j^B | \varphi_{j'}^B \rangle = \delta_{i,i'} \delta_{j,j'} \quad (107)$$

Moreover, $\mathcal{H}_A \otimes \mathcal{H}_B$ is a Hilbert space itself of dimension $N_A \times N_B$. Similarly, we may extend operators to tensor product spaces. Let A and B be two operators defined on \mathcal{H}_A and \mathcal{H}_B , respectively. Then

$$(A \otimes B) | \varphi_{i'}^A, \varphi_{j'}^B \rangle := | A \varphi_{i'}^A, B \varphi_{j'}^B \rangle. \quad (108)$$

Returning to our example of two spin- $\frac{1}{2}$ particles, we can then define the operators

$$S_z^{(1)} = S_z \otimes I \quad \text{and} \quad S_z^{(2)} = I \otimes S_z \quad (109)$$

and the total spin operator in z -direction

$$S_z^{\text{tot}} = S_z^{(1)} + S_z^{(2)} \quad (110)$$

Similarly, we can define S_x^{tot} and S_y^{tot} , as the corresponding spin magnitude operator

$$(S^{\text{tot}})^2 = (S_x^{\text{tot}})^2 + (S_y^{\text{tot}})^2 + (S_z^{\text{tot}})^2 \quad (111)$$

We observe that

$$\begin{aligned}
S_z^{\text{tot}}|+z,+z\rangle &= S_z \otimes I|+z,+z\rangle + I \otimes S_z|+z,+z\rangle = \frac{1}{2}|+z,+z\rangle + \frac{1}{2}|+z,+z\rangle = |+z,+z\rangle \\
S_z^{\text{tot}}|-z,-z\rangle &= S_z \otimes I|-z,-z\rangle + I \otimes S_z|-z,-z\rangle = -\frac{1}{2}|-z,-z\rangle - \frac{1}{2}|-z,-z\rangle = -|-z,-z\rangle \\
S_z^{\text{tot}}|+z,-z\rangle &= S_z \otimes I|+z,-z\rangle + I \otimes S_z|+z,-z\rangle = \frac{1}{2}|+z,-z\rangle - \frac{1}{2}|+z,-z\rangle = |0\rangle \\
S_z^{\text{tot}}|-z,+z\rangle &= S_z \otimes I|-z,+z\rangle + I \otimes S_z|-z,+z\rangle = -\frac{1}{2}|-z,+z\rangle + \frac{1}{2}|-z,+z\rangle = |0\rangle
\end{aligned} \tag{112}$$

Hence all basis vectors of the tensor product basis are eigenstates of S_z^{tot} . We now want to investigate if S_z^{tot} and the corresponding total spin square operator commute. To that end, we first note that

$$\begin{aligned}
(S_z^{\text{tot}})^2 &= (S_z \otimes I)^2 + (S_z \otimes I)(I \otimes S_z) + (I \otimes S_z)(S_z \otimes I) + (I \otimes S_z)^2 \\
&= \frac{1}{2}I \otimes I + 2S_z \otimes S_z
\end{aligned} \tag{113}$$

which yields

$$(S^{\text{tot}})^2 = \frac{3}{2}I \otimes I + 2S_x \otimes S_x + 2S_y \otimes S_y + 2S_z \otimes S_z \tag{114}$$

and therefore

$$\begin{aligned}
[(S^{\text{tot}})^2, S_z^{\text{tot}}] &= \frac{3}{2}[I \otimes I, S_z^{\text{tot}}] + 2[S_x \otimes S_x, S_z^{\text{tot}}] + 2[S_y \otimes S_y, S_z^{\text{tot}}] + 2[S_z \otimes S_z, S_z^{\text{tot}}] \\
&= 2[S_x \otimes S_x, S_z^{\text{tot}}] + 2[S_y \otimes S_y, S_z^{\text{tot}}] + 2[S_z \otimes S_z, S_z^{\text{tot}}]
\end{aligned} \tag{115}$$

We then note that

$$\begin{aligned}
[S_\alpha \otimes S_\alpha, S_z^{\text{tot}}] &= [S_\alpha \otimes S_\alpha, S_z^{(1)}] + [S_\alpha \otimes S_\alpha, S_z^{(2)}] \\
&= [S_\alpha, S_z] \otimes S_\alpha + S_\alpha \otimes [S_\alpha, S_z]
\end{aligned} \tag{116}$$

Since the spin operators for single particles are cyclic we find

$$\begin{aligned}
[S_x \otimes S_x, S_z^{\text{tot}}] &= -iS_y \otimes S_x - iS_x \otimes S_y \\
[S_y \otimes S_y, S_z^{\text{tot}}] &= iS_x \otimes S_y + iS_y \otimes S_x \\
[S_z \otimes S_z, S_z^{\text{tot}}] &= 0
\end{aligned} \tag{117}$$

which yields that

$$[(S^{\text{tot}})^2, S_z^{\text{tot}}] = 0. \tag{118}$$

Therefore $(S^{\text{tot}})^2$ and S_z^{tot} can be simultaneously diagonalized. More specifically,

$$(S^{\text{tot}})^2|+z,+z\rangle = 2|+z,+z\rangle \quad \text{and} \quad (S^{\text{tot}})^2|-z,-z\rangle = 2|-z,-z\rangle \tag{119}$$

On the other hand, we note that

$$\begin{aligned}
(S^{\text{tot}})^2 |+z, -z\rangle &= \frac{3}{2} I \otimes I |+z, -z\rangle + 2S_x \otimes S_x |+z, -z\rangle \\
&\quad + 2S_y \otimes S_y |+z, -z\rangle + 2S_z \otimes S_z |+z, -z\rangle \\
&= \frac{3}{2} |+z, -z\rangle + \frac{1}{2} |-z, +z\rangle + \frac{1}{2} |-z, +z\rangle - \frac{1}{2} |+z, -z\rangle \\
&= |+z, -z\rangle + |-z, +z\rangle
\end{aligned} \tag{120}$$

and similar

$$(S^{\text{tot}})^2 |-z, +z\rangle = |-z, +z\rangle + |+z, -z\rangle \tag{121}$$

Hence, $|\pm_z, \mp_z\rangle$ are not eigenstates of $(S^{\text{tot}})^2$. Instead,

$$\begin{aligned}
(S^{\text{tot}})^2 &\left(\frac{1}{\sqrt{2}} |+z, -z\rangle + \frac{1}{\sqrt{2}} |-z, +z\rangle \right) \\
&= \frac{1}{\sqrt{2}} (S^{\text{tot}})^2 |+z, -z\rangle + \frac{1}{\sqrt{2}} (S^{\text{tot}})^2 |-z, +z\rangle \\
&= \frac{1}{\sqrt{2}} (|+z, -z\rangle + |-z, +z\rangle) + \frac{1}{\sqrt{2}} (|-z, +z\rangle + |+z, -z\rangle) \\
&= 2 \left(\frac{1}{\sqrt{2}} |+z, -z\rangle + \frac{1}{\sqrt{2}} |-z, +z\rangle \right)
\end{aligned} \tag{122}$$

and similar

$$(S^{\text{tot}})^2 \left(\frac{1}{\sqrt{2}} |+z, -z\rangle - \frac{1}{\sqrt{2}} |-z, +z\rangle \right) = 0. \tag{123}$$

Thus, the operator $(S^{\text{tot}})^2$ can be used to distinguish the 0-eigenspace of S_z^{tot} . So we can summarize the eigenstates as

State	Type	$(S^{\text{tot}})^2$	S_z^{tot}
$\frac{1}{\sqrt{2}} (+z, -z\rangle - -z, +z\rangle)$	singlet	0	0
$ +z, +z\rangle$			1
$\frac{1}{\sqrt{2}} (+z, -z\rangle + -z, +z\rangle)$	triplet	2	0
$ -z, -z\rangle$			-1

4.1 The hydrogen molecule ion H_2^+

We now consider the H_2^+ molecule consisting of two hydrogen atoms, but with one electron removed. The Hamiltonian then reads

$$H = \frac{1}{2}\Delta_{\mathbf{r}} - \frac{1}{\|\mathbf{r}\|} - \frac{1}{\|\mathbf{r} - \mathbf{R}\|}. \quad (124)$$

A crude approximation to the solution would be a linear combination of two 1s orbitals centered at 0 and \mathbf{R} respectively, i.e.,

$$\psi(\mathbf{r}) = c_1\psi_{1,0,0}(\mathbf{r}) + c_2\psi_{1,0,0}(\mathbf{r} - \mathbf{R}). \quad (125)$$

In this framework, we assume the Galerkin space

$$\mathcal{H} = \text{Span}\{\psi_{1,0,0}(\cdot), \psi_{1,0,0}(\cdot - \mathbf{R})\} \quad (126)$$

which yields the matrix eigenvalue problem

$$\begin{bmatrix} \varepsilon & -t \\ -t & \varepsilon \end{bmatrix} \mathbf{c} = E \begin{bmatrix} 1 & s \\ s & 1 \end{bmatrix} \mathbf{c} \quad (127)$$

where

$$\begin{aligned} \varepsilon &= \int_{\mathbb{R}^3} \psi_{1,0,0}(\mathbf{r})(H\psi_{1,0,0})(\mathbf{r})d\mathbf{r} \\ -t &= \int_{\mathbb{R}^3} \psi_{1,0,0}(\mathbf{r})(H\psi_{1,0,0})(\mathbf{r} - \mathbf{R})d\mathbf{r} \\ s &= \int_{\mathbb{R}^3} \psi_{1,0,0}(\mathbf{r})\psi_{1,0,0}(\mathbf{r} - \mathbf{R})d\mathbf{r} \end{aligned} \quad (128)$$

Given that $s, t > 0$, the state eigenvalue and eigenfunctions are

$$E_g = \frac{\varepsilon - t}{1 + s}, \quad \mathbf{c}_g = \frac{1}{\sqrt{2(1 + s)}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (129)$$

and

$$E_e = \frac{\varepsilon + t}{1 - s}, \quad \mathbf{c}_e = \frac{1}{\sqrt{2(1 - s)}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (130)$$

5 Slater Determinant

6 Hartree–Fock Theory

$$\min_{|\psi\rangle \in \mathcal{H}_N} \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} \leqslant \min_{|\psi\rangle \in \mathcal{S}_N} \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} \quad (131)$$

where

$$\mathcal{S}_N = \left\{ \psi \in \mathcal{H}_N \mid \exists (\phi_1, \dots, \phi_N) \in \mathcal{C}_N ; \psi = \frac{1}{\sqrt{N!}} \bigwedge_{i=1}^N \phi_i \right\},$$

is the set of Slater determinants and the set of N -frames

$$\mathcal{C}_N = \{(\phi_1, \dots, \phi_N) \mid \phi_i \in H^1 \text{ and } \langle\phi_i, \phi_j\rangle_{L^2} = \delta_{i,j} \text{ for } 1 \leq j, i \leq N\}$$

is the *Stiefel manifold*. Note that the functional in Eq. (131) over \mathcal{S}_N is invariant with respect to unitary transformations. To remedy this non-uniqueness we introduce the Grassmannian as suitable quotient space of \mathcal{C}_N , i.e.,

$$\mathcal{G}_N = \mathcal{C}_N|_{\sim}, \quad (132)$$

where the equivalence classes are given by

$$[\Phi]_{\sim} = \left\{ \tilde{\Phi} \in \mathcal{C}_N \mid \exists U \in \mathcal{U}(N) \text{ s.t. } \Phi = U\tilde{\Phi} \right\}. \quad (133)$$

In other words, we identify all orthonormal N -frames spanning the same subspace. This yields the following minimization problem

$$E_{\text{HF}} = \inf_{\Psi \in \mathcal{S}_N} \mathcal{J}(\Psi) = \inf_{\Phi \in \mathcal{G}_N} \varepsilon(\Phi), \quad (134)$$

6.1 One-body part

$$\langle\Psi|h|\Psi\rangle = \dots = \sum_{i=1}^N \int_{\mathbb{R}^3 \times \{\pm\frac{1}{2}\}} \frac{1}{2} |\nabla_r \phi_i(\mathbf{x})|^2 + V_{\text{ext}}(\mathbf{r}) |\phi_i(\mathbf{x})|^2 d\mathbf{x} \quad (135)$$

6.2 Two-body part

Recall that the two-body interaction term is given by

$$H_I := \sum_{i < j} g(i, j) := \sum_{i < j} \frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|} \quad (136)$$

Again, let $\Psi = \Phi[j_1, \dots, j_N]$ be a Slater determinant. Then

$$\begin{aligned} \langle\Psi|H_I|\Psi\rangle &= \sum_{i < j} \left\langle \Psi \left| \frac{1}{\|\mathbf{r}_i - \mathbf{r}_j\|} \right| \Psi \right\rangle \\ &= \binom{N}{2} \left\langle \Psi \left| \frac{1}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \right| \Psi \right\rangle \end{aligned} \quad (137)$$

We note that

$$\begin{aligned}
\left\langle \Psi \left| \frac{1}{\|\mathbf{r}_1 - \mathbf{r}_2\|} \right| \Psi \right\rangle &= \int_{(\mathbb{R}^3 \times \{\pm \frac{1}{2}\})^{\otimes N}} \Psi^*(\mathbf{x}_1, \dots, \mathbf{x}_N) \frac{1}{\|r_1 - r_2\|} \Psi(\mathbf{x}_1, \dots, \mathbf{x}_N) d\mathbf{x}_1 \dots d\mathbf{x}_N \\
&= \frac{1}{N!} \sum_{\pi, \pi' \in S_N} \text{sgn}(\pi) \text{sgn}(\pi') \int_{(\mathbb{R}^3 \times \{\pm \frac{1}{2}\})^{\otimes 2}} \phi_{\pi(1)}^*(\mathbf{x}_1) \phi_{\pi(2)}^*(\mathbf{x}_2) \frac{1}{\|r_1 - r_2\|} \phi_{\pi'(1)}(\mathbf{x}_1) \phi_{\pi'(2)}(\mathbf{x}_2) \\
&\quad \times \prod_{i=3}^N \int_{\mathbb{R}^3 \times \{\pm \frac{1}{2}\}} \phi_{\pi(i)}^*(\mathbf{x}) \phi_{\pi'(i)}(\mathbf{x}) d\mathbf{x}
\end{aligned}$$

Note that

$$\prod_{i=3}^N \int_{\mathbb{R}^3 \times \{\pm \frac{1}{2}\}} \phi_{\pi(i)}^*(\mathbf{x}) \phi_{\pi'(i)}(\mathbf{x}) d\mathbf{x} = \prod_{i=3}^N \delta_{\pi(i), \pi'(i)} \quad (138)$$

which implies that either

$$\pi(1) = \pi'(1) = i_1 \quad \text{and} \quad \pi(2) = \pi'(2) = i_2 \quad (139)$$

or

$$\pi(1) = \pi'(2) = i_1 \quad \text{and} \quad \pi(2) = \pi'(1) = i_2 \quad (140)$$

where $i_1, i_2 \in \{1, \dots, N\}$. In case of Eq. (139) we have that

$$\text{sgn}(\pi) = \text{sgn}(\pi') \quad \Rightarrow \quad \text{sgn}(\pi) \text{sgn}(\pi') = 1$$

Whereas in case of Eq. (140) we have that

$$\text{sgn}(\pi) = -\text{sgn}(\pi') \quad \Rightarrow \quad \text{sgn}(\pi) \text{sgn}(\pi') = -1$$

Thus

$$\begin{aligned}
\left\langle \Psi \left| \frac{1}{\|r_1 - r_2\|} \right| \Psi \right\rangle &= \frac{(N-2)!}{N!} \sum_{i=1}^N \sum_{i \neq j=1}^N \left(\int_{(\mathbb{R}^3 \times \{\pm \frac{1}{2}\})^{\otimes 2}} \phi_i^*(\mathbf{x}_1) \phi_j^*(\mathbf{x}_2) \frac{1}{\|r_1 - r_2\|} \phi_i(\mathbf{x}_1) \phi_j(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \right. \\
&\quad \left. - \int_{(\mathbb{R}^3 \times \{\pm \frac{1}{2}\})^{\otimes 2}} \phi_i^*(\mathbf{x}_1) \phi_j^*(\mathbf{x}_2) \frac{1}{\|r_1 - r_2\|} \phi_j(\mathbf{x}_1) \phi_i(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \right) \\
&:= \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{i \neq j=1}^N \langle ij || ij \rangle - \langle ij || ji \rangle
\end{aligned}$$

Hence

$$\langle \Psi | H_I | \Psi \rangle = \frac{1}{2} \sum_{i,j=1}^N \langle ij || ij \rangle - \langle ij || ji \rangle \quad (141)$$

Putting it all together

$$\begin{aligned}
\mathcal{E}_{\text{HF}}(\{\phi_i\}_{i=1}^N) &= \sum_{i=1}^N \int_{\mathbb{R}^3 \times \{\pm \frac{1}{2}\}} \frac{1}{2} |\nabla_r \phi_i(x)|^2 + V_{\text{ext}}(r) |\phi_i(x)|^2 dx \\
&\quad + \frac{1}{2} \sum_{i,j} \int_{(\mathbb{R}^3 \times \{\pm \frac{1}{2}\})^{\otimes 2}} \frac{|\phi_i(x)|^2 |\phi_j(x')|^2}{\|r - r'\|} dx dx' \\
&\quad - \frac{1}{2} \sum_{i,j} \int_{(\mathbb{R}^3 \times \{\pm \frac{1}{2}\})^{\otimes 2}} \frac{\phi_i^*(x) \phi_j^*(x') \phi_j(x) \phi_i(x')}{\|r - r'\|} dx dx' \\
&= \sum_{i=1}^N \langle i | i \rangle + \frac{1}{2} \sum_{i,j} \langle ij | ij \rangle - \langle ij | ji \rangle
\end{aligned}$$

How do we optimize this energy function?

We assume a given set of spin-orbitals $\{\chi_i\}_{i=1}^K$ where $N \leq K$. Note that in this context, χ_i may be inspired by atomic orbital functions and are not necessarily orthogonal, i.e.,

$$S_{i,j} := \langle \chi_i | \chi_j \rangle \quad (142)$$

and $\mathbf{S} \neq \mathbf{I}_K$. We then make the ansatz

$$\phi_i = \sum_{j=1}^K \mathbf{C}_{j,i} \chi_j \quad (143)$$

defining $\mathbf{C} \in \mathbb{C}^{K \times N}$. Applying this to the kinetic energy part we find

$$\begin{aligned}
&\frac{1}{2} \sum_{i=1}^N \int_{\mathbb{R}^3 \times \{\pm \frac{1}{2}\}} (\nabla_r \phi_i(\mathbf{x}))^* \nabla_r \phi_i(\mathbf{x}) d\mathbf{x} \\
&= \frac{1}{2} \sum_{i=1}^N \sum_{j,k=1}^K \mathbf{C}_{j,i}^* \mathbf{C}_{k,i} \int_{\mathbb{R}^3 \times \{\pm \frac{1}{2}\}} (\nabla_r \chi_j(\mathbf{x}))^* \nabla_r \chi_k(\mathbf{x}) d\mathbf{x} \\
&= \sum_{i=1}^N \frac{1}{2} \sum_{j,k=1}^K \mathbf{C}_{j,i}^* \mathbf{C}_{k,i} \langle \nabla \chi_j, \nabla \chi_k \rangle
\end{aligned} \quad (144)$$

Overall, this yields

$$\begin{aligned}
\mathcal{E}_{\text{HF}}(\mathbf{C}) &= \sum_{i=1}^N \frac{1}{2} \sum_{j,k=1}^K \mathbf{C}_{j,i}^* \mathbf{C}_{k,i} \langle \nabla \chi_j, \nabla \chi_k \rangle \\
&\quad - \sum_{i=1}^N \sum_{j,k=1}^K \mathbf{C}_{j,i}^* \mathbf{C}_{k,i} \langle \chi_j, V_{\text{ext}} \chi_k \rangle \\
&\quad + \frac{1}{2} \sum_{i,j=1}^N \sum_{k,\ell,m,n=1}^K \mathbf{C}_{k,i}^* \mathbf{C}_{\ell,j}^* \mathbf{C}_{m,i} \mathbf{C}_{n,j} \langle \chi_k \chi_\ell | | \chi_m \chi_n \rangle \\
&\quad - \frac{1}{2} \sum_{i,j=1}^N \sum_{k,\ell,m,n=1}^K \mathbf{C}_{k,i}^* \mathbf{C}_{\ell,j}^* \mathbf{C}_{m,j} \mathbf{C}_{n,i} \langle \chi_k \chi_\ell | | \chi_m \chi_n \rangle
\end{aligned}$$

Introducing the tensors

$$\begin{aligned}\mathbf{T}_{j,k} &= \frac{1}{2} \langle \nabla \phi_j | \nabla \phi_k \rangle \\ \mathbf{V}_{j,k} &= \langle \phi_j | V | \phi_k \rangle \\ \mathbf{h} &= \frac{1}{2} \mathbf{T} - \mathbf{V} \\ \mathbf{V}_{i,j,k,l} &= \langle \chi_i \chi_j | | \chi_k \chi_l \rangle\end{aligned}$$

we find

$$\begin{aligned}\mathcal{E}_{\text{HF}}(\mathbf{C}) &= \text{Tr} (\mathbf{C}^\dagger \mathbf{h} \mathbf{C}) \\ &+ \frac{1}{2} \sum_{i,j=1}^N \sum_{k,\ell,m,n=1}^K \mathbf{C}_{k,i}^* \mathbf{C}_{\ell,j}^* \mathbf{C}_{m,i} \mathbf{C}_{n,j} \mathbf{V}_{k,\ell,m,n} \\ &- \frac{1}{2} \sum_{i,j=1}^N \sum_{k,\ell,m,n=1}^K \mathbf{C}_{k,i}^* \mathbf{C}_{\ell,j}^* \mathbf{C}_{m,j} \mathbf{C}_{n,i} \mathbf{V}_{k,\ell,m,n}\end{aligned}$$

We define the one-particle reduced density matrix (1-RDM) $\mathbf{D} \in \mathbb{C}^{K \times K}$ element-wise as

$$\mathbf{D}_{i,j} = [\mathbf{C} \mathbf{C}^\dagger]_{i,j} = \sum_{k=1}^N \mathbf{C}_{i,k} \mathbf{C}_{j,k}^*. \quad (145)$$

which yields

$$\begin{aligned}\mathcal{E}_{\text{HF}}(\mathbf{D}) &= \text{Tr} (\mathbf{h} \mathbf{D}) \\ &+ \frac{1}{2} \sum_{k,\ell,m,n=1}^K \mathbf{D}_{m,k} \mathbf{D}_{n,\ell} \mathbf{V}_{k,\ell,m,n} \\ &- \frac{1}{2} \sum_{k,\ell,m,n=1}^K \mathbf{D}_{n,k} \mathbf{D}_{m,\ell} \mathbf{V}_{k,\ell,m,n} \\ &=: \text{Tr} (\mathbf{h} \mathbf{D}) + \frac{1}{2} \text{Tr} (\mathbf{J}(\mathbf{D}) \mathbf{D}) - \frac{1}{2} \text{Tr} (\mathbf{K}(\mathbf{D}) \mathbf{D})\end{aligned} \quad (146)$$

where

$$\mathbf{J}(\mathbf{D})_{k,m} = \sum_{\ell,n=1}^K \mathbf{D}_{n,\ell} \mathbf{V}_{k,\ell,m,n} \quad (147)$$

is the Hartree or Coulomb operator leading to

$$\text{Tr} (\mathbf{J}(\mathbf{D}) \mathbf{D}) = \sum_{k,m=1}^K \mathbf{J}(\mathbf{D})_{k,m} \mathbf{D}_{m,k} = \sum_{k,\ell,m,n=1}^K \mathbf{D}_{n,\ell} \mathbf{V}_{k,\ell,m,n} \mathbf{D}_{m,k} \quad (148)$$

and

$$\mathbf{K}(\mathbf{D})_{k,n} = \sum_{\ell,m=1}^K \mathbf{D}_{m,\ell} \mathbf{V}_{k,\ell,m,n} \quad (149)$$

is the Fock or exchange operator leading to

$$\text{Tr} (\mathbf{K}(\mathbf{D}) \mathbf{D}) = \sum_{k,n=1}^K \mathbf{K}(\mathbf{D})_{k,n} \mathbf{D}_{n,k} = \sum_{k,\ell,m,n=1}^K \mathbf{D}_{m,\ell} \mathbf{V}_{k,\ell,m,n} \mathbf{D}_{n,k} \quad (150)$$

Given the compact form of the energy functional, we will now take a closer look at \mathbf{C} or equivalently \mathbf{D} . We first note that we have used that

$$\delta_{i,j} = \langle \phi_i | \phi_j \rangle = \sum_{k,\ell} C_{k,i}^* \langle \chi_k | \chi_\ell \rangle C_{\ell,j} \quad (151)$$

which yields the condition that

$$\mathbf{C}^\dagger \mathbf{S} \mathbf{C} = \mathbf{I}_N \quad (152)$$

Hence we can write the Hartree–Fock method as a constraint minimization, namely,

$$\begin{aligned} & \min_{\mathbf{C} \in \mathbb{C}^{K \times N}} \mathcal{E}_{\text{HF}}(\mathbf{C}) \\ & \text{subject to } \mathbf{C}^\dagger \mathbf{S} \mathbf{C} = \mathbf{I}_N \end{aligned} \quad (153)$$

This yields the Lagrangian

$$\mathcal{L}[\mathbf{C}, \Lambda] = \mathcal{E}_{\text{HF}}(\mathbf{C}) - \text{Tr}[\Lambda(\mathbf{I} - \mathbf{C}^\dagger \mathbf{S} \mathbf{C})] \quad (154)$$

We find that

$$\begin{aligned} 0 &\stackrel{(!)}{=} \nabla_{\mathbf{C}} \mathcal{L}[\mathbf{C}, \Lambda] \\ &= 2\mathbf{h}\mathbf{C} + \frac{1}{2}4\mathbf{J}(\mathbf{C}\mathbf{C}^\dagger)\mathbf{C} - \frac{1}{2}4\mathbf{K}(\mathbf{C}\mathbf{C}^\dagger)\mathbf{C} + 2\mathbf{S}\mathbf{C}\Lambda \\ \Leftrightarrow \quad \mathbf{F}(\mathbf{C})\mathbf{C} &:= (\mathbf{h} + \mathbf{J}(\mathbf{C}\mathbf{C}^\dagger) - \mathbf{K}(\mathbf{C}\mathbf{C}^\dagger))\mathbf{C} = \mathbf{S}\mathbf{C}\Lambda \end{aligned} \quad (155)$$

where $F(\mathbf{C})$ is the Fock matrix constructed from the $\mathbf{C} \in \mathbb{C}^{K \times N}$. Diagonalizing Λ , i.e.,

$$\Lambda = \mathbf{U} \mathcal{E} \mathbf{U}^* \quad (156)$$

yields the non-linear generalized eigenvalue problem (GEVP)

$$\begin{aligned} \mathbf{F}(\mathbf{C}\mathbf{U})\mathbf{C}\mathbf{U} &= \mathbf{S}\mathbf{C}\mathbf{U}\mathcal{E} \\ \Leftrightarrow \quad \mathbf{F}(\mathbf{C})\mathbf{C} &= \mathbf{S}\mathcal{E} \end{aligned} \quad (157)$$

This is known as the Roothan Hall equations, and the self-consistent field (SCF) iteration method solves this problem self-consistently, i.e.,

$$\mathbf{C}^{(k)} \rightarrow \mathbf{F}(\mathbf{C}^{(k)}) \rightarrow \mathbf{C}^{(k+1)} \quad (158)$$

until convergence is reached. Note that we here used a certain notational ambiguity, namely, solving the GEVP yields a matrix $\mathbf{C} \in \mathbb{C}^{K \times K}$. These are *all* molecular orbitals. Those corresponding the lowest N eigenvalues ε are called occupied molecular orbitals and are used to generate the Fock matrix for the next iteration which is consistent with the Aufbau principle. The remaining orbitals are the virtual molecular orbitals.

Having the 1-RDM as the central object, we make the following observations: First, we see that \mathbf{D} is hermitian since

$$\mathbf{D}^\dagger = (\mathbf{C}\mathbf{C}^\dagger)^\dagger = \mathbf{C}\mathbf{C}^\dagger = \mathbf{D} \quad (159)$$

Second, we note that

$$\mathrm{Tr}[\mathbf{SD}] = \mathrm{Tr}[\mathbf{C}^\dagger \mathbf{SC}] = \mathrm{Tr}[\mathbf{I}_N] = N \quad (160)$$

Third, we find

$$\mathbf{DSD} = \mathbf{CC}^\dagger \mathbf{SCC}^\dagger = \mathbf{CC}^\dagger = \mathbf{D} \quad (161)$$

$$\begin{aligned} & \min_{\mathbf{D} \in \mathbb{C}^{K \times K}} \mathcal{E}_{\mathrm{HF}}(\mathbf{D}) \\ & \text{subject to } \mathbf{D} \in \mathcal{D} \end{aligned} \quad (162)$$

where

$$\mathcal{D} = \{\mathbf{D} \in \mathbb{C}^{K \times K} \mid \mathbf{DSD} = \mathbf{D}, \mathrm{Tr}[\mathbf{SD}] = N, \mathbf{D}^\dagger = \mathbf{D}\} \quad (163)$$

Note that $\mathbf{0} \prec \mathbf{S}$, i.e., \mathbf{S} is positive definite and when the orbitals $\{\chi_i\}_{i=1}^K$ are orthonormalized, e.g.,

$$\tilde{\chi}_i = \sum_{j=1}^K [\mathbf{S}^{-\frac{1}{2}}]_{i,j} \chi_j \quad (164)$$

then

$$\langle \tilde{\chi}_i | \tilde{\chi}_j \rangle = \sum_{k,\ell=1}^K [\mathbf{S}^{-\frac{1}{2}}]_{i,k} \langle \chi_k | \chi_\ell \rangle [\mathbf{S}^{-\frac{1}{2}}]_{j,\ell} = [\mathbf{S}^{-\frac{1}{2}} \mathbf{S} \mathbf{S}^{-\frac{1}{2}}]_{i,j} = \delta_{i,j}, \quad (165)$$

we have

$$\mathcal{D} = \{\mathbf{D} \in \mathbb{C}^{K \times K} \mid \mathbf{D}^2 = \mathbf{D}, \mathrm{Tr}[\mathbf{D}] = N, \mathbf{D}^\dagger = \mathbf{D}\} \quad (166)$$

This set is equivalent to the Grassmann manifold. Therefore we may also directly minimize the energy functional $\mathcal{E}_{\mathrm{HF}}(\mathbf{D})$ over the Grassmann manifold. Note that this argument may also be made for $\mathcal{E}_{\mathrm{HF}}(\mathbf{C})$ via $\mathbf{C} \in \mathbb{C}^{K \times N}$. For obvious reasons, methods that follow this optimization procedures are called direct minimization approaches to the Hartree–Fock problem.

6.3 Spin symmetries (Fokotume)

Up to this point, we assumed that $\{\chi_i\}_{i=1}^K \subseteq H^1(\mathbb{R}^3 \times \{\pm \frac{1}{2}\})$ is a very general set of functions. However, this approach is more general than it needs to be. Recall that the Hamiltonian does not explicitly depend on the spin degree of freedom. In particular the ansatz

$$\chi_i(\mathbf{x}) = \xi_i(\mathbf{r}) m_i(s) \quad (167)$$

where ξ are the spatial orbitals and m are the spinors, is justified. The different spin symmetry approaches then corresponds to different ansätze in the LCAO formalism, namely,

Spin unrestricted HF

$$\phi_{i,s} = \sum_{p=1}^{K/2} C_{p,i}^{(s)} \xi_i \otimes m_s \quad (168)$$

here $\mathbf{C} \in \mathbb{C}^{K \times N}$ is block diagonal, with not necessarily equal blocks.

Spin restricted HF

$$\phi_{i,s} = \sum_{p=1}^{K/2} C_{p,i} \xi_i \otimes m_s \quad (169)$$

here $\mathbf{C} \in \mathbb{C}^{K \times N}$ is block diagonal, with equal blocks.

Spin generalized HF

$$\phi_i = \sum_{p=1}^{K/2} \sum_{s=1}^2 C_{p,i} \xi_i \otimes m_s \quad (170)$$

here $\mathbf{C} \in \mathbb{C}^{K \times N}$ is block diagonal, with not necessarily equal blocks.

Note that these symmetries have effect the optimization functional. In case of spin restricted Hartree–Fock we find that the Fock matrix is given by

$$F(\mathbf{C}) = \mathbf{h} + \mathbf{J}(\mathbf{CC}^\dagger) - \frac{1}{2}\mathbf{K}(\mathbf{CC}^\dagger) \quad (171)$$

where $\mathbf{C} \in \mathbb{C}^{k/2 \times N/2}$ is only one of the diagonal blocks, and the ERI using in the Coulomb and exchange term are coming from $K/2$ spatial orbitals. The corresponding RHF energy is then given by

$$E_{RHF} = \text{Tr}[\mathbf{D}(\mathbf{h} + \mathbf{F}(\mathbf{C}))] \quad (172)$$

where $\mathbf{D} = \mathbf{CC}^\dagger$.

7 Second Quantization

7.1 The Fermionic Fock space

Starting point are the atomic spin orbitals $\{\phi_i\}$ which are sufficiently smooth functions defined on $\mathbb{R}^3 \times \{\pm 1/2\}$. The set of atomic orbitals is denoted \mathcal{B} and spans the finite-dimensional space h . In our case, $|\mathcal{B}| = n \gg d$. In physics and chemistry parlance, h is commonly referred to as the *single-particle Hilbert space*.

The M -particle (Hilbert) space, denoted $\mathcal{H}^{(M)}$, is then the M -th exterior power of h equipped with the natural orthonormal basis, denoted $\mathfrak{B}^{(M)}$, coming from wedge-products of elements of \mathcal{B} (these are the Slater determinants). We define the fermionic Fock space as the Grassmann algebra on $h = \mathcal{H}^{(1)}$:

$$\mathcal{F} = \bigoplus_{M=0}^n \mathcal{H}^{(M)}. \quad (173)$$

All vector spaces defined above are real vector spaces. Recall that by Pauli's exclusion principle, atomic spin orbitals can either be occupied or unoccupied, hence, any element in \mathcal{F} can be expressed by means of an occupation vector, i.e., a K -tuple of zeros and ones.

Example 7.0.1. *We may write an occupation vector using Dirac notation: Consider*

$$|s_1, \dots, s_n\rangle = \frac{1}{\sqrt{M!}} \phi_1^{s_1} \wedge \phi_2^{s_2} \wedge \dots \wedge \phi_n^{s_n} \quad (174)$$

where $M = \sum_i s_i \leq n$ and $s_i \in \{0, 1\}$ for all $i = 1, \dots, n$. A general element in \mathcal{F} , is then given as

$$|\Psi\rangle = \sum_{s_1, \dots, s_n \in \{0, 1\}} \Psi(s_1, \dots, s_n) |s_1, \dots, s_n\rangle \quad (175)$$

where $\Psi(s_1, \dots, s_n) \in \mathbb{C}$.

The central objects in the second quantization are the fermionic creation and annihilation operators, i.e.,

$$\begin{aligned} a_p^\dagger : \mathcal{F} &\rightarrow \mathcal{F}; |s_1, \dots, s_n\rangle \mapsto (-1)^{\sigma(p)} (1 - s_p) |s_1, \dots, s_{p-1}, 1 - s_p, s_{p+1}, \dots, s_n\rangle \\ a_p : \mathcal{F} &\rightarrow \mathcal{F}; |s_1, \dots, s_n\rangle \mapsto (-1)^{\sigma(p)} s_p |s_1, \dots, s_{p-1}, 1 - s_p, s_{p+1}, \dots, s_n\rangle \end{aligned} \quad (176)$$

where $\sigma(p) = \sum_{q=1}^{p-1} s_q$. As the name suggests, these operators create or annihilate a “particle” in the p th atomic spin-orbital.

7.2 The Jordan-Wigner transformation

One particularly useful way of thinking of occupation vectors is by means of tensor products. To that end, we denote

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv \text{unoccupied} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv \text{occupied}$$

Note that this is an arbitrary choice, but this is the convention used in the community. Then, we may identify

$$|s_1, \dots, s_n\rangle = \begin{pmatrix} 1 - s_1 \\ s_1 \end{pmatrix} \otimes \dots \otimes \begin{pmatrix} 1 - s_n \\ s_n \end{pmatrix}. \quad (177)$$

Example 7.0.2. Let $n = 2$, then

$$|01\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad (178)$$

This perspective is particularly useful since it allows us to derive a discretization of the Hamiltonian in terms of exponentially large but extremely sparse matrices. We begin by defining the matrix

$$a = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

and note that

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = a \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} = a^\dagger \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

In other words, a annihilates and a^\dagger creates a particle. Moreover, we define the matrix

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which is also known as the *Pauli-z matrix*. Note that

$$(-1) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \sigma_z \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \sigma_z \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

In other words, the Pauli-z matrix multiplies by the factor -1 if the orbital is occupied, and by 1 if the state is unoccupied. We can therefore use the Pauli-z matrix to get the parity in Eq. (176) correct.

Then, the fermionic creation and annihilation operators are sparse matrices of the form

$$a_p^\dagger = \underbrace{\sigma_z \otimes \dots \otimes \sigma_z}_{p-1 \text{ times}} \otimes a^\dagger \otimes \underbrace{I \otimes \dots \otimes I}_{n-p-1 \text{ times}} \quad \text{and} \quad a_p = \underbrace{\sigma_z \otimes \dots \otimes \sigma_z}_{p-1 \text{ times}} \otimes a \otimes \underbrace{I \otimes \dots \otimes I}_{n-p-1 \text{ times}} \quad (179)$$

where

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (180)$$

Example 7.0.3. Let $n = 3$, then

$$a_2 = \sigma_z \otimes a \otimes I = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (181)$$

Clearly, one does not want to store the creation and annihilation matrices explicitly, as they scale as $\mathcal{O}(2^n)$, yet their very structured sparsity pattern can be leveraged.

Before venturing further into the Hamiltonian description, we want to highlight the most important property of the fermionic creation and annihilation matrices.

Theorem 7.1. *the fermionic creation and annihilation operators obey the canonical anti-commutation relation (CAR):*

$$[a_p, a_q]_+ = [a_p^\dagger, a_q^\dagger]_+ = 0 \quad \text{and} \quad [a_p, a_q^\dagger]_+ = \delta_{p,q}. \quad (182)$$

Proof. Let $p < q$. Then

$$\begin{aligned} a_p a_q &= I \otimes \dots \otimes I \otimes \underbrace{a \sigma_z}_{=-a} \otimes \sigma_z \otimes \dots \otimes \sigma_z \otimes a \otimes I \otimes \dots \otimes I \\ &= -I \otimes \dots \otimes I \otimes \underbrace{\sigma_z a}_{=a} \otimes \sigma_z \otimes \dots \otimes \sigma_z \otimes a \otimes I \otimes \dots \otimes I \\ &= -a_q a_p \end{aligned} \quad (183)$$

For $p = q$ we find $a^2 = 0$ tensored with identities. Similarly, we find that $[a_p^\dagger, a_q^\dagger]_+ = 0$. \square

For Slater determinants there holds

$$|s_1, \dots, s_n\rangle = \prod_{i=1^n} (a_i^\dagger)^{s_i} |-\rangle = a_k^\dagger \dots a_1^\dagger |-\rangle$$

7.3 Hamiltonian in second quantization

Given the fermionic creation and annihilation operators, the Coulomb Hamiltonian can be reformulated in second quantized form it reads

$$H = \sum_{p,q=1}^n h_{p,q} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s=1}^n v_{p,q,r,s} a_p^\dagger a_r^\dagger a_s a_q, \quad (184)$$

where

$$h_{p,q} = \int_X \phi_p^*(x_1) \left(-\frac{\Delta}{2} - \sum_j \frac{Z_j}{|r_1 - R_j|} \right) \phi_q(x_1) dx_1 \quad (185)$$

and

$$v_{p,q,r,s} = \int_{X \times X} \frac{\phi_p^*(x_1) \phi_q(x_1) \phi_r^*(x_2) \phi_s(x_2)}{|r_1 - r_2|} dx_1 dx_2, \quad (186)$$

An important observation is that although the matrix $H \in \mathbb{C}^{2^n \times 2^n}$, the number of coefficients scales much lower, namely, $h = [h_{p,q}] \in \mathbb{C}^{n \times n}$ and $v = [v_{p,q,r,s}] \in \mathbb{C}^{n \times n \times n \times n}$. Moreover, we note that

$$h = h^\dagger$$

or in the case of real-valued atomic spin orbitals

$$h = h^\top$$

and that

$$v_{p,q,r,s} = v_{r,s,p,q} = v_{q,p,s,r}^* = v_{s,r,q,p}^*$$

or in the case of real-valued atomic spin orbitals

$$v_{p,q,r,s} = v_{r,s,p,q} = v_{q,p,s,r} = v_{s,r,q,p} = v_{q,p,r,s} = v_{s,r,p,q} = v_{p,q,s,r} = v_{r,s,q,p}$$

8 Contractions

In quantum mechanics, the eigenstate is a nice quantity to have, but of interest are *expectation values*

$$\langle \Psi, \mathcal{O} \Psi \rangle \quad (187)$$

since they correspond to measurable quantities, called observables. In the case of the Hamiltonian, the observable is the system's energy. Remember that

$$[a_p, a_q]_+ = [a_p^\dagger, a_q^\dagger]_+ = 0 \quad \text{and} \quad [a_p^\dagger, a_q]_+ = \delta_{p,q}. \quad (188)$$

Then

$$\langle ij | \sum_{p,q} h_{p,q} a_p^\dagger a_q | kl \rangle = \sum_{p,q} h_{p,q} \langle -| a_j a_i a_p^\dagger a_q a_l^\dagger a_k^\dagger | - \rangle \quad (189)$$

and

$$\begin{aligned} \langle -| a_j a_i a_p^\dagger a_q a_l^\dagger a_k^\dagger | - \rangle &= (-1) \langle -| a_j a_p^\dagger a_i a_q a_l^\dagger a_k^\dagger | - \rangle + \delta_{p,i} \langle -| a_j a_q a_l^\dagger a_k^\dagger | - \rangle \\ &= \underbrace{(-1)^2 \langle -| a_p^\dagger a_j a_i a_q a_l^\dagger a_k^\dagger | - \rangle}_{=0} + (-1) \delta_{p,j} \langle -| a_i a_q a_l^\dagger a_k^\dagger | - \rangle \\ &\quad + (-1) \delta_{p,i} \langle -| a_j a_l^\dagger a_q a_k^\dagger | - \rangle + \delta_{p,i} \delta_{l,q} \langle -| a_j a_k^\dagger | - \rangle \\ &= (-1)^2 \delta_{p,j} \langle -| a_i a_l^\dagger a_q a_k^\dagger | - \rangle + (-1) \delta_{p,j} \delta_{q,l} \langle -| a_i a_k^\dagger | - \rangle \\ &\quad + \underbrace{(-1)^2 \delta_{p,i} \langle -| a_l^\dagger a_j a_q a_k^\dagger | - \rangle}_{=0} + (-1) \delta_{p,i} \delta_{l,j} \langle -| a_q a_k^\dagger | - \rangle \\ &\quad + \underbrace{(-1) \delta_{p,i} \delta_{l,q} \langle -| a_k^\dagger a_j | - \rangle}_{=0} + \delta_{p,i} \delta_{l,q} \underbrace{\delta_{k,j} \langle -| I | - \rangle}_{=1} \\ &= (-1)^3 \delta_{p,j} \underbrace{\langle -| a_l^\dagger a_i a_q a_k^\dagger | - \rangle}_{=0} + (-1)^2 \delta_{p,j} \delta_{i,l} \langle -| a_q a_k^\dagger | - \rangle \quad (190) \\ &\quad + \underbrace{(-1)^2 \delta_{p,j} \delta_{q,l} \langle -| a_k^\dagger a_i | - \rangle}_{=0} + (-1) \delta_{p,j} \delta_{q,l} \underbrace{\delta_{i,k} \langle -| I | - \rangle}_{=1} \\ &\quad + \underbrace{(-1)^2 \delta_{p,i} \delta_{l,j} \langle -| a_k^\dagger a_q | - \rangle}_{=0} + (-1) \delta_{p,i} \delta_{l,j} \delta_{q,k} \underbrace{\langle -| I | - \rangle}_{=1} \\ &\quad + \delta_{p,i} \delta_{l,q} \delta_{k,j} \\ &= (-1)^3 \delta_{p,j} \delta_{i,l} \underbrace{\langle -| a_k^\dagger a_q | - \rangle}_{=0} + (-1)^2 \delta_{p,j} \delta_{i,l} \underbrace{\delta_{q,k} \langle -| I | - \rangle}_{=1} \\ &\quad + (-1) \delta_{p,j} \delta_{q,l} \delta_{i,k} + (-1) \delta_{p,i} \delta_{l,j} \delta_{q,k} + \delta_{p,i} \delta_{l,q} \delta_{k,j} \\ &= \delta_{p,j} \delta_{i,l} \delta_{q,k} - \delta_{p,j} \delta_{q,l} \delta_{i,k} - \delta_{p,i} \delta_{l,j} \delta_{q,k} + \delta_{p,i} \delta_{l,q} \delta_{k,j} \end{aligned}$$

Hence

$$\begin{aligned} \langle ij | \sum_{p,q} h_{p,q} a_p^\dagger a_q | kl \rangle &= \sum_{p,q} h_{p,q} (\delta_{p,j} \delta_{i,l} \delta_{q,k} - \delta_{p,j} \delta_{q,l} \delta_{i,k} - \delta_{p,i} \delta_{l,j} \delta_{q,k} + \delta_{p,i} \delta_{l,q} \delta_{k,j}) \\ &= h_{j,k} \delta_{i,l} - h_{j,l} \delta_{i,k} - h_{i,k} \delta_{l,j} + h_{i,l} \delta_{k,j} \end{aligned}$$

Which can be evaluated in $\mathcal{O}(1)$. For comparison, a naïve VecMatVec evaluations in first quantization costs $\mathcal{O}\left(\binom{n}{d}^3\right)$. To reach the final formula of the matrix elements following the procedure in Eq. (190) is very tedious. A shorter path is Wick's theorem.

8.1 Wick's contraction theorem

At the core of Wick's theorem stands the normal ordering of an operator string O which we shall denote by $\{O\}$. The normal ordering of an operator string means that all creation operators are to the left of the string and all annihilation operators are to the right. The procedure is to move all creation operators are to the left as if they anti-commuted.

Example 8.0.1. Let's investigate the normal ordering of operator strings of length two:

$$\{a_p a_q\} = a_p a_q \quad \{a_p^\dagger a_q^\dagger\} = a_p^\dagger a_q^\dagger \quad \{a_p^\dagger a_q\} = a_p^\dagger a_q \quad \{a_p a_q^\dagger\} = -a_q^\dagger a_p \quad (191)$$

Example 8.0.2. Let's look at a second, slightly more complicated example

$$\{a_s a_r a_p^\dagger a_q a_t^\dagger a_u^\dagger\} = \{a_p^\dagger a_s a_r a_q a_t^\dagger a_u^\dagger\} = -\{a_p^\dagger a_t^\dagger a_s a_r a_q a_u^\dagger\} = a_p^\dagger a_t^\dagger a_u^\dagger a_s a_r a_q \quad (192)$$

A key observation is that the vacuum expectation value of a normal ordered string of operators is zero, i.e.,

$$\langle -|\{O\}|-\rangle = 0. \quad (193)$$

Note that this is an immediate consequence of

$$\mathbf{0} = a_i |-\rangle = \left(\langle -| a_i^\dagger \right)^\dagger. \quad (194)$$

We now define contractions of two operators as

$$\overline{\square}_{xy} = xy - \{xy\} \quad (195)$$

For the creation and annihilation operators there are now four options in total, namely

$$\begin{aligned} \overline{\square}_{p q} &= a_p^\dagger a_q^\dagger - \{a_p^\dagger a_q^\dagger\} = a_p^\dagger a_q^\dagger - a_p^\dagger a_q^\dagger = 0 \\ \overline{\square}_{p q} &= a_p a_q - \{a_p a_q\} = a_p a_q - a_p a_q = 0 \\ \overline{\square}_{p q} &= a_p^\dagger a_q - \{a_p^\dagger a_q\} = a_p^\dagger a_q - a_p^\dagger a_q = 0 \\ \overline{\square}_{p q} &= a_p a_q^\dagger - \{a_p a_q^\dagger\} = a_p a_q^\dagger + a_q^\dagger a_p = \delta_{p,q} \end{aligned} \quad (196)$$

We see that the only non-zero contribution occurs when an annihilation operator appears to the left of a creation operator.

We are now able to formulate Wick's theorem:

Theorem 8.1. An operator string can be written as a linear combination of normal-ordered strings, i.e.,

$$\begin{aligned}
ABC\dots XYZ &= \{ABC\dots XYZ\} \\
&\quad + \sum_{\text{singles}} \{\overbrace{ABC\dots XYZ}^{\square}\} \\
&\quad + \sum_{\text{doubles}} \{\overbrace{\overbrace{ABC\dots XYZ}^{\square}}^{\square}\} \\
&\quad + \dots
\end{aligned} \tag{197}$$

Corollary 8.1.1. *Only fully contracted terms contribute to vacuum expectation values.*

With these results, we may now return to the problem of evaluating vacuum expectation values. Recall that by direct – yet tedious – application of the CAR we found

$$\langle -|a_j a_i a_p^\dagger a_q a_l^\dagger a_k^\dagger |-\rangle = \delta_{p,j} \delta_{i,l} \delta_{q,k} - \delta_{p,j} \delta_{q,l} \delta_{i,k} - \delta_{p,i} \delta_{l,j} \delta_{q,k} + \delta_{p,i} \delta_{l,q} \delta_{k,j} \tag{198}$$

Alternatively to the bruit force derivation, we may apply Wick's theorem to the expectation value. Noting that only fully contracted terms contribute to the vacuum expectation value we find

$$\begin{aligned}
\langle -|a_j a_i a_p^\dagger a_q a_l^\dagger a_k^\dagger |-\rangle &= \langle -|a_j a_i a_p^\dagger a_q a_l^\dagger a_k^\dagger |-\rangle \\
&\quad + \langle -|a_j a_i a_p^\dagger a_q a_l^\dagger a_k^\dagger |-\rangle \\
&\quad + \langle -|a_j a_i a_p^\dagger a_q a_l^\dagger a_k^\dagger |-\rangle \\
&\quad + \langle -|a_j a_i a_p^\dagger a_q a_l^\dagger a_k^\dagger |-\rangle \\
&= \delta_{p,j} \delta_{i,l} \delta_{q,k} - \delta_{p,j} \delta_{q,l} \delta_{i,k} - \delta_{p,i} \delta_{l,j} \delta_{q,k} + \delta_{p,i} \delta_{l,q} \delta_{k,j}
\end{aligned} \tag{199}$$

Clearly, this is much more tractable than the brut force derivation. However, it is still very tedious when we are considering a general N -particle state. Consider the expectation value

$$\langle i_1 i_2 \dots i_N | a_p^\dagger a_q^\dagger a_r a_s | j_1 j_2 \dots j_N \rangle = \langle -|a_{i_N} \dots a_{i_1} a_p^\dagger a_q^\dagger a_r a_s a_{j_N}^\dagger \dots a_{j_1}^\dagger |-\rangle \tag{200}$$

which rapidly yields an intractable number of Wick contractions.

8.2 Particle-Hole formalism

Fortunately, Wick's theorem extends to the Fermi-vaccuum, that is, a single reference Slater determinant. To that end, we consider the Slater determinant

$$|\phi_0\rangle = |\mathbf{0}\rangle = a_{j_N}^\dagger \dots a_{j_1}^\dagger |-\rangle, \tag{201}$$

and without loss of generality we assume $j_i = i$. As before, we denote the orbitals $i \in \llbracket N \rrbracket$ as occupied and $a \in \llbracket K \rrbracket \setminus \llbracket N \rrbracket$ as virtual. Before proceeding, we make the observation that the vacuum state $|-\rangle$ is the only occupation number vector for which

$$a_p |-\rangle = \mathbf{0} \quad \forall p \in \llbracket K \rrbracket \tag{202}$$

holds. Interestingly, we observe that

$$a_a |\phi_0\rangle = a_i^\dagger |\phi_0\rangle = \mathbf{0} \quad \forall i \in [\![N]\!] \text{ and } a \in [\![K]\!] \setminus [\![N]\!] \quad (203)$$

This suggests that we can interpret the reference state $|\phi_0\rangle$ as ‘‘vacuum’’ state in a quasi-particle picture, called the Fermi-vacuum. We can then define the ladder operators with respect to this quasi-particle picture, i.e.,

$$c_p := \begin{cases} a_p^\dagger, & \text{if } p \in [\![N]\!] \\ a_p, & \text{else.} \end{cases} \quad (204)$$

The physical interpretation of this is that $c_a^\dagger = a_a^\dagger$ creates a particle (electron) and $c_i^\dagger = a_i$ creates a hole (vacancy). We can then extend Wick’s theorem to this new framework where we are seeking normal ordering with respect to the Fermi-vacuum, denoted by $\{\cdot\}_0$. For the one-body Hamiltonian we find

$$\sum_{p,q=1}^K h_{pq} a_p^\dagger a_q = \sum_{p,q=1}^K h_{pq} \left(\{a_p^\dagger a_q\}_0 + \overline{\{a_p^\dagger a_q\}_0} \right) \quad (205)$$

Recall that the only non-zero contribution appears when an annihilation operator appears to the left of a creation operator, i.e., $c_p c_q^\dagger$. Given the definition of the particle-hole operators this only occurs if $p, q \in [\![N]\!]$, leading to

$$\overline{\{a_p^\dagger a_q\}_0} = \delta_{p,q} \mathbb{1}_{[\![N]\!]}(p). \quad (206)$$

Hence,

$$\sum_{p,q=1}^K h_{pq} a_p^\dagger a_q = \sum_{p,q=1}^K h_{pq} \{a_p^\dagger a_q\}_0 + \sum_{i=1}^N h_{ii} \quad (207)$$

For the two-electron part, we note that the non-zero contractions are

$$\begin{aligned} a_p^\dagger a_r^\dagger a_s a_q &= \{a_p^\dagger a_r^\dagger a_s a_q\}_0 \\ &\quad + \overline{\{a_p^\dagger a_r^\dagger a_s a_q\}_0} + \overline{\{a_p^\dagger a_r^\dagger a_s a_q\}_0} + \{a_p^\dagger a_r^\dagger a_s a_q\}_0 + \overline{\{a_p^\dagger a_r^\dagger a_s a_q\}_0} \\ &\quad + \overline{\{a_p^\dagger a_r^\dagger a_s a_q\}_0} + \overline{\{a_p^\dagger a_r^\dagger a_s a_q\}_0} \end{aligned} \quad (208)$$

Evaluating the contraction for the double contractions leads to

$$\begin{aligned} \overline{\{a_p^\dagger a_r^\dagger a_s a_q\}_0} &= -\delta_{p,s} \mathbb{1}_{[\![N]\!]}(p) \delta_{r,q} \mathbb{1}_{[\![N]\!]}(q) \\ \overline{\{a_p^\dagger a_r^\dagger a_s a_q\}_0} &= \delta_{p,q} \mathbb{1}_{[\![N]\!]}(p) \delta_{r,s} \mathbb{1}_{[\![N]\!]}(r) \end{aligned} \quad (209)$$

which results in

$$\begin{aligned}
& \frac{1}{2} \sum_{pqrs=1}^K v_{p,q,r,s} \left(\{a_p^\dagger a_r^\dagger a_s a_q\}_0 + \{a_p^\dagger a_r^\dagger a_s a_q\}_0 \right) \\
& = \frac{1}{2} \sum_{pqrs=1}^K v_{p,q,r,s} (\delta_{p,q} \mathbb{1}_{[N]}(p) \delta_{r,s} \mathbb{1}_{[N]}(r) - \delta_{p,s} \mathbb{1}_{[N]}(p) \delta_{r,q} \mathbb{1}_{[N]}(q)) \\
& = \frac{1}{2} \sum_{p,r=1}^N (v_{p,p,r,r} - v_{p,r,r,p})
\end{aligned} \tag{210}$$

Similarly, for the singles we find

$$\begin{aligned}
\{a_p^\dagger a_r^\dagger a_s a_q\}_0 &= -\{a_p^\dagger a_s a_r^\dagger a_q\}_0 = -\delta_{p,s} \mathbb{1}_{[N]}(p) \{a_r^\dagger a_q\}_0 \\
\{a_p^\dagger a_r^\dagger a_s a_q\}_0 &= \{a_p^\dagger a_q a_r^\dagger a_s\}_0 = \delta_{p,q} \mathbb{1}_{[N]}(p) \{a_r^\dagger a_s\}_0 \\
\{a_p^\dagger a_r^\dagger a_s a_q\}_0 &= \delta_{r,s} \mathbb{1}_{[N]}(r) \{a_p^\dagger a_q\}_0 \\
\{a_p^\dagger a_r^\dagger a_s a_q\}_0 &= -\{a_p^\dagger a_r^\dagger a_q a_s\}_0 = -\delta_{r,q} \mathbb{1}_{[N]}(r) \{a_p^\dagger a_s\}_0
\end{aligned} \tag{211}$$

which results in

$$\begin{aligned}
& \frac{1}{2} \sum_{pqrs=1}^K v_{p,q,r,s} \left(\{a_p^\dagger a_r^\dagger a_s a_q\}_0 + \{a_p^\dagger a_r^\dagger a_s a_q\}_0 + \{a_p^\dagger a_r^\dagger a_s a_q\}_0 + \{a_p^\dagger a_r^\dagger a_s a_q\}_0 \right) \\
& = \frac{1}{2} \sum_{pqrs=1}^K v_{p,q,r,s} \left(\delta_{p,q} \mathbb{1}_{[N]}(p) \{a_r^\dagger a_s\}_0 - \delta_{p,s} \mathbb{1}_{[N]}(p) \{a_r^\dagger a_q\}_0 \right. \\
& \quad \left. + \delta_{r,s} \mathbb{1}_{[N]}(r) \{a_p^\dagger a_q\}_0 - \delta_{r,q} \mathbb{1}_{[N]}(r) \{a_p^\dagger a_s\}_0 \right) \\
& = \frac{1}{2} \left(\sum_{rs=1}^K \sum_{i=1}^N v_{i,i,r,s} \{a_r^\dagger a_s\}_0 - \sum_{qr=1}^K \sum_{i=1}^N v_{i,q,r,i} \{a_r^\dagger a_q\}_0 \right. \\
& \quad \left. + \sum_{pq=1}^K \sum_{i=1}^N v_{p,q,i,i} \{a_p^\dagger a_q\}_0 - \sum_{ps=1}^K \sum_{i=1}^N v_{p,i,i,s} \{a_p^\dagger a_s\}_0 \right) \\
& = \frac{1}{2} \sum_{pq=1}^K \sum_{i=1}^N \left(v_{i,i,p,q} \{a_p^\dagger a_q\}_0 - v_{i,q,p,i} \{a_p^\dagger a_q\}_0 \right. \\
& \quad \left. + v_{p,q,i,i} \{a_p^\dagger a_q\}_0 - v_{p,i,i,q} \{a_p^\dagger a_q\}_0 \right) \\
& = \sum_{pq=1}^K \sum_{i=1}^N (v_{p,q,i,i} - v_{p,i,i,q}) \{a_p^\dagger a_q\}_0
\end{aligned} \tag{212}$$

where used that

$$v_{p,q,r,s} = v_{r,s,p,q}. \tag{213}$$

Therefore, the final form of the electronic structure Hamiltonian is given by

$$H = E^{\text{HF}} + \sum_{p,q=1}^K \left(h_{pq} + \sum_{i=1}^N (v_{p,q,i,i} - v_{p,i,i,q}) \right) \{a_p^\dagger a_q\}_0 + \frac{1}{2} \sum_{pqrs} v_{pqrs} \{a_p^\dagger a_r^\dagger a_s a_q\}_0 \quad (214)$$

where the Hartree–Fock energy appears naturally

$$E^{(\text{HF})} = \sum_{i=1}^N h_{ii} + \frac{1}{2} \sum_{p,r=1}^N (v_{p,p,r,r} - v_{p,r,r,p}) = \langle \phi_0 | H | \phi_0 \rangle. \quad (215)$$

The normal ordered Hamiltonian is then given by

$$H_N := H - \langle \phi_0 | H | \phi_0 \rangle = \sum_{p,q=1}^K f_{p,q} \{a_p^\dagger a_q\}_0 + \frac{1}{2} \sum_{pqrs} v_{pqrs} \{a_p^\dagger a_r^\dagger a_s a_q\}_0 \quad (216)$$

where

$$f_{p,q} := h_{pq} + \sum_{i=1}^N (v_{p,q,i,i} - v_{p,i,i,q}). \quad (217)$$