Atomic Physics

lectures

University of Amsterdam

J.T.M. Walraven

August 18, 2023

Contents

Fundamental constants

The values given below for the fundamental physical constants correspond to the recommended values (CODATA-2018).

$\Delta\nu_{\mathrm{Cs}}$	$9\,192\,631\,770$	Hz	^{133}Cs hyperfine transition frequency	def
c	$299\,792\,458$	$\mathrm{m\,s^{-1}}$	speed of light in vacuum	def
h	$6.626\,070\,15 \times 10^{-34}$	$\mathrm{J\,Hz^{-1}}$	Planck constant	def
$\hbar$	$1.054\,571\,817 \cdots \times 10^{-34}$	$\mathrm{J\,s}$	Planck constant (reduced)	$\hbar = h/2\pi$
e	$1.602\,176\,634 \times 10^{-19}$	C	elementary charge	def
N_{A}	$6.022\,140\,76 \times 10^{23}$	$\mathrm{mol^{-1}}$	Avogadro constant	def
k_B	$1.380\,649 \times 10^{-23}$	$\mathrm{J\,K^{-1}}$	Boltzmann constant	def
K_{cd}	683	$\mathrm{lm\,W^{-1}}$	Luninous efficacy	def
K_J	$4.835\,978\,484 \cdots \times 10^{14}$	$\mathrm{Hz\,V^{-1}}$	Josephson constant	$K_J = 2e/h$
R_K	$25\,812\,807\,45 \cdots$	Ω	Von Klitzing constant	$R_K = h/e^2$
R	$8.314\,462\,618 \cdots$	$\mathrm{J\,mol^{-1}K^{-1}}$	Molar gas constant	$R = N_A k_B$
σ	$5.670\,374\,419 \cdots \times 10^{-8}$	$\mathrm{W\,m^{-2}K^{-4}}$	Stephan-Boltzmann constant	$\sigma = \pi^2 k_B^4/60\hbar^3 c^2$
μ_0	$1.256\,637\,062\,12(19) \times 10^{-6}$	$\mathrm{N\,A^{-2}}$	magnetic permeability of vacuum	$\mu_0 = 2\alpha h/e^2 c$
ε_0	$8.854\,187\,8128(13) \times 10^{-12}$	$\mathrm{F\,m^{-1}}$	electric permitivity of vacuum	$\varepsilon_0 = 1/\mu_0 c^2$
α	$7.297\,352\,5693(11) \times 10^{-3}$		fine-structure constant	$\alpha = e^2/2\varepsilon_0 hc$
α^{-1}	$137.035\,999\,084(21)$		inverse fine-structure constant	$\alpha^{-1} = 2\varepsilon_0 hc/e^2$
a_0	$0.529\,177\,210\,92(17) \times 10^{-10}$	m	Bohr radius	$a_0 = \hbar/\alpha m_e c$
m_e	$9.109\,383\,7015(28) \times 10^{-31}$	kg	electron mass	
m_p	$1.007\,276\,466\,621(53)$	m_u	proton mass	
m_p/m_e	$1836.152\,673\,43(11)$		proton-electron mass ratio	
R_∞	$1.097\,373\,156\,8160(21) \times 10^{7}$	$\mathrm{m^{-1}}$	Rydberg constant	$R_\infty = \alpha^2 m_e c/2h$
Ry	$3.289\,841\,960\,2508(64) \times 10^{15}$	Hz	Rydberg frequency	$R_\infty c$
λ_C	$3.861\,592\,6796(12) \times 10^{-13}$	m	Compton wavelength (reduced)	$\lambda_C = \hbar/m_e c = \alpha a_0$
μ_B	$9.274\,010\,0783(28) \times 10^{-24}$	$\mathrm{J\,T^{-1}}$	Bohr magneton	$\mu_B = e\hbar/2m_e$
g_e	$2.002\,319\,304\,362\,56(35)$		electron g factor [1]	$g_e = 2(1 + a_e)$
a_e	$1.159\,652\,181\,28(18) \times 10^{-3}$		electron anomaly	$a_e = \lvert\mu_e\rvert/\mu_B - 1$
m_d	$2.013\,553\,212\,745(40)$	m_u	deuteron mass	
μ_N	$5.050\,783\,699(31) \times 10^{-27}$	$\mathrm{J\,T^{-1}}$	nuclear magneton	$\mu_N = e\hbar/2m_p$
g_p	$5.585\,694\,689\,3(16)$		proton g factor	$g_p = 2\mu_p/\mu_N$
g_d	$0.857\,438\,233\,8(22)$		deuteron g factor	$g_d = 2\mu_d/\mu_N$
r_p	$0.87\,51(61) \times 10^{-15}$	m	proton rms charge radius	
m_u	$1.660\,539\,066\,60(50) \times 10^{-27}$	kg	atomic mass constant	
M_u	$0.999\,999\,999\,65(30) \times 10^{-3}$	$\mathrm{kg\,mol^{-1}}$	molar mass constant	$M_u = m_u N_{\mathrm{A}}$
u	1	m_u	unified atomic mass unit (dalton)	$\mathrm{u} = \frac{1}{12} m(^{12}\mathrm{C})$

[1] Here we deviate from the CODATA recommendation in which g_e is defined as a negative number.

Introduction

The course *Atomic Physics* offers a systematic introduction in *atomic structure theory.* Ironically, it deals with the *internal* structure of what was once was considered to be "the smallest unit of matter" [102]. The exchange of energy between atoms and their environment is *not* part of the course as this forms a huge topic in itself. Furthermore, the course is *not* simply about the physics of the hydrogen atom. Hydrogen is only the first of the more than 100 elements listed in the periodic system (see Appendix A). Moreover, many more atomic forms can be created by ionizing the elements or by combining charged particles into exotic atoms like positronium or muonium. In many cases the understanding of such atoms calls upon physical principles which play no role in hydrogen.

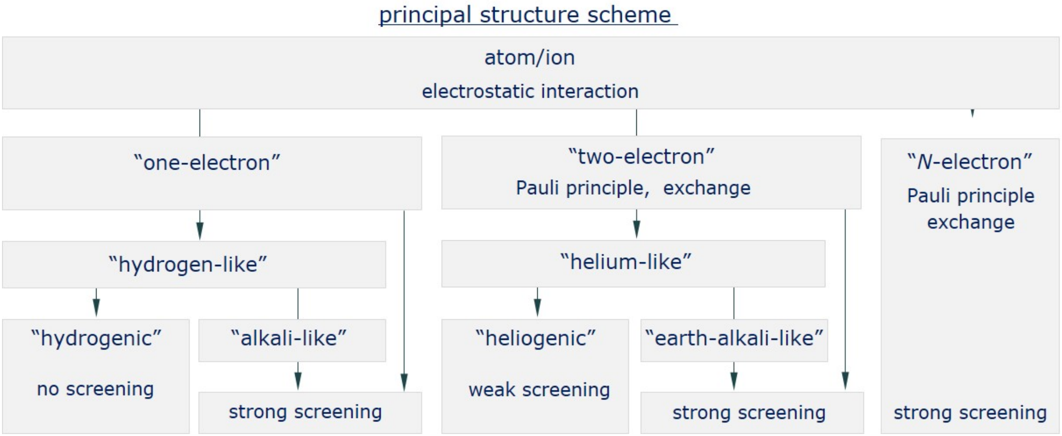
Atomic Physics builds on the great scientific advances of the 19^{th} century in the theory of *Classical Mechanics* and *Electromagnetism.* It took the discovery of the *electron* (β particle) by Joseph Thomson in [29] and the helium *nucleus* (α particle) by Ernest Rutherford in [96] before the foundation of the modern atomic structure theory could be established by Niels Bohr [17, 18, 19]. The Bohr atom was proposed in the language of the "old quantum theory". We use the well-known modern quantum mechanics as developed by Erwin Schrödinger and Paul Dirac [100, 33]. Actually, the advances in atomic physics over the passed century mark, more in general, the discovery of deep new insights in physics, often achieved in close dialogue with experiment. For *Atomic Physics* we made a selection of some of the finest of these advances, keeping in mind the main goal of the course:

Understanding the structure and properties of all atomic systems.

This goal may seem ambitious at first sight but fortunately many-electron atoms have inherited a lot from the simple hydrogen atom. For instance, all atoms remain close to spherical in symmetry because the electrostatic interaction is dominant. So, by adding only a few new ingredients our goal can be reached for all but a few notable exceptions. Moreover these ingredients are already well-know for most students: *quantum statistics* and the *Pauli principle.* The two electrons of the helium atom teach us how these concepts emerge from experiment. Furthermore, we learn about *exchange*, which is a new symmetry - absent in classical physics. Along the way we say farewell to exact solutions, but in return we learn beautiful new concepts, such as the *selfconsistent field* of Hartree and Fock [56, 57, 43, 103] and the *many-body formalism* of Jordan and Wigner [63]. Without entering into numerical details we learn how to analyze the energy levels of valence electrons, in particular the theoretical underpinning of the *three Hund rules,* which formulate the criteria for the atomic ground state [59]. In this way we succeed in our goal for most of the elements of the periodic system and many more exotic atomic variations. To understand the intriguing exceptions we need the Hartree-Fock numerics but this falls outside the scope of the course.

Outline

The course consists of 12 lectures covering 9 chapters, 4 of which focus on the mathematics of the material. Our task is organized while keeping in mind the hierarchy of electrostatic interaction phenomena as sketched in the *principal structure scheme* shown on the next page. In all cases the electrostatic interaction is dominant but the theory depends strongly on the presence or absence of



Hierarchy of electrostatic interactions in atoms/ions as a guideline for their description.

Pauli exclusion and *exchange* as well as the *level of screening* of the nuclear charge by core electrons. The first 5 lectures focus on *hydrogen-like, one-electron atoms*, the rest of the course covers the *many-electron* case with increasing levels of complexity. The pivoting point is the helium atom (Chapter 6), which perfectly serves to introduce the role of the Pauli principle and the exchange interaction, which dominate the later chapters. To clearly separate the story line from the background material we include a set of appendices with supplemental information.

1

Quantum motion in a central potential field

The motion of particles in a *central* potential plays an important role in atomic and molecular physics. To understand the properties of the hydrogen atom we rely on careful analysis of the motion of the electron in the Coulomb field of the nucleus. Likewise, many properties related to interactions between atoms, such as collisional properties, can be understood by analyzing the relative atomic motion under the influence of central forces.

In view of the importance of central forces we summarize in this chapter the derivation of the Schrödinger equation for the motion of two particles, say of masses m_1 and m_2, interacting through a *central* potential $\mathcal{V}(r)$, $r = |\mathbf{r}_1 - \mathbf{r}_2|$ being the radial distance between the particles. For such potentials, purely depending on the *relative* distance between the particles, it is (in the absence of externally applied fields) practical to eliminate the center of mass motion of the pair and represent the relative motion by a single particle of *reduced mass* $m_r = m_1 m_2/(m_1+m_2)$ in the same potential field (see Appendix D.7). To further exploit the symmetry we can separate the radial motion from the rotational motion, obtaining the radial and angular momentum operators as well as the hamiltonian operator in spherical coordinates (Section 1.1). Knowing the Hamiltonian we can write down the Schrödinger equation (Section 1.2) and specializing to specific angular momentum values we obtain the *radial wave equation.* The radial wave equation is the central equation for the description of the radial motion associated with specific angular momentum states.

The approach just described amounts mathematically to the method of separation of variables for solving differential equations. This suggests to extend the discussion to cylindrical coordinates, as this opens - with little additional effort - a window onto related problems like quantization of electronic orbitals into Landau levels as well as the description of the flow fields of quantized vortices. In these cases the central potential is absent but the solutions are rotational in character; hence, show a form of central symmetry.

1.1 Hamiltonian

The classical Hamiltonian for the motion of a particle of (reduced) mass m_r in the central potential $\mathcal{V}(r)$ is given by the sum of the kinetic and potential energy,

$$H = \tfrac{1}{2} m_r \mathbf{v}^2 + \mathcal{V}(r), \tag{1.1}$$

where $\mathbf{v} = \dot{\mathbf{r}}$ is the velocity of the particle with $\mathbf{r}$ its position relative to the potential center. In the absence of externally applied fields $\mathbf{p} = m_r \mathbf{v}$ is the canonical momentum of the particle and the Hamiltonian can be written as[1]

$$H_0 = \frac{\mathbf{p}^2}{2m_r} + \mathcal{V}(r). \tag{1.2}$$

---

[1] In the presence of an external electromagnetic field the *non-relativistic* momentum of a *charged* particle of mass m and charge q is given by $\mathbf{p} = m\mathbf{v} + q\mathbf{A}$, with $m\mathbf{v}$ its *kinetic* momentum and $q\mathbf{A}$ its *electromagnetic* momentum.

To exploit the central symmetry we separate the radial motion from the angular motion by writing the Hamiltonian in the form (see Problem 1.1)

$$H_0 = \frac{1}{2m_r}\left(p_r^2 + \frac{\mathbf{L}^2}{r^2}\right) + \mathcal{V}(r) \quad (r \neq 0). \tag{1.3}$$

Here $p_r = \hat{\mathbf{r}} \cdot \mathbf{p}$ (see Fig. 1.1) is the *radial momentum*, with $\hat{\mathbf{r}} = \mathbf{r}/r$ the unit vector in radial direction, and $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ the *orbital angular momentum* with respect to the potential center. As Eq. (1.3) is well-defined only *outside* the origin any result based on this expression should be tested for its validity at the origin.

Problem 1.1. Derive the vector relation (see also Appendix N).

$$(\hat{\mathbf{r}} \cdot \mathbf{p})^2 + (\hat{\mathbf{r}} \times \mathbf{p})^2 = \mathbf{p}^2. \tag{1.4}$$

Solution. In the Einstein notation with summation over repeating indices the cartesian components of $\hat{\mathbf{r}} \times \mathbf{p}$ are given by $(\hat{\mathbf{r}} \times \mathbf{p})_i = \varepsilon_{ijk}\hat{r}_j p_k$, where $i, j, k \in \{x, y, z\}$ and ε_{ijk} is the Levi-Civita tensor

$$\varepsilon_{ijk} = \begin{cases} 1 & \text{for } \textit{even} \text{ permutations of } x, y, z \\ 0 & \text{for } i = j \text{ or } i = k \text{ or } j = k \\ -1 & \text{for } \textit{odd} \text{ permutations of } x, y, z. \end{cases} \tag{1.5}$$

Using the summation convention, the contraction of the Levi-Civita tensor is given by

$$\varepsilon_{ijk}\varepsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}.$$

Since $\mathbf{L}^2 = L_i L_i$ we obtain

$$\begin{aligned} (\hat{\mathbf{r}} \times \mathbf{p})^2 &= (\varepsilon_{ijk}\hat{r}_j p_k)(\varepsilon_{ilm}\hat{r}_l p_m) = (\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl})\,\hat{r}_j p_k \hat{r}_l p_m \\ &= (\hat{r}_j \hat{r}_j)(p_k p_k) - \hat{r}_j p_j \hat{r}_k p_k = \mathbf{p}^2 - (\hat{\mathbf{r}} \cdot \mathbf{p})^2. \end{aligned}$$

□

1.1.1 Quantization of the Hamiltonian - basic commutation relations

The transition from classical mechanics to quantum mechanics is made by postulating that the dynamical variables for position and momentum correspond to the following *hermitian operators* in the *position representation*, $\mathbf{p} \to -i\hbar\boldsymbol{\nabla}$ and $\mathbf{r} \to \underline{\mathbf{r}}$.[1] This is known as *quantization by correspondence.* With this quantization rule, Eq. (1.2) becomes the quantum mechanical Hamiltonian and takes the familiar form of the *Schrödinger Hamiltonian* for a particle in a central field,

$$\mathcal{H}_0 = -\frac{\hbar^2}{2m_r}\Delta + \mathcal{V}(r). \tag{1.6}$$

Although the quantization went smoothly in this case, in general we should watch out for ambiguities in the application of the correspondence rules. For instance, whereas in classical mechanics the expressions $p_r = \hat{\mathbf{r}} \cdot \mathbf{p}$ and $p_r = \mathbf{p} \cdot \hat{\mathbf{r}}$ are equivalent this does not hold for $p_r = -i\hbar(\boldsymbol{\nabla} \cdot \hat{\mathbf{r}})$ and $p_r = -i\hbar(\hat{\mathbf{r}} \cdot \boldsymbol{\nabla})$ because $\hat{\mathbf{r}} = \mathbf{r}/r$ and $-i\hbar\boldsymbol{\nabla}$ do not commute.

Up to this point we did not make any choice of coordinate system (metric). To deal with non-commutativity for a given coordinate system the operator algebra has to be completed with commutation relations for the operators. In arbitrary *orthogonal* curvilinear coordinates $\mathbf{r} = \mathbf{r}(r_1, \cdots r_d)$ of a d-dimensional euclidean vector space the gradient vector is given by

$$\boldsymbol{\nabla} = \{h_1^{-1}\partial_1, \cdots, h_d^{-1}\partial_d\} = \hat{\mathbf{e}}_1 h_1^{-1}\partial_1 + \cdots + \hat{\mathbf{e}}_d h_d^{-1}\partial_d, \tag{1.7}$$

---

[1] Here we emphasized in the notation that $\underline{\mathbf{r}}$ is the position *operator* rather than the position $\mathbf{r}$. As this distinction rarely leads to confusion the underscore will be omitted in most of the text.

where the unit vectors are defined by $\hat{\mathbf{u}} \equiv \hat{\mathbf{e}}_u = \partial_u \mathbf{r}/|\partial_u \mathbf{r}|$ and the scale factors by $h_u \equiv |\partial_u \mathbf{r}|$, with $u \in \{r_1, \cdots r_d\}$. Here $\partial_u \equiv \partial/\partial r_u$ is a shorthand notation for the partial derivative operator. Note that $\partial_u \mathbf{r} = h_u \hat{\mathbf{u}}$.

In *cartesian* coordinates we have $\mathbf{r} = (r_1, \cdots r_d) = (x, y, z)$. As the radius vector is given by $\mathbf{r} = \hat{\mathbf{x}}x + \hat{\mathbf{y}}y + \hat{\mathbf{z}}z$ it follows that $\partial_x \mathbf{r} = \hat{\mathbf{x}}$, $\partial_y \mathbf{r} = \hat{\mathbf{y}}$, $\partial_z \mathbf{r} = \hat{\mathbf{z}}$ and $h_x = h_y = h_z = 1$. Note that the property $h_i = 1$, with $i \in \{1, \cdots, d\}$, is valid for an euclidean vector space of arbitrary dimension d. The commutation relations for the components of the operators r_i and $p_j = -i\hbar\partial_j$ are obtained by evaluating the action of the operator $[r_i, p_j]$ on a smooth test function of position $\phi(r_x, r_y, r_z)$,

$$[r_i, p_j]\,\phi = -i\hbar\,(r_i\partial_j - \partial_j r_i)\,\phi = -i\hbar\,(r_i\partial_j\phi - r_i\partial_j\phi - \phi\delta_{ij}) = i\hbar\delta_{ij}\phi. \tag{1.8}$$

Thus we derived the commutation relations

$$[r_i, p_j] = i\hbar\delta_{ij}. \tag{1.9}$$

These commutation relations hold for cartesian coordinates. In general, the direction of the unit vectors depends on position ($h_i \neq 1$) and the commutation relations *do not have this simple form.*

A consequence of the commutation relations (1.9) is that $\mathbf{r}$ and $\mathbf{p}$ do not commute with the Hamiltonian $\mathcal{H}_0$: for $\mathbf{p}$ we have $[p_i, \mathcal{H}_0] = [p_i, \mathcal{V}(r)] \neq 0$, for $\mathbf{r}$ we find an important relation between $\mathbf{p}$ and $\mathbf{r}$

$$[r_i, \mathcal{H}_0] = \left[r_i, \mathbf{p}^2/2m_r\right] = i(\hbar/m_r)p_i, \tag{1.10}$$

which can be written in the form

$$\mathbf{p} = -i\,(m_r/\hbar)\,[\mathbf{r}, \mathcal{H}_0]\,. \tag{1.11}$$

1.1.1.1 Laplacian in spherical coordinates

To explore the central symmetry of our problem we wish to write the laplacian from the Schrödinger Hamiltonian in *spherical coordinates* $\{r, \theta, \phi\}$. The relation between the cartesian coordinates and the spherical coordinates of a point $\mathbf{P} = \mathbf{P}(x, y, z) = \mathbf{P}\,(r, \theta, \phi)$ is given by (see Fig. 1.1)

$$x = r\sin\theta\cos\phi,\;\; y = r\sin\theta\sin\phi,\;\; z = r\cos\theta. \tag{1.12}$$

Using $\partial_a \mathbf{P} = h_a \hat{\mathbf{a}}$, with $a \in (\{r, \theta, \phi\}$, the angular dependence of the unit vectors is given by[1]

$$\hat{\mathbf{r}} = \hat{\mathbf{x}}\sin\theta\cos\phi + \hat{\mathbf{y}}\sin\theta\sin\phi + \hat{\mathbf{z}}\cos\theta \tag{1.13a}$$

$$\hat{\boldsymbol{\theta}} = \hat{\mathbf{x}}\cos\theta\cos\phi + \hat{\mathbf{y}}\cos\theta\sin\phi - \hat{\mathbf{z}}\sin\theta \tag{1.13b}$$

$$\hat{\boldsymbol{\phi}} = -\hat{\mathbf{x}}\sin\phi + \hat{\mathbf{y}}\cos\phi. \tag{1.13c}$$

Using $\mathbf{P} = \hat{\mathbf{r}}r = \mathbf{r}$ we calculate $h_r = |\partial\mathbf{r}/\partial r| = 1$, $h_\phi = |\partial\mathbf{r}/\partial\phi| = r(\sin^2\theta\sin^2\phi + \sin^2\theta\cos^2\phi)^{1/2} = r\sin\theta$ and $h_\theta = |\partial\mathbf{r}/\partial\theta| = r(\cos^2\theta\cos^2\phi + \cos^2\theta\sin^2\phi + \sin^2\theta)^{1/2} = r$.

Hence, in spherical coordinates the gradient operator (1.7) becomes

$$\boldsymbol{\nabla} = \hat{\mathbf{r}}\,\partial_r + \hat{\boldsymbol{\theta}}\frac{1}{r}\partial_\theta + \hat{\boldsymbol{\phi}}\frac{1}{r\sin\theta}\partial_\phi. \tag{1.14}$$

Evaluating the inner product $\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}$ we obtain for the Laplace operator (see Problem 1.2)

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}. \tag{1.15}$$

---

[1] In these lecture notes we use interchangeable notations for the radial unit vector: $\hat{\mathbf{r}} \equiv \Omega \equiv (\theta, \phi)$.

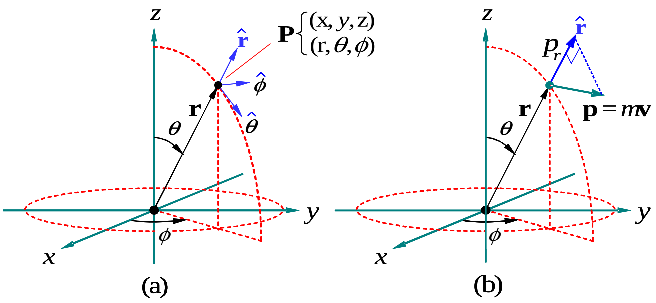


Figure 1.1: Illustration of spherical coordinates of a point $\mathbf{P} = \mathbf{P}(x, y, z) = \mathbf{P}\,(r, \theta, \phi)$: (a) unit vector convention; (b) vector diagram indicating the direction $\hat{\mathbf{r}}$ and amplitude p_r of the radial momentum vector.

Problem 1.2. Derive the expression (1.15) for the laplacian in spherical coordinates.

Solution. Starting from the vector expression (1.14) for the gradient operator the laplacian is obtained by evaluating the inner product

$$\begin{aligned}\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\psi = \hat{\mathbf{r}}\cdot\left(\partial_r\hat{\mathbf{r}}\,\partial_r + [\partial_r\hat{\boldsymbol{\theta}}]\frac{1}{r}\partial_\theta + [\partial_r\hat{\boldsymbol{\phi}}]\frac{1}{r\sin\theta}\partial_\phi\right)\psi \\ + \frac{1}{r}\hat{\boldsymbol{\theta}}\cdot\left([\partial_\theta\hat{\mathbf{r}}]\partial_r + \partial_\theta\hat{\boldsymbol{\theta}}\frac{1}{r}\partial_\theta + [\partial_\theta\hat{\boldsymbol{\phi}}]\frac{1}{r\sin\theta}\partial_\phi\right)\psi \\ + \frac{1}{r\sin\theta}\hat{\boldsymbol{\phi}}\cdot\left([\partial_\phi\hat{\mathbf{r}}]\partial_r + [\partial_\phi\hat{\boldsymbol{\theta}}]\frac{1}{r}\partial_\theta + \partial_\phi\hat{\boldsymbol{\phi}}\frac{1}{r\sin\theta}\partial_\phi\right)\psi.\end{aligned}$$

Here we dismissed many terms that yield zero due to the orthogonality of the unit operators. Using $\partial_r\hat{\mathbf{r}} = \partial_r\hat{\boldsymbol{\theta}} = \partial_r\hat{\boldsymbol{\phi}} = 0$ and $\partial_\theta\hat{\mathbf{r}} = \hat{\boldsymbol{\theta}}$, $\partial_\theta\hat{\boldsymbol{\theta}} = -\hat{\mathbf{r}}$, $\partial_\theta\hat{\boldsymbol{\phi}} = 0$ and $\partial_\phi\hat{\mathbf{r}} = \hat{\boldsymbol{\phi}}\,\sin\theta$, $\partial_\phi\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\phi}}\,\cos\theta$, $\partial_\phi\hat{\boldsymbol{\phi}} = -\,\hat{\mathbf{r}}\sin\theta - \hat{\boldsymbol{\theta}}\cos\theta$ the expression further simplifies to

$$\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\psi = \partial_r^2\boldsymbol{\psi} + \frac{1}{r}\left(\partial_r + \frac{1}{r}\partial_\theta^2\right)\boldsymbol{\psi} + \frac{1}{r\sin\theta}\left(\sin\theta\partial_r + \cos\theta\frac{1}{r}\partial_\theta + \frac{1}{r\sin\theta}\partial_\phi^2\right)\psi.$$

Collecting the terms we obtain for the Laplace operator

$$\Delta = \partial_r^2 + \frac{2}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2 + \frac{\cos\theta}{r^2\sin\theta}\partial_\theta + \frac{1}{r^2\sin^2\theta}\partial_\phi^2,$$

which can be written in the form (1.15). □

1.1.1.2 Laplacian in cylindrical coordinates

To describe central symmetry in *two-dimensional systems* or systems with *cylindrical symmetry* it is important to be aware of the expression for the laplacian in *cylindrical coordinates* $\mathbf{P} = \mathbf{P}(x, y, z) = \mathbf{P}\,(r_\perp, \phi, z)$. As illustrated in Fig. 1.2 the relation between the cartesian coordinates and the cylindrical coordinates is given by the expressions

$$x = r_\perp\cos\phi,\;\; y = r_\perp\sin\phi,\;\; z = z. \tag{1.16}$$

Using $\partial_a\mathbf{P} = h_a\hat{\mathbf{a}}$, with $a \in \{r_\perp, \phi, z\}$, the angular dependence of the unit vectors is given by

$$\hat{\mathbf{r}}_\perp = +\hat{\mathbf{x}}\cos\phi + \hat{\mathbf{y}}\sin\phi \tag{1.17a}$$

$$\hat{\boldsymbol{\phi}} = -\hat{\mathbf{x}}\sin\phi + \hat{\mathbf{y}}\cos\phi \tag{1.17b}$$

$$\hat{\mathbf{z}} = \hat{\mathbf{z}}. \tag{1.17c}$$

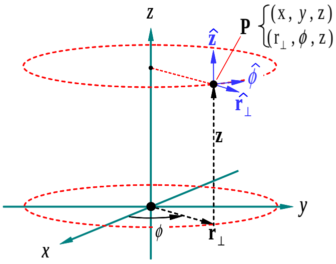


Figure 1.2: Illustration of cylindrical coordinates of a point $\mathbf{P} = \mathbf{P}(x, y, z) = \mathbf{P}\,(r_\perp, \phi, z)$ with unit vector convention.

Using $\mathbf{P} = \hat{\mathbf{r}}_\perp\, r_\perp + \hat{\mathbf{z}}\, z$ we have $\mathbf{P}^2 = r_\perp^2 + z^2$ and calculate $h_{r_\perp} = |\partial\mathbf{P}/\partial r_\perp| = 1$, $h_\phi = |\partial\mathbf{P}/\partial\phi| = r_\perp(\sin^2\phi + \cos^2\phi)^{1/2} = r_\perp$ and $h_z = |\partial\mathbf{P}/\partial z| = 1$.

Hence, in cylindrical coordinates the gradient operator (1.7) is given by

$$\boldsymbol{\nabla} = \hat{\mathbf{r}}_\perp\, \partial_{r_\perp} + \hat{\boldsymbol{\phi}}\frac{1}{r_\perp}\partial_\phi + \hat{\mathbf{z}}\, \partial_z. \tag{1.18}$$

Evaluating the inner product we obtain for the laplacian (see Problem 1.3)

$$\Delta = \frac{\partial^2}{\partial r_\perp^2} + \frac{1}{r_\perp}\frac{\partial}{\partial r_\perp} + \frac{1}{r_\perp^2}\frac{\partial^2}{\partial\phi^2} + \frac{\partial^2}{\partial z^2}. \tag{1.19}$$

Fixing the value of z this expression also serves to describe two-dimensional systems. In the form (1.19) the laplacian is used to describe the flow field of quantized vortices in superfluids.

Problem 1.3. Derive the expression (1.19) for the laplacian in cylindrical coordinates.

Solution. Starting from the vector expression (1.18) for the gradient operator the laplacian is obtained by evaluating the inner product (here we set $r_\perp \!\to \mathbf{r}$ for compactness of notation)

$$\begin{aligned}\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\psi = \hat{\mathbf{r}}\cdot\left(\partial_r\hat{\mathbf{r}}\,\partial_r + [\partial_r\hat{\boldsymbol{\phi}}]\frac{1}{r}\partial_\phi + [\partial_r\hat{\mathbf{z}}]\partial_z\right)\psi \\ + \frac{1}{r}\hat{\boldsymbol{\phi}}\cdot\left([\partial_\phi\hat{\mathbf{r}}]\partial_r + \partial_\phi\hat{\boldsymbol{\phi}}\frac{1}{r}\partial_\phi + [\partial_\phi\hat{\mathbf{z}}]\partial_z\right)\psi \\ + \hat{\mathbf{z}}\cdot\left([\partial_z\boldsymbol{r}]\,\partial_r + [\partial_z\hat{\boldsymbol{\phi}}]\frac{1}{r}\partial_\phi + \partial_z\hat{\mathbf{z}}\,\partial_z\right)\psi.\end{aligned}$$

Here we dismissed the many terms that yield zero due to the orthogonality of the unit operators. Using $\partial_r\hat{\mathbf{r}} = \partial_r\hat{\boldsymbol{\phi}} = \partial_r\hat{\mathbf{z}} = \partial_z\hat{\mathbf{r}} = \partial_z\hat{\boldsymbol{\phi}} = \partial_z\hat{\mathbf{z}} = 0$ and $\partial_\phi\hat{\mathbf{r}} = \hat{\boldsymbol{\phi}}$, $\partial_\phi\hat{\boldsymbol{\phi}} = -\hat{\mathbf{r}} + \hat{\mathbf{z}}\, z$, $\partial_\phi\hat{\mathbf{z}} = 0$ the expression for $\Delta\psi$ further simplifies to

$$\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\psi = \partial_r^2\psi + \frac{1}{r}\left(\partial_r + \partial_\phi\frac{1}{r}\partial_\phi\right)\psi + \partial_z^2\psi.$$

Thus, the Laplace operator can be written in the form (1.19). □

1.1.2 Angular momentum operator L

To obtain the operator expression for the angular momentum $\mathbf{L} = \mathbf{r}\times\mathbf{p}$ in the position representation we use the correspondence rules $\mathbf{p} \to -i\hbar\boldsymbol{\nabla}$ and $\mathbf{r} \to \underline{\mathbf{r}}$. Importantly, although $\mathbf{r}$ and $\mathbf{p}$ do not commute the transition to the quantum mechanical expression,

$$\mathbf{L} = -i\hbar\,(\mathbf{r}\times\boldsymbol{\nabla})\,, \tag{1.20}$$

can be made without ambiguity because the correspondence rules yield the same result for $\mathbf{L} = \mathbf{r}\times\mathbf{p}$ and for $\mathbf{L} = -\mathbf{p}\times\mathbf{r}$. This is easily verified by using the cartesian vector components of $\mathbf{r}$ and $\mathbf{p}$ and their commutation relations,

$$-\left(\mathbf{p}\times\mathbf{r}\right)_i = -\varepsilon_{ijk}p_j r_k = -\varepsilon_{ijk}r_k p_j = \varepsilon_{ikj}r_k p_j = \varepsilon_{ijk}r_j p_k = \left(\mathbf{r}\times\mathbf{p}\right)_i . \tag{1.21}$$

Note that for $j \neq k$ the operators r_j and p_k commute and for $j = k$ one has $\varepsilon_{ijk} = 0$.

Having identified Eq. (1.20) as the proper operator expression for the orbital angular momentum we can turn to arbitrary *orthogonal* curvilinear coordinates $\mathbf{r} = \mathbf{r}(u, v, w)$. In this case the gradient vector is given by $\boldsymbol{\nabla} = \{h_u^{-1}\partial_u, h_v^{-1}\partial_v, h_w^{-1}\partial_w\}$ and the angular momentum operator can be decomposed in the following form

$$\mathbf{L} = -\,i\hbar\left(\mathbf{r}\times\boldsymbol{\nabla}\right) = -\,i\hbar\begin{vmatrix} \hat{\mathbf{u}} & \hat{\mathbf{v}} & \hat{\mathbf{w}} \\ r_u & r_v & r_w \\ h_u^{-1}\partial_u & h_v^{-1}\partial_v & h_w^{-1}\partial_w \end{vmatrix}. \tag{1.22}$$

For *spherical coordinates* the components of the radius vector are $r_r = r$ and $r_\theta = r_\phi = 0$. Working out the determinant in Eq. (1.22), while *respecting the order* of the vector components r_u and $h_u^{-1}\partial_u$, we find for the angular momentum operator in spherical coordinates

$$\mathbf{L} = -i\hbar\left(\mathbf{r}\times\boldsymbol{\nabla}\right) = i\hbar\left(\hat{\boldsymbol{\theta}}\frac{1}{\sin\theta}\frac{\partial}{\partial\phi} - \hat{\boldsymbol{\phi}}\frac{\partial}{\partial\theta}\right). \tag{1.23}$$

Importantly, as was to be expected for a rotation operator in a spherical coordinate system, $\mathbf{L}$ depends only on the angles θ and ϕ and not on the radial distance r.

1.1.3 The operator L_z

The operator for the angular momentum along the z direction is a differential operator obtained by taking the inner product of $\mathbf{L}$ with the unit vector along the z direction, $L_z = \hat{\mathbf{z}}\cdot\mathbf{L}$. From Eq. (1.23) we see that

$$L_z = i\hbar\left((\hat{\mathbf{z}}\cdot\hat{\boldsymbol{\theta}})\frac{1}{\sin\theta}\frac{\partial}{\partial\phi} - (\hat{\mathbf{z}}\cdot\hat{\boldsymbol{\phi}})\frac{\partial}{\partial\theta}\right). \tag{1.24}$$

Because the unit vector $\hat{\boldsymbol{\phi}} = -\hat{\mathbf{x}}\sin\phi + \hat{\mathbf{y}}\cos\phi$ has no z component, only the θ component of $\mathbf{L}$ will give a contribution to L_z. Substituting the unit vector decomposition $\hat{\boldsymbol{\theta}} = \hat{\mathbf{x}}\cos\theta\cos\phi + \hat{\mathbf{y}}\cos\theta\sin\phi - \hat{\mathbf{z}}\sin\theta$ we obtain

$$L_z = -i\hbar\frac{\partial}{\partial\phi}. \tag{1.25}$$

The eigenvalues and eigenfunctions of L_z are obtained by solving the equation

$$-i\hbar\frac{\partial}{\partial\phi}\Phi_m(\phi) = m\hbar\;\Phi_m(\phi). \tag{1.26}$$

Here, the eigenvalue m is called the *magnetic quantum number* for the projection of the angular momentum $\mathbf{L}$ on the z axis.[1] The eigenfunctions are

$$\Phi_m(\phi) = a_m e^{im\phi}. \tag{1.27}$$

Assuming the solutions of the Schrödinger equation to be single-valued functions of position,[2] the wavefunction has to be invariant under rotation over 2π about the z axis; i.e., we have to

---

[1] In this chapter we use the shorthand notation m for the magnetic quantum numbers m_l corresponding to states with orbital quantum number l. When other forms of angular momentum appear we will use the subscript notation to discriminate between the different magnetic quantum numbers; e.g., lm_l, sm_s, jm_j, etc..

[2] For a discussion of the single-valuedness see [73, 48]

impose the boundary condition $e^{im\phi} = e^{im(\phi+2\pi)}$. Thus we require $e^{im2\pi} = 1$, which implies $m \in \{0, \pm1, \pm2, \ldots\}$. In other words we quantized the rotation about the z axis. As the orientation of the coordinate system was not defined up to this point, the chosen z direction is called the *quantization axis*. With the normalization

$$\int_0^{2\pi} |\Phi_m(\phi)|^2 \, d\phi = 1 \tag{1.28}$$

we find the same normalization coefficient for all values of the m quantum number, $a_m = (2\pi)^{-1/2}$.

1.1.4 Commutation relations for L_x, L_y, L_z and $\mathbf{L}^2$

The three cartesian components of the angular momentum operator are differential operators satisfying the following commutation relations

$$[L_i, L_j] = i\hbar\varepsilon_{ijk}L_k \quad \Leftrightarrow \quad [L_x, L_y] = i\hbar L_z,\ [L_y, L_z] = i\hbar L_x \text{ and } [L_z, L_x] = i\hbar L_y. \tag{1.29}$$

These expressions are readily derived with the help of some elementary commutator algebra (see Appendix L.8). We derive the relation $[L_x, L_y] = i\hbar L_z$ explicitly; the other commutators are obtained by cyclic permutation of x, y and z. Starting from the definition $L_i = \varepsilon_{ijk}r_j p_k$ we use subsequently the distributive rule (L.29b), the multiplicative rule (L.29d) and the commutation relation (1.9),

$$\begin{aligned}[L_x, L_y] &= [yp_z - zp_y, zp_x - xp_z] = [yp_z, zp_x] + [zp_y, xp_z] \\ &= y\,[p_z, z]\,p_x - x\,[p_z, z]\,p_y = i\hbar(xp_y - yp_x) = i\hbar L_z.\end{aligned} \tag{1.30}$$

A scalar operator always commutes with itself. This well-known commutation rule does *not* hold for vector operators. Two vector operators $\mathbf{A}$ and $\mathbf{B}$ only commute if *all components* of $\mathbf{A}$ commute with all components of $\mathbf{B}$.[1] Hence, $\mathbf{L}$ does *not* commute with itself (see Problem 1.4).

The components of $\mathbf{L}$ commute with $\mathbf{L}^2$,

$$[L_x, \mathbf{L}^2] = 0,\ [L_y, \mathbf{L}^2] = 0,\ [L_z, \mathbf{L}^2] = 0. \tag{1.31}$$

We verify this explicitly for L_z. Using the relation

$$\mathbf{L}^2 = \mathbf{L} \cdot \mathbf{L} = L_x^2 + L_y^2 + L_z^2 \tag{1.32}$$

we obtain with the aid of the multiplicative rule (L.29c)

$$\begin{aligned}&[L_z, L_z^2] = 0 \\ &[L_z, L_y^2] = [L_z, L_y]L_y + L_y[L_z, L_y] = -i\hbar(L_xL_y + L_yL_x) \\ &[L_z, L_x^2] = [L_z, L_x]L_x + L_x[L_z, L_x] = +i\hbar(L_yL_x + L_xL_y).\end{aligned}$$

By adding these terms we find $[L_z, \mathbf{L}^2] = 0$ as well as $[L_z, L_x^2 + L_y^2] = 0$.

Problem 1.4. Vector *operators* differ from classical vectors. Show that

$$\mathbf{L} \times \mathbf{L} = i\hbar\mathbf{L}, \quad [\mathbf{L}, \mathbf{L}] \neq 0, \quad [\mathbf{L}^2, \mathbf{L}^2] = 0.$$

---

[1] The commutator of two *vectors* is a second order tensor. This becomes evident in the Einstein notation $[\mathbf{A}, \mathbf{B}] \Leftrightarrow [A_i, B_j] = A_iB_j - B_jA_i \Leftrightarrow \mathbf{AB}^T - (\mathbf{BA}^T)^T$, where $()^T$ represents matrix transposition.

1.1.5 The operators $L_\pm$

The operators

$$L_\pm = L_x \pm iL_y \tag{1.33}$$

are obtained by taking the inner products of $\mathbf{L}$ with the unit vectors along the x and y direction, $L_\pm = (\hat{\mathbf{x}}\cdot\mathbf{L}) \pm i\,(\hat{\mathbf{y}}\cdot\mathbf{L})$. In spherical coordinates this results in

$$L_\pm = i\hbar\left(\left[(\hat{\mathbf{x}}\cdot\hat{\boldsymbol{\theta}}) \pm i(\hat{\mathbf{y}}\cdot\hat{\boldsymbol{\theta}})\right]\frac{1}{\sin\theta}\frac{\partial}{\partial\phi} - \left[(\hat{\mathbf{x}}\cdot\hat{\boldsymbol{\phi}}) \pm i(\hat{\mathbf{y}}\cdot\hat{\boldsymbol{\phi}})\right]\frac{\partial}{\partial\theta}\right), \tag{1.34}$$

as follows directly with Eq. (1.23). Substituting the unit vector decompositions $\hat{\boldsymbol{\phi}} = -\hat{\mathbf{x}}\sin\phi + \hat{\mathbf{y}}\cos\phi$ and $\hat{\boldsymbol{\theta}} = \hat{\mathbf{x}}\cos\theta\cos\phi + \hat{\mathbf{y}}\cos\theta\sin\phi - \hat{\mathbf{z}}\sin\theta$ we obtain

$$L_\pm = \hbar e^{\pm i\phi}\left(i\cot\theta\frac{\partial}{\partial\phi} \pm \frac{\partial}{\partial\theta}\right). \tag{1.35}$$

These operators are known as *shift operators* and more specifically as *raising* (L_+) and *lowering* (L_-) operators because their action is to raise or to lower the angular momentum along the quantization axis by one quantum of angular momentum (see Section 1.1.6).

Several useful relations for $L_\pm$ follow straightforwardly. Using the commutation relations (1.29) we obtain

$$[L_z, L_\pm] = [L_z, L_x] \pm i\,[L_z, L_y] = i\hbar L_y \pm \hbar L_x = \pm\hbar L_\pm. \tag{1.36}$$

Further we have

$$L_+L_- = L_x^2 + L_y^2 - i\,[L_x, L_y] = L_x^2 + L_y^2 + \hbar L_z = \mathbf{L}^2 - L_z^2 + \hbar L_z \tag{1.37a}$$

$$L_-L_+ = L_x^2 + L_y^2 + i\,[L_x, L_y] = L_x^2 + L_y^2 - \hbar L_z = \mathbf{L}^2 - L_z^2 - \hbar L_z, \tag{1.37b}$$

where we used again one of the commutation relations (1.29). Subtracting these equations we obtain

$$[L_+, L_-] = 2\hbar L_z \tag{1.38}$$

and by adding Eqs. (1.37) we find

$$\mathbf{L}^2 = L_z^2 + \tfrac{1}{2}\left(L_+L_- + L_-L_+\right). \tag{1.39}$$

1.1.6 The operator $\mathbf{L}^2$

To derive an expression for the operator $\mathbf{L}^2$ we use the operator relation (1.39). Substituting Eqs. (1.25) and (1.35) we obtain after some straightforward manipulation

$$\mathbf{L}^2 = -\hbar^2\left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\right]. \tag{1.40}$$

The eigenfunctions and eigenvalues of $\mathbf{L}^2$ are obtained by solving the equation

$$-\hbar^2\left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\right]Y(\theta,\phi) = \lambda\hbar^2 Y(\theta,\phi). \tag{1.41}$$

Because the operators $\mathbf{L}^2$ and L_z commute they share a complete set of eigenstates (see Problem F.1); i.e., the shared eigenfunctions $Y(\theta,\phi)$ must be of the form $Y(\theta,\phi) = P(\theta,\phi)\Phi_m(\phi)$, where the function $\Phi_m(\phi)$ is an eigenfunction of the L_z operator. Because of Eq. (1.26) this implies that $L_zP(\theta,\phi) \equiv 0$, which can only be satisfied for arbitrary value of θ if the variables θ and ϕ separate:

$P(\theta,\phi) = P(\theta)Q(\phi)$. In turn this requires $L_z Q(\phi) \equiv 0$, which can only be satisfied if $Q(\phi)$ is a constant. Thus, we conclude that the shared eigenfunctions $Y(\theta,\phi)$ must be of the form

$$Y(\theta,\phi) = P(\theta)\Phi_m(\phi). \tag{1.42}$$

Evaluating the second order derivative $\partial^2/\partial\phi^2$ in Eq. (1.41) we obtain

$$\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} - \frac{m^2}{\sin^2\theta} + \lambda\right] P(\theta) = 0. \tag{1.43}$$

As this equation is real its solutions are real functions of the variable θ. Introducing the notation $\lambda = l(l+1)$ and $u \equiv \cos\theta$ (with θ restricted to the interval $0 \le \theta \le \pi$) this equation takes the form of the associated Legendre differential equation (L.36),

$$\left[\left(1-u^2\right)\frac{d^2}{du^2} - 2u\frac{d}{du} + l(l+1) - \frac{m^2}{1-u^2}\right] P_l^m(u) = 0. \tag{1.44}$$

Since $0 \le \theta \le \pi$ we have $\sin\theta = \sqrt{1-u^2} \ge 0$. The solutions are determined up to a constant factor, which has to be fixed by convention. For $m = 0$ the Eq. (1.44) reduces to the Legendre differential equation and its normalized solutions are the *Legendre polynomials*, defined by

$$P_l(u) = \frac{1}{2^l l!}\frac{d^l}{du^l}(u^2-1)^l. \tag{1.45}$$

The solution $P_l(u)$ is a real polynomial of *degree* $l \in \{0,1,2\cdots\}$ with l zeros in the interval $-1 \le u \le 1$. The Legendre polynomials of lowest degree are

$$P_0(u) = 1, \quad P_1(u) = u, \quad P_2(u) = \tfrac{1}{2}(3u^2-1). \tag{1.46}$$

Note that the sign of the highest power term is chosen to be positive. For $m \ne 0$ the solutions are the *associated Legendre functions* $P_l^m(u)$, real functions which can be written as the product of a positive function $(1-u^2)^{m/2}$ and a polynomial of degree $(l-m)$, parity $(-1)^{l-m}$ with $(l-m)$ zeros in the interval $-1 \le u \le 1$. For $m = 0$ we define $P_l^0(u) \equiv P_l(u)$. For $m > 0$ the $P_l^m(u)$ are obtained by differentiation of the Legendre polynomials,[1]

$$P_l^m(u) = (-1)^m(1-u^2)^{m/2}\frac{d^m}{du^m}P_l(u). \tag{1.47}$$

The *parity* of the $P_l^m(u)$ is given by

$$P_l^m(-u) = (-1)^{l-m}P_l^m(u) \tag{1.48}$$

and the above definitions fix the normalization,

$$\int_{-1}^{1}\left[P_l^m(u)\right]^2 du = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}. \tag{1.49}$$

As Eq. (1.44) depends on m^2 we also can define solutions for $m < 0$. Obviously, $P_l^{-m}(u)$ and $P_l^m(u)$ are in essence the same but for the $P_l^{-m}(u)$ we still have to define sign and normalization. Unfortunately, several competing conventions can be found in the literature. In this course we use a positive sign and adhere to the convention in which the normalization (1.49) is valid for both positive and negative m. This is realized by extending the $P_l^m(u)$ to negative m by the relation

$$P_l^{-m}(u) \equiv (-1)^m\frac{(l-m)!}{(l+m)!}P_l^m(u), \tag{1.50}$$

where $0 \le m \le l$. The inclusion of the phase factor $(-1)^m$ in *both* Eq. (1.47) and (1.50) is referred to as the *Condon and Shortley phase convention* [28]. It implies that the phase factor $(-1)^m$ is present for positive m but absent for negative m. Note that the $P_l^m(u)$ are nonzero only for $-l \le m \le l$; i.e., the index m can assume $2l+1$ possible values for a given value of l.

---

[1] Beware of other phase conventions for the $P_l^m(u)$, they affect the recursion relations.

1.1.6.1 Spherical harmonics - Condon and Shortley phase convention

At this point we define the *spherical harmonics* of *degree* l and *order* m (cf. Section L.9.1)

$$Y_l^m(\theta,\phi) = A_m P_l^m(\cos\theta)e^{im\phi} \tag{1.51}$$

as the joint eigenfunctions of $\mathbf{L}^2$ and L_z in the position representation,

$$\mathbf{L}^2\, Y_l^m(\theta,\phi) = l(l+1)\hbar^2\, Y_l^m(\theta,\phi) \tag{1.52}$$

$$L_z\, Y_l^m(\theta,\phi) = m\hbar\, Y_l^m(\theta,\phi). \tag{1.53}$$

The prefactor A_m is a normalization constant. Combining the normalization integrals (1.28) and (L.46) we obtain

$$Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}\,P_l^m(\cos\theta)e^{im\phi}, \tag{1.54}$$

with $-l \leq m \leq l$. The Condon and Shortley phase convention assures that the shift operators satisfy the following relations

$$\begin{aligned} L_\pm Y_l^m(\theta,\phi) &= \sqrt{(l\mp m)\,(l\pm m+1)}\,\hbar\, Y_l^{m\pm1}(\theta,\phi) \\ &= \sqrt{l\,(l+1)-m(m\pm1)}\,\hbar\, Y_l^{m\pm1}(\theta,\phi), \end{aligned} \tag{1.55}$$

with a *positive sign* in front of the square root for *all* values of l and m. Eqs. (1.55) are readily obtained with the aid of Eqs. (1.35) and the recursion relations (L.49). The *parity* of the $Y_l^m(\theta,\phi)$ under inversion in 3D, $\hat{\mathbf{r}} = (\theta,\phi) \to -\hat{\mathbf{r}} = (\pi-\theta,\phi+\pi)$, is independent of m and given by

$$Y_l^m(-\hat{\mathbf{r}}) = (-1)^l Y_l^m(\hat{\mathbf{r}}) \tag{1.56}$$

as follows with Eqs. (1.54) and (1.48); i.e., the parity is even for l even and odd for l odd. This makes *parity* into a property defining *selection rules* in atomic transitions.

1.1.7 Orbital angular momentum in Dirac notation

The observables of the orbital angular momentum are represented by the operators $\mathbf{L}^2$ and L_z. In Dirac notation (cf. Appendix F.1.1) their shared basis is defined by

$$\mathbf{L}^2\,|l,m\rangle = l(l+1)\hbar^2\,|l,m\rangle \tag{1.57a}$$

$$L_z\,|l,m\rangle = m\hbar\,|l,m\rangle\,, \tag{1.57b}$$

where the $|l,m\rangle$ are abstract state vectors in Hilbert space, with l and m the *rotational quantum numbers.* The spherical harmonics

$$Y_l^m(\theta,\phi) \equiv Y_l^m(\hat{\mathbf{r}}) = \langle\hat{\mathbf{r}}|l,m\rangle \tag{1.58}$$

are the corresponding wavefunctions in the position representation using spherical coordinates. The action of the shift operators $L_\pm$ is given by

$$L_\pm\,|l,m\rangle = \sqrt{l\,(l+1)-m(m\pm1)}\,\hbar\,|l,m\pm1\rangle\,, \tag{1.59}$$

with a *positive sign* in front of the square root for *all* values of l and m. The latter requirement constitutes the Condon and Shortley phase convention for the eigenstates $|l,m\rangle$ (cf. Section 1.1.6.1).

1.1.8 Radial momentum operator p_r

Thus far we succeeded in quantizing the Schrödinger Hamiltonian $\mathcal{H}_0$ and the angular momentum $\mathbf{L}$. Let us now turn to the radial momentum $p_r = \hat{\mathbf{r}} \cdot \mathbf{p}$. Here we have a difficulty because the correspondence rules are not unambiguous in this case. In classical mechanics the expressions $p_r = \hat{\mathbf{r}} \cdot \mathbf{p}$ and $p_r = \mathbf{p} \cdot \hat{\mathbf{r}}$ are identities but since $\hat{\mathbf{r}} = \mathbf{r}/r$ and $-i\hbar\boldsymbol{\nabla}$ do not commute this is not the case for $p_r = -i\hbar\,(\boldsymbol{\nabla} \cdot \hat{\mathbf{r}})$ and $p_r = -i\hbar\,(\hat{\mathbf{r}} \cdot \boldsymbol{\nabla})$. There is a way around this problem. Since we know how to quantize $\mathbf{p}^2$ and $(\hat{\mathbf{r}} \times \mathbf{p})^2$, we infer with the aid of Eq. (1.4) that the radial momentum must be given by

$$p_r^2 = (\hat{\mathbf{r}} \cdot \mathbf{p})^2 = \mathbf{p}^2 - (\hat{\mathbf{r}} \times \mathbf{p})^2 = -\hbar^2\Delta - \mathbf{L}^2/r^2. \tag{1.60}$$

Substituting Eqs. (1.15) and (1.40) this yields

$$p_r^2\psi = -\hbar^2 \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} \right) \psi = -\hbar^2 \left(\frac{\partial}{\partial r} + \frac{1}{r} \right)^2 \psi = -\hbar^2 \frac{1}{r} \frac{\partial^2}{\partial r^2} (r\psi) . \tag{1.61}$$

Hence, up to a sign, the radial momentum in spherical coordinates is given by

$$p_r\psi = -i\hbar \left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \psi = -i\hbar \frac{1}{r} \frac{\partial}{\partial r} (r\psi). \tag{1.62}$$

We shall use the sign convention of Eq. (1.62). This implies the commutation relation

$$[r, p_r] = i\hbar. \tag{1.63}$$

Importantly, since L is independent of r and p_r is independent of θ and ϕ, we find that p_r^2 commutes with both L_z and $\mathbf{L}^2$,

$$\left[p_r^2, L_z\right] = 0 \quad \text{and} \quad \left[p_r^2, \mathbf{L}^2\right] = 0. \tag{1.64}$$

In Problem 1.5 it is shown that p_r is only *hermitian* if one restricts oneself to the sub-class of normalizable wavefunctions which are *regular* in the origin; i.e.,

$$\lim_{r \to 0} r\psi(\mathbf{r}) = 0.$$

This additional condition is essential to select *physically relevant solutions* for the (radial) wavefunction. Here we mean by physically relevant that the wavefunction satisfies not only the laplacian in spherical coordinates (which is not defined in the origin) but also the laplacian in cartesian coordinates (which is defined throughout space).

Problem 1.5. Show that p_r is hermitian for square-integrable functions $\psi(r)$ only if they are regular at the origin, i.e. $\lim_{r\to 0} r\psi(r) = 0$.

Solution. For p_r to be hermitian we require the following expression to be zero for any wavefunction ψ within its Hilbert space:

$$\begin{aligned} \langle \psi, p_r\psi \rangle - \langle \psi, p_r\psi \rangle^* &= -i\hbar \int \left[\psi^* \frac{1}{r} \frac{\partial}{\partial r} (r\psi) + \frac{1}{r} \psi \frac{\partial}{\partial r} (r\psi^*) \right] r^2 dr d\Omega \\ &= -i\hbar \int \left[r\psi^* \frac{\partial}{\partial r} (r\psi) + r\psi \frac{\partial}{\partial r} (r\psi^*) \right] dr d\Omega \\ &= -i\hbar \int \frac{\partial}{\partial r} \left| r\psi \right|^2 dr d\Omega. \end{aligned}$$

For this to be zero we require

$$\int \frac{\partial}{\partial r} \left| r\psi \right|^2 dr = \left[\left| r\psi \right|^2 \right]_0^\infty = 0.$$

Because $\psi(r)$ is taken to be a square-integrable function; i.e., $\int \left| r\psi \right|^2 dr = \mathcal{N}$ with $\mathcal{N}$ finite, we have $\lim_{r\to\infty} r\psi(r) = 0$ and $\lim_{r\to 0} r\psi(r) = \chi_0$, where χ_0 is (in general) finite. Thus, for p_r to be hermitian we require $\psi(r)$ to be regular in the origin ($\chi_0 = 0$) on top of being square-integrable. However, square-integrable eigenfunctions of p_r can also be irregular at the origin and have complex eigenvalues, e.g.

$$p_r \frac{\exp[-\alpha r]}{r} = -\frac{i\hbar}{r}\frac{\partial}{\partial r} r \frac{\exp[-\alpha r]}{r} = i\hbar\alpha \frac{\exp[-\alpha r]}{r}. \qquad \square$$

Problem 1.6. Show that the radial momentum operator can be written in the form

$$p_r = \tfrac{1}{2}\left(\hat{\mathbf{r}}\cdot\mathbf{p} + \mathbf{p}\cdot\hat{\mathbf{r}}\right) = -\frac{i\hbar}{2}\left[\frac{\mathbf{r}}{r}\cdot\boldsymbol{\nabla} + \boldsymbol{\nabla}\cdot\left(\frac{\mathbf{r}}{r}\right)\right]. \tag{1.65}$$

Verify that in two dimensions p_r cannot be written in the form given on the r.h.s..

1.2 Schrödinger equation

1.2.1 Schrödinger equation in spherical coordinates

We are now in a position to write down the Schrödinger equation of a (reduced) mass m_r moving at energy E in a *central* potential field $\mathcal{V}(r)$

$$\left[\frac{1}{2m_r}\left(p_r^2 + \frac{\mathbf{L}^2}{r^2}\right) + \mathcal{V}(r)\right]\psi(r,\theta,\phi) = E\psi(r,\theta,\phi). \tag{1.66}$$

Because the operators $\mathbf{L}^2$ and L_z commute with the Hamiltonian[1] they share a complete set of eigenstates with that Hamiltonian (See Problem F.1); i.e., the shared eigenfunctions $\psi(r,\theta,\phi)$ must be of the form $\psi = R(r,\theta,\phi)Y_l^m(\theta,\phi)$, which implies $\mathbf{L}^2 R(r,\theta,\phi) \equiv 0$ in view of Eq. (1.52). This can only be satisfied for arbitrary values of r if the radial variable can be separated from the angular variables, $R(r,\theta,\phi) = R(r)X(\theta,\phi)$. In turn this requires $\mathbf{L}^2 X(\theta,\phi) \equiv 0$, which implies that $X(\theta,\phi)$ must be a constant. Thus, we conclude that the shared eigenfunctions $\psi(r,\theta,\phi)$ must be of the form

$$\psi(r,\theta,\phi) = R_{nl}(r)Y_l^m(\theta,\phi), \tag{1.67}$$

where the quantum number n appears to allow for more than one solution of given l. Hence, using Eq. (1.52) and substituting Eqs. (1.61) and (1.67) into Eq. (1.66) we obtain

$$\left[\frac{\hbar^2}{2m_r}\left(-\frac{\partial^2}{\partial r^2} - \frac{2}{r}\frac{\partial}{\partial r} + \frac{l(l+1)}{r^2}\right) + \mathcal{V}(r)\right] R_{nl}(r)Y_l^m(\theta,\phi) = ER_{nl}(r)Y_l^m(\theta,\phi). \tag{1.68}$$

Here the term

$$\mathcal{V}_{\mathrm{rot}}(r) \equiv \frac{l(l+1)\hbar^2}{2m_r r^2} \tag{1.69}$$

is called the *rotational energy barrier* and represents the *centrifugal energy* at a given distance from the origin and for a given value of the angular momentum. Because the operator on the left of Eq. (1.68) is independent of θ and ϕ we can eliminate the functions $Y_l^m(\theta,\phi)$ from this equation. The remaining equation takes the form of the *radial wave equation.*

$$\left[\frac{\hbar^2}{2m_r}\left(-\frac{d^2}{dr^2} - \frac{2}{r}\frac{d}{dr} + \frac{l(l+1)}{r^2}\right) + \mathcal{V}(r)\right] R_{nl}(r) = ER_{nl}(r), \tag{1.70}$$

where the solutions $R_{nl}(r)$ must depend on r but be independent of θ and ϕ. Note that the solutions do not depend on m because the Hamiltonian does not depend on L_z. This is a property of central potentials.

---

[1] Note that L_z commutes with $\mathbf{L}^2$ (see Section 1.1.6); L_z and $\mathbf{L}^2$ commute with r and p_r (see Section 1.1.8).

Reduction to one-dimensional Schrödinger equation

Eq. (1.70) is the starting point for the description of the relative radial motion of any particle in a central potential field. Introducing the quantities

$$\varepsilon = 2m_r E/\hbar^2 \quad \text{and} \quad U(r) = 2m_r \mathcal{V}(r)/\hbar^2, \tag{1.71}$$

Suppressing the quantum number n, Eq. (1.70) can be written in the compact form

$$R_l'' + \frac{2}{r}R_l' + \left[\varepsilon - U(r) - \frac{l(l+1)}{r^2}\right] R_l = 0, \tag{1.72}$$

where the prime refers to a derivative with respect to r. Eq. (1.61) suggests to introduce so-called *reduced* radial wavefunctions

$$\chi_l(r) = rR_l(r), \tag{1.73}$$

which allows us to reduce the radial wave equation (1.70) to the form of a one-dimensional Schrödinger equation

$$\chi_l'' + \left[\frac{2m_r}{\hbar^2}(E - \mathcal{V}) - \frac{l(l+1)}{r^2}\right] \chi_l = 0. \tag{1.74}$$

The 1D-Schrödinger equation is a second-order differential equation of the following general form

$$\chi'' + F(r)\chi = 0. \tag{1.75}$$

Equations of this type satisfy some general properties. These are related to the Wronskian theorem, which is derived and discussed in appendix L.13.

Not all solutions of the 1D Schrödinger equation are physically acceptable. The physical solutions must be normalizable; i.e., for bound states

$$\int r^2 \left|R(r)\right|^2 dr = \int \left|\chi(r)\right|^2 dr = \mathcal{N}, \tag{1.76}$$

where $\mathcal{N}$ is a finite number. However, there is an additional requirement. Because the Hamiltonian (1.3) is only valid outside the origin ($r \neq 0$) the solutions of the radial wave equation are not necessarily valid *at* the origin. To be valid for all values of r the solutions must, in addition to be being normalizable, also be *regular* in the origin; i.e., $\lim_{r\to 0} rR(r) = \lim_{r\to 0} \chi(r) = 0$. Although this is stated without proof we demonstrate in Problem 1.7 that normalizable wavefunctions $\psi(\mathbf{r})$ scaling like $R(r) \sim 1/r$ near the origin do not satisfy the Schrödinger equation in the origin. All this being said, only wavefunctions based on the *regular* solutions of Eqs. (1.70) and (1.74) can be valid solutions for all values of r, *including the origin.*

Problem 1.7. Show that a normalizable radial wavefunction scaling like $R(r) \sim 1/r$ for $r \to 0$ does not satisfy the Schrödinger equation in the origin.

Solution. Next we turn to solutions $\psi(\mathbf{r}) = R_l(r)Y_l^m(\theta, \phi)$ of the Schrödinger equation for the motion of a particle in a central field. We presume that the wavefunction is well behaved everywhere but diverges like $R_l(r) \sim 1/r$ for $r \to 0$. We ask ourselves whether this is a problem because - after all - the wavefunction is normalizable. However, the divergent wavefunction $R_l(r)$ is defined everywhere except in the origin. This is more than a technicality because it implies that the Schrödinger equation is *not* satisfied in the origin. Using Problem 1.9 we find

$$\left(-\frac{\hbar^2}{2m_r}\Delta + \mathcal{V}(r) - E\right)\psi(\mathbf{r}) = -\frac{4\pi\hbar^2}{2m_r}\delta(\mathbf{r}),$$

which is zero everywhere *except in the origin.* Apparently, by solving the Schrödinger equation after separation in radial and angular variables we have generated a solution that does not satisfy the original equation (which is valid everywhere in space - including the origin). □

Problem 1.8. Show that for a normalizable radial wavefunction scaling like $R(r) \sim 1/r$ for $r \to 0$ the kinetic energy diverges in the origin.

Solution. We first write the radial wavefunction in the form $R_0(r) = (\chi_0(r)/r)$, where $\chi_0(r)$ is nonzero in the origin, $\lim_{r\to 0} \chi_0(r) = \chi_0(0) \neq 0$. Calculating the kinetic energy we find

$$\begin{aligned} -\int R_0(r)Y_0^0(\theta,\phi)\,\frac{\hbar^2}{2m_r}\Delta R_0(r)Y_0^0(\theta,\phi)\,d\mathbf{r} &> -\frac{\hbar^2}{2m_r}\chi_0^2(0)\lim_{\epsilon\to 0}\int_V \frac{1}{4\pi r}\Delta\frac{1}{r}d\mathbf{r} \\ &= -\frac{\hbar^2}{2m_r}\chi_0^2(0)\lim_{\epsilon\to 0}\int_V \frac{1}{r}\delta(\mathbf{r})\,d\mathbf{r} \to \infty. \end{aligned}$$ □

Problem 1.9. Use the Gauss theorem to demonstrate the relation $\Delta(1/r) = -4\pi\delta(\mathbf{r})$.

Solution. We first integrate this expression on both sides over a small sphere V of radius ϵ centered at the origin,

$$\int_V \Delta\frac{1}{r}d\mathbf{r} = -4\pi.$$

Here we used $\int_V \delta(\mathbf{r})\,d\mathbf{r} = 1$ for an arbitrarily small sphere at the origin. The l.h.s. also yields -4π as follows with the divergence theorem (Gauss theorem)

$$\lim_{\epsilon\to 0}\int_V \Delta\frac{1}{r}d\mathbf{r} = \lim_{\epsilon\to 0}\oint_S d\mathbf{S}\cdot\boldsymbol{\nabla}\frac{1}{r} = \lim_{\epsilon\to 0}\oint_S d\mathbf{S}\cdot\hat{\mathbf{r}}\left(-\frac{1}{r^2}\right) = \lim_{\epsilon\to 0}4\pi\epsilon^2\left(-\frac{1}{\epsilon^2}\right) = -4\pi.$$ □

1.2.2 Schrödinger equation in cylindrical coordinates

In systems with cylindrical symmetry about the z axis (see Fig. 1.2) the motion separates into motion parallel to the z axis and motion in planes orthogonal to the z axis,

$$\mathbf{p}^2 = (\hat{\mathbf{z}}\cdot\mathbf{p})^2 + (\hat{\mathbf{r}}_\perp\cdot\mathbf{p}_\perp)^2 + (\hat{\mathbf{r}}_\perp\times\mathbf{p}_\perp)^2, \tag{1.77}$$

where $\hat{\mathbf{r}}_\perp$ is defined in Eq. (1.17a) and $\hat{\mathbf{z}}$ is independent of position; $\mathbf{p}_\perp$ is the momentum in the xy plane. The quantization rule for the linear momentum in the z direction, $p_z\phi = \hat{\mathbf{z}}\cdot\mathbf{p}\phi \to -i\hbar\hat{\mathbf{z}}\cdot\boldsymbol{\nabla}\phi = -i\hbar\partial_z\phi$, can be applied without ambiguity and we have for the p_z^2 operator

$$p_z^2 = (\hat{\mathbf{z}}\cdot\mathbf{p})^2 = -\hbar^2\partial^2/\partial z^2. \tag{1.78}$$

With regard to the motion in planes orthogonal to the z axis we run into the same dilemma as we encountered with spherical coordinates. The quantization rules for $\hat{\mathbf{r}}_\perp\cdot\mathbf{p}_\perp = \mathbf{p}_\perp\cdot\hat{\mathbf{r}}_\perp$ are not unambiguous since $\hat{\mathbf{r}}_\perp\cdot\boldsymbol{\nabla}_\perp\phi \neq \boldsymbol{\nabla}_\perp\cdot\hat{\mathbf{r}}_\perp\phi$ (the direction of the unit vector $\hat{\mathbf{r}}_\perp$ depends on position). This dilemma can be circumvented in the same way as we did for p_r in Section 1.1.8. By rewriting Eq. (1.77) in the form

$$p_{r_\perp}^2 = (\hat{\mathbf{r}}_\perp\cdot\mathbf{p}_\perp)^2 = \mathbf{p}^2 - (\hat{\mathbf{r}}_\perp\times\mathbf{p}_\perp)^2 - (\hat{\mathbf{z}}\cdot\mathbf{p})^2 \tag{1.79}$$

we obtain an expression where $p_{r_\perp}$ is expressed in quantities for which the correspondence rules can be applied without ambiguity: $\mathbf{p}^2 \to -\hbar^2\Delta$ and $\hat{\mathbf{r}}_\perp\times\mathbf{p}_\perp \to -i\hbar\hat{\mathbf{r}}_\perp\times\boldsymbol{\nabla}_\perp$. Note that for cylindrical coordinates the angular momentum along the symmetry axis, $\mathbf{L}_\perp = \mathbf{r}_\perp\times\mathbf{p}_\perp$, satisfies the property

$$\mathbf{L}_\perp = -i\hbar\frac{\partial}{\partial\phi} = L_z. \tag{1.80}$$

Replacing the dynamical variables by their operators Eq. (1.79) takes the form

$$p_{r_\perp}^2 = -\hbar^2\left(\Delta - \frac{1}{r_\perp^2}\frac{\partial^2}{\partial\phi^2} - \frac{\partial^2}{\partial z^2}\right). \tag{1.81}$$

Substituting Eq. (1.19) for the laplacian we obtain

$$p_{r_\perp}^2 = -\hbar^2\left(\frac{\partial^2}{\partial r_\perp^2} + \frac{1}{r_\perp}\frac{\partial}{\partial r_\perp}\right). \tag{1.82}$$

At this point we can quantize the Hamiltonian and starting from Eq. (1.77) the Schrödinger equation takes the form

$$\left[-\frac{\hbar^2}{2m_r}\left(\frac{\partial^2}{\partial r_\perp^2}+\frac{1}{r_\perp}\frac{\partial}{\partial r_\perp}+\frac{\partial^2}{\partial z^2}\right)+\frac{L_z^2}{2m_r r_\perp^2}+\mathcal{V}(r_\perp,\phi,z)\right]\psi(r_\perp,\phi,z)=E\psi(r_\perp,\phi,z). \tag{1.83}$$

If the potential only depends on the $r_\perp$ variable, $\mathcal{V}(r_\perp,\phi,z)=\mathcal{V}(r_\perp)$, we have separation of variables for the axial, radial and angular motion and we find for the radial Schrödinger equation in this case

$$\frac{\hbar^2}{2m_r}\left(-\frac{d^2}{dr_\perp^2}-\frac{1}{r_\perp}\frac{d}{dr_\perp}+\frac{l^2}{r_\perp^2}+\mathcal{V}(r_\perp)\right)R_{nl}(r_\perp)=ER_{nl}(r_\perp). \tag{1.84}$$

Reduction to one-dimensional Schrödinger equation

Suppressing the quantum number n and using the definitions (1.71) the radial wave equation (1.84) takes the compact form

$$R_l''+\frac{1}{r_\perp}R_l'+\left[\varepsilon-U(r_\perp)-\frac{m^2}{r_\perp^2}\right]R_l=0, \tag{1.85}$$

where the prime refers to a derivative with respect to $r_\perp$. Introducing the functions

$$y_l(r_\perp)=r_\perp^{1/2}R_l(r_\perp), \tag{1.86}$$

also in this case the radial wave equation is reduced to the form of a one-dimensional Schrödinger equation,

$$y_l''+\left[\frac{2m_r}{\hbar^2}(E-\mathcal{V})+\frac{1/4-m^2}{r_\perp^2}\right]y_l=0. \tag{1.87}$$

1.3 Symmetry properties, conserved quantities and good quantum numbers

To conclude this chapter we return to the Schrödinger Hamiltonian,

$$\mathcal{H}_0=-\frac{\hbar^2}{2m_r}\Delta+\mathcal{V}(r), \tag{1.88}$$

and discuss the main findings of the chapter against the background of the symmetry properties of this Hamiltonian. In general, the symmetries of a physical quantity are defined by the set of coordinate transformations that leave this quantity invariant. With each symmetry we associate an operator that generates the corresponding coordinate transformation. The transformations may be continuous (e.g., translation and rotation) or discrete (e.g., reflection and inversion). For instance, an equilateral triangle is invariant under rotation over 120° about an axis through the center of mass and perpendicular to the plane of the triangle; this reveals a three-fold symmetry under rotation (a finite rotation symmetry in this case). The mathematical discipline for the investigation of symmetry properties is called *Group theory* [111].

What are the symmetries of the Schrödinger Hamiltonian? For the potential energy term this is self evident from the notation because the central potential $\mathcal{V}(r)$ depends only on the radial distance $r=|\mathbf{r}|$ to the atomic center of mass. This manifestly being the case, $\mathcal{V}(r)$ is said to be *manifestly invariant* under any transformation that does not affect $|\mathbf{r}|$. This is the case for all *rotations* about the origin, all *reflections* about a plane through the origin and for *inversion* about the origin. For the kinetic energy term the symmetry under such operations is certainly not manifest and we need to have a closer look.

As an example we consider the *space (or parity) inversion operator* $\mathcal{P}$. This operator transforms the position $\mathbf{r}$ into position $-\mathbf{r}$, which is equivalent to sign reversal of the cartesian coordinates ($x\to -x$, $y\to -y$, $z\to -z$). As this operation conserves $|\mathbf{r}|$, the potential energy is invariant under space inversion. To determine the inversion symmetry of the kinetic energy operator we write the laplacian in cartesian coordinates,

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{1.89}$$

As this expression only contains second derivatives with regard to position it also is invariant under sign reversal. In other words the *Schrödinger Hamiltonian conserves parity for any interaction that conserves parity.* Note that by inverting twice we obtain the unit operator, $\mathcal{P}^2 = \mathbb{1}$. In group theory $\mathcal{P}$ and $\mathcal{P}^2$ are said to be the elements of a group of order 2: the inversion group ($\mathcal{P}^2$ is the identity element). In the language of quantum mechanics we say that $\mathcal{P}$ commutes with the Hamiltonian. This implies that $\mathcal{P}$ is a hermitian operator; hence, its eigenvalues are real. As $\mathcal{P}^2 = \mathbb{1}$ these eigenvalues are 1 (*even parity*) or -1 (*odd parity*). The radius vector is odd under parity (such vectors are called *polar vectors*). Angular momentum is even under parity (such vectors are called *axial vectors* or *pseudo vectors*). As $\mathcal{P}$ commutes with $\mathcal{H}_0$, also the energy eigenstates must be parity eigenstates. This property was already noticed in Section 1.1.6.

What about rotational symmetry? From the commutation of L_z with $\mathcal{H}_0$ we find by using the product rule for differentiation

$$\left(\frac{\partial}{\partial\phi}\mathcal{H}_0 - \mathcal{H}_0\frac{\partial}{\partial\phi}\right)\psi(r,\theta,\phi) = \left(\frac{\partial\mathcal{H}_0}{\partial\phi}\right)_{\psi}\psi(r,\theta,\phi) = 0. \tag{1.90}$$

Because this relation is valid for any function $\psi(r,\theta,\phi)$ it implies the *invariance* of $\mathcal{H}_0$ under an *infinitesimal* rotation about the z axis,

$$\frac{\partial\mathcal{H}_0}{\partial\phi} = 0. \tag{1.91}$$

As the quantization axis was chosen in an arbitrary direction, this invariance of $\mathcal{H}_0$ holds for *any* infinitesimal rotation about the origin. Inversely, it is this invariance that makes L_z commute with $\mathcal{H}_0$, which implies that L_z is a conserved quantity (see Appendix F.2.4). The operator L_z can be identified with the operator for an infinitesimal rotation about the z axis as introduced above. The Hamiltonian is also invariant under any *finite* rotation about the origin because any such rotation can be realized by an infinite sequence of infinitesimal rotations about the origin. The infinite set of all rotations about the origin constitutes the elements of a continuous group: the *full rotation group.*

In this course we raise awareness for the symmetry properties but do not enter into the systematics of group theory. With regard to rotation we emphasize that for the Hamiltonian $\mathcal{H}_0$ the expectation values of $\mathbf{L}^2$ and L_z are *conserved* whatever the radial motion, showing that $\mathbf{L}^2$ and L_z are *observables* (observable constants of the motion). This means that the corresponding eigenvalues and eigenstates can be measured simultaneously to arbitrary precision and are uniquely determined by the quantum numbers l and m_l. Whenever a quantum number is conserved, it is called a *good quantum number* of the Hamiltonian under consideration. As a counter example we consider p_r^2. As p_r^2 does not commute with r (see Section 1.1.8), it does not commute with the Hamiltonian. This means that p_r^2 is not a conserved quantity and no good quantum number can be identified with the radial kinetic energy. This is no surprise from the physical point of view because the radial motion is oscillatory.

2

Hydrogenic atoms

The notion *hydrogenic* refers for atomic systems in which exactly one elementary charge is orbiting the nucleus [14]. Aside from the hydrogen isotopes *hydrogen* (H), *deuterium* (D) and *tritium* (T), ions like singly ionized *helium* (He^+), doubly ionized *lithium* (Li^{2+}) and exotic atoms like *positronium* and *muonic hydrogen* belong to this class. *Hydrogen-like* behavior is more common. It occurs in *one-electron atoms.* This is the class of atoms in which a single electron dominates the atomic properties. This class includes of course the *hydrogenic atoms* but more typically involves *many-electron* atoms with one distinct outer electron, the *valence electron.* Many-electron atoms behave at best *hydrogen-like* because the nuclear charge is partially *screened* by one or more shells of *core electrons.* For instance, hydrogen-like behavior is observed in one-electron *Rydberg atoms.* These are atoms in which the valence electron is excited to a weakly bound state. Other examples are the *alkali-like* atoms. Aside from the alkali (group I) atoms *lithium* (Li), *sodium* (Na), *potassium* (K), *rubidium* (Rb) and *cesium* (Cs) also the singly charged ions of "two-electron" atoms (group II) such as the ions of *magnesium* (Mg^+), *calcium* (Ca^+), *strontium* (Sr^+) and *barium* (Ba^+) are alkali-like.

In the present chapter we focus on the *non-relativistic* theory of hydrogenic atoms. We analyze the electronic motion under influence of the Coulomb interaction with a point-like nucleus; i.e., for the so-called *Bohr atom.* We obtain the *principal structure* of the atom. The typical energy level separation turns out to be a factor α^2 smaller than the electron *rest mass energy*, which justifies the non-relativistic Ansatz ($\alpha \approx 1/137$ is the *fine-structure constant*).

The subsequent task will be to include *relativistic* corrections. This is the subject of Chapter 4. Relativistic effects give rise to the atomic *fine structure* and manifest themselves most prominently as *magnetic forces.* As these forces are much weaker than Coulomb forces they may be neglected in describing the principal structure. Importantly, being weak does not mean unimportant. Magnetism is crucial for the understanding of many physical phenomena and of seminal importance for major technological applications. As we show in Chapter 4 the weakness of the magnetic interaction is convenient from the theoretical point of view because it allows an analytic description of many magnetic properties with the aid of perturbation theory. As magnetism is intimately related to angular momentum, the properties of angular momentum in its quantum mechanical context are summarized in Chapter 3. Before turning to Chapter 4, also a refreshment of the operational skills on perturbation theory (Appendix H) may prove well invested.

2.1 Hydrogenic atoms

Hydrogenic atoms consist of a single electronic charge orbiting a positively charged nucleus. It was demonstrated by Ernest Rutherford in 1911 that the nucleus is orders of magnitude smaller than the size of the atom [96]. After this discovery the atom was visualized as a miniature planetary system in which the nucleus could be approximated by a point charge. The conceptual step to the modern atom was made by Niels Bohr in 1913 [17, 18, 19]. By quantization of the planetary motion

Bohr formulated the famous Bohr model for the atom which explains the optical spectrum of atomic hydrogen and defines the *principal structure* of the atom. In this section we solve the Bohr atom with the method developed by Erwin Schrödinger in 1926 [100]. To start the discussion we note that the Coulomb interaction between electron and nucleus gives rise to a *central* potential. Hence, the hamiltonian operator to calculate the energy eigenvalues in the absence of externally applied fields is of the general form (1.6) and will be referred to as the *Schrödinger* Hamiltonian for the atom,

$$\mathcal{H}_0 = \frac{\mathbf{p}^2}{2m_r} + \mathcal{V}(r) = -\frac{\hbar^2}{2m_r}\Delta - \frac{Ze^2}{4\pi\varepsilon_0 r}. \tag{2.1}$$

Here r is the electron-nuclear radial distance and m_r the *reduced mass,*

$$m_r = m_e/(1 + m_e/M), \tag{2.2}$$

with $m_e \simeq 9.1 \times 10^{-31}$ Kg the electronic and M the nuclear rest mass. The quantity

$$\mathcal{V}(r) = -e\varphi(r) = -\frac{Ze^2}{4\pi\varepsilon_0 r} \tag{2.3}$$

is the *electrostatic energy*; i.e., the *potential energy* of the *electronic charge*, $-e$, in the electromagnetic *scalar potential* (Coulomb potential) of the nucleus

$$\varphi(r) = \frac{Ze}{4\pi\varepsilon_0 r}, \tag{2.4}$$

where Ze is the nuclear charge, with e being the elementary charge and Z the number of protons in the nucleus; Z is called the *atomic number* or *nuclear charge number.* Furthermore, $\varepsilon_0 = 1/\mu_0 c^2$ is the *vacuum permittivity,*[1] $\mu_0 = 4\pi\alpha\hbar/e^2 c$ the *vacuum permeability* and $c = 299\,792\,458\ \mathrm{m\,s^{-1}}$ the defined value of the *speed of light in vacuum.* The reduced mass has typically a value close to the mass of the free electron but can be substantially smaller for exotic atoms like positronium.

As $\mathcal{V}(r)$ represents a *central* potential energy field, the motion of the electron can be described by a Schrödinger equation of the type (1.66)

$$\left[\frac{1}{2m_r}\left(p_r^2 + \frac{\mathbf{L}^2}{r^2}\right) + \mathcal{V}(r)\right]\psi(r,\theta,\phi) = E\psi(r,\theta,\phi), \tag{2.5}$$

where p_r is the radial momentum operator, $\mathbf{L}$ the angular momentum operator and E the total energy of the system (see Section 1.1). The Hamiltonian commutes simultaneously with $\mathbf{L}^2$ and L_z. Therefore, these operators share a complete set of eigenfunctions with the full Hamiltonian and, as discussed in Section 1.1, can be written in the form of a product of the type

$$\psi_{nlm}(r,\theta,\phi) = R_{nl}(r)Y_l^m(\theta,\phi), \tag{2.6}$$

where n is the *principal quantum number* for the radial motion to be determined later in this chapter. The wavefunctions (2.6) are called *atomic orbitals.* Substituting this expression in Eq. (2.5) we obtain the *radial wave equation*

$$\left[\frac{\hbar^2}{2m_r}\left(-\frac{d^2}{dr^2} - \frac{2}{r}\frac{d}{dr} + \frac{l(l+1)}{r^2}\right) + \mathcal{V}(r)\right]R_{nl}(r) = ER_{nl}(r). \tag{2.7}$$

The term $l(l+1)/r^2$ represents the centrifugal energy in electronic motion (see Section 1.1) and $R_{nl}(r)$ is called the radial wave function. There is a standard convention to refer to the value of the orbital quantum number l

$$\begin{array}{lcccccccccccccccccc} l = & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & \cdots \\ & s & p & d & f & g & h & i & k & l & m & n & o & q & r & t & u & \cdots \end{array} \tag{2.8}$$

---

[1] Since the redefinition of the SI in 2019, the *vacuum permeability* and *vacuum permittivity* are no longer constants but depend on a measured quantity, the fine-structure constant α. This made the names *magnetic constant* and *electric constant* obsolete.

Table 2.1: Comparison of Hartree (27.2 eV) and Rydberg (13.6 eV) atomic units.

	SI	Hartree[†]	Rydberg
length unit:	m	$a = \dfrac{4\pi\varepsilon_0\hbar^2}{m_r e^2}$	$a = \dfrac{4\pi\varepsilon_0\hbar^2}{m_r e^2}$
energy unit:	J	$E_H = \dfrac{\hbar^2}{m_r a^2} = 2hcR_M$	$E_R = \dfrac{\hbar^2}{2m_r a^2} = hcR_M$
length:	r	$\rho = r/a$	$\rho = r/a$
energy:	E	$\varepsilon = E/E_H$	$\varepsilon = E/Ry$
$\mathcal{V}(r)$:	$-\dfrac{Ze^2}{4\pi\varepsilon_0 r}$	$U(\rho) = -\dfrac{Z}{\rho}$	$U(\rho) = -2\dfrac{Z}{\rho}$
$\mathcal{V}_{\rm rot}(r)$	$\dfrac{\hbar^2}{2m_r}\dfrac{l(l+1)}{r^2}$	$U_{\rm rot}(\rho) = \dfrac{l(l+1)}{2\rho^2}$	$U_{\rm rot}(\rho) = \dfrac{l(l+1)}{\rho^2}$
$\mathcal{H}_{\rm rad}(r)$	$-\dfrac{\hbar^2}{2m_r}\left(\dfrac{d^2}{dr^2} + \dfrac{2}{r}\dfrac{d}{dr}\right)$	$\mathcal{H}_{\rm rad}(\rho) = -\dfrac{1}{2\rho}\dfrac{d^2}{d\rho^2}\rho$	$\mathcal{H}_{\rm rad}(\rho) = -\dfrac{1}{\rho}\dfrac{d^2}{d\rho^2}\rho$

[†] Hartree a.u. correspond to setting $m_r = e = \hbar = 1/4\pi\varepsilon_0 = 1$; this implies $a = E_H = 1$.

In this convention one refers to s orbitals ($l = 0$), p orbitals ($l = 1$), d orbitals ($l = 2$), *etc.*. This nomenclature has its origin in atomic spectroscopy, with the first four orbital labels referring to the terms sharp, principal, diffuse, and fundamental. When it adds to the readability of equations one sometimes writes $R_{ns}(r), R_{np}(r), R_{nd}(r), \cdots$, rather than $R_{n0}(r), R_{n1}(r), R_{n2}(r), \cdots$.

2.1.1 Atomic units

The 1D Schrödinger equation is made dimensionless by turning to atomic units (a.u.). For this purpose Eq. (1.74) is multiplied by a^2, where a is a characteristic length to be determined. If we choose a such that

$$\frac{m_r a^2}{\hbar^2}\mathcal{V}(r) = -\frac{m_r a}{\hbar^2}\frac{Ze^2 a}{4\pi\varepsilon_0 r} = -\frac{Z}{\rho} \equiv U(\rho), \tag{2.9}$$

where $\rho = r/a$ is the dimensionless radius, we find that

$$a = \frac{4\pi\varepsilon_0\hbar^2}{m_r e^2} = (m_e/m_r)a_0 = (1 + m_e/M)a_0. \tag{2.10}$$

where $a_0 = 4\pi\varepsilon_0\hbar^2/m_e e^2\,(= \alpha/4\pi R_\infty) \simeq 5.2 \times 10^{-11}$ m is the *Bohr radius.* Note that the proton charge radius, $r_p \simeq 8.8 \times 10^{-16}$ m, is almost 5 orders of magnitude smaller. The characteristic length a obtained with this procedure is called the *atomic unit of length.* The corresponding *Hartree atomic unit of energy* is defined as

$$E_H = \frac{\hbar^2}{m_r a^2} = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{a} = \alpha^2 m_r c^2 = 2hcR_M = \frac{2hcR_\infty}{1 + m_e/M}, \tag{2.11}$$

where the dimensionless constant $\alpha = e^2/4\pi\varepsilon_0\hbar c = \hbar/m_e c a_0 \approx 1/137$ is the *fine-structure constant,* and $R_M = R_\infty/(1 + m_e/M)$ and $R_\infty = \alpha^2 m_e c/2h$ the finite-mass/infinite-mass *Rydberg constants,* respectively. Note that $4\pi\alpha^2 a = \mu_0 e^2/m_r$. The *Hartree energy* is defined as $E_h = \alpha^2 m_e c^2 \simeq$

4.36×10^{-18} J $\simeq 27.2$ eV.[1] This shows that the characteristic energy scale of the atom is α^2 times smaller the rest mass energy of the electron. In other words, relativistic phenomena play a minor role in the description of the orbital motion of the electron. Sometimes the *Rydberg atomic unit* is used,[2] 1Ry $\simeq 13.6$ eV, twice as small as the Hartree and corresponding to the ionization energy of the hydrogen ground state. Hartree and Rydberg atomic units are compared in Table 2.1. The operator for the dimensionless radial kinetic energy $\mathcal{H}_{\mathrm{rad}}(r)$ is given by

$$\frac{m_r a^2}{\hbar^2}\mathcal{H}_{\mathrm{rad}}(r) = \frac{m_r a^2}{\hbar^2}\frac{\hbar^2}{2m_r}\left(-\frac{d^2}{dr^2} - \frac{2}{r}\frac{d}{dr}\right) = -\frac{1}{2\rho}\frac{d^2}{d\rho^2}\rho. \tag{2.12}$$

Defining the energy of bound states as negative, $\varepsilon = -\kappa^2$ (with $\kappa > 0$), we find for the energy in the dimensionless form

$$\varepsilon = -\kappa^2 = \frac{m_r a^2}{\hbar^2}E. \tag{2.13}$$

Hence, in Hartree atomic units the radial wave equation can be written compactly as

$$\tilde{R}_l''/2 + \tilde{R}_l'/\rho + [Z/\rho - l(l+1)/2\rho^2 - \kappa^2]\tilde{R}_l = 0, \tag{2.14}$$

where we suppressed (for the time being) the quantum number n and write

$$\tilde{R}_l(\rho) = \tilde{R}_l(r/a) \equiv a^{3/2}R_l(r) = a^{3/2}R_l(\rho a) \tag{2.15}$$

for the radial wavefunctions redefined as unit-normalized dimensionless functions of ρ,

$$\int R_l^2(r)r^2dr = \int \tilde{R}_l^2(\rho)\rho^2 d\rho = 1. \tag{2.16}$$

The primes in Eq. (2.14) represent derivatives with respect to the variable ρ. Note that in Hartree atomic units ($a \to 1$, $\rho = r/a \to r$) the expressions for $R_l(r)$ and $\tilde{R}_l(\rho)$ coincide. A similar equation can be based on the Rydberg atomic unit of energy, 1Ry $= hcR_M$.

Problem 2.1. Show that in Hartree atomic units $a = 1$, $E_H = 1$, $c = 1/\alpha$, $R_M = \alpha/4\pi$.

2.1.2 Solving the radial wave equation

In this section we solve the radial wave equation (2.14) in dimensionless form using Hartree atomic units. Recalling the substitution

$$\tilde{\chi}_l(\rho) = \rho\tilde{R}_l(\rho), \tag{2.17}$$

we first reformulate Eq. (2.14) into the form of a 1D Schrödinger equation,

$$\tilde{\chi}_l'' + 2\left[Z/\rho - l(l+1)/2\rho^2 - \kappa_l^2\right]\tilde{\chi}_l = 0. \tag{2.18}$$

For $\rho \to 0$ the rotational term is dominant and the 1D equation may be approximated by

$$\tilde{\chi}_l'' + \left[-l(l+1)/\rho^2\right]\tilde{\chi}_l = 0, \tag{2.19}$$

with solutions $\tilde{\chi}_l(\rho) \sim \rho^{l+1}$, regular in the origin for all values of l. Likewise, for $\rho \to \infty$ we may neglect the Z/ρ and $l(l+1)/\rho^2$ terms and obtain (for all values of l)

$$\tilde{\chi}_l'' - 2\kappa_l^2\tilde{\chi}_l = 0, \tag{2.20}$$

---

[1] Note the difference between the Hartree a.u., $E_H = \hbar^2/m_r a^2 = E_h/(1 + m_e/M)$, and the Hartree energy, E_h.

[2] Typically we use the Hartree a.u.; when occasionally using the Rydberg a.u. this is explicitly indicated.

with solutions $\tilde{\chi}_l(\rho) \sim e^{-\kappa_l \rho\sqrt{2}}$, which are finite for large ρ. This suggest to write

$$\tilde{\chi}_l(\rho) = \rho^{l+1} e^{-\kappa_l \rho\sqrt{2}} \tilde{w}_l(\rho) \tag{2.21}$$

and search for a function $\tilde{w}_l(\rho)$ that allows us to connect the short-range expression of the wave function to the long-range one. The desired function must satisfy the condition $\tilde{w}_l(0) = 1$ and be algebraic for $r \to \infty$. Substituting the expression for $\tilde{\chi}_l(\rho)$ into Eq. (2.18) we obtain the following *nonlinear* differential equation for $\tilde{w}_l(\rho)$,

$$\rho \tilde{w}_l'' + 2[(l+1) - \sqrt{2}\,\kappa_l \rho]\tilde{w}_l' + 2[Z - (l+1)\sqrt{2}\,\kappa_l]\tilde{w}_l = 0. \tag{2.22}$$

Hence, the condition $\tilde{w}_l(0) = 1$ implies $\tilde{w}_l'(0) = \sqrt{2}\,\kappa_l \gamma/(l+1)$, where $\gamma \equiv (l+1) - Z/(\sqrt{2}\,\kappa_l)$. Dividing Eq. (2.22) by $2\sqrt{2}\,\kappa_l$ and turning to new variables, $\tilde{w}_l \to w_l$ and $\rho \to x$, where

$$w_l(x) = \tilde{w}_l(\rho) \quad \text{and } x = 2\sqrt{2}\,\kappa_l \rho, \tag{2.23}$$

we obtain the *Kummer equation*,

$$x w_l'' + [\beta - x] w_l' - \gamma w_l = 0, \tag{2.24}$$

where $\beta = 2(l+1)$. The derivatives are now with respect to the variable x and the boundary conditions become $w_l(0) = 1$ and $w_l'(0) = \gamma/\beta$. For non-positive values of γ the solutions of Eq. (2.24) are confluent hypergeometric series (Kummer functions) [1]

$$w_l(x) = {}_1F_1(\gamma|\beta|x) = \sum_{p=0}^{\infty} \frac{\Gamma(\gamma+p)\Gamma(\beta)}{\Gamma(\gamma)\Gamma(\beta+p)} \frac{x^p}{p!} = 1 + \frac{\gamma}{\beta}\frac{x}{1!} + \frac{\gamma(\gamma+1)}{\beta(\beta+1)}\frac{x^2}{2!} + \cdots . \tag{2.25}$$

If γ is a non-positive integer, $\gamma(\kappa_{n'l}) = -n'$, this series turns into a polynomial of degree $n' \geq 0$

$$w_l(x) = {}_1F_1(-n'|\beta|x) = \sum_{p=0}^{n'} \frac{\Gamma(p-n')\Gamma(\beta)}{\Gamma(-n')\Gamma(\beta+p)} \frac{x^p}{p!}. \tag{2.26}$$

For non-integer values of γ the series is not truncated and has an essential singularity for $x \to \infty$. As we are only interested in normalizable wave functions it is conventional to use Laguerre polynomials of degree n' (see Appendix L.11) rather than Kummer functions,

$$w_l(x) = \frac{\Gamma(n'+1)\Gamma(\beta)}{\Gamma(\beta+n')} L_{n'}^{\beta-1}(x). \tag{2.27}$$

The truncation condition for the series, $\gamma(\kappa_{n'l}) = -n' \leq 0$, allows us to normalize the radial wave function and provides us with the quantization condition

$$\gamma(\kappa_{n'l}) = (l+1) - \frac{Z}{\sqrt{2}\,\kappa_{n'l}} = -n' \leq 0 \quad \Leftrightarrow \quad \sqrt{2}\,\kappa_{n'l} = \frac{Z}{n'+l+1}. \tag{2.28}$$

Note that the energy eigenvalues depend on two quantum numbers, n' and l. The integer n' is known as the *radial quantum number*. Its value corresponds to the number of zeros of the Laguerre polynomial; i.e., the number of nodes in the radial wavefunction outside the origin. In view of the experimental practice since the discovery of the *Balmer formula* in 1885 and the prominent role of the *Bohr theory* [17, 18, 19] since 1913, the radial quantum number is *not* used in the common scientific literature; preference is given to the *principal quantum number*, a historical mix of the quantum numbers n' and l,

$$n \equiv n' + l + 1. \tag{2.29}$$

Recalling the definitions (2.23) we find with the aid of Eq. (2.28) $x = 2Z\rho/n$. The radial wavefunction corresponding to the quantum numbers n and l is of the form

$$\tilde{R}_{nl}(\rho) = A_{nl}\rho^l e^{-Z\rho/n} w_l(2Z\rho/n), \tag{2.30}$$

where A_{nl} is a normalization constant. As $w_l(x)$ is a polynomial of degree $n' = n - l - 1 \geq 0$, the orbital angular momentum quantum number satisfies the condition

$$l \leq n - 1. \tag{2.31}$$

Since the formulation of the Bohr theory, electrons with the same principal quantum number are said to belong to the same *electron shell.* The shells are labeled $K, L, M, N, O, P, Q, \cdots$ in order of increasing principal quantum number $n = 1, 2, 3, \cdots$. Electrons sharing the same quantum numbers for n and l are called *equivalent electrons* and occupy a *(sub)shell.* Specifying the number of electrons per subshell one obtains the *electron configuration.* For example, with one electron in the $1s$ shell the configuration of hydrogenic atoms in their ground state is $(1s)^1$, usually abbreviated to $1s$.

2.2 Energy levels and degeneracy

In terms of the principal quantum number the quantization condition (2.28) takes the famous form of the *Bohr formula* [17],

$$\varepsilon = -\frac{Z^2}{2n^2} \text{ Hartree}, \tag{2.32}$$

with $n \geq 1$. Restoring the dimensions we obtain for the energy eigenvalues

$$E_n = -\alpha^2 m_r c^2 \frac{Z^2}{2n^2} = -\frac{\alpha^2 m_e c^2 Z^2}{1 + m_e/M} \frac{1}{2n^2} = -hcR_M \frac{Z^2}{n^2} = -\frac{Z^2}{n^2} \text{ Ry}. \tag{2.33}$$

Note that the atom has an infinite number of bound states. Writing $E_n \simeq \frac{1}{2} m_e \left\langle v^2 \right\rangle$ we infer that $\alpha^2 Z^2 \simeq \left\langle (v/c)^2 \right\rangle$. States with a large principal quantum number are called *Rydberg states* and atoms excited to those states are called *Rydberg atoms.* Eq. (2.33) defines the complete spectrum of the hydrogen atom according to the Schrödinger theory. It shows *degeneracy* of the energy levels, which means that different states have the same energy. For states of given l the degeneracy is $2l+1$ because the energy eigenvalues are independent of the the quantum number m_l. Degeneracies of this type are called *essential.* In contrast, degeneracies of states with different quantum numbers (e.g., states of given n but differing in l) are called *accidental.* With the condition $l \leq n - 1$ the degeneracy of the level E_n is given by the arithmetic series

$$\sum_{l=0}^{n-1} (2l + 1) = \tfrac{1}{2}[1 + (2n - 1)]n = n^2. \tag{2.34}$$

The atomic energy levels are illustrated in Fig. 2.1*a* along with the names of some well-known atomic transitions between hydrogenic levels: Lyman α (L_α) and Balmer-α (H_α). In the case of optically induced transitions the change in state of the atom is accompanied by a simultaneous change of the light field. This happens through exchange of a *photon* between the atom and field in such a way that the total energy and momentum are conserved. Energy-level diagrams discriminating between various quantum numbers (like in Fig. 2.1*a*) are often referred to as *Term diagrams.*[1] This practice finds its origin in the assignment of the energy levels in optical spectroscopy (see Fig. 2.1*b*) in which spectral lines are assigned to *terms* in a *series* of transitions sharing the same final (or

---

[1] Term diagrams were introduced by the astrophysicist Walter Grotrian in 1928. The formal definition of the notion *Term* is given in Section 4.6.2.

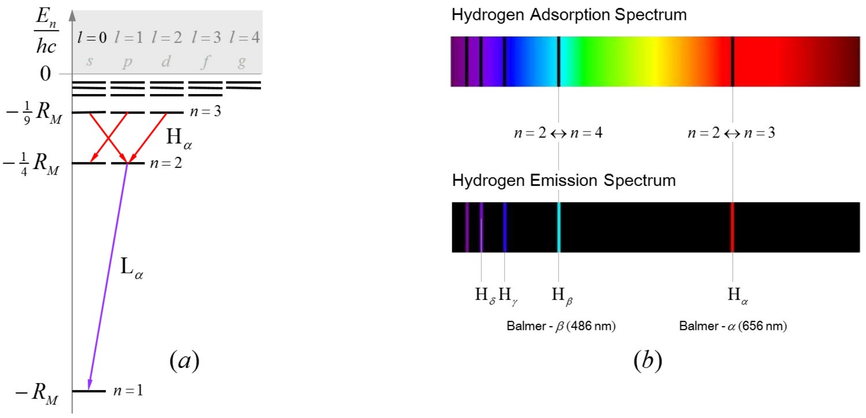


Figure 2.1: (a) Atomic energy level diagram (*Term diagram*) of hydrogen. Indicated are the first transitions of both the Lyman series (L_α) and the Balmer series (H_α). (b) Balmer spectrum of hydrogen observed in absorption (upper) and emission (lower).

initial) state. The first series of this type was discovered empirically by Balmer in 1885 for the spectral lines of hydrogen in the visible part of the spectrum. The *Balmer series* is defined by all optical excitations from (or decay to) the $n = 2$ level. Analogously, the *Lyman series* involves the transitions to (or from) the hydrogen ground state ($n = 1$). In 1888 Rydberg generalized the expression for the Balmer series into the *Rydberg formula*, which yields the wavelength (in vacuum) of any transition between two hydrogenic energy levels,

$$\frac{1}{\lambda_n} = -R_M\left[\left(\frac{1}{n}\right)^2 - \left(\frac{1}{n_0}\right)^2\right], \tag{2.35}$$

where $n > n_0$, with n_0 being the principal quantum number of the common and lowest level in all transitions considered; i.e., the defining level of the series. The corresponding energy splitting is given by

$$\Delta E_n = E_n - E_0 = hc/\lambda_n. \tag{2.36}$$

If n_0 refers to an initial state, the atom is *excited* from the common level $E_0 \equiv E_{n_0}$ to one of the levels E_n under *absorption* of a photon of energy $\hbar\omega_n = E_n - E_0$.[1] In astronomy, such transitions are observed as dark lines in the spectrum of stellar light (see Fig. 2.1*b-upper*) and demonstrate the presence of *cold* (dark) interstellar gas between the observer and a distant star. Inversely, if n_0 refers to a final state, the atom is *de-excited* from one of the levels E_n to the level E_0 under *emission* of a photon of energy $\hbar\omega_n = E_n - E_0$. This is observed in spectroscopic studies of glowing *hot* interstellar clouds (in regions of star formation), where the emission appears as bright lines against a dark background in the spectrum of the glow (see Fig. 2.1*b-lower*).

2.3 Eigenfunctions of the bound states

Combining Eqs. (2.30) and (2.27) we find for the radial wavefunctions

$$\tilde{R}_{nl}(\rho) = \mathcal{N}_{nl}^{-1/2}(2Z/n)^{l+3/2}\rho^l e^{-Z\rho/n} L_{n-l-1}^{2l+1}(2Z\rho/n). \tag{2.37}$$

Note that the normalization factor A_{nl} used in Eq. (2.30) has been replaced by the normalization factor $\mathcal{N}_{nl}^{-1/2}$. Substituting this expression into Eq. (2.16) and changing to the variable $x = 2Z\rho/n$

---

[1] Here we neglect the *recoil shift*, a small shift of the optical transition frequency with respect to the value ω_n. It results from the conservation of linear momentum in the emission of a photon.

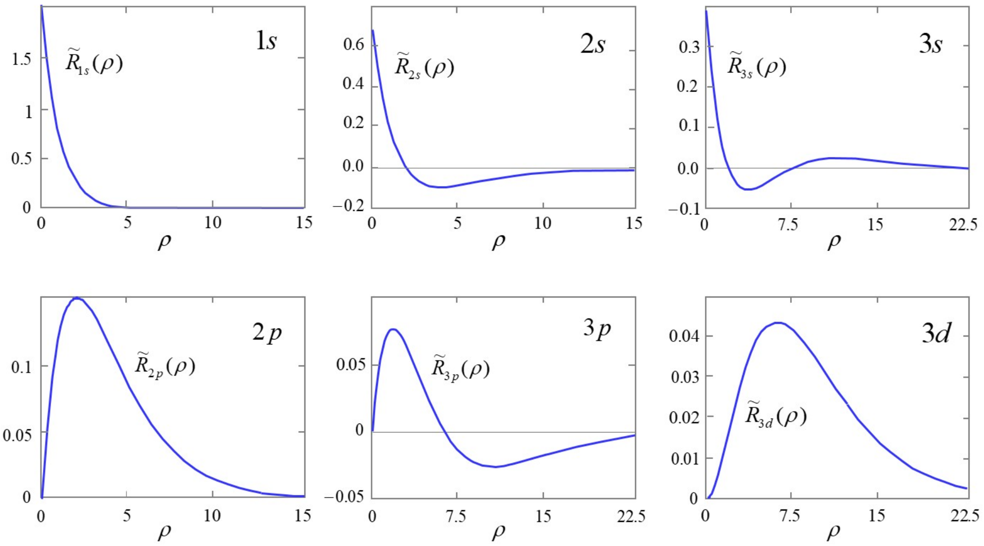


Figure 2.2: The lowest order hydrogenic radial wavefunctions plotted in atomic units as a function of radial distance. Note that for a given value of the orbital quantum number l the number of nodes increases with the principal quantum number n.

we find with the aid of Eq. (L.83)

$$\mathcal{N}_{nl} = \int_0^\infty x^{2l+2} e^{-x} \left[L_{n-l-1}^{2l+1}(x) \right]^2 dx = J_1(n-l-1, 2l+1) = \frac{2n\Gamma(n+l+1)}{\Gamma(n-l)}. \tag{2.38}$$

A few important cases are $\mathcal{N}_{1s} = 2$, $\mathcal{N}_{2p} = 24$, $\mathcal{N}_{3d} = 720$ and $\mathcal{N}_{4f} = 40320$. The integral $J_1(n-l-1, 2l+1)$ is a generalized normalization integral as defined by Eq. (L.81). Inspecting the solution Eq. (2.37), we find that $\tilde{\chi}_l(\rho) \sim \rho^n e^{-Z\rho/n}$ for $\rho \to \infty$. This is independent of l because asymptotically the Coulomb term, $2Z/\rho$, always dominates over the $-l(l+1)/\rho^2$ term in the Hamiltonian (all radial wave functions behave asymptotically as s waves). For $\rho \to 0$ the radial wavefunction vanishes for all values $l > 0$. For the case $l = 0$ the *probability density* (see Appendix F.1.1) of the radial wavefunction at the origin is

$$\tilde{R}_{ns}^2(0) = 4(Z/n)^3. \tag{2.39}$$

As an example we calculate the lowest hydrogenic radial wavefunctions for the cases $n = 1, 2, 3$ and arbitrary Z. The results for the hydrogen atom are obtained by setting $Z = 1$. Substituting the quantum numbers n and l into Eqs. (2.37) and (2.38) we find

$$\tilde{R}_{1s}(\rho) = Z^{3/2} 2e^{-Z\rho} \tag{2.40a}$$

$$\tilde{R}_{2s}(\rho) = (Z/2)^{3/2} (1 - Z\rho/2) 2e^{-Z\rho/2} \tag{2.40b}$$

$$\tilde{R}_{2p}(\rho) = (Z/2)^{3/2} \sqrt{1/3} (Z\rho/2) 2e^{-Z\rho/2} \tag{2.40c}$$

$$\tilde{R}_{3s}(\rho) = (Z/3)^{3/2} \left[1 - 2(Z\rho/3) + \tfrac{2}{3}(Z\rho/3)^2 \right] 2e^{-Z\rho/3} \tag{2.40d}$$

$$\tilde{R}_{3p}(\rho) = (Z/3)^{3/2} \sqrt{8/9} \left[1 - \tfrac{1}{2}(Z\rho/3) \right] 2e^{-Z\rho/3} \tag{2.40e}$$

$$\tilde{R}_{3d}(\rho) = (Z/3)^{3/2} \sqrt{2/45} (Z\rho/3)^2 2e^{-Z\rho/3} \tag{2.40f}$$

In all cases $\int \tilde{R}_{nl}^2(\rho)\rho^2 d\rho = 1$. Some examples are shown (for $Z = 1$) in Fig. 2.2.

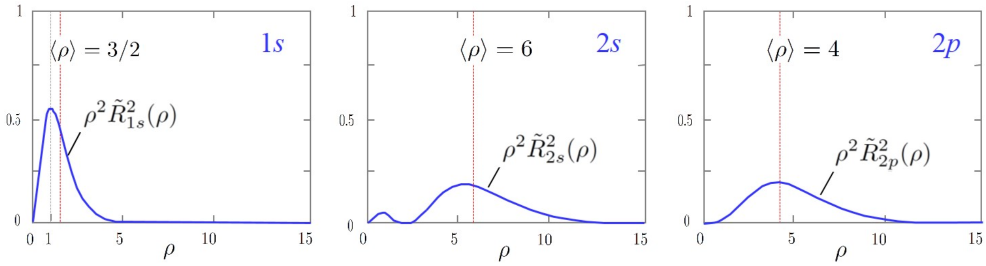


Figure 2.3: Radial distribution functions for the lowest hydrogenic wavefunctions. For the $1s$ wavefunction the radial distribution reaches a maximum at the Bohr radius ($\rho = 1$).

2.3.1 Dirac notation

In many cases it is convenient to adopt the Dirac notation $|nlm\rangle$ for the electronic orbital eigenstates of hydrogenic atoms. In this notation the Schrödinger equation (2.5) takes the compact form

$$\mathcal{H}_0\,|nlm\rangle = E_n\,|nlm\rangle\,. \tag{2.41}$$

The relation with the eigenfunctions in the position representation is given by

$$\psi_{nlm}(\mathbf{r}) = \langle\mathbf{r}|nlm\rangle = \langle r|nl\rangle\langle\hat{\mathbf{r}}|lm\rangle, \tag{2.42}$$

with $\langle r|nl\rangle = rR_{nl}(r)$ and $\langle\hat{\mathbf{r}}|lm\rangle = Y_l^m(\hat{\mathbf{r}})$.

2.4 Diagonal matrix elements

2.4.1 Radial averages

The expectation value of an operator A for an atom in eigenstate $|nlm\rangle$ is given by

$$\langle A\rangle \equiv \langle nlm|A|nlm\rangle\,. \tag{2.43}$$

In many cases of practical importance the operator A depends analytically on the radial distance to the nucleus, $A = A(\rho)$. Operators of this type are *diagonal* in the representation $\{|nlm\rangle\}$. Importantly, $\langle A\rangle$ separates into the product of a radial and an angular integral,[1]

$$\langle A\rangle_{nlm} = \int d\rho\,\rho^2\tilde{R}_{nl}(\rho)A(\rho)\tilde{R}_{nl}(\rho)\int |Y_l^m(\Omega)|^2\,d\Omega, \tag{2.44}$$

which reduces, given the normalization of the spherical harmonics, to a purely radial integral,

$$\langle A\rangle_{nlm} = \int d\rho\,A(\rho)\tilde{\chi}^2_{nl}(\rho) = \langle A\rangle_{nl}. \tag{2.45}$$

Here $\tilde{\chi}_{nl}(\rho)$ is a *reduced radial wavefunction* as introduced in Section 1.2.1. By separating the radial from the angular integral we obtain a *reduced matrix element*; i.e., it depends on n and l but *not* on m (see Problem 2.2). Like an ordinary matrix element it is an integral over the radial variable but differs by the presence of the additional weight factor ρ^2. The function

$$\tilde{\chi}^2_{nl}(\rho) = \rho^2\tilde{R}^2_{nl}(\rho) \tag{2.46}$$

is called the *radial distribution function* of the state $|nl\rangle$ and expresses the radial distribution of the probability density to find the electron at distance ρ from the origin. For the lowest radial wavefunctions these are shown in Fig. 2.3.

---

[1] In these lecture notes we use interchangeable notations for the radial unit vector: $\hat{\mathbf{r}} \equiv \Omega \equiv (\theta, \phi)$.

2.4.1.1 Atomic size

As a first example we calculate the average radius of the hydrogen ground state $|1s\rangle$ with the aid of Eq. (2.40a),

$$\langle\rho\rangle_{1s} = \int \rho[2\rho e^{-\rho}]^2 d\rho = 4\int \rho^3 e^{-2\rho} d\rho = \tfrac{1}{4}\int x^3 e^{-x} dx = \tfrac{1}{4}\Gamma(4) = \tfrac{3}{2}. \tag{2.47}$$

Comparing with Eq. (2.10) we find that $\langle r\rangle_{1s} = \frac{3}{2}a$, which is 50% larger than the Bohr radius, $a_0 \simeq 5.2\times 10^{-11}$ m. As the proton size is orders of magnitude smaller, $r_p \simeq 8.8\times 10^{-16}$ m, this justifies in hindsight our Ansatz of treating the nucleus as a point charge. Compact relations for the radial averages $\langle\rho^k\rangle_{nl}$ in atomic units can be obtained by expressing Eq. (2.44) in terms of the integrals $J_\nu(n', 2l+1)$ defined in Eq. (L.81),

$$\begin{aligned}\langle\rho^k\rangle_{nl} &= \frac{1}{\mathcal{N}_{nl}}\left(\frac{n}{2Z}\right)^k \int_0^\infty x^{2l+2+k} e^{-x}\left[L_{n-l-1}^{2l+1}(x)\right]^2 dx \\ &= \left(\frac{n}{2Z}\right)^k \frac{J_{k+1}(n-l-1, 2l+1)}{J_1(n-l-1, 2l+1)}.\end{aligned} \tag{2.48}$$

In particular we obtain for a number of important special cases with the aid of Eqs. (L.82)-(L.85)

$$\langle\rho\rangle_{nl} = \frac{1}{2Z}[3n^2 - l(l+1)] \qquad \langle\rho^{-1}\rangle_{nl} = \frac{Z}{n^2} \tag{2.49a}$$

$$\langle\rho^2\rangle_{nl} = \frac{1}{2Z^2}n^2[5n^2+1-3l(l+1)] \qquad \langle\rho^{-2}\rangle_{nl} = \frac{Z^2}{n^3}\frac{1}{l+1/2} \tag{2.49b}$$

$$\langle\rho^{-3}\rangle_{nl} = \frac{Z^3}{n^3}\frac{1}{(l+1)(l+1/2)\,l}. \tag{2.49c}$$

By substituting $n = 1$, $l = 0$, $Z = 1$ we regain the result $\langle r\rangle = 3/2a$ for the average radius of the hydrogen ground state in atomic units. With Eqs. (2.49a) and (2.49b) the variance in the radial position is calculated to be

$$\langle(\rho - \langle\rho\rangle)^2\rangle_{nl} = \langle\rho^2\rangle_{nl} - \langle\rho\rangle_{nl}^2 = \frac{1}{4Z^2}\left[n^2(n^2+2) - l^2(l+1)^2\right]. \tag{2.50}$$

As another example of a radial average we calculate the average radial kinetic energy for the hydrogen ground state $|100\rangle$ in Hartree atomic units

$$\begin{aligned}\left\langle -\frac{1}{2\rho}\frac{\partial^2}{\partial\rho^2}\rho\right\rangle_{1s} &= -2\int d\rho\,\rho e^{-\rho}\frac{d^2}{d\rho^2}(\rho e^{-\rho}) \\ &= 4\int \rho e^{-2\rho}d\rho - 2\int \rho^2 e^{-2\rho} d\rho = \Gamma(2) - \tfrac{1}{4}\Gamma(3) = \tfrac{1}{2}.\end{aligned} \tag{2.51}$$

This expression may be generalized for arbitrary hydrogenic eigenstates (see Problem 2.3).

Problem 2.2. Show that Eq. (2.45) is a reduced matrix element in the sense of the Wigner-Eckart theorem (cf. Appendix K.2).

Solution. The scalar operator $A(\rho)$ is an irreducible tensor operator of rank zero. Using the sum rule of Appendix K.2.1 the reduced matrix element is given by

$$|\langle nl||A||nl\rangle|^2 = \sum_{m=-l}^{l} |\langle nlm|A|nlm\rangle|^2 .$$

Substituting Eq. (2.45) we have $2l+1$ equal terms and obtain

$$\langle A\rangle_{nl}^2 = |\langle nl||A||nl\rangle|^2/(2l+1),$$

which is independent of m and, as such, a reduced matrix element in the sense of the Wigner-Eckart theorem. For zero-order tensors the use of this theorem is overkill because the angular integral not only factors out of the integral but evaluates to unity - see derivation of Eq. (2.45). □

Problem 2.3. Show that the radial kinetic energy of the state $|nlm\rangle$ is given by

$$\left\langle -\frac{1}{2\rho}\frac{\partial^2}{\partial\rho^2}\rho\right\rangle_{nl} = \frac{Z^2}{n^2}\left[\frac{1}{2} - \frac{l\,(l+1)}{n\,(2l+1)}\right] \text{ a.u.}. \tag{2.52}$$

Solution. Changing to the variable $x = 2Z\rho/n$ we have

$$\tilde{R}_{nl}^2(\rho)\rho^2 d\rho = \mathcal{N}_{nl}^{-1} x^{2l+2} e^{-x}\left[L_{n-l-1}^{2l+1}(x)\right]^2 dx$$

and after angular integration we are left with the radial average,

$$\left\langle -\frac{1}{2\rho}\frac{\partial^2}{\partial\rho^2}\rho\right\rangle_{nl} = -\frac{1}{2\mathcal{N}_{nl}}\left(\frac{2Z}{n}\right)^2 \int_0^\infty x^{l+1}e^{-x/2}L_{n-l-1}^{2l+1}(x)\frac{d^2}{dx^2}[x^{l+1}e^{-x/2}L_{n-l-1}^{2l+1}(x)]dx.$$

In view of the orthogonality relation (L.76) we have to retain only those derivatives which are proportional to $\left[L_{n-l-1}^{2l+1}(x)\right]^2$,

$$\left[L_{n-l-1}^{2l+1}(x)\right]^2 x^{l+1}e^{-x/2}\tfrac{d^2}{dx^2}x^{l+1}e^{-x/2} = x^{2l+1}e^{-x}\left(l(l+1)x^{-1} - (l+1) + \tfrac{1}{4}x\right)\left[L_{n-l-1}^{2l+1}(x)\right]^2$$

$$-L_{n-l-1}^{2l+1}(x)\textstyle\sum_{m=0}^{n-l-2} L_m^{2l+1}(x)\left[2x^{l+1}e^{-x/2}\tfrac{d}{dx}x^{l+1}e^{-x/2}\right] \to -x^{2l+1}e^{-x}\,(n-l-1)\left[L_{n-l-1}^{2l+1}(x)\right]^2$$

$$L_{n-l-1}^{2l+1}(x)\textstyle\sum_{m=0}^{n-l-2} lL_m^{2l+1}(x)\left(x^{l+1}e^{-x/2}\right)^2 \to 0.$$

Here we used the expansions (L.80b) and (L.80c). Collecting the selected terms we obtain

$$\begin{aligned}\left\langle -\frac{1}{2\rho}\frac{\partial^2}{\partial\rho^2}\rho\right\rangle_{nl} &= -\frac{2}{\mathcal{N}_{nl}}\frac{Z^2}{n^2}\int_0^\infty x^{2l+1}e^{-x}\left[L_{n-l-1}^{2l+1}(x)\right]^2\left(l(l+1)x^{-1} - n + \frac{1}{4}x\right)dx \\ &= \frac{Z^2}{n^2}\left\{-2l(l+1)\frac{J_{-1}(n-l-1,2l+1)}{J_1(n-l-1,2l+1)} + 2n\frac{J_0(n-l-1,2l+1)}{J_1(n-l-1,2l+1)} - \frac{1}{2}\right\}.\end{aligned}$$

Substituting the expressions for the $J_\nu(m,\alpha)$ from Appendix L.11 we find Eq. (2.52). □

2.4.2 Angular averages

Angular distributions can be written as an expansion over spherical harmonics $Y_k^q(\hat{\mathbf{r}})$ for which the angular averages are easily calculated using the *Gaunt integral* (see Problem K.4) [49],

$$\begin{aligned}\langle Y_k^q(\hat{\mathbf{r}})\rangle_{lm} &\equiv \langle lm|Y_k^q(\hat{\mathbf{r}})|lm\rangle \\ &= \int Y_k^q(\hat{\mathbf{r}})\,|Y_l^m(\hat{\mathbf{r}})|^2\,d\hat{\mathbf{r}} \qquad (2.53)\\ &= (-1)^m\,(2l+1)\sqrt{\frac{(2k+1)}{4\pi}}\begin{pmatrix} l\,k\,l \\ 0\,0\,0\end{pmatrix}\begin{pmatrix} l & k & l \\ -m & q & m\end{pmatrix}. \qquad (2.54)\end{aligned}$$

This average is *non-zero only* for $q=0$ and $k=even$ in the interval $0\le k\le 2l$. See Eq. (J.21) and the selection rules for $3j$ symbols.

2.4.2.1 Examples

The angular averages of the lowest-order spherical harmonics are found by using the expressions for the Wigner $3j$ symbols - see Eq. (L.59),

$$\langle lm|Y_0^0(\hat{\mathbf{r}})|lm\rangle = \sqrt{\frac{1}{4\pi}} \qquad (2.55a)$$

$$\langle lm|Y_1^0(\hat{\mathbf{r}})|lm\rangle = 0 \qquad (2.55b)$$

$$\langle lm|Y_2^0(\hat{\mathbf{r}})|lm\rangle = \frac{l\,(l+1)-3m^2}{(2l+3)(2l-1)}\sqrt{\frac{5}{4\pi}}. \qquad (2.55c)$$

Problem 2.4. Show that for s orbitals $\langle lm|Y_k^q(\hat{\mathbf{r}})|lm\rangle$ is non-zero only for $k=q=0$.

Solution. To calculate the angular average of $Y_k^q(\hat{\mathbf{r}})$ we use Eq. (2.54). The angular average is nonzero if the $3j$ symbols are nonzero; i.e., for $0\le k\le 2l$. For s orbitals ($l=0$) this implies $0\le k\le 0$. Hence, for $k>0$ the angular averages all vanish. For $k=0$ we find

$$\langle 00|Y_0^0(\hat{\mathbf{r}})|00\rangle = \sqrt{1/4\pi}\langle 00|00\rangle = \sqrt{1/4\pi}. \qquad \square$$

Problem 2.5. Show that the s-wave angular average of $\sin^2\theta$ is $\langle 00|\sin^2\theta|00\rangle = 2/3$.

Solution. First we express $\sin^2\theta$ in spherical harmonics

$$\sin^2\theta = \frac{2}{3}-\frac{2}{3}\sqrt{\frac{4\pi}{5}}Y_2^0(\theta,\phi)$$

Then the angular average follows with Eq. (2.55c),

$$\langle lm|\sin^2\theta|lm\rangle = \frac{2}{3}-\frac{2}{3}\frac{l\,(l+1)-3m^2}{(2l+3)(2l-1)}.$$

For s waves ($l=m=0$) this yields 2/3. It may speak for itself that there are simpler ways to average $\sin^2\theta$ over a sphere. This problem merely serves to demonstrate the general case. □

2.5 Off-diagonal matrix elements

2.5.1 Transition dipole matrix elements

In this section we investigate the matrix elements $\langle n'l'm'|\,\mathbf{d}\,|nlm\rangle$ of the *electric-dipole operator*

$$\mathbf{d} = -\,e\mathbf{r} = -\,ea\rho\,\hat{\mathbf{r}} \qquad (2.56)$$

between eigenstates of the Schrödinger Hamiltonian. The electric-dipole operator has odd *parity*; i.e., the operator changes sign when replacing the radius vector $\mathbf{r}$ by $-\mathbf{r}$. Since the electronic eigenstates carry the parity of the $Y_l^m(\hat{\mathbf{r}})$ - see Eq. (L.53) - the expectation value of the electric-dipole operator averages to zero, $\langle nlm|\mathbf{d}|nlm\rangle = 0$. This reflects the absence of a *permanent electric-dipole moment* for hydrogenic eigenstates.[1] Hence, the atomic-dipole matrix is off-diagonal.

The off-diagonal matrix element

$$\mathbf{D}_{eg} = \langle e|\,\mathbf{d}\,|g\rangle \tag{2.57}$$

is called the *transition-dipole moment* of the $e \leftrightarrow g$ transition between atomic states. As $\mathbf{D}_{eg}$ is generally a complex vector it may be written as the product of a *generally complex* unit vector $\hat{\mathbf{u}}$ and a *real* prefactor D_{eg},

$$\mathbf{D}_{eg} = D_{eg}\hat{\mathbf{u}}. \tag{2.58}$$

Because the dipole operator is hermitian,[2] $\mathbf{D}_{eg} = \mathbf{D}_{ge}^*$, we have

$$|\mathbf{D}_{eg}|^2 = |\mathbf{D}_{ge}|^2 = D_{eg}^2. \tag{2.59}$$

With regard to the eigenstates $|g\rangle = |nlm\rangle$ and $|e\rangle = |n'l'm'\rangle$ of hydrogenic atoms, the transition dipole moment is given by

$$\mathbf{D}_{n'l'm',nlm} = -ea\langle n'l'm'|\rho\,\hat{\mathbf{r}}\,|nlm\rangle\,. \tag{2.60}$$

Decomposing the hydrogenic wavefunctions into the product of radial and angular states, $\psi_g(\mathbf{r}) = R_{nl}(r)Y_l^m(\hat{\mathbf{r}})$, the matrix elements of the transition dipole separate into a radial and an angular contribution,

$$\mathbf{D}_{n'l'm',nlm} = -ea\mathcal{R}_{n'l',nl}\,\langle l'm'|\,\hat{\mathbf{r}}\,|lm\rangle\,. \tag{2.61}$$

The radial contribution is given by the integral - compare with Eq. (2.45)

$$\mathcal{R}_{n'l',nl} \equiv \int_0^\infty \tilde{\chi}_{n'l'}\rho\tilde{\chi}_{nl}(\rho)d\rho. \tag{2.62}$$

Note the property

$$\mathcal{R}_{n'l',nl} = \mathcal{R}_{nl,n'l'}. \tag{2.63}$$

The angular contribution is given by

$$\langle l'm'|\,\hat{\mathbf{r}}\,|lm\rangle = \int Y_{l'}^{m'*}(\hat{\mathbf{r}})\,\hat{\mathbf{r}}\,Y_l^m(\hat{\mathbf{r}})d\hat{\mathbf{r}}. \tag{2.64}$$

Note that this integral is zero if the parity of the integrand is odd. This reveals that the electric-dipole transition requires a *change of parity* of the electronic state.

2.5.2 Angular matrix element - spherical basis

To find an explicit expression for the angular matrix element we decompose the radial unit vector $\hat{\mathbf{r}}$, as defined in Eq. (1.13a), in the *spherical basis*[3]

$$\hat{\mathbf{r}} = \sqrt{4\pi/3}\left[\hat{\mathbf{u}}_{+1}^* Y_1^1(\hat{\mathbf{r}}) + \hat{\mathbf{u}}_{-1}^* Y_1^{-1}(\hat{\mathbf{r}}) + \hat{\mathbf{u}}_0^* Y_1^0(\hat{\mathbf{r}})\right], \tag{2.65}$$

where the *spherical* unit vectors are defined by

$$\hat{\mathbf{u}}_{+1} = -\sqrt{\tfrac{1}{2}}(\hat{\mathbf{x}} + i\hat{\mathbf{y}}), \quad \hat{\mathbf{u}}_{-1} = +\sqrt{\tfrac{1}{2}}(\hat{\mathbf{x}} - i\hat{\mathbf{y}}), \quad \hat{\mathbf{u}}_0 = \hat{\mathbf{z}}. \tag{2.66}$$

---

[1] States sharing the same principal quantum number but differing in orbital angular momentum are degenerate in the case of the Schrödinger Hamiltonian. Linear combinations of such states do have a permanent electric dipole moment. This moment becomes nonstationary when the degeneracy is lifted by some perturbation.

[2] Note that $\langle \mathbf{r}|\underline{\mathbf{r}}|\mathbf{r}'\rangle^* = \mathbf{r}'\delta(\mathbf{r}' - \mathbf{r}) = \mathbf{r}\delta(\mathbf{r} - \mathbf{r}') = \langle \mathbf{r}'|\underline{\mathbf{r}}|\mathbf{r}\rangle$.

[3] This is readily verified: $\hat{\mathrm{u}}_{\pm1} \cdot \mathbf{r} = \mp\sqrt{1/2}\,[(\hat{\mathrm{x}} \cdot \mathbf{r}) \pm i(\hat{\mathrm{y}} \cdot \mathbf{r})] = \mp r\sqrt{1/2}\,e^{\pm i\phi}\sin\theta$; $\hat{\mathrm{u}}_0 \cdot \mathbf{r} = \hat{\mathrm{z}} \cdot \mathbf{r} = r\cos\theta$.

The decomposition (2.65) is called the *standard (or spherical) decomposition* of the radial unit vector $\hat{\mathbf{r}}$. Substituting Eqs. (2.66) into Eq. (2.65) we regain the *cartesian decomposition* of $\hat{\mathbf{r}}$ as given in Eq. (1.13a). The orthonormality relations for the spherical unit vectors are

$$\hat{\mathbf{u}}_i^* \cdot \hat{\mathbf{u}}_j = \delta_{ij} \tag{2.67}$$

with $i, j \in \{-1, 0, 1\}$. Note that $\hat{\mathbf{u}}_{+1} = -\hat{\mathbf{u}}_{-1}^*$, $\hat{\mathbf{u}}_{-1} = -\hat{\mathbf{u}}_{+1}^*$ and $\hat{\mathbf{u}}_0 = \hat{\mathbf{u}}_0^*$, which is summarized by the expression

$$\hat{\mathbf{u}}_q = (-1)^q \hat{\mathbf{u}}_{-q}^*. \tag{2.68}$$

The reverse transformation is given by

$$\hat{\mathbf{x}} = -\sqrt{\tfrac{1}{2}}(\hat{\mathbf{u}}_{+1} - \hat{\mathbf{u}}_{-1}), \quad \hat{\mathbf{y}} = +\sqrt{\tfrac{1}{2}}(\hat{\mathbf{u}}_{+1} + \hat{\mathbf{u}}_{-1})\, i, \quad \hat{\mathbf{z}} = \hat{\mathbf{u}}_0. \tag{2.69}$$

The angular matrix element is readily evaluated in the spherical decomposition. Using the *Gaunt integral* for the product of three spherical harmonics, see Eq. (L.59), we obtain ,

$$\begin{aligned} \langle l'm'|\, \hat{\mathbf{r}}\, |lm\rangle &= \textstyle\sum_{q=-1}^{1} \hat{\mathbf{u}}_q^* \, \langle l'm'|\, \sqrt{4\pi/3}\, Y_1^q(\hat{\mathbf{r}})\, |lm\rangle \\ &= \textstyle\sum_{q=-1}^{1} \hat{\mathbf{u}}_q^* \, (-1)^{m'+\max(l,l')} \sqrt{\max(l,l')} \begin{pmatrix} l' & 1 & l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l' & 1 & l \\ -m' & q & m \end{pmatrix} \\ &= \hat{\mathbf{u}}_{m'-m}^* \mathcal{A}_{l'm',lm}, \end{aligned} \tag{2.70}$$

where the modulus of the angular integral is given by

$$\mathcal{A}_{l'm',lm} = (-1)^{m'+\max(l,l')} \sqrt{\max(l,l')} \begin{pmatrix} l' & 1 & l \\ -m' & m'-m & m \end{pmatrix} \delta_{l',l\pm1}. \tag{2.71}$$

Note the property

$$\mathcal{A}_{l'm',lm} = (-1)^{m'-m} \mathcal{A}_{lm,l'm'}. \tag{2.72}$$

The last step in Eq. (2.70) follows from the projection rule of $3j$ symbols, $q + m - m' = 0$, which implies that *only one* of the three terms of the spherical decomposition gives a nonzero contribution to the matrix element $\langle l'm'|\, \hat{\mathbf{r}}\, |lm\rangle$. This selection rule expresses the conservation of angular momentum along the quantization axis. The factor $\delta_{l',l\pm1}$ expresses the *electric-dipole parity selection rule* for one-electron atoms,

$$\begin{pmatrix} l' & 1 & l \\ 0 & 0 & 0 \end{pmatrix} = \delta_{l',l\pm1} \quad \Rightarrow \quad l' = l \pm 1. \tag{2.73}$$

2.5.3 Transition dipole and transition strength

Thus we established that the transition dipole can be written in the general form (2.58),

$$\mathbf{D}_{n'l'm',nlm} = D_{n'l'm',nlm} \hat{\mathbf{u}}_{m'-m}^*, \tag{2.74}$$

where $\hat{\mathbf{u}}_{m'-m}^*$ is a *spherical unit vector* and

$$D_{n'l'm',nlm} = -ea\, \mathcal{R}_{n'l',nl} \mathcal{A}_{l'm',lm}. \tag{2.75}$$

The quantity

$$|\mathbf{D}_{n'l'm',nlm}|^2 = D_{n'l'm',nlm}^2 = e^2 a^2 \mathcal{R}_{n'l',nl}^2 \mathcal{A}_{l'm',lm}^2 \tag{2.76}$$

is called the $n'l'm' \leftrightarrow nlm$ (atomic-dipole) *transition strength*. Note that $D_{n'l'm',nlm}$ can be written in the form

$$D_{n'l'm',nlm} = -ea\, \langle n'l'm'|\, \sqrt{4\pi/3}\, \rho\, Y_1^{(m'-m)}(\hat{\mathbf{r}})\, |nlm\rangle\, . \tag{2.77}$$

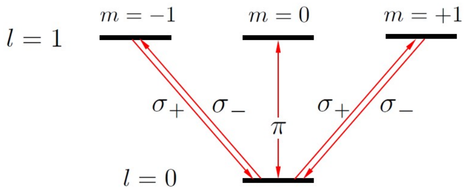


Figure 2.4: Depending on the change in magnetic quantum number of the atom, electric-dipole transitions are referred to as *sigma-plus* ($\Delta m = 1$), *sigma-minus* ($\Delta m = -1$) or *pi* ($\Delta m = 0$) transitions.

As the operator $\sqrt{4\pi/3}\, r\, Y_1^{(m'-m)}(\hat{\mathbf{r}})$ is a standard component of the irreducible vector operator $\mathbf{r}$ (cf. Section 2.5.2 and Appendix K.1.2) we can apply the Wigner-Eckart theorem (K.24),

$$D_{n'l'm',nlm} = -ea\,(-1)^{l'-m'}\langle n'l'||\sqrt{4\pi/3}\,\rho\, Y_1(\hat{\mathbf{r}})||nl\rangle \begin{pmatrix} l' & 1 & l \\ -m' & m'-m & m \end{pmatrix}. \tag{2.78}$$

Comparing Eqs. (2.78) with (2.75) and (2.71) we obtain *once and for all* the expression for the reduced matrix element in the standard representation $\{L^2, L_z\}$

$$\langle n'l'||\sqrt{4\pi/3}\,\rho\, Y_1(\hat{\mathbf{r}})||nl\rangle = \pm\sqrt{\max(l,l')}\,\delta_{l',l\pm1}\mathcal{R}_{n'l',nl}. \tag{2.79}$$

2.5.4 Selection rules for electric-dipole transitions - spin of the photon

Since $e^2a^2\mathcal{R}^2_{n'l',nl}$ is positive-definite, the transition-dipole is only non-zero if $\mathcal{A}_{l'm',lm}$ is non-zero; i.e., if the condition

$$\begin{pmatrix} l' & 1 & l \\ -m' & m'-m & m \end{pmatrix}\delta_{l',l\pm1} \neq 0 \tag{2.80}$$

is satisfied. The parity selection rule (2.73) is enforced by the Kronecker symbol $\delta_{l',l\pm1}$. Together with the properties of the $3j$ symbol Eq. (2.80) yields the *selection rules* for electric-dipole transitions

$$\Delta l = \pm1, \qquad \Delta m = 0, \pm1, \tag{2.81}$$

where $\Delta l = l' - l$ and $\Delta m = m' - m$. Transitions in which the magnetic quantum number increases (decreases) are called σ_+ (σ_-) transitions; transitions with $\Delta m = 0$ are called π transitions (see Fig. 2.4).

If the transition is induced by the electromagnetic field, the change in angular momentum of the atom is compensated by a change in angular momentum of the field (to conserve the total angular momentum). This happens in the exchange of a *photon* between atom and field. From the selection rule $\Delta l = \pm1$ we infer that the emitted (or absorbed) photon always carries a single unit of angular momentum. This angular momentum is called the *spin of the photon* ($s = 1$). From the selection rule $\Delta m = 0, \pm1$ we infer that the quantum number m_s for the *projection of the photon spin* on the quantization axis of the atom is given by $m_s = \Delta m = 0, \pm1$. The probability to observe the photon as a particle with *polarization* unit vector $\hat{\epsilon}$ is given by $|\hat{\epsilon}\cdot\hat{\mathbf{u}}^*_{m'-m}|^2$. Further discussion of optical transitions is best given after the interaction of atoms with the electromagnetic field has been introduced.

2.5.5 Examples of electric-dipole transitions in hydrogen:

2.5.5.1 Lyman transitions:

Let us calculate the radial and angular integrals for a couple of important cases. First we consider excitation from (or decay to) the electronic ground state ($np \leftrightarrow 1s$). These transitions define the

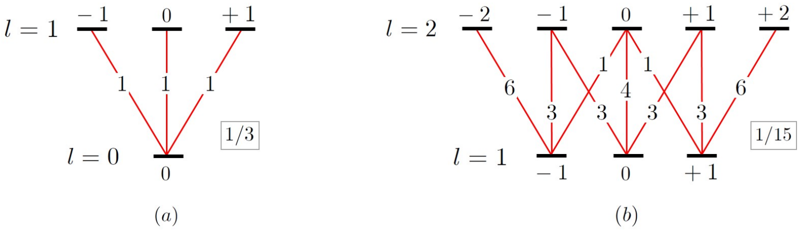


Figure 2.5: Angular contribution to the transition strength for: (*a*) *sp* transitions; (*b*) *pd* transitions. The squared angular matrix element of the reference transition (the weakest of the set) is 1/3 (for *sp*) and 1/15 (for *dp*). Note that the sum of the transition probabilities is the same for all upper (or lower) levels.

Lyman series. Setting $l = 0$ in Eq. (2.80) the selection rule for these transitions becomes

$$\begin{pmatrix} l' & 1 & 0 \\ -m' & m' & 0 \end{pmatrix} \neq 0. \tag{2.82}$$

In other words we require $l' = 1$ and $m' = -1, 0, 1$.

Substituting the expressions for $\tilde{R}_{1s}(\rho)$ and $\tilde{R}_{np}(\rho)$ from Eq. (2.37) into the radial integral (2.62) and changing to the variable $x = 2Z\rho/n$ the radial integral becomes

$$\mathcal{R}_{np,1s} \equiv \frac{1}{Z}\sqrt{\frac{(n-2)!}{(n+1)!}}\left(\frac{n}{2}\right)^2 \int x^4 e^{-(n+1)x/2} L^3_{n-2}(x)dx. \tag{2.83}$$

This integral is evaluated by repetitive use (3×) of the recursion relation (L.80a),

$$\int x^4 e^{-(n+1)x/2} L^3_{n-2}(x)dx = \sum_{n'=0}^{n-2}\sum_{n''=0}^{n'}\sum_{k=0}^{n''}\int x^4 e^{-(n+1)x/2} L_k(x)dx. \tag{2.84}$$

Note that the number of recursions required is equal to the upper index of the Laguerre polynomial. Evaluating the integral and summation*s* we obtain

$$\mathcal{R}_{np,1s} = (1/Z)2^4 n^{7/2}\frac{(n-1)^{n-5/2}}{(n+1)^{n+5/2}} \qquad n \geq 2. \tag{2.85}$$

The angular integral yields (see also Fig. 2.5*a*)

$$\mathcal{A}_{p\to s} = \mathcal{A}_{1m',00} = (-1)^{m'+1}\begin{pmatrix} 1 & 1 & 0 \\ -m' & m' & 0 \end{pmatrix} = \sqrt{\tfrac{1}{3}}. \tag{2.86}$$

In particular we find for the Lyman α transition strength ($2p_{m'} \leftrightarrow 1s$) - see Problem 2.6

$$D^2_{2p,1s} = 0.554\,93\, e^2 a^2/Z^2. \tag{2.87}$$

Note that this results holds for transitions is independent of the choice of m'.

2.5.5.2 Balmer transitions:

The second example is the Balmer series, which is defined by *excitation from* (or decay to) the first excited electronic state ($n = 2$ level).

$np \leftrightarrow 2s$ transitions $(n \geq 2)$
Substituting the expressions for $\tilde{R}_{2s}(\rho)$ and $\tilde{R}_{np}(\rho)$ from Eq. (2.37) into the radial integral (2.62) and changing to the variable $x = 2Z\rho/n$ the radial integral becomes

$$\mathcal{R}_{np,2s} \equiv \frac{1}{2Z\sqrt{2}}\sqrt{\frac{(n-2)!}{(n+1)!}}\left(\frac{n}{2}\right)^2 \int x^4(1-nx/4)e^{-(n+2)x/4}L^3_{n-2}(x)dx. \tag{2.88}$$

The integral is evaluated by repetitive use (3×) of the recursion relation (L.80a),

$$\int x^4(1-nx/4)e^{-(n+2)x/4}L^3_{n-2}(x)dx = \sum_{n'=0}^{n-2}\sum_{n''=0}^{n'}\sum_{k=0}^{n''}\int x^4(1-nx/4)e^{-(n+2)x/4}L_k(x)dx. \tag{2.89}$$

Evaluating the integral and summations we obtain

$$\mathcal{R}_{np,2s} = \begin{cases} -(1/Z)3\sqrt{3} & n=2 \\ +(1/Z)2^8\sqrt{2}\,n^{7/2}(n^2-1)^{1/2}\dfrac{(n-2)^{n-3}}{(n+2)^{n+3}} & n>2. \end{cases} \tag{2.90}$$

The angular integral yields for all *sp* transitions (see also Fig. 2.5*a*)

$$\mathcal{A}_{p\to s} = \mathcal{A}_{1m',00} = (-1)^{m'+1}\begin{pmatrix} 1 & 1 & 0 \\ -m' & m' & 0 \end{pmatrix} = \sqrt{\tfrac{1}{3}}. \tag{2.91}$$

In particular we find for $n = 3$ the $2s \leftrightarrow 3p$ contribution to the Balmer-α transition strength,

$$D^2_{3p,2s} = 3.131\, e^2a^2/Z^2. \tag{2.92}$$

$ns \leftrightarrow 2p$ transitions
Substituting the expressions for $\tilde{R}_{2p}(\rho)$ and $\tilde{R}_{ns}(\rho)$ from Eq. (2.37) into the radial integral (2.62) and changing to the variable $x = 2Z\rho/n$ the radial integral becomes

$$\mathcal{R}_{ns,2p} \equiv \frac{1}{32Z\sqrt{6}}n^{5/2}\int x^4 e^{-(n+2)x/4}L^1_{n-1}(x)dx. \tag{2.93}$$

The integral is evaluated with the aid of the recursion relation (L.80a),

$$\int x^4 e^{-(n+2)x/4}L^1_{n-1}(x)dx = \sum_{k=0}^{n-1}\int x^4 e^{-(n+2)x/4}L_k(x)dx. \tag{2.94}$$

Evaluating the integral and summations we obtain[1]

$$\mathcal{R}_{ns,2p} = \begin{cases} -(1/Z)3\sqrt{3} & n=2 \\ +(1/Z)2^7\sqrt{2/3}\,n^{9/2}\dfrac{(n-2)^{n-3}}{(n+2)^{n+3}} & n\neq 2. \end{cases} \tag{2.95}$$

The angular integral yields for all *sp* transitions (see also Fig. 2.5*a*)

$$\mathcal{A}_{00,1m} = (-1)^{m+1}\begin{pmatrix} 0 & 1 & 1 \\ 0 & -m & m \end{pmatrix} = \sqrt{\tfrac{1}{3}}. \tag{2.96}$$

Note that for $n = 1$ we regain the $1s \leftrightarrow 2p_m$ (Lyman α) transition strength - see Eq. (2.87)

$$D^2_{2p,1s} = e^2a^2\mathcal{R}^2_{1s,2p}\mathcal{A}^2_{00,1m} = 0.554\,93\, e^2a^2/Z^2. \tag{2.97}$$

---

[1] For $n = 1$ the transition is part of the Lyman series - compare with Eq. (2.85).

Likewise, we find for $n = 3$ the $3s \leftrightarrow 2p_{m'}$ Balmer-α transition strength,

$$D^2_{3s,2p_m} = e^2a^2\mathcal{R}^2_{3s,2p}\mathcal{A}^2_{00,1m} = 0.293\,53\,e^2a^2/Z^2. \tag{2.98}$$

A new aspect comes into play when dealing with incoherent sums over (initial or final) states. A well-known example is the calculation of spontaneous emission rates from a given magnetic sublevel. This involves a sum over the final states contributions. Summing over the magnetic sublevels of the p level we find

$$\mathcal{A}^2_{s\to p} \equiv \sum_{m=-1}^{1} \mathcal{A}^2_{00,1m} = 1. \tag{2.99}$$

Using this sum rule we obtain for the total transition probability

$$D^2_{3s,2p} = e^2a^2\mathcal{R}^2_{3s,2p}\sum_{m=-1}^{1}\mathcal{A}^2_{00,1m} = 0.880\,60\,e^2a^2/Z^2. \tag{2.100}$$

$nd \leftrightarrow 2p$ transitions
Substituting the expressions for $\tilde{R}_{2p}(\rho)$ and $\tilde{R}_{nd}(\rho)$ from Eq. (2.37) into the radial integral (2.62) and changing to the variable $x = 2Z\rho/n$ the radial integral becomes

$$\mathcal{R}_{nd,2p} \equiv (1/Z)\frac{1}{32\sqrt{6}}\sqrt{\frac{(n-3)!}{(n+2)!}}n^3\int x^6e^{-(n+2)x/4}L^5_{n-3}(x)dx. \tag{2.101}$$

The integral is evaluated by repetitive use (5×) of the recursion relation (L.80a)

$$\int x^6e^{-(n+2)x/4}L^5_{n-3}(x)dx = \sum_{n'=0}^{n-3}\sum_{n''=0}^{n'}\sum_{n'''=0}^{n''}\sum_{n''''=0}^{n'''}\sum_{k=0}^{n''''}\int x^6e^{-(n+2)x/4}L_k(x)dx. \tag{2.102}$$

Evaluating the integral and summation we obtain

$$\mathcal{R}_{nd,2p} = (1/Z)2^9\sqrt{2/3}\,n^5(n^2-1)\frac{(n-2)^{n-3}}{(n+2)^{n+3}}\sqrt{\frac{(n-3)!}{(n+2)!}} \qquad n > 2. \tag{2.103}$$

The angular integral yields for $d \leftrightarrow p$ transitions

$$\mathcal{A}_{2m',10} = (-1)^{-m'}\sqrt{2}\begin{pmatrix}1 & 1 & 2\\ 0 & m' & -m'\end{pmatrix} = (-1)^{-m'}\sqrt{\frac{4-m'^2}{15}}, \tag{2.104a}$$

$$\mathcal{A}_{2m',1\pm1} = (-1)^{-m'}\sqrt{2}\begin{pmatrix}1 & 1 & 2\\ \pm1 & m'\mp1 & m'\end{pmatrix} = (-1)^{-m'}\begin{cases}\sqrt{\frac{2+3m'+m'^2}{30}} & 0 \le m' \le 2\\ \sqrt{\frac{2-3m'+m'^2}{30}} & -2 \le m' \le 0.\end{cases} \tag{2.104b}$$

For the case $n = 3$ we find for the $3d_{m'} \leftrightarrow 2p_m$ Balmer-α transition strengths - see Problem 2.6

$$D^2_{3dm',2pm} = e^2a^2\mathcal{R}^2_{3d,2p}\mathcal{A}^2_{2m',1m} = 22.543\,\mathcal{A}^2_{2m',1m}e^2a^2/Z^2. \tag{2.105}$$

Summing over the magnetic quantum numbers of the p or d level we obtain the sum rules

$$\sum_{m=-1}^{1}\mathcal{A}^2_{2m',1m} = \sum_{m=-2}^{2}\mathcal{A}^2_{2m,1m'} = \tfrac{2}{5}. \tag{2.106}$$

Interestingly, this sum is *independent* of the choice of the initial magnetic sublevel m' as is illustrated in Fig. 2.5. We return to the use of sum rules when discussing electric-dipole transitions in the presence of fine structure (Section 4.6.4) and hyperfine structure (Section 5.3.7). For the case $n = 3$ we find for the $3d_m \to 2p$ and $2p_m \to 3d$ Balmer-α transition probabilities

$$D^2_{3d_m,2p} = D^2_{3d,2p_m} = 9.0172\,e^2a^2/Z^2. \tag{2.107}$$

Problem 2.6. Verify the general expressions for $1s \leftrightarrow np$, $2s \leftrightarrow np$ and $2p \leftrightarrow nd$ transitions by direct substitution of Eqs. (2.40) into Eq. (2.62) for the $1s \leftrightarrow 2p$ (*Lyman* α), the $2s \leftrightarrow 2p$ and $2p \leftrightarrow 3d$ transitions in hydrogen.

Solution. By direct substitution of Eqs. (2.40) into Eq. (2.62) we calculate

$$\begin{aligned}
\mathcal{R}_{2p,1s} &\equiv \int_0^\infty \rho^3 \tilde{R}_{2p}(\rho)\tilde{R}_{1s}(\rho)d\rho = +1.29027/Z \\
\mathcal{R}_{2p,2s} &\equiv \int_0^\infty \rho^3 \tilde{R}_{2p}(\rho)\tilde{R}_{2s}(\rho)d\rho = -5.19616/Z \\
\mathcal{R}_{2p,3d} &\equiv \int_0^\infty \rho^3 \tilde{R}_{3d}(\rho)\tilde{R}_{2p}(\rho)d\rho = +4.74799/Z.
\end{aligned}$$

These values are used in the formulas for $1s \leftrightarrow np$, $2s \leftrightarrow np$ and $2p \leftrightarrow nd$ transitions. □

3

Angular Momentum

In Chapter 1 we solved the Schrödinger equation for the motion of a particle in a central potential field. Exploiting the central symmetry we found separate equations of motion as a function of r, θ and ϕ and established that the radial and angular motions are quantized. For the angular motion this was expressed by the eigenvalue equations for the operators $\mathbf{L}^2$ and L_z. In this second chapter on the mathematics of the atom we develop another line of reasoning. We start by demonstrating that the properties of quantized angular momentum follow directly from the commutation relations (1.29). In particular, the eigenvalues of $\mathbf{L}^2$ and L_z and the properties of L_+ and L_- are obtained *without* solving the Schrödinger equation. This leads us to generalize the discussion: whenever we meet a vector operator $\mathbf{J}$ of which the cartesian components, J_x, J_y and J_z, are hermitian operators satisfying commutation relations of the type (1.29), the quantization properties of its observables are immediately known. We shall find that these commutation rules define an algebra that not only reproduces the properties obtained in Chapter 1 but also allows for different (non-classical) kinds of angular momentum. In Section 3.2 we introduce the matrix representation for angular momentum operators in vector spaces of arbitrary dimension. For the two-dimensional case we obtain the Pauli matrices, identifying spin as a $s = \frac{1}{2}$ form of angular momentum. In Section 3.3 we turn to the measurement of angular momentum. We introduce the vector model and the concept of polarization; i.e., the expectation value of the atomic angular momentum averaged over an ensemble of atoms. In Section 3.4 we introduce the vector addition of two angular momenta to form a total angular momentum. We define the coupled and the uncoupled representation and introduce the Clebsch-Gordan transformation between these two. We are free to chose one of these representation but as soon we have coupling between the angular momenta this choice is leveraged by conservation rules (see Section 3.5). In Section 3.6 we investigate the relation between rotations in real space and unitary transformations in Hilbert spaces. We find for the spin $s = \frac{1}{2}$ case that the angular momentum operators S_x, S_y and S_z can be written as differential operators representing infinitesimal rotations about the x, y and z axis, respectively. We arrive in Section 3.7 at the *formal definition* of the angular momentum operator J_a as an infinitesimal rotation about the direction $\hat{\mathbf{a}}$. Without the above justification this is the starting for many applications of angular momentum theory in the literature as it implies the proper commutation relations and with this all properties of quantized angular momentum. We find that half-integral angular momenta have rotation properties that do not exist in classical physics.

Introductions in the theory of quantized angular momentum are given by Albert Messiah [75, 76] and in dedicated books on *Angular Momentum* by Morris E. Rose [94] and A.R. Edmonds [38]. The classics by Ugo Fano and Giulio Racah on *Irreducible Tensorial Sets* [40] and by Eugene P. Wigner on *Group Theory* [120] put emphasis on the mathematical basis of the theory.

3.1 Angular momentum algebra

In this section we demonstrate that the vector operator $\mathbf{J}$ has the properties of an *angular momentum operator* if its cartesian components (J_x, J_y and J_z in some frame of reference $\mathcal{S}$) are *hermitian* operators satisfying the commutation relations

$$[J_x, J_y] = i\hbar J_z, \;\; [J_y, J_z] = i\hbar J_x \;\text{ and }\; [J_z, J_x] = i\hbar J_y. \tag{3.1}$$

Using the *inner product rule*

$$\mathbf{J}^2 = J_x^2 + J_y^2 + J_z^2, \tag{3.2}$$

it is straightforward to show that $\mathbf{J}^2$ commutes with J_z - see Eq. (1.31). Therefore, $\mathbf{J}^2$ and J_z share a complete set of eigenstates (see Problem F.1). Adopting the Dirac notation we denote this joint basis by $\{|\lambda, m\rangle\}$ and the corresponding eigenvalue equations take the form

$$\mathbf{J}^2 |\lambda, m\rangle = \lambda\hbar^2 |\lambda, m\rangle \;\text{ and }\; J_z |\lambda, m\rangle = m\hbar |\lambda, m\rangle\,, \tag{3.3}$$

where the eigenvalues λ and m are *real* numbers still to be determined. Note that, equally well we could have chosen the joint basis of $\mathbf{J}^2$ and J_x (or $\mathbf{J}^2$ and J_y). By selecting $\mathbf{J}^2$ and J_z we adopted the z axis of our coordinate system $\mathcal{S}$ as the *quantization axis* in the real space of observation; the basis $\{|\lambda, m\rangle\}$ that diagonalizes $\mathbf{J}^2$ and J_z simultaneously defines the *standard representation* $\{\mathbf{J}^2, J_z\}$.

Like in the case of orbital angular momentum we proceed by introducing *shift operators*

$$J_\pm \equiv J_x \pm iJ_y. \tag{3.4}$$

For convenience of reference we also give the inverse relations

$$J_x = \tfrac{1}{2}(J_+ + J_-) \quad \text{and} \quad J_y = -i\tfrac{1}{2}(J_+ - J_-). \tag{3.5}$$

Note that J_+ and J_- are *hermitian conjugates*, as follows from the hermiticity of J_x and J_y,

$$\begin{aligned}\langle\chi'| J_+ |\chi\rangle &= \langle\chi'| J_x |\chi\rangle + \langle\chi'| iJ_y |\chi\rangle \\ &= \langle\chi| J_x |\chi'\rangle^* - i\,\langle\chi| J_y |\chi'\rangle^* = \langle\chi| J_- |\chi'\rangle^*\,,\end{aligned} \tag{3.6}$$

where $|\chi\rangle$ and $|\chi'\rangle$ represent arbitrary angular momentum states. With the aid of the commutation relations (3.1) it is straightforward to derive commutation relations for the shift operators - compare Eqs. (1.36) and (1.38),

$$[J_z, J_\pm] = \pm\hbar J_\pm \text{ and } [J_+, J_-] = 2\hbar J_z. \tag{3.7}$$

Furthermore, we can derive the following operator identities - cf. Eqs. (1.37),

$$J_-J_+ = \mathbf{J}^2 - J_z^2 - \hbar J_z \tag{3.8a}$$

$$J_+J_- = \mathbf{J}^2 - J_z^2 + \hbar J_z. \tag{3.8b}$$

Adding these equations we obtain the *inner product rule*

$$\mathbf{J}^2 = J_z^2 + \tfrac{1}{2}[J_+J_- + J_-J_+]. \tag{3.9}$$

The shift operators are introduced because they enable us to construct the subspace V_λ of the angular momentum Hilbert space corresponding to the eigenvalue λ. To elucidate this point we note that J_+ is an operator that *raises* the eigenvalue $m\hbar$ by one unit of angular momentum about the quantization axis; i.e., if $|\lambda, m\rangle$ is an eigenstate of J_z with the eigenvalue $m\hbar$ also $J_+ |\lambda, m\rangle$ will be an eigenstate of J_z but with eigenvalue $(m+1)\hbar$,

$$J_zJ_+ |\lambda, m\rangle = (J_+m\hbar + \hbar J_+) |\lambda, m\rangle = (m+1)\hbar J_+ |\lambda, m\rangle\,. \tag{3.10}$$

This follows directly from the commutation relations (3.7). Comparing Eq. (3.10) with the generic eigenvalue relation (3.3) we find

$$J_+ \left|\lambda, m\right\rangle = c_+ \left(\lambda, m\right) \hbar \left|\lambda, m+1\right\rangle , \tag{3.11}$$

where $c_+ \left(\lambda, m\right)$ is a (generally complex) constant to be determined. Likewise, we find that J_- is an operator that *lowers* the eigenvalue by $\hbar$,

$$J_- \left|\lambda, m\right\rangle = c_- \left(\lambda, m\right) \hbar \left|\lambda, m-1\right\rangle . \tag{3.12}$$

Thus we established that for a given eigenvalue λ the operators $J_\pm$ act as *construction operators* by which new angular momentum eigenstates of the subspace V_λ can be generated.

We are now in a position to determine the constants $c_\pm \left(\lambda, m\right)$. As we shall see this leads us to the quantized spectrum of λ and m. First we derive a relation between c_+ and c_- using the property that J_+ and J_- are hermitian conjugates,

$$c_+ \left(\lambda, m\right) \equiv \left\langle j, m+1\right| \left(J_+/\hbar\right) \left|j, m\right\rangle = \left\langle j, m\right| \left(J_-/\hbar\right) \left|j, m+1\right\rangle^* = c_-^* \left(\lambda, m+1\right) . \tag{3.13}$$

We proceed by deriving two expressions for the expectation value $\left\langle\lambda, m\right| J_- J_+ \left|\lambda, m\right\rangle$. Using Eqs. (3.11), (3.12) and (3.13) we obtain

$$\left\langle\lambda, m\right| J_- J_+ \left|\lambda, m\right\rangle = c_- \left(\lambda, m+1\right) c_+ \left(\lambda, m\right) \hbar^2 = \left|c_+ \left(\lambda, m\right)\right|^2 \hbar^2. \tag{3.14}$$

On the other hand from Eq. (3.8a) it follows that

$$\left\langle\lambda, m\right| J_- J_+ \left|\lambda, m\right\rangle = \left\langle\lambda, m\right| \mathbf{J}^2 - J_z^2 - \hbar J_z \left|\lambda, m\right\rangle = \left[\lambda - m(m+1)\right] \hbar^2. \tag{3.15}$$

Equating Eqs. (3.14) and (3.15) we find the condition

$$0 \leq \left|c_+ \left(\lambda, m\right)\right|^2 = \lambda - m(m+1). \tag{3.16}$$

Repeating the derivation starting from $J_+ J_-$ we obtain in a similar way

$$0 \leq \left|c_- \left(\lambda, m\right)\right|^2 = \lambda - m(m-1). \tag{3.17}$$

To assure that the conditions (3.16) and (3.17) are not violated, the construction of the subspace V_λ by the operators $J_\pm$ has to be terminated at some point; i.e., we require the existence of a maximum and a minimum value of m, $-\bar{j} \leq m \leq j$. The only way in which these bounds follow logically from the algebra is if the conditions $c_+ \left(\lambda, j\right) = 0$ and $c_- \left(\lambda, -\bar{j}\right) = 0$ are simultaneously satisfied. Together with Eqs. (3.16) and (3.17) this leads to the condition

$$\lambda = j(j+1) = \bar{j}(\bar{j}+1). \tag{3.18}$$

Hence, we find $\bar{j} = j$ and since m has to vary in integral steps, the condition $-j \leq m \leq j$ implies that j and m have to be either both *integers* $(0, 1, 2, \cdots)$ or *half-integers* $(\frac{1}{2}, \frac{3}{2}, \cdots)$. Other values of j and m are not allowed because the construction procedure would not be truncated by the algebra and result in violation of the conditions (3.16) and (3.17). For the coefficients we find the relation

$$\left|c_\pm \left(j(j+1), m\right)\right|^2 = j(j+1) - m(m \pm 1). \tag{3.19}$$

Simplifying the notation, $c_\pm \left(j(j+1), m\right) \rightarrow C_{j,\pm m}$, and choosing the phase of $C_{j,m}$ such that it always represents a *non-negative real* number we obtain the coefficients in the *Condon and Shortley phase convention* [28],

$$C_{j,m} \equiv \sqrt{j(j+1) - m(m+1)} = \sqrt{(j-m)\,(j+m+1)}. \tag{3.20}$$

This convention represents a generalization of the phase convention introduced in Section 1.1.6.1 to the case of arbitrary angular momenta. In view of Eqs. (3.11) and (3.12) this convention assures that the basis states $\{\left|j, m\right\rangle\}$ are constructed with the same relative phase.

3.1.0.1 Summary:

A vector operator $\mathbf{J}$ is called *angular momentum operator* if its cartesian components (J_x, J_y and J_z in some frame of reference $\mathcal{S}$) are *hermitian* operators satisfying the following commutation relations

$$[J_i, J_j] = i\hbar\varepsilon_{ijk}J_k, \tag{3.21}$$

with $i,j,k \in \{x,y,z\}$. Choosing the quantization axis along the z direction of a cartesian coordinate system in the euclidean space of observation, the eigenvalue equations of an arbitrary angular momentum operator $\mathbf{J}$ can be written in the form,

$$\mathbf{J}^2 \left|j,m\right\rangle = j(j+1)\hbar^2 \left|j,m\right\rangle \tag{3.22a}$$

$$J_z \left|j,m\right\rangle = m\hbar \left|j,m\right\rangle . \tag{3.22b}$$

The quantum numbers j and m are called *rotational quantum numbers.* They are either both integral or half-integral. The quantum number m is often referred to as the *magnetic quantum number* and is restricted to the interval

$$-j \le m \le j. \tag{3.23}$$

The kets $|j,m\rangle$ correspond to the basis vectors in the *standard representation* $\{\mathbf{J}^2, J_z\}$ of a $d = 2j+1$ dimensional subspace V^d of Hilbert space, with $d = 1,2,3,\cdots$. Using the Condon and Shortley phase convention the shift relations (3.11) and (3.12) become

$$J_\pm \left|j,m\right\rangle = C_{j,\pm m}\hbar \left|j,m\pm 1\right\rangle \tag{3.24a}$$

where the *shift coefficients* $C_{j,\pm m}$ are *non-negative* and *real*,

$$C_{j,\pm m} \equiv \sqrt{j\,(j+1) - m(m\pm 1)}, \tag{3.24b}$$

with the *symmetry properties*

$$C_{j,\pm m} = C_{j,\mp m-1}. \tag{3.24c}$$

In this convention all sublevels $|j,m\rangle$ can be constructed from $|j,j\rangle$ by repeated action of the lowering operator,

$$\left|j,m\right\rangle = \sqrt{\frac{(j+m)!}{(2j)!\,(j-m)!}}\left(\frac{J_-}{\hbar}\right)^{j-m}\left|j,j\right\rangle . \tag{3.24d}$$

This expression is readily verified by induction. Note that the shift operators conserve the phase. Therefore, for *given* j, all $|j,m\rangle$ carry the *same phase.* Furthermore, the shift operators satisfy the following commutation relations:

$$[J_\pm, J_\pm] = 0, \quad [J_\pm, J_\mp] = \pm 2\hbar J_z, \quad [J_z, J_\pm] = \pm\hbar J_\pm. \tag{3.25}$$

3.1.1 Shift operators versus standard components of vector operators

With respect to the *cartesian basis* the angular momentum operator $\mathbf{J}$ can be written as

$$\mathbf{J} = \hat{\mathbf{x}}J_x + \hat{\mathbf{y}}J_y + \hat{\mathbf{z}}J_z, \tag{3.26}$$

where J_x, J_y, J_z are the cartesian components. An inconvenience of the cartesian decomposition is that the properties of the angular momentum are best accessed by the operators J_z, J_+ and J_-. Thus we are lead to re-express J_x and J_y in terms of the shift operators. Substituting Eqs. (3.5) into (3.26) we find

$$\mathbf{J} = \tfrac{1}{2}\left(\hat{\mathbf{x}} - i\hat{\mathbf{y}}\right)J_+ + \tfrac{1}{2}\left(\hat{\mathbf{x}} + i\hat{\mathbf{y}}\right)J_- + \hat{\mathbf{z}}J_z. \tag{3.27}$$

Here we recognize the *spherical* unit vectors (2.66),

$$\mathbf{J} = -\sqrt{\tfrac{1}{2}}\hat{\mathbf{u}}_{+1}^{*} J_{+} + \hat{\mathbf{u}}_{0}^{*} J_{z} + \sqrt{\tfrac{1}{2}}\hat{\mathbf{u}}_{-1}^{*} J_{-}. \tag{3.28}$$

Rewriting this expression in the form

$$\mathbf{J} = \hat{\mathbf{u}}_{+1}^{*} J_{+1} + \hat{\mathbf{u}}_{0}^{*} J_{0} + \hat{\mathbf{u}}_{-1}^{*} J_{-1} \tag{3.29}$$

we obtain the *standard (or spherical) decomposition* of $\mathbf{J}$. The *standard components* J_{+1}, J_0, J_{-1} are defined by [1]

$$J_0 = J_z; \qquad J_{\pm 1} = \mp\sqrt{\tfrac{1}{2}}(J_x \pm iJ_y) = \mp\sqrt{\tfrac{1}{2}}J_{\pm}. \tag{3.30}$$

As the standard decomposition of $\mathbf{J}$ maps one-to-one onto the standard decomposition of the radial unit vector $\hat{\mathbf{r}}$, see Eq. (2.65), the standard components J_m have the same transformation properties as the $Y_1^m(\hat{\mathbf{r}})$,

$$J_m = \hat{\mathbf{u}}_m \cdot \mathbf{J} \sim \hat{\mathbf{u}}_m \cdot \hat{\mathbf{r}} = Y_1^m(\hat{\mathbf{r}}). \tag{3.31}$$

With the standard decomposition we implicitly adopt the *Condon and Shortley* phase convention. More generally, operators that transform like the $Y_l^m(\hat{\mathbf{r}})$ are called *spherical tensor* operators of *rank* $l = 1$ [93]. Tensor operators of rank 1 are called *vector* operators, those of rank $l = 0$ *scalar* operators. The standard decomposition can be generalized to hold for spherical tensor operators of any rank and is used in unified approaches for the calculation of matrix elements in systems with angular momentum (see Appendix K). We return to the transformation properties of $\mathbf{J}$ in Section 3.6.

3.2 Matrix representation of angular momentum

The dimension of a $d = 2j + 1$ dimensional subspace V^d becomes most explicit in the matrix notation, where an arbitrary state of angular momentum is represented by a normalized column vector of dimension d and the angular momentum operators by $d \times d$ matrices. In the case of *orbital* angular momentum this dimension is an *odd* integer because the quantum number j is an integer. In general, the dimension of the vector space can be *odd* (integral angular momentum) or *even* (half-integral angular momentum). In the present section we shall determine some of these matrices starting from the eigenvalue equations (3.22). In Section 3.7 they will be rederived using a formal generating procedure.

3.2.1 Example: the case $l = 1$ - spherical basis

First we discuss an example of integral angular momentum. We shall use the *spherical* basis $\{|l, m_l\rangle\}$, well known from the eigenstates of orbital angular momentum. In the matrix representation the states are represented by column vectors of *odd* dimension $d = 2l + 1$. The example is given for the case $l = 1$ where the basis vectors correspond to $m_l \in \{-1, 0, 1\}$ and are given by

$$|1,1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |1,0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |1,-1\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{3.32}$$

The angular momentum operators L_a, with $a \in \{\pm, z\}$, are represented by 3×3 matrices of elements $\langle l', m_l' | L_a | l, m_l \rangle$, which can be determined with the aid of Eqs. (3.22b) and (3.24). Using Eq. (3.22b) we find for the L_z operator

$$L_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{3.33}$$

---

[1] Beware of the subtle difference between $J_{\pm 1}$ and $J_{\pm}$

As the L_z operator is diagonal in this representation, the spherical representation is called the *diagonal representation.* Using the same approach we find with Eqs. (3.24) for the shift operators

$$L_+ = \sqrt{2}\,\hbar \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad L_- = \sqrt{2}\,\hbar \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{3.34}$$

The expressions for L_x and L_y follow from the definitions of the shift operators,

$$L_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad L_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \tag{3.35}$$

It is readily verified that these 3×3 matrices indeed satisfy the commutation relations (3.1), (1.36) and (1.38). Note that the operators L_u are *traceless*, $\operatorname{tr} L_u = 0$, with $u \in \{x, y, z\}$, where the trace of the matrix is defined by Eq. (M.5). Furthermore, it is straightforward to show with Eq. (1.39) that $\mathbf{L}^2 = 2\hbar^2 \mathbb{1}$, where $\mathbb{1}$ is the unit matrix. Hence, also Eq. (1.57a) is satisfied; note that $\mathbf{L}^2$ and L_z are both diagonal.

3.2.2 Example: the case $l = 1$ - cartesian basis

Of course we are free to choose any set of three orthogonal unit vectors to span the $l = 1$ angular momentum Hilbert space. For the case $l = 1$ the *cartesian* basis deserves our attention,[1]

$$|x\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |y\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |z\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

This basis is obtained from spherical basis by the same unitary transformation as the one that separates the spherical harmonic $Y_1^m(\hat{\mathbf{r}})$ into its real and imaginary part - see also Eq. (L.55)

$$|x\rangle \equiv -\frac{1}{\sqrt{2}}(|1,1\rangle - |1,-1\rangle), \qquad |y\rangle \equiv \frac{i}{\sqrt{2}}(|1,1\rangle + |1,-1\rangle), \qquad |z\rangle \equiv |1,0\rangle. \tag{3.36}$$

Note that these vectors are normalized and orthogonal. The angular momentum operators L_a, with $a \in \{\pm, z\}$, are represented by 3×3 matrices of elements $\langle u'|L_a|u\rangle$, with $u, u' \in \{|x\rangle, |y\rangle, |z\rangle\}$. Using Eq. (3.22b) we find for the L_z operator:

$$L_z|x\rangle = i|y\rangle, \quad L_z|y\rangle = -i|x\rangle, \quad L_z|z\rangle = 0. \tag{3.37}$$

Having these relations it is straightforward to determine L_z operator in the cartesian representation

$$L_z = \hbar \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.38}$$

Note that the representation is no longer diagonal. Likewise, using Eqs. (3.24) we find the following properties for the shift operators

$$L_+|x\rangle = +|z\rangle, \quad L_+|y\rangle = i|z\rangle, \quad L_+|z\rangle = \sqrt{2}|1,+1\rangle = -|x\rangle - i|y\rangle \tag{3.39a}$$

$$L_-|x\rangle = -|z\rangle, \quad L_-|y\rangle = i|z\rangle, \quad L_-|z\rangle = \sqrt{2}|1,-1\rangle = +|x\rangle - i|y\rangle. \tag{3.39b}$$

---

[1] Beware that these three unit vectors represent a basis of Hilbert space and *have nothing to do* with the cartesian basis of the real space (e.g. the laboratory-fixed frame) in which the angular momentum is observed.

From these relations we find for the shift operators in the cartesian representation

$$L_+ = \hbar \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & -i \\ 1 & i & 0 \end{pmatrix}, \quad L_- = \hbar \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -i \\ -1 & i & 0 \end{pmatrix}. \tag{3.40}$$

Using the definitions of the shift operators we obtain the matrix representations of L_x and L_y,

$$L_x = \hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad L_y = \hbar \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}. \tag{3.41}$$

It is readily verified that these 3×3 matrices indeed satisfy the commutation relations (3.1), (1.36) and (1.38). Furthermore, using Eq. (1.32) it is straightforward to demonstrate the operator identity $\mathbf{L}^2 = 2\hbar^2 \mathbb{1}$. Hence, also Eq. (1.57a) is satisfied. Note that the matrices L_u are again *traceless*, $\operatorname{tr} L_u = 0$, with $u \in \{x, y, z\}$, as they should be because the trace of a matrix is invariant under unitary transformation.

3.2.3 Example: the case $s = 1/2$ - Pauli spin matrices

For the case of half-integral angular momentum we shall use (in this example) the notation $|s, m_s\rangle$. In the matrix representation the states are represented by column vectors of *even* dimension $d = 2s+1$. In particular, for the case $s = \frac{1}{2}$ the basis vectors correspond to $m_s \in \{-\frac{1}{2}, \frac{1}{2}\}$. We shall meet this case when discussing the intrinsic angular momentum of the electron, the electron spin $\mathbf{S}$. It is good to emphasize already at this point that the case $s = \frac{1}{2}$ is of more general importance because it can be used to describe any quantum mechanical two-level system. The basis vectors are[1]

$$|+\rangle = |\uparrow\rangle \equiv |\tfrac{1}{2}, \tfrac{1}{2}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |-\rangle = |\downarrow\rangle \equiv |\tfrac{1}{2}, -\tfrac{1}{2}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.42}$$

The operators S_+, S_- and S_z are given by 2×2 matrices, which are easy to determine using Eqs. (3.22b) and (3.24) in accordance with the phase convention (3.20)

$$S_z = \hbar \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}, \quad S_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad S_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{3.43}$$

Defining

$$\mathbf{S} = \tfrac{1}{2}\hbar \boldsymbol{\sigma}, \tag{3.44}$$

the corresponding matrices for the cartesian components of $\boldsymbol{\sigma}$ are given by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.45}$$

These matrices are called the *Pauli spin matrices.* They are *traceless*, $\operatorname{tr} \sigma_u = 0$, *idempotent*, $\sigma_u^2 = \mathbb{1}$, have *negative determinant*, $\det \sigma_u = -1$, and satisfy the relation

$$\sigma_u \sigma_v = \mathbb{1}\, \delta_{uv} + i\varepsilon_{uvw}\sigma_w, \text{ with } u, v, w \in \{x, y, z)\}. \tag{3.46}$$

This combination is unique for $s = \frac{1}{2}$. We can also decompose the Pauli matrix $\boldsymbol{\sigma}$ along a quantization axis in the arbitrary direction $\hat{\mathbf{r}} = (\theta, \phi) = (x, y, z)$,

$$\sigma_r \equiv \hat{\mathbf{r}} \cdot \boldsymbol{\sigma} = x\sigma_x + y\sigma_y + z\sigma_z, \tag{3.47}$$

---

[1] In quantum information science one uses the notation $|0\rangle \equiv |\uparrow\rangle$ and $|1\rangle \equiv |\downarrow\rangle$ for the basis vectors that define the possible superposition states of the *qubit.*

where $x = \sin\theta\cos\phi$, $y = \sin\theta\sin\phi$ and $z = \cos\theta$ are the cartesian components of the unit vector $\hat{\mathbf{r}}$, cf. Eq. (1.13a). The operator σ_r is said to measure the angular momentum of the $s = \frac{1}{2}$ system in the direction $\hat{\mathbf{r}}$. Note that with the substitution $\hat{\mathbf{r}} \to \hat{\mathbf{z}}$ we have $x = y = 0$ and $z = 1$, regaining σ_z as we should. In terms of the spherical components σ_+, σ_- and σ_z the component σ_r can be written in the form

$$\sigma_r = \tfrac{1}{2}(x - iy)\sigma_+ + \tfrac{1}{2}(x + iy)\sigma_- + z\sigma_z. \tag{3.48}$$

In matrix notation σ_r becomes

$$\sigma_r = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta \end{pmatrix}, \tag{3.49}$$

as follows immediately by substitution of the Pauli matrices into Eq. (3.47). Note that

$$\sigma_r^2 = \mathbb{1}. \tag{3.50}$$

This does not come as a surprise because by choosing the quantization axis along the direction $\hat{\mathbf{r}}$ we have $\sigma_r \to \sigma_z$ and Eq. (3.46). A valuable operator identity is (see Problem 3.1)

$$(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B}) = \mathbf{A}\cdot\mathbf{B} + i\,\boldsymbol{\sigma}\boldsymbol{\cdot}(\mathbf{A}\times\mathbf{B}), \tag{3.51}$$

where $\mathbf{A}$ and $\mathbf{B}$ are vector operators that commute with $\boldsymbol{\sigma}$.

Problem 3.1. Prove the relation (3.51)

Solution. Since $\mathbf{A}$ and $\mathbf{B}$ commute with $\boldsymbol{\sigma}$ we have in Einstein notation

$$(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B}) = \sigma_u\sigma_v A_u B_v.$$

Using the property (3.46) this becomes

$$(\boldsymbol{\sigma}\cdot\mathbf{A})(\boldsymbol{\sigma}\cdot\mathbf{B}) = \mathbb{1}\delta_{uv}A_uB_v + i\sigma_w\varepsilon_{uvw}A_uB_v = \mathbb{1}A_uB_u + i\sigma_w\left(\mathbf{A}\times\mathbf{B}\right)_w,$$

which can be rewritten in the desired form. □

3.3 Vector model and polarization

In this section we elaborate on the measurement of angular momentum. We already established that the angular momentum of a physical system is a vector quantity, which, in a cartesian frame of reference, can be written in the form

$$\mathbf{J} = \hat{\mathbf{x}}J_x + \hat{\mathbf{y}}J_y + \hat{\mathbf{z}}J_z. \tag{3.52}$$

In classical physics the state of angular momentum can be determined to arbitrary precision by measuring the magnitude, direction and sense of the vector $\mathbf{J}$. For closed mechanical systems this quantity is conserved in time (cf. Appendix D.5.4). In quantum mechanics, the best we can do is prepare the system in an eigenstate, $|j, m\rangle$, where j defines the magnitude and m the projection (including sense) of the angular momentum with respect to the quantization axis. The *direction* remains undetermined in this process. Formally, it makes no sense to ask for the direction because the components of $\mathbf{J}$ do not commute. This makes J_x, J_y and J_z into *incompatible* operators, which means that (by lack of a common basis) the components of the vector *operator* $\mathbf{J}$ cannot be determined simultaneously to arbitrary precision (cf. Appendix F.1.1).

A way to visualize the difference between angular momentum in classical physics and in quantum mechanics is the semi-classical *vector model* illustrated in Fig. 3.1*a*. In this model, the angular momentum is represented by a *semi-classical vector* of *quantized magnitude*, $\sqrt{\langle\mathbf{J}^2\rangle} = \sqrt{j(j+1)}\hbar$, and *quantized projection* on the z axis, $\langle J_z\rangle = m\hbar$. Importantly, the same projection is obtained for

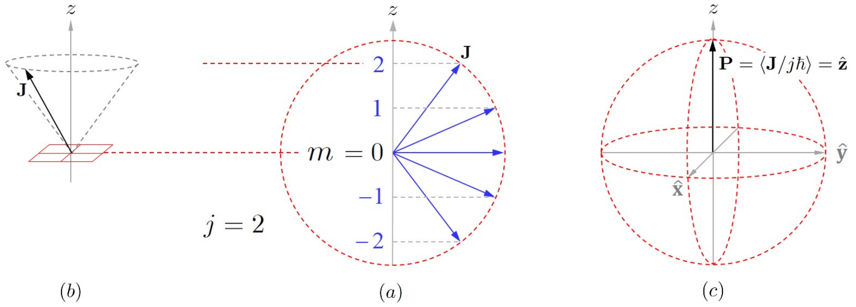


Figure 3.1: (a) Vector model for a system with angular momentum **J** for the case $j=2$; (b) Preparing the system in the state $|2,2\rangle$ the direction of the vector **J** is restricted to a cone of possible values, all sharing the same *quantized* projection onto the z axis. The quantum numbers j and m are conserved - the direction on the cone is uncertain. (c) The polarization of the state $|2,2\rangle$ is given by $\mathbf{P}=\langle\mathbf{J}/j\hbar\rangle=\hat{\mathbf{z}}$.

all vectors on the cone of possibilities shown in Fig. 3.1*b*. This indicates that, whereas any vector on the cone corresponds to a *classically* distinguishable state, quantum mechanically the direction on the cone is undetermined (all directions are equally probable). One may argue that it makes little sense to draw a vector for a direction that cannot be determined. However, as we shall see later, the *operator* **J** transforms under rotation exactly like the classical angular momentum vector. Although the direction itself has no observable meaning, changes in the direction (e.g., caused by *precession* of **J** about the quantization axis) can be (and have been) observed experimentally (cf. Section 3.7.2). Moreover, we shall find that also the addition of angular momentum operators proceeds as in the case of classical vectors (under the constraint of quantization - see Section 3.4.2). All this being said, vector diagrams offer a valuable geometric tool for visualizing the addition and rotation of angular momenta under the constraint of quantization.

In a typical (state-selective) measurement, the state in which a system is prepared is destroyed by projection on the state of observation. So, to gain more information about the state of a system, the best we can do with a state-selective detector (a detector that discriminates between the eigenstates) is repeat the measurement with identically prepared systems until the desired precision is reached. In this way we can determine the expectation values $\langle J_x\rangle$, $\langle J_y\rangle$ and $\langle J_z\rangle$, which define the components of the *polarization vector* of the angular momentum state prepared,

$$\mathbf{P}=\langle\mathbf{J}/j\hbar\rangle\equiv(\langle J_x/j\hbar\rangle,\langle J_y/j\hbar\rangle,\langle J_z/j\hbar\rangle)\,. \tag{3.53}$$

If the operators J_x, J_y and J_z commute with the Hamiltonian of the system the polarization is conserved in time (cf. Appendix F.2).

Let us have a look at a few examples. For particles prepared in the state $|2,2\rangle$ with respect to the quantization axis $\hat{\mathbf{z}}$, we calculate $\mathbf{P}=(0,0,1)=\hat{\mathbf{z}}$ and the particles are said to be *fully polarized* in the z direction (see Fig. 3.1*c*). When prepared in the state $|2,1\rangle$ we find $\mathbf{P}=(0,0,\frac{1}{2})=\frac{1}{2}\hat{\mathbf{z}}$; i.e., the particles are *partially polarized* in the z direction. For the state $|2,0\rangle$ we calculate $\mathbf{P}=(0,0,0)=0$. In this case the particles are called *unpolarized.*

Problem 3.2. Calculate for the angular momentum state $|j,m\rangle$ the following quantities:

$$\langle J_x^2\rangle,\ \langle J_y^2\rangle,\ \langle J_z^2\rangle,\ \langle J_x\rangle,\ \langle J_y\rangle \text{ and } \langle J_z\rangle.$$

Discuss the relation $\langle J_x^2\rangle+\langle J_z^2\rangle=\langle\mathbf{J}^2\rangle$ in the context of the semi-classical vector model.

3.3.1 Ensemble average and density matrix

Next we have a closer look at the measurement of the polarization vector. Typically one creates a beam of particles, for the above example all prepared in the state $|\psi\rangle=|2,2\rangle$. Ideally, the particles are prepared one by one and observed sequentially using a detector that discriminates between the eigenstates $\{|j,m\rangle\}$ with respect to the quantization axis $\hat{\mathbf{z}}$. To determine $\langle J_z/j\hbar\rangle$ we measure the *probability density* to observe the particle in the eigenstate $|j,m\rangle$,

$$\mathcal{P}_m = |\langle j,m|\psi\rangle|^2, \tag{3.54}$$

subject to the Parseval relation

$$\sum_m \mathcal{P}_m = 1. \tag{3.55}$$

This type of measurement is called a *state-selective measurement* (or *filtration*). Repeating the measurement many times we find the polarization from the *ensemble average*, which is the *statistical average* over a lot of quantum measurements,

$$\mathbf{P} = \langle\!\langle \mathbf{J}/j\hbar\rangle\!\rangle \equiv (\langle\!\langle J_x/j\hbar\rangle\!\rangle, \langle\!\langle J_y/j\hbar\rangle\!\rangle, \langle\!\langle J_z/j\hbar\rangle\!\rangle)\,. \tag{3.56}$$

If all particles are prepared in the same state (as in the above example) we speak of a *pure ensemble.* Pure ensembles satisfy the property $\langle\!\langle \mathbf{J}/j\hbar\rangle\!\rangle = \langle \mathbf{J}/j\hbar\rangle$. The magnitude of the polarization vector is called the *degree of polarization*,

$$p = \sqrt{\langle\!\langle J_x/j\hbar\rangle\!\rangle^2 + \langle\!\langle J_y/j\hbar\rangle\!\rangle^2 + \langle\!\langle J_z/j\hbar\rangle\!\rangle^2}. \tag{3.57}$$

This degree can vary between zero and unity, $0 \le p \le 1$. For a pure ensemble prepared in the state $|\psi\rangle = |2,q\rangle$ the detector will only measure particles in the channel $|2,q\rangle$; i.e., $\mathcal{P}_m = \delta_{m,q}$, which implies $\langle\!\langle J_z/j\hbar\rangle\!\rangle = q/2$ and $\langle\!\langle J_x/j\hbar\rangle\!\rangle = \langle\!\langle J_y/j\hbar\rangle\!\rangle = 0$, and confirms the values obtained in the examples given above. The degree of polarization is $p = q/2$.

More generally one can create beams with a mixture of particles in various angular momentum states. In this case we speak of a *mixed ensemble.* The mixture may contain a bias towards certain states or be completely random. The state may vary from particle to particle but (for a given mixture) the statistical average remains well defined. For *random ensembles* all states are equally probable, which implies $\langle\!\langle J_x/j\hbar\rangle\!\rangle = \langle\!\langle J_y/j\hbar\rangle\!\rangle = \langle\!\langle J_z/j\hbar\rangle\!\rangle = 0$, whatever the quantization axis. In this case the degree of polarization is zero, $p = 0$, and the beam is called *unpolarized.*

Density matrix and statistical operator

The density matrix formalism provides an alternative method for calculating the polarization of angular momentum systems (cf. Appendix F.3). For a pure ensemble of particles prepared in the state $|\psi\rangle$ the *probability density* for observing a particle in the eigenstate $|j,m\rangle$ is given by

$$\mathcal{P}_m = |\langle j,m|\psi\rangle|^2 = \langle \rho_m\rangle, \tag{3.58}$$

where $\rho_m = |j,m\rangle\langle j,m|$ is the density operator subject to the closure relation $\mathbb{1} = \sum_m \rho_m$ and Parseval relation (cf. Appendix F.3),

$$\operatorname{tr}\rho = 1. \tag{3.59}$$

For the eigenstate $|\psi\rangle = |j,j\rangle$ the expectation value $\langle J_z\rangle$ is given by

$$\langle J_z\rangle = \operatorname{tr}\rho J_z = \sum_m \langle\psi|j,m\rangle\langle jm|J_z|\psi\rangle = j\hbar.$$

When dealing with a mixed ensemble of angular momentum eigenstates, the density operator is replaced by the *statistical operator* of the ensemble $\{|j,m\rangle\}$

$$\varrho = \sum_m w_m \rho_m, \tag{3.60}$$

where w_m is the statistical weight of the state $|j,m\rangle$ from the ensemble. Without active preparation of a specific state the ensemble is random and w_m has the same value for all m in the interval $-j < m < j$. For ensemble averages the Parseval relation relation is replaced by the trace over the statistical operator,

$$\mathrm{tr}\,\varrho = 1. \tag{3.61}$$

For a random ensemble of angular momentum states the degree of polarization is zero, $p = 0$ (see Problem 3.3).

Problem 3.3. Show that for a random ensemble of angular momentum states the degree of polarization is zero, $p = 0$.

Solution. Consider a random ensemble of angular momentum states. Since the ensemble is random, the probability to observe the system in the state $|j,m\rangle$ is the same for any quantization axis and for all values of m, $w_m = 1/(2j+1)$. Then, the degree of polarization is zero because the ensemble average of J_z vanishes for any quantization axis,

$$\langle\!\langle J_z\rangle\!\rangle = \mathrm{tr}\,\varrho J_z = w_m \sum_{m,q} \langle j,q|j,m\rangle\langle jm|J_z|j,q\rangle = w_m \sum_m m = 0. \qquad \square$$

3.3.2 The case $s = 1/2$ as a two-level system - relation with the Bloch sphere

It is appropriate to pay special attention to polarization in $s = \frac{1}{2}$ systems, the most prominent among the angular momentum systems. The $s = \frac{1}{2}$ system is a two-level systems, the simplest type of discrete quantum system. Many properties of these systems are derived in Appendix G. An arbitrary state $|\chi\rangle$ of a two-level system is determined by four parameters, three of which correspond to observable quantities. Denoting these by θ, ϕ and normalization $\mathcal{N} = 1$ the state $|\chi\rangle$ can be written in the form - see Appendix G.2.2

$$|\chi\rangle = \cos(\theta/2)\,|\!\uparrow\rangle + e^{i\phi}\sin(\theta/2)\,|\!\downarrow\rangle\,, \tag{3.62}$$

where the probability amplitudes to observe the system in the eigenstates $|\!\uparrow\rangle$ and $|\!\downarrow\rangle$ are given by

$$a_1 = \langle\uparrow|\,\chi\rangle = \cos(\theta/2) \quad \text{and} \quad a_2 = \langle\downarrow|\,\chi\rangle = e^{i\phi}\sin(\theta/2). \tag{3.63}$$

Here ϕ represents the relative phase of the two eigenstates. Furthermore, the amplitude of the spin-up state is chosen to be real in the convention of Appendix G.2.2. The angles θ and ϕ correspond to the polar and azimuthal angles of the polarization vector $\mathbf{P}$, as follows by calculating the cartesian components of its components, $P_i = \langle 2S_i/\hbar\rangle = \langle\sigma_i\rangle$ with $i \in \{x,y,z\}$,

$$P_x = \langle\sigma_x\rangle = \tfrac{1}{2}\langle\chi|\sigma_- + \sigma_+|\chi\rangle = a_1a_2^* + a_1^*a_2 = \sin\theta\cos\phi \tag{3.64a}$$

$$P_y = \langle\sigma_y\rangle = i\tfrac{1}{2}\langle\chi|\sigma_- - \sigma_+|\chi\rangle = i(a_1a_2^* - a_1^*a_2) = \sin\theta\sin\phi \tag{3.64b}$$

$$P_z = \langle\sigma_z\rangle = \langle\chi|\sigma_z|\chi\rangle = a_1a_1^* - a_2a_2^* = \cos\theta. \tag{3.64c}$$

The polarization vector $\mathbf{P}$ is a real three-dimensional vector pointing to a point on the surface of the *Bloch sphere* - see Fig. G.1. In this context and $\mathbf{P}$ is called the *Bloch vector* of the $s = \frac{1}{2}$ system. Each point on the surface of the sphere corresponds to one of the spin states described by Eq. (3.62).

3.3.3 Two level system $(s = 1/2)$ in the density matrix formalism

Let us start with the pure state $|\chi\rangle$ defined by Eq. (3.62). This state can be used to represent an arbitrary state of a two-level system. The density operator for $|\chi\rangle$ is given by the 2×2 matrix

$$\rho = |\chi\rangle\langle\chi| = \frac{1}{2}\begin{pmatrix} 1+\cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & 1-\cos\theta \end{pmatrix}. \tag{3.65}$$

Note that $\langle\chi|\rho|\chi\rangle = 1$. This reflects the normalization of the state. It is evident that this matrix can be separated into a unit matrix and a traceless part,

$$\rho = |\chi\rangle\langle\chi| = \frac{1}{2}\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} +\cos\theta & \sin\theta\, e^{-i\phi} \\ \sin\theta\, e^{i\phi} & -\cos\theta \end{pmatrix}\right]. \tag{3.66}$$

Recalling Eq. (3.49) we find that the density operator can be written in the compact form

$$\rho = \tfrac{1}{2}\left(\mathbb{1} + \hat{\mathbf{r}}\cdot\boldsymbol{\sigma}\right). \tag{3.67}$$

Using the vector relation (3.51) we find

$$\rho^2 = \rho. \tag{3.68}$$

This *idempotence* of the density operator is a property of pure states (see Appendix F.3). The components of the polarization vector are given by the trace

$$\langle P_i\rangle = \mathrm{tr}\,(\rho\sigma_i) = \langle\uparrow|\,\rho\sigma_i\,|\uparrow\rangle + \langle\downarrow|\,\rho\sigma_i\,|\downarrow\rangle\,, \tag{3.69}$$

where $i \in \{x, y, z\}$. Substituting the density and Pauli matrices we regain Eq. (3.64) for the polarization vector of $|\chi\rangle$. Eq. (3.67) is the general form for the density matrix of a two-level system.

To illustrate the use of the expression (3.67) we recalculate $\mathcal{P}_+$ for the arbitrary spin state (3.62). For this we need the density matrix ρ_+ of the state $|+\rangle$, for which the polarization is given by $\mathbf{P} = (0, 0, 1) = \hat{\mathbf{z}}$, which means that

$$\rho_+ = \tfrac{1}{2}\left(\mathbb{1} + \sigma_z\right). \tag{3.70}$$

Using this operator we calculate with the aid of Eqs. (G.6) and (3.64c)

$$\mathcal{P}_+ = \langle\chi|\rho_+\,|\chi\rangle = \tfrac{1}{2}\left(1 + \langle\sigma_z\rangle\right) = \tfrac{1}{2}\left(1 + |a_1|^2 - |a_2|^2\right) = |a_1|^2. \tag{3.71}$$

Indeed this result coincides with that of Eq. (3.75), as it should.

Specializing Eq. (3.62) to the following special directions on the Bloch sphere, $x, y, -x, -y \leftrightarrow (\theta = \pi/2,\ \phi = 0, \pi/2, \pi, 3\pi/2)$ we obtain (see Fig. G.1)

$$\begin{aligned} |\uparrow\rangle_x &= \sqrt{\tfrac{1}{2}}\left(|\uparrow\rangle + |\downarrow\rangle\right) \text{ and } |\uparrow\rangle_{-x} = \sqrt{\tfrac{1}{2}}(|\uparrow\rangle - |\downarrow\rangle) \\ |\uparrow\rangle_y &= \sqrt{\tfrac{1}{2}}(|\uparrow\rangle + i\,|\downarrow\rangle) \text{ and } |\uparrow\rangle_{-y} = \sqrt{\tfrac{1}{2}}(|\uparrow\rangle - i\,|\downarrow\rangle). \end{aligned} \tag{3.72}$$

For the negative z direction the result depends on the angle ϕ chosen for the plane of rotation. Rotating about the positive y axis we have $-z \leftrightarrow (\theta = \pi, \phi = 0)$ and obtain

$$|\uparrow\rangle_{-z} \equiv |\downarrow\rangle\,. \tag{3.73}$$

The density matrices $\rho_{x\uparrow}$ and $\rho_{y\uparrow}$ follow directly from Eqs. (3.72) and are given by

$$\rho_{x\uparrow} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad\text{and}\quad \rho_{y\uparrow} = \frac{1}{2}\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}. \tag{3.74}$$

Note that these density matrices are idempotent and have unit trace as required for pure states.

The probability $\mathcal{P}_+$ can be expressed as the expectation values of the density operators $\rho_+ = |\uparrow\rangle\,\langle\uparrow|$,

$$\mathcal{P}_+ = |a_1|^2 = \langle\chi|\rho_+|\chi\rangle = \langle\rho_+\rangle, \tag{3.75}$$

and a similar relation holds for $\mathcal{P}_-$.

Problem 3.4. Consider the state $|\chi\rangle$ of a $s=\frac{1}{2}$ angular momentum system with polarization vector $\mathbf{P}$. Show that $|\chi\rangle$ is an eigenstate of the operator σ_r,

$$\sigma_r|\chi\rangle = |\chi\rangle,$$

where $\sigma_r = \mathbf{P}\cdot\boldsymbol{\sigma}$ is the operator measuring the angular momentum in the direction $\hat{\mathbf{r}} = \mathbf{P}$.

Solution. Rewriting Eq. (3.67) we have

$$\mathbf{P}\cdot\boldsymbol{\sigma} = 2\rho_\chi - \mathbb{1} \Rightarrow \sigma_r = 2|\chi\rangle\langle\chi| - \mathbb{1}.$$

Hence, $\sigma_r|\chi\rangle = (2|\chi\rangle\langle\chi| - \mathbb{1})\,|\chi\rangle = |\chi\rangle$. □

3.4 Addition of two angular momenta

In many cases of practical importance, quantum mechanical systems are characterized by more than one type of angular momentum; for instance the orbital and spin angular momenta of a single electron, or the angular momenta of two electrons in a many-electron atom. To analyze such cases we consider a closed mechanical system of two *independent* angular momenta, $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$, as well as the *total* angular moment of the system,

$$\mathbf{J} = \boldsymbol{j}_1 + \boldsymbol{j}_2. \tag{3.76}$$

As $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ act in independent sectors of Hilbert space, they satisfy the commutation relations

$$[\boldsymbol{j}_1, \boldsymbol{j}_2] = 0. \tag{3.77}$$

The eigenstates of the angular momentum $\boldsymbol{j}_i$ are specified as $|j_i m_i\rangle$ in the standard representation $\{\boldsymbol{j}_i^2, j_{i\,z}\}$,with $i \in \{1,2\}$.

3.4.1 The uncoupled basis of $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$

In view of the commutation of the operators $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ commute, also the operators $\boldsymbol{j}_1^2$, $j_{1\,z}$ and $\boldsymbol{j}_2^2$, $j_{2\,z}$ have to commute and this implies the existence of a joint basis (see Problem F.1) of eigenstates which can be written in the form of an ordered product, the Kronecker-product (a tensor product)

$$|j_1 m_1; j_2 m_2\rangle \equiv |j_1 m_1\rangle_1 \otimes |j_2 m_2\rangle_2, \tag{3.78}$$

where the eigenstates of $\boldsymbol{j}_1$ are denoted by $|j_1 m_1\rangle_1$ and those of $\boldsymbol{j}_2$ by $|j_2 m_2\rangle_2$. This convention is illustrated by the following examples:

$$\boldsymbol{j}_i^2\,|j_1 m_1; j_2 m_2\rangle = j_i(j_i+1)\hbar^2|j_1 m_1; j_2 m_2\rangle \tag{3.79a}$$

$$j_{iz}\,|j_1 m_1; j_2 m_2\rangle = m_i\hbar|j_1 m_1; j_2 m_2\rangle, \tag{3.79b}$$

with $i \in \{1,2\}$. The basis $\{|j_1 m_1; j_2 m_2\rangle\}$ is called the *uncoupled basis* of the Hilbert space of the pair. For given values of j_1 and j_2 this basis spans a $(2j_1+1)(2j_2+1)$-dimensional space, which represents the Hilbert space of the tensor $\boldsymbol{j}_1 \otimes \boldsymbol{j}_2$. In general, the states do *not* commute under the Kronecker-product,

$$|j_1 m_1\rangle_1 \otimes |j_2 m_2\rangle_2 \neq |j_2 m_2\rangle_2 \otimes |j_1 m_1\rangle_1, \tag{3.80}$$

which means that we have to specify the order of the product, which is called *coupling order* of $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$. Importantly, for the special case $m_1 = j_1$ and $m_2 = j_2$ the Kronecker product is independent of the coupling order,

$$|j_1 j_1\rangle_1 \otimes |j_2 j_2\rangle_2 = |j_2 j_2\rangle_2 \otimes |j_1 j_1\rangle_1. \tag{3.81}$$

This is a property of so-called *stretched* states, which will be introduced in Section 3.4.2. For the special case $j_1 = \frac{1}{2}$ and $j_2 = 1$ their independence of the coupling order is demonstrated in the example below. Note that, for given states $|j_1 m_1\rangle_1$ and $|j_2 m_2\rangle_2$, Eqs. (3.79) are independent of the coupling order.

Example 3.1. The product space in matrix notation

Consider two angular momenta, $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$, acting in vector spaces of dimension $d_1 = 2$ and $d_2 = 3$, respectively (i.e., $j_1 = \frac{1}{2}$ and $j_2 = 1$). The corresponding eigenstates are given by

$$|\tfrac{1}{2}, +\tfrac{1}{2}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |\tfrac{1}{2}, -\tfrac{1}{2}\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.82a}$$

and

$$|1, +1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |1, 0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |1, -1\rangle = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}. \tag{3.82b}$$

In the uncoupled representation the eigenstate $|\frac{1}{2}, +\frac{1}{2}; 1, 0\rangle$ is given by the Kronecker product $|\frac{1}{2}, +\frac{1}{2}\rangle \otimes |1, 0\rangle$. Note that $|\frac{1}{2}, +\frac{1}{2}\rangle$ and $|1, 0\rangle$ do *not* commute under the Kronecker product,

$$|\tfrac{1}{2}, +\tfrac{1}{2}\rangle \otimes |1, 0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |1, 0\rangle \otimes |\tfrac{1}{2}, +\tfrac{1}{2}\rangle. \tag{3.82c}$$

For the doublet state $|\frac{1}{2}, +\frac{1}{2}\rangle$ the Kronecker product $|\frac{1}{2}, +\frac{1}{2}\rangle \otimes |1, 0\rangle$ opens a three-dimensional subspace for the triplet state $|1, 0\rangle$. Likely, for the triplet state $|1, 0\rangle$, the Kronecker product $|1, 0\rangle \otimes |\frac{1}{2}, +\frac{1}{2}\rangle$ opens a two-dimensional subspace for the doublet state $|\frac{1}{2}, +\frac{1}{2}\rangle$. In both ways we obtain a 6-dimensional Hilbert space (for the tensors $\boldsymbol{j}_1 \otimes \boldsymbol{j}_2$ and $\boldsymbol{j}_2 \otimes \boldsymbol{j}_1$, respectively). Note that the states $|\frac{1}{2}, +\frac{1}{2}\rangle$ and $|1, 1\rangle$ are special because for $m_1 = j_1$ and $m_2 = j_2$ the Kronecker product is independent of the coupling order, a property of *stretched* states (see next section

$$|\tfrac{1}{2}, +\tfrac{1}{2}\rangle \otimes |1, 1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |1, 1\rangle \otimes |\tfrac{1}{2}, +\tfrac{1}{2}\rangle. \tag{3.83}$$

3.4.2 The coupled basis of $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$

Realizing that $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ individually satisfy commutation relations of the type (3.1) and commute with each other it is straightforward to show that also the *total angular momentum*, given by the vector sum

$$\mathbf{J} = \boldsymbol{j}_1 + \boldsymbol{j}_2, \tag{3.84}$$

satisfies the commutation relations (3.1). This immediately implies that $\mathbf{J}$ satisfies all properties of quantized angular momentum as summarized in Section 3.1. In particular we have

$$\mathbf{J}^2 |JM\rangle = J(J+1)\hbar^2 |JM\rangle \tag{3.85a}$$

$$J_z |JM\rangle = M\hbar |JM\rangle, \tag{3.85b}$$

with M restricted to the interval

$$-J \leq M \leq J. \tag{3.86}$$

At this point we have established that the vector addition is subject to the constraint that also the resultant vector be quantized. Since $\boldsymbol{j}_1^2$ and $\boldsymbol{j}_2^2$ commute with both J_z and $\mathbf{J}^2$ (see Problem 4.6) we infer that the quantum numbers j_1, j_2, J and M are *compatible* (see Section F.1.1); i.e., they define the basis $\{|(j_1 j_2)JM\rangle\}$, the so-called *coupled* basis of the standard representation $\{J^2, J_z\}$. With the notation $(j_1 j_2)$ we specify the values as well as the *coupling order* of the two angular momenta (see Section J.3.1). Often one writes $\{|j_1 j_2 JM\rangle\}$ or, simply $\{|JM\rangle\}$ when the coupling order and values of j_1 and j_2 are irrelevant or known from the context. For given values of j_1, j_2

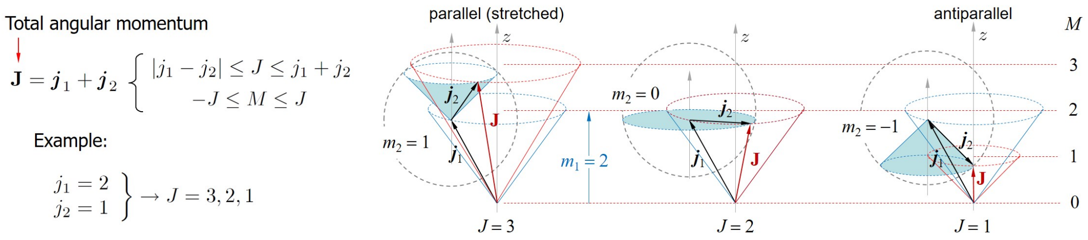


Figure 3.2: Vector addition of quantized angular momenta. The quantization conditions limit the number of possible resulting vectors. This is illustrated for the example of the angular momenta $j_1 = 2$ and $j_2 = 1$ (compare with Fig. 3.2). The state with $m_1 = j_1$ and $m_2 = j_2$ (i.e., $M = j_1 + j_2 = 3$) is called the state of *parallel coupling* or the *stretched* state as it represents the largest total angular momentum projection that can be constructed by quantized vector addition.

and J, the operator $\mathbf{J}$ acts in a $(2J+1)$-dimensional subspace of the product space of the tensor $\boldsymbol{j}_1 \otimes \boldsymbol{j}_2$, with the quantum number J restricted to the interval (see Problem 3.5),

$$|j_1 - j_2| \le J \le j_1 + j_2. \tag{3.87}$$

Denoting the lesser of j_1 and j_2 by $j_< \equiv \min\{j_1, j_2\}$, the quantum number J can take $2j_< + 1$ values (see Problem 3.5). This is illustrated geometrically in Fig. 3.2. Geometry helps to swiftly analyze the coupling options. More subtle properties (like phase relations) require the algebra of commutation relations - see Section J.2.1.

Since angular momentum states of different J are orthogonal, the dimension of the $2j_< + 1$ allowed values of J add up to $(2j_1 + 1)(2j_2 + 1)$ as is demonstrated in Problem 3.6. Formally, the product space of $\boldsymbol{j}_1 \otimes \boldsymbol{j}_2$, represented by the *uncoupled basis*, $\{|j_1 m_1; j_2 m_2\rangle\}$, is called *reducible* because it can be decomposed into the *direct sum* of $2j_< + 1$ *irreducible* subspaces, one for each allowed value of J and represented in the coupled basis $\{|(j_1 j_2) JM\rangle\}$. As we shall find in Section 3.6 the irreducible subspaces of the coupled representation are invariant under rotations about a point. Importantly, the coupled basis and the uncoupled basis are defined irrespective of the presence (or absence) of a *coupling mechanism.*

Example 3.2. Angular momentum addition in matrix notation

Let us return to the example of two angular momenta, $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$, acting in vector spaces of dimension $d_1 = 2$ and $d_2 = 3$, respectively ($j_1 = \frac{1}{2}$ and $j_2 = 1$) - see Example 3.1. Before we can add the operators $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ in the matrix representation they have to be written as operators in the $d_1 \times d_2 = 6$-dimensional Hilbert space of the Kronecker product $\boldsymbol{j}_1 \otimes \boldsymbol{j}_2$,

$$\mathbf{J} = \boldsymbol{j}_1 \otimes \mathbb{1}_{j_2} + \mathbb{1}_{j_1} \otimes \boldsymbol{j}_2 = \underline{\boldsymbol{j}}_1 + \underline{\boldsymbol{j}}_2,$$

where $\mathbb{1}_{j_1}$ and $\mathbb{1}_{j_2}$ are the identity matrices of the two and three dimensional subspaces in which $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ act, respectively. In this notation the operator J_z takes the form

$$J_z = \underline{j}_{1z} + \underline{j}_{2z} = \hbar \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \hbar \begin{pmatrix} \frac{3}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{3}{2} \end{pmatrix}.$$

Triangle inequality - stretching of angular momenta and pure states

The principle of quantized vector addition is illustrated in Fig. 3.2. The quantum numbers of the vector **J** satisfy the *triangle inequality.* The largest possible value of J,

$$J_{\max} = j_1 + j_2, \tag{3.88}$$

is the result of so-called *parallel coupling* of the angular momenta, pictorially referred to as the state of *stretched* angular momentum. Likewise, the smallest possible value,

$$J_{\min} = |j_1 - j_2|, \tag{3.89}$$

corresponds to the state of *antiparallel coupling.* Hence, the allowed total angular momenta differ in the level of stretching. Antiparallel coupling corresponds to *minimal stretching* and becomes *zero stretching* for $j_1 = j_2$. Parallel coupling corresponds to *maximal stretching. Complete stretching* is ruled out by the quantization condition, it is only possible in the classical limit. If either j_1 or j_2 (not both) is half-integral, J also has to be half-integral and, therefore, *nonzero.* In all other cases J can also take the value zero. Note that, $j_1 + j_2 + J$ is always *integral.*

Importantly, the eigenstates of the uncoupled basis $\{|j_1, m_1; j_2, m_2\rangle\}$ are also eigenstates of J_z. This implies the selection rule

$$M = m_1 + m_2. \tag{3.90}$$

The maximum and minimum values of M, $M = \pm(j_1 + j_2)$, correspond to *pure states* (unique eigenstates of the uncoupled representation with unit relative phase imposed by convention). In particular this holds for the so-called *stretched* state, which is the state of largest total angular momentum and projection along the quantization axis,

$$|j_1 + j_2, j_1 + j_2\rangle = |j_1, j_1; j_2, j_2\rangle. \tag{3.91}$$

Problem 3.5. Let $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ be two commuting angular momentum operators and $\mathbf{J} = \boldsymbol{j}_1 + \boldsymbol{j}_2$ the resulting total angular momentum. Show that the $2j_< + 1$ allowed values of J satisfy the inequality

$$|j_1 - j_2| \le J \le j_1 + j_2.$$

Solution. The maximum value of J corresponds to the stretched state $J = j_1 + j_2$; hence, $M \le J = j_1 + j_2$. Next we search for the minimum value of J. Without loss of generality we presume that $j_1 \ge j_2$. Considering first the case of integral M, we know that for any allowed value of J the minimal value of $|M|$ must be zero, $M = m_1 + m_2 = 0$. Since we have $2j_2 + 1$ possible values of m_2 we find for the degeneracy of the value $M = 0$

$$N(J) = 2j_2 + 1 = 2j_< + 1.$$

This corresponds to all possible vector additions (i.e, allowed values of J) consistent with the quantization rules. Counting down $2j_1 + 1$ values, starting from $J = j_1 + j_2$, we arrive at the minimum value $J = j_1 - j_2$. This corresponds to the condition $J \ge (j_1 + j_2) - 2j_2 = j_1 - j_2$. The same result is obtained in a similar fashion for half-integral M. □

Problem 3.6. Show that the coupled basis and uncoupled basis have the same dimension.

Solution. First we recall that the uncoupled basis $\{|j_1 m_1; j_2 m_2\rangle\}$ spans a product space of dimension $d = (2j_1 + 1)(2j_2 + 1)$. For a given value of J the coupled basis $\{|(j_1 j_2)J, M\rangle\}$ spans a $(2J + 1)$-dimensional subspace of this product space. Without loss of generality we presume that $j_1 \ge j_2$. Summing over the $2j_2 + 1$ subspaces of this type we regain

$$d = \sum_{n=0}^{2j_2} [2(j_1 + j_2 - n) + 1] = (2j_2 + 1) \times [2(j_1 + j_2) + 1] - \sum_{n=0}^{2j_2} n = (2j_2 + 1)(2j_1 + 1). \qquad \square$$

3.4.3 Clebsch-Gordan basis transformation

The coupled basis is related to the uncoupled basis by a basis transformation known as the *Clebsch-Gordan transformation.* For a system of two spin-$\frac{1}{2}$ particles this is demonstrated in Problem 3.7. For arbitrary angular momenta this transformation is given by

$$|j_1 j_2 JM\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |j_1 m_1; j_2 m_2\rangle\langle j_1 m_1; j_2 m_2|j_1 j_2 JM\rangle, \tag{3.92}$$

where the use of the closure relation

$$\mathbb{1} = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |j_1 m_1; j_2 m_2\rangle\langle j_1 m_1; j_2 m_2| \tag{3.93}$$

does not go unnoticed. Recall that the pair state is subject to an ordering convention. The coefficients $\langle j_1 m_1; j_2 m_2|j_1 j_2 JM\rangle$ are called *Clebsch-Gordan coefficients* (CGCs). As $J_Z = j_{1z} + j_{2z}$ is diagonal in both the coupled and the uncoupled representation we find the selection rule

$$M = m_1 + m_2 \tag{3.94}$$

as a property of the CGCs - the so-called *projection rul*e. The Clebsch-Gordan transformation can be represented by a square matrix. This matrix is unitary and can be written in block-diagonal form with one block for each physically relevant value of M (see Problem 3.7). The inverse transformation is given by

$$|j_1 m_1; j_2 m_2\rangle = \sum_{J=|j_1-j_2|}^{j_1+j_2} \sum_{M=-J}^{J} |j_1 j_2 JM\rangle\langle j_1 j_2 JM|j_1 m_1; j_2 m_2\rangle, \tag{3.95}$$

where we recognize

$$\mathbb{1} = \sum_{J=|j_1-j_2|}^{j_1+j_2} \sum_{M=-J}^{J} |j_1 j_2 JM\rangle\langle j_1 j_2 JM|. \tag{3.96}$$

The CGCs are usually written in the shorthand form

$$\langle j_1 m_1; j_2 m_2|j_1 j_2 JM\rangle \Leftrightarrow \langle j_1 m_1; j_2 m_2|JM\rangle, \tag{3.97}$$

where the coupling order is implicitly defined by the properties of the Kronecker product (we return ordering issues in Section J.3.1).

From the inner products of Eqs. (3.92) and (3.95) with their hermitian conjugates we obtain

$$\sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \langle J'M'|j_1 m_1; j_2 m_2\rangle\langle j_1 m_1; j_2 m_2\,|JM\rangle = \delta_{J'J}\,\delta_{M'M}\,\delta(j_1 j_2 J) \tag{3.98a}$$

$$\sum_{J=|j_1-j_2|}^{j_1+j_2} \sum_{M=-J}^{J} \langle j_1 m_1'; j_2 m_2'|JM\rangle\langle JM|j_1 m_1; j_2 m_2\rangle = \delta_{m_1' m_1}\delta_{m_2' m_2}, \tag{3.98b}$$

where $\delta(j_1 j_2 J)$ is a logical function of the property $\Delta(j_1 j_2 J)$ - see Eq. (J.12); it is *true,* $\delta(j_1 j_2 J) = 1$, if the triangle inequalities are satisfied, i.e., for $\Delta(j_1 j_2 J) > 0$, and *false*, $\delta(j_1 j_2 J) = 0$, if they are violated, i.e., for $\Delta(j_1 j_2 J) = 0$. Taking into account the projection rule (3.94) the Eqs. (3.98) reduce to a set of $2J + 1$ subexpressions, one for each physically relevant value of M,

$$\sum_{m_1=-j_1}^{j_1} \langle J'M|j_1 m_1; j_2, M - m_1\rangle\langle j_1 m_1; j_2, M - m_1\,|JM\rangle = \delta_{J'J}\,\delta(j_1 j_2 J), \tag{3.99a}$$

$$\sum_{J=|j_1-j_2|}^{j_1+j_2} \langle j_1 m_1'; j_2, M - m_1'|JM\rangle\langle JM|j_1 m_1; j_2, M - m_1\rangle = \delta_{m_1' m_1}. \tag{3.99b}$$

Problem 3.7. Find the unitary matrix U transforming the uncoupled basis $\{|j_1m_1;j_2m_2\rangle\}$ into the coupled basis $\{|j_1j_2JM\rangle\}$ for a system of two spin-$\frac{1}{2}$ particles. Write the matrix in block-diagonal form, with one block for each physically relevant value of $M \in \{-1,0,1\}$.

Solution. Starting from the Clebsch-Gordan transformation (3.92) and using Table J.2 we find the unitary matrix ($UU^\dagger = \mathbb{1}$ - see Appendix M.2.1)

$$\begin{pmatrix} |1,1\rangle \\ |1,0\rangle \\ |0,0\rangle \\ |1,-1\rangle \end{pmatrix} = U \begin{pmatrix} |\uparrow\uparrow\rangle \\ |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \\ |\downarrow\downarrow\rangle \end{pmatrix} \quad \text{where } U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1/2} & \sqrt{1/2} & 0 \\ 0 & \sqrt{1/2} & -\sqrt{1/2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \qquad \square$$

3.5 Addition in the presence of coupling - conservation rules

Let us continue with the closed mechanical system of two angular momenta introduced above. Let $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$, with $[\boldsymbol{j}_1, \boldsymbol{j}_2] = 0$, be these angular momenta and $\mathcal{H}_1$ and $\mathcal{H}_2$ the Hamiltonians of the orthogonal subsystems. As discussed in Appendix F.2.4 the angular momenta are conserved quantities if they commute with their Hamiltonian,

$$[\boldsymbol{j}_1, \mathcal{H}_1] = 0 \text{ and } [\boldsymbol{j}_2, \mathcal{H}_2] = 0. \tag{3.100}$$

When $\mathcal{H}_0 = \mathcal{H}_1 + \mathcal{H}_2$ is the Hamiltonian of the combined system, it follows from the commutation relations (3.100) that also the total angular momentum $\mathbf{J} = \boldsymbol{j}_1 + \boldsymbol{j}_2$ is a conserved quantity,

$$[\mathbf{J}, \mathcal{H}_0] = 0. \tag{3.101}$$

This is no surprise: given the homogeneity and isotropy of free space, the linear and angular momenta of a closed mechanical system are always conserved. Importantly, note that $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ commute with each other, $[\boldsymbol{j}_1, \boldsymbol{j}_2] = 0$, but *not* with $\mathbf{J}$ (only the components along the quantization axis do). This shows that the three conserved quantities $\boldsymbol{j}_1$, $\boldsymbol{j}_2$ and $\mathbf{J}$ do *not* share a joint basis. For this reason the properties of $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ are most conveniently evaluated in the uncoupled basis, $\{|j_1m_1;j_2m_2\rangle\}$, whereas the coupled basis, $\{|j_1j_2JM\rangle\}$, is the preferred basis to analyze $\mathbf{J}$.

In typical situations, the simultaneous conservation of $\boldsymbol{j}_1$, $\boldsymbol{j}_2$ and $\mathbf{J}$ is broken when a so-called *coupling term*, $\mathcal{H}'$, enters the Hamiltonian

$$\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2 + \mathcal{H}'. \tag{3.102}$$

Let us analyze two examples. First we consider a case of *mutual* coupling,

$$\mathcal{H}' = \alpha\, \boldsymbol{j}_1 \cdot \boldsymbol{j}_2. \tag{3.103}$$

Being mutual, this coupling is *internal* and the system remains mechanically closed, so we know that $\mathbf{J}$ has to remain conserved. This is confirmed by the commutation relation $[\mathbf{J}, \mathcal{H}] = 0$, which is valid because $\mathbf{J}$ commutes with $\boldsymbol{j}_1 \cdot \boldsymbol{j}_2$, although, individually, $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ do not commute with $\boldsymbol{j}_1 \cdot \boldsymbol{j}_2$ (see Problem 4.7). This shows that $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ are no longer conserved, although the magnitude of these vectors (i.e., the quantum numbers j_1 and j_2) as well as their vector sum, $\mathbf{J}$, still are. This is expressed by the commutation relations

$$[j_i^2, j_1 \cdot j_2] = 0, \;\; [J_z, j_1 \cdot j_2] = 0, \;\; [\mathbf{J}^2, j_1 \cdot j_2] = 0 \text{ and } [j_i^2, \mathbf{J}^2] = 0, \tag{3.104}$$

with $i \in \{1,2\}$ (see Problems 4.3-4.6). Hence this coupling is best analyzed in the coupled basis, $\{|j_1j_2JM\rangle\}$. The behavior of the physical system is sketched in Fig. 3.3*a*: $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ precess about $\mathbf{J}$. In the semi-classical picture, the angle between $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ is conserved because the system cannot lower its interaction energy (hamiltonian evolution is dissipation free).

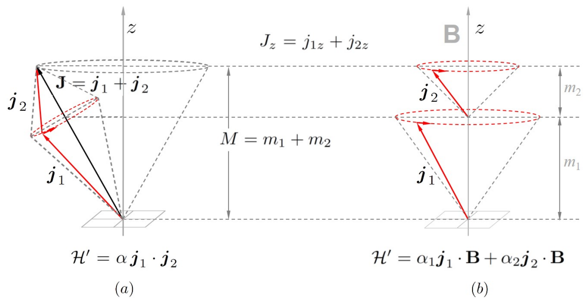


Figure 3.3: Two examples of coupling of the angular momenta $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$: (a) In the presence of the *mutual coupling mechanism* $\mathcal{H}' = \alpha\,\boldsymbol{j}_1 \cdot \boldsymbol{j}_2$, the angular momenta $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ precess about $\mathbf{J} = \boldsymbol{j}_1 + \boldsymbol{j}_2$. This gives rise to time-varying projections of $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ on the z-axis but $\mathbf{J}$ as well as the projections $\boldsymbol{j}_1 \cdot \mathbf{J}$ and $\boldsymbol{j}_2 \cdot \mathbf{J}$ remain conserved. In other words: j_1, j_2, J and M are good quantum numbers; (b) In the presence of *coupling to an external field in the z-direction*, $\mathcal{H}' = \alpha_1 \boldsymbol{j}_1 \cdot \mathbf{B} + \alpha_2 \boldsymbol{j}_2 \cdot \mathbf{B}$, the angular momenta $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ precess individually about the z-axis. Since the projections j_{1z} and j_{2z} remain conserved, j_1, m_1, j_2 and m_2 are good quantum numbers. Note that only *in the absence of any coupling* the three angular momenta $\boldsymbol{j}_1$, $\boldsymbol{j}_2$ and $\mathbf{J}$ are simultaneously conserved.

As a second example we consider a coupling term of the Zeeman type (see Section 4.3.3),

$$\mathcal{H}' = \alpha_1 \boldsymbol{j}_1 \cdot \mathbf{B} + \alpha_2 \boldsymbol{j}_2 \cdot \mathbf{B}. \tag{3.105}$$

In this case the angular momenta are individually coupled to the *external* magnetic field $\mathbf{B}$. The external field breaks the spherical symmetry of free space and by choosing the quantization axis along $\mathbf{B}$ we find that only the z components of $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ remain conserved,

$$[j_{1z}, \mathcal{H}] = 0 \;\text{ and }\; [j_{2z}, \mathcal{H}] = 0. \tag{3.106}$$

The physical system is sketched in Fig. 3.3*b*: the angular momenta $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ precess individually about the $\mathbf{B}$ field. Since $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ are no longer conserved also $\mathbf{J}$ is no longer conserved. In this case only the magnitude of the vectors $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ (i.e., the quantum numbers j_1 and j_2) as well as their projections (i.e., the quantum numbers m_1 and m_2) are constants of the motion. This behavior is best captured by the uncoupled basis, $\{|j_1 m_1; j_2 m_2\rangle\}$.

Summary

In the absence of any coupling all three angular momenta $\boldsymbol{j}_1$, $\boldsymbol{j}_2$ and $\mathbf{J}$ are simultaneously conserved; i.e., j_1, m_1, j_2, m_2 as well as J and M are good quantum numbers - whatever the quantization axis. In the presence of pure $\boldsymbol{j}_1 \cdot \boldsymbol{j}_2$ coupling both $\mathbf{J}$ and J_z are conserved; i.e., J and M are good quantum numbers - whatever the quantization axis. In the presence of pure Zeeman coupling only j_{1z}, j_{2z} and J_z are conserved; i.e., m_1, m_2 and $M = m_1 + m_2$ are good quantum numbers with respect to the direction of the symmetry-breaking field.

Coupling of the J levels

Note that in both of the above coupling cases the commutation relation $[J_z, \mathcal{H}]$ holds. So, also in the simultaneous presence of both couplings,

$$\mathcal{H}' = \alpha_1 \boldsymbol{j}_1 \cdot \mathbf{B} + \alpha_2 \boldsymbol{j}_2 \cdot \mathbf{B} + \alpha\,\boldsymbol{j}_1 \cdot \boldsymbol{j}_2, \tag{3.107}$$

the operator $J_z = j_{1z} + j_{2z}$ corresponds to a conserved quantity; i.e., M is a good quantum number. For $B \to 0$ the eigenstates are given by the coupled representation, $\{|j_1 j_2 J M\rangle\}$, whereas for $B \to \infty$

this becomes the uncoupled representation, $\{|j_1m_1;j_2m_2\rangle\}$, with $M = m_1 + m_2$. For intermediate fields the basis can be denoted by $\{|j_1j_2m\rangle\}$, where $\mathcal{H}'|j_1j_2m\rangle = \varepsilon_{jm}(B)|j_1j_2m\rangle$ and the definition of $|j_1j_2m\rangle$ depends on the field. Decomposing $|j_1j_2m\rangle$ with respect to the coupled basis we find

$$|j_1j_2m\rangle = \sum_{J,M} |j_1j_2JM\rangle\langle j_1j_2JM|j_1j_2m\rangle = \sum_{J} \alpha_J(B)\,|j_1j_2Jm\rangle, \tag{3.108}$$

where the $\alpha_J(B) = \langle j_1j_2JM|j_1j_2m\rangle$ are field-dependent coefficients. This shows that for intermediate fields $|j_1j_2jm\rangle$ can be written as a linear combination of all coupled states $|j_1j_2JM\rangle$ for which $M = m$. This is called mixing of the J levels by the Zeeman coupling.

Problem 3.8. Let us denote two angular momenta by $\boldsymbol{j}_1 = \mathbf{L}$ and $\boldsymbol{j}_2 = \mathbf{S}$, and their vector sum by $\mathbf{J} = \mathbf{L} + \mathbf{S}$. Derive the following three inner product rules,

$$\mathbf{L}\cdot\mathbf{S} = L_xS_x + L_yS_y + L_zS_z \tag{3.109a}$$

$$\mathbf{L}\cdot\mathbf{S} = L_zS_z + \tfrac{1}{2}\left(L_+S_- + L_-S_+\right) \tag{3.109b}$$

$$\mathbf{L}\cdot\mathbf{S} = \tfrac{1}{2}(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2). \tag{3.109c}$$

Solution. Rule (a) is simply the definition of the inner product in cartesian coordinates. Using the definitions for the shift operators we have

$$L_+S_- = (L_x + iL_y)(S_x - iS_y) = L_xS_x + L_yS_y + i\,(L_yS_x - L_xS_y)$$
$$L_-S_+ = (L_x - iL_y)(S_x + iS_y) = L_xS_x + L_yS_y - i\,(L_yS_x - L_xS_y)\,.$$

Adding these expressions and using rule (a) we obtain after rearrangement of terms rule (b). Note that the L_zS_z operator as well as the operators L_+S_- and L_-S_+ conserve the total angular momentum along the quantization axis $m_j = m_l + m_s$. Rule (c) follows straightforwardly from the inner product $\mathbf{J}^2 = (\mathbf{L}+\mathbf{S})^2 = \mathbf{L}^2 + \mathbf{S}^2 + 2\mathbf{L}\cdot\mathbf{S}$ because $\mathbf{L}$ and $\mathbf{S}$ commute (they act in independent sectors of Hilbert space). □

3.6 Angular momentum and infinitesimal rotations

3.6.1 Rotations versus unitary transformations

In Section 3.1.1 we found that the decomposition of the angular momentum operator $\mathbf{J}$ maps uniquely onto that of the radius vector $\mathbf{r}$. To further explore the properties of $\mathbf{J}$ we ask ourselves how this vector operator transforms into an operator $\mathbf{J}'$ by changing from one quantization axis to another. Such a change can be implemented by a *proper* rotation of the coordinate system about the origin. This is a rotation in which the handedness of the coordinate system is conserved.[1] A rotation of the coordinate system is called a *passive* rotation to distinguish it from *active* (i.e., *physical*) rotations, in which the physical system is rotated (i.e., the "contours" of the wavefunction) and the coordinate system is fixed.

So let us consider the proper rotation R by which the right-handed cartesian coordinate system $\mathcal{S}$ transforms into $\mathcal{S}'$; i.e., the direction $\hat{\mathbf{r}} = (x, y, z)$ of a vector with respect to $\mathcal{S}$ changes into $\hat{\mathbf{r}}' = (x', y', z')$ with respect to $\mathcal{S}'$ under conservation of the handedness of the coordinate system,

$$\hat{\mathbf{r}}' = R\,\hat{\mathbf{r}}. \tag{3.110}$$

This transformation holds, in particular, for the quantization axis. Equivalently, we can fix the coordinate system and apply the inverse rotation to $\mathbf{J}$,

$$\mathbf{J}' = R^{-1}\mathbf{J}. \tag{3.111}$$

---

[1] A proper rotation, R, is an orthogonal transformation ($R^{-1} = R^T$) with unit determinant ($\det R = 1$) - see Appendix M.

The existence of the inverse is evident from a physical point of view. Mathematically, it follows from the nonzero determinant ($\det R = 1$ for proper rotations) - see Appendix M. Comparing the two approaches we obtain the following equivalent expressions for the passive rotation R:

$$\mathbf{J}'(\hat{\mathbf{r}}) = R^{-1}\mathbf{J}(\hat{\mathbf{r}}) = \mathbf{J}(R\hat{\mathbf{r}}) = \mathbf{J}(\hat{\mathbf{r}}'). \tag{3.112}$$

In particular we have

$$\hat{\mathbf{r}}' \cdot \mathbf{J} = \hat{\mathbf{r}} \cdot \mathbf{J}'. \tag{3.113}$$

Since the vector $\mathbf{J}$ stands for an operator we ask for the effect of a change of quantization axis on the angular momentum states. A new quantization axis comes with a new set of basis states, $\{|j,m\rangle'\}$. These are the joint eigenstates of $\mathbf{J}'^2$ and J_z' and are related to the original basis of the operators $\mathbf{J}^2$ and J_z, $\{|j,m\rangle\}$, by a *norm-conserving* basis transformation,

$$|\chi\rangle' = \underline{u}|\chi\rangle. \tag{3.114}$$

Here $|\chi\rangle$ is an arbitrary angular momentum state within the invariant subspace V^{2j+1} of the operators $\mathbf{J}'$ and $\mathbf{J}$. Note that the operator $\underline{u}$ represent a *passive* transformation. It is norm conserving because $|\chi\rangle$ and $|\chi\rangle'$ correspond to the *same* state with respect to two *different* bases (corresponding to the quantization axes $\hat{\mathbf{r}}$ and $\hat{\mathbf{r}}'$, respectively),

$$\langle\chi|\chi\rangle = {}'\langle\chi|\chi\rangle' = \langle\chi|\underline{u}^\dagger\underline{u}|\chi\rangle \to \underline{u}^\dagger\underline{u} = \mathbb{1}. \tag{3.115}$$

This identity shows that the transformation must be unitary, $\underline{u}^\dagger = \underline{u}^{-1}$, which implies that the inverse transformation is given by

$$\underline{u}^\dagger|\chi\rangle = |\chi\rangle' \tag{3.116}$$

As the expectation values of physical quantities must be invariant under *passive* rotations, we require

$${}'\langle\chi|\mathbf{J}'|\chi\rangle' \equiv \langle\chi|\mathbf{J}|\chi\rangle. \tag{3.117}$$

In particular, this holds for observables. Substituting Eq. (3.116) in the r.h.s. of Eq. (3.117) we obtain

$${}'\langle\chi|\mathbf{J}'|\chi\rangle' = {}'\langle\chi|\underline{u}\mathbf{J}\underline{u}^\dagger|\chi\rangle'. \tag{3.118}$$

As this expression holds for arbitrary $|\chi\rangle'$ we arrive at the conclusion that under the rotation (3.110) the transformation of the operator $\mathbf{J}$ involves the same unitary operator as used for the states $|\chi\rangle$,[1]

$$\mathbf{J}' = \underline{u}\,\mathbf{J}\,\underline{u}^\dagger. \tag{3.119}$$

We now return to the two inner products of Eq. (3.112). Substituting Eq. (3.110) into the l.h.s. of Eq. (3.112) and Eq. (3.119) into the r.h.s., we arrive at an expression relating the rotation R of the cartesian coordinate system $\mathcal{S}$ (in real space) to the corresponding unitary transformation $\underline{u}$ (in Hilbert space),

$$\mathbf{J}\cdot R\,\hat{\mathbf{r}} = \underline{u}\,\mathbf{J}\,\underline{u}^\dagger\cdot\hat{\mathbf{r}}. \tag{3.120}$$

For two subsequent rotations this expression becomes

$$\mathbf{J}\cdot R_2(R_1\,\hat{\mathbf{r}}) = \underline{u}_2(\underline{u}_1\,\mathbf{J}\,\underline{u}_1^\dagger)\underline{u}_2^\dagger\cdot\hat{\mathbf{r}}. \tag{3.121}$$

So, once we have an expression for R (see Section 3.6.2), Eqs. (3.120) and (3.121) can serve to determine $\underline{u}$. This is the subject of Section 3.6.3.

---

[1] Some authors prefer not to introduce passive rotations for the states and use Eq. (3.114) to indicate an active transformation. In this equivalent approach Eq. (3.117) is replaced by $\langle\chi|\mathbf{J}'|\chi\rangle \equiv {}'\langle\chi|\mathbf{J}|\chi\rangle'$ and the similarity transformation (3.119) takes the form $\mathbf{J}' = \underline{u}^\dagger\,\mathbf{J}\,\underline{u}$.

Towards generating representations

Eq. (3.120) is the stepping stone for finding actual expressions for the (unknown) unitary transformations u in Hilbert space by linking them to (known and often familiar) expressions for proper rotations R in the real space of the laboratory. For instance, using *cartesian* components, $J'_\nu = \underline{u}\, J_\nu\, \underline{u}^\dagger$, the transformation is given by

$$J'_\nu = \hat{\mathbf{r}}'_\nu \cdot \mathbf{J} = R\hat{\mathbf{r}}_\nu \cdot \mathbf{J} = \sum_\mu [R^{-1}]_{\nu\mu} \hat{\mathbf{r}}_\mu \cdot \mathbf{J} = \sum_\mu [R^{-1}]_{\nu\mu} J_\mu = \sum_\mu [R]_{\mu\nu} J_\mu, \tag{3.122}$$

with $\mu, \nu \in \{x, y, z\}$. Here we used the property (M.13) of orthogonal matrices, $R^{-1} = R^T$. Hence, u and R are related by

$$J'_\nu = \underline{u}\, J_\nu\, \underline{u}^\dagger = \sum_\mu [R]_{\mu\nu} J_\mu. \tag{3.123}$$

The unitary transformations u can be represented by square $d \times d$ matrices, where $d = 2j + 1$ is the dimension of the subspace of $\mathbf{J}$ in Hilbert space. The rotations R are represented by 3×3 matrices $[R]_{\mu\nu}$. So, we could explore the correspondence between u and R as expressed by Eq. (3.123) starting from the 9 elements of the matrix $[R]_{\mu\nu}$. Although this approach is a good idea for systems with cubic symmetry (as happens in crystalline solids), for atoms (with their dominant central symmetry) we can do better by using a representation of R in terms of only 3 elements, the *Euler angles* α, β, γ (see Section 3.6.2). In Section 3.7.5 we will show that in this case the transformation is given by

$$J'_q = \underline{u}\, J_q\, \underline{u}^\dagger = \sum_{q'=-1}^{+1} J_{q'} \mathscr{D}^1_{q'q}(\alpha, \beta, \gamma). \tag{3.124}$$

In the language of *group theory* it is said that the representation $[R]_{\mu\nu}$ can be reduced to the representation $\mathscr{D}^1_{q'q}(\alpha, \beta, \gamma)$. In this representation $J_q \in \{J_{-1}, J_0, J_{+1}\}$ are the standard components of $\mathbf{J}$. Since an arbitrary rotation about a point cannot be represented by a set of less than 3 elements, the representation $\mathscr{D}^1_{q'q}(\alpha, \beta, \gamma)$ is called *irreducible*. .

3.6.2 Rotation in the euclidean space - Euler angles

First we discuss rotations of the coordinate system about the origin as expressed by Eq. (3.110). In matrix notation this passive rotation takes the form

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{3.125}$$

where x, y, z and x', y', z' are the cartesian coordinates of the position vector before and after the transformation, respectively. According to the *Euler rotation theorem* any proper rotation of the coordinate system $\mathcal{S}$ can be decomposed into three subsequent proper rotations about specified angles and axes, as illustrated in Fig. 3.4. Throughout these lectures we adopt the $z - y - z$ rotation sequence convention of Rose [94] (see Fig. 3.4). First, a *positive* rotation $R_z(\alpha)$ of $\mathcal{S}$ to $\mathcal{S}'$ over the angle α about the *positive* z direction is given by[1]

$$R_z(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.126}$$

---

[1] The sense of rotation in the direction $\hat{\mathbf{r}}$ is called positive if the *coordinate system*, $\mathcal{S}$, rotates in the same way as a right-handed screw advances in the direction $\hat{\mathbf{r}}$. By convention, the rotation angle *increases* for a *positive* rotation.

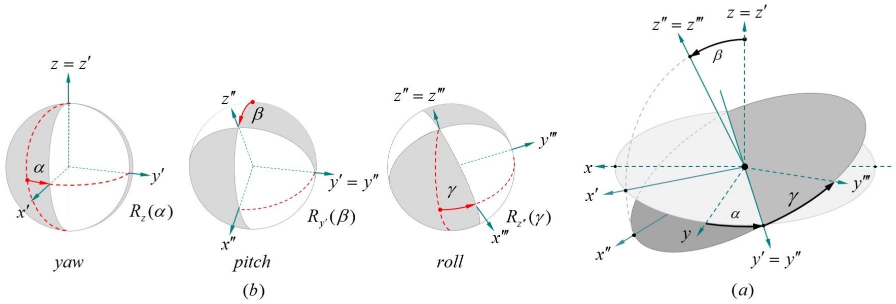


Figure 3.4: (a) an arbitrary *passive* rotation can be decomposed into three subsequent positive rotations of a coordinate system over the Euler angles α, β and γ called *yaw*, *pitch* and *roll* in aviation; (b) the same three Euler rotations shown for a coordinate system attached to an *imaginary* ball for clarity of illustration. All figures show the orientation *after* rotation over the angle indicated by the arrow. The right-handedness of coordinate system and positive sense of rotation, in combination with the $z-y-z$ rotation sequence for the Euler angles correspond to the convention of Rose [94] (beware of other conventions in the literature).

In aviation this rotation is called *yaw* and corresponds to a change of *heading*. Second, a positive rotation $R_{y'}(\beta)$ of $\mathcal{S}'$ to $\mathcal{S}''$ over the angle β about the y' axis (the y axis of $\mathcal{S}'$) is given by

$$R_{y'}(\beta) = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix}. \tag{3.127}$$

In aviation β is called the *pitch* angle. Finally, a rotation $R_{z''}(\gamma)$ of $\mathcal{S}''$ to $\mathcal{S}'''$ over the angle γ about the z'' axis (the z axis of $\mathcal{S}''$) is, analogously to $R_z(\alpha)$, given by

$$R_{z''}(\gamma) = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.128}$$

In aviation γ is called the *roll* angle. Note that $R_z(\alpha)$, $R_{y'}(\beta)$ and $R_{z''}(\gamma)$ are orthogonal matrices with determinant +1 (see Appendix M). Hence, in using the Euler angles, the rotations are specified with respect to the coordinate system fixed to an observer (the *pilot*) experiencing the rotations,

$$\hat{\mathbf{r}}' = R(\alpha,\beta,\gamma)\,\hat{\mathbf{r}} = R_{z''}(\gamma)R_{y'}(\beta)R_z(\alpha)\,\hat{\mathbf{r}}. \tag{3.129}$$

Evaluating the matrix product we find

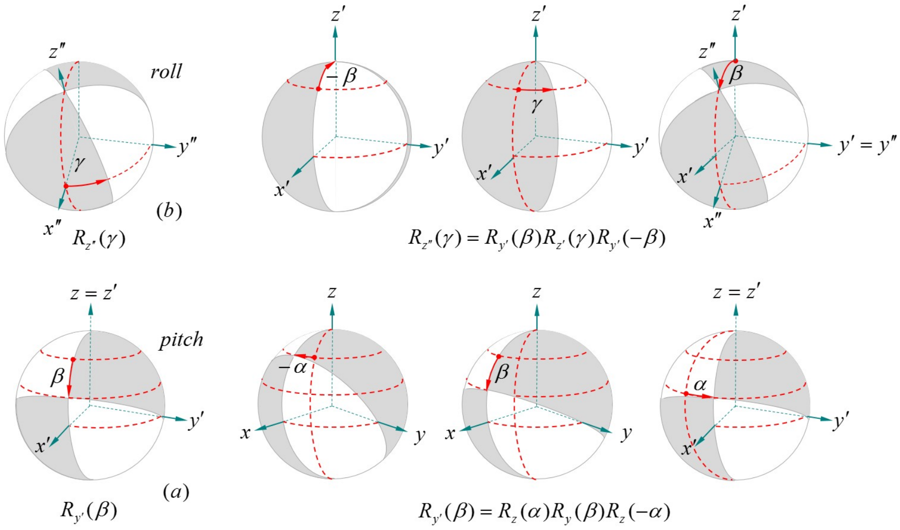
$$R = \begin{pmatrix} -\sin\alpha\sin\gamma + \cos\alpha\cos\beta\cos\gamma & \cos\alpha\sin\gamma + \sin\alpha\cos\beta\cos\gamma & -\sin\beta\cos\gamma \\ -\sin\alpha\cos\gamma - \cos\alpha\cos\beta\sin\gamma & \cos\alpha\cos\gamma - \sin\alpha\cos\beta\sin\gamma & \sin\beta\sin\gamma \\ \cos\alpha\sin\beta & \sin\alpha\sin\beta & \cos\beta \end{pmatrix}. \tag{3.130}$$

Interestingly, a pure pitch β about the y' axis can be decomposed into three subsequent rotations in the laboratory-fixed frame $\mathcal{S}$ (see Fig. 3.5a),

$$R_{y'}(\beta) = R_z(\alpha)R_y(\beta)R_z(-\alpha). \tag{3.131}$$

Likewise, a pure roll γ about the z'' axis can be decomposed into three subsequent rotations in the frame $\mathcal{S}'$ (see Fig. 3.5b),

$$R_{z''}(\gamma) = R_{y'}(\beta)R_{z'}(\gamma)R_{y'}(-\beta). \tag{3.132}$$



$R_{z''}(\gamma)$

$R_{z''}(\gamma)=R_{y'}(\beta)R_{z'}(\gamma)R_{y'}(-\beta)$

$R_{y'}(\beta)$

$R_{y'}(\beta)=R_{z}(\alpha)R_{y}(\beta)R_{z}(-\alpha)$

Figure 3.5: (a) the Euler rotation $R(\beta,y')$, i.e., pure *pitch* of the coordinate system $\mathcal{S}'=(x',y',z')$, is equivalent with three subsequent rotations of the frame $\mathcal{S}=(x,y,z)$; (b) the Euler rotation $R(\gamma,z'')$, i.e., pure *roll* of the coordinate system $\mathcal{S}''=(x'',y'',z'')$, is equivalent with three subsequent rotations of the frame $\mathcal{S}'$. All figures show the orientation *after* rotation over the angle indicated by the arrow.

Substituting Eqs. (3.131) and (3.132) into Eq. (3.129) and using the commutation of $R_z(\alpha)$ and $R_{z'}(\gamma)$ (note that the z axis coincides with the z' axis) we find that the rotation R can also be realized by three subsequent rotations about laboratory-fixed axes,

$$R(\alpha,\beta,\gamma)=R_z(\alpha)R_y(\beta)R_z(\gamma). \tag{3.133}$$

Comparing Eqs. (3.133) and (3.129) we note that the same angles appear (α,β,γ) but *in reverse order*.

To conclude this section we point out that the set of all passive rotations in real space of a sphere about its center constitutes a *group*. If the general element of this group is given by $R(\alpha,\beta,\gamma)$, we find that $R(-\gamma,-\beta,-\alpha)$ is its inverse and $R(0,0,0)$ is the unit element. Furthermore, the set is closed under the group operation and successive operations are associative. This group is called the *special orthogonal group* SO(3); i.e., the group of all orthogonal coordinate transformations with determinant +1 in the three-dimensional real space.

3.6.3 Unitary transformation in Hilbert space for the case $s=1/2$

Let us now leave real space and turn to the unitary transformations (3.114) and (3.119) in Hilbert space corresponding to the rotations R (the z axis being the quantization axis). In view of the special importance of two-level systems we demonstrate this for the case $s=\frac{1}{2}$. In Section 3.7.1 our findings will be generalized to the case of arbitrary angular momentum. For $s=\frac{1}{2}$ Eq. (3.120) can be written as a transformation of the Pauli matrices,

$$\boldsymbol{\sigma}\cdot R\,\hat{\mathbf{r}}=\underline{u}\,\boldsymbol{\sigma}\,\underline{u}^{\dagger}\cdot\hat{\mathbf{r}}. \tag{3.134}$$

To start the discussion we point out that the most general unitary 2×2 matrix is of the form (see Problem 3.9)

$$\underline{u}\equiv\sqrt{\Delta}\,U=\pm\sqrt{\Delta}\begin{pmatrix} a & b\\ -b^{*} & a^{*}\end{pmatrix}, \tag{3.135}$$

where $aa^* + bb^* = 1$ and $\Delta \equiv \det \underline{u}$ can be expressed in the form of a phase factor ($\Delta = e^{i\xi}$). The matrix U, with $\det U = 1$, is called the *special unitary matrix*. It is fully defined by the parameters a and b (*Cayley-Klein* parameters). Substituting Eq. (3.135) into the r.h.s. of Eq. (3.134) we obtain an expression for the most general unitary transformation of $\sigma_r = \boldsymbol{\sigma} \cdot \hat{\mathbf{r}}$,

$$\underline{u}\,\boldsymbol{\sigma}\,\underline{u}^\dagger \cdot \hat{\mathbf{r}} = \begin{pmatrix} A_x x + A_y y + A_z z & B_x x + B_y y + B_z z \\ (B_x x + B_y y + B_z z)^* & -(A_x x + A_y y + A_z z) \end{pmatrix}, \tag{3.136}$$

where the coefficients A_i and B_i, with $i \in \{x, y, z\}$, are defined in terms of the matrix elements a and b by the following fundamental expressions

$$\begin{array}{lll} A_x = ab^* + ba^* & A_y = -i\,(ab^* - ba^*) & A_z = aa^* - bb^* \\ B_x = a^2 - b^2 & B_y = -i\left(a^2 + b^2\right) & B_z = -2ab. \end{array} \tag{3.137}$$

Note from Eq. (3.136) that the phase factor $\Delta = e^{i\xi}$ has dropped out of the analysis. This means that we are free to choose Δ. The obvious choice is $\Delta = 1$, which means that the analysis can be restricted to the special unitary matrix U.

To determine the $U_z(\alpha)$ corresponding to the Euler rotation $R_z(\alpha)$ we substitute Eq. (3.126) into the matrix equation (3.134) and obtain after some matrix manipulation

$$\boldsymbol{\sigma} \cdot R_z(\alpha)\hat{\mathbf{r}} = \begin{pmatrix} z & e^{i\alpha}x - ie^{i\alpha}y \\ e^{-i\alpha}x + ie^{-i\alpha}y & -z \end{pmatrix}. \tag{3.138}$$

It is instructive to compare this expression with the non-rotated form of σ_r as given in Eq. (3.49). Equating the matrices (3.138) and (3.136) gives $A_x = A_y = 0$, $A_z = 1$ en $B_x = e^{i\alpha}, B_y = -ie^{i\alpha}, B_z = 0$. Further comparison with the fundamental expression for A_z in (3.137) yields $A_z = aa^* - bb^* = 1$. In combination with the property $aa^* + bb^* = 1$ we find $b = 0$. Substituting this result into the fundamental expression for B_x we find $B_x = a^2 = e^{i\alpha}$. Thus, we arrive at $a = e^{i\alpha/2}$ and obtain for the special unitary matrix corresponding to $R_z(\alpha)$

$$U_z(\alpha) = \pm \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix}. \tag{3.139}$$

Note that the positive and negative solution are connected by a rotation over 2π,

$$-U_z(\alpha) = U_z(2\pi)U_z(\alpha) = U_z(\alpha + 2\pi) = U_{-z}(-\alpha - 2\pi). \tag{3.140}$$

Apparently, after rotating over 2π we evolve from one solution to the other.

Likewise, for the Euler rotation $R_y(\beta)$ we obtain $A_x = \sin\beta$, $A_y = 0$, $A_z = \cos\beta$ and $B_x = \cos\beta$, $B_y = -i$, $B_z = -\sin\beta$. Comparing the result for B_x and B_y with the fundamental expressions in (3.137) we find $\cos\beta = 2a^2 - 1$, which implies $a = \cos(\beta/2)$. The comparison of A_x and A_y with the fundamental expressions in (3.137) yields $\sin\beta = 2ab^*$, which after the substitution of the result for a leads to $b = \sin(\beta/2)$. Thus we obtain for the special unitary matrix corresponding to $R_y(\beta)$

$$U_y(\beta) = \pm \begin{pmatrix} \cos(\beta/2) & \sin(\beta/2) \\ -\sin(\beta/2) & \cos(\beta/2) \end{pmatrix}, \tag{3.141}$$

which satisfies the same rotation property as Eq. (3.140),

$$-U_y(\beta) = U_y(2\pi)U(2\pi) = U_y(\beta + 2\pi) = U_{-y}(-\beta - 2\pi). \tag{3.142}$$

Next we turn to the general case. From Eq. (3.133) we know that an arbitrary rotation of the coordinate system can be written as the product of three proper rotations over the Euler angles α, β and γ,

$$R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma). \tag{3.143}$$

The corresponding unitary transformation is found by applying the product rule (3.121) to Eq. (3.134),

$$U(\alpha,\beta,\gamma) = U_z(\alpha)U_y(\beta)U_z(\gamma) \tag{3.144}$$

$$= \pm\begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2}\end{pmatrix}\begin{pmatrix} \cos(\beta/2) & \sin(\beta/2) \\ -\sin(\beta/2) & \cos(\beta/2)\end{pmatrix}\begin{pmatrix} e^{i\gamma/2} & 0 \\ 0 & e^{-i\gamma/2}\end{pmatrix}. \tag{3.145}$$

The set of all 2×2 unitary matrices with determinant $+1$ constitutes a group: the special unitary group, SU(2). Writing the general element of this group as $\pm U(\alpha,\beta,\gamma)$, we find for $U(\alpha,\beta,\gamma)$ that $U(-\gamma,-\beta,-\alpha)$ is its inverse and $U(0,0,0)$ is the unit element; for $-U(\alpha,\beta,\gamma)$ the inverse is $-U(-\gamma,-\beta,-\alpha)$ and $-U(2\pi,2\pi,2\pi)$ the unit element. For both branches, the set is closed under the group operation and successive operations are associative. As the unitary operators $U(\alpha,\beta,\gamma)$ and $-U(\alpha,\beta,\gamma)$ are connected by Euler rotations over 2π we can *equally well work with one branch*, using either $U(\alpha,\beta,\gamma)$ or $-U(\alpha,\beta,\gamma)$ to represent the rotation, *provided we double the domain of the Euler angles* from an interval of 2π to an interval of 4π. This is sometimes compared to transforming a circular band into a *Möbius band*. From here on we shall use $+U(\alpha,\beta,\gamma)$ along with the 4π domain. One may argue that there is a certain elegance in using $U(\alpha,\beta,\gamma)$ rather than $-U(\alpha,\beta,\gamma)$ because the former connects to the unit matrix for $\alpha,\beta,\gamma\to 0$ (rather than for $\alpha,\beta,\gamma\to 2\pi$) but this is already a matter of taste.

Eq. (3.134) defines a double-valued function on the domain defined by the elements of SO(3), $R(\alpha,\beta,\gamma)\mapsto \pm U(\alpha,\beta,\gamma)$. In this sense, SU(2) has "twice as many elements" as SO(2). More precisely, there exists a two-to-one homomorphism from SU(2) onto SO(3),[1]

$$\pm U(\alpha,\beta,\gamma)\mapsto R(\alpha,\beta,\gamma).$$

The group SO(3) is said to be doubly covered by SU(2). The double covering can be reduced to a single cover by extending the domain of the Euler angles to 4π. We return to the physical significance of this double covering in Section 3.7.3.

Problem 3.9. Show that any unitary 2×2 matrix $\underline{u}$ can be written in the form

$$\underline{u} = \pm\sqrt{\Delta}\begin{pmatrix} a & b \\ -b^* & a^*\end{pmatrix},$$

where $aa^* + bb^* = 1$ and $\Delta \equiv \det\underline{u} = e^{i\xi}$ with ξ an arbitrary phase factor.

Solution. Since $\underline{u}$ is unitary the determinant of $\underline{u}^*$ can be expressed as $\Delta^* = \det\underline{u}^* = \det\underline{u}^\dagger = \det\underline{u}^{-1} = \Delta^{-1}$. This implies $|\Delta|^2 = 1$ and, hence $\Delta = e^{i\xi}$, generally an arbitrary phase factor. To demonstrate that the most general unitary 2×2 matrix can be expressed in the given form we start from

$$\underline{u} = \begin{pmatrix} a & b \\ c & d\end{pmatrix}.$$

Using Eq. (M.23) for the inverse we can equate $\underline{u}^\dagger$ and $\underline{u}^{-1}$

$$\begin{pmatrix} a^* & c^* \\ b^* & d^*\end{pmatrix} = \underline{u}^\dagger = \underline{u}^{-1} = \frac{1}{\Delta}\begin{pmatrix} d & -b \\ -c & a\end{pmatrix}.$$

Thus we find $d = a^*\Delta$ and $c = -b^*\Delta$ and

$$\underline{u} = \begin{pmatrix} a & b \\ -b^*\Delta & a^*\Delta\end{pmatrix} = \pm\sqrt{\Delta}\begin{pmatrix} \pm a\Delta^{-1/2} & \pm b\Delta^{-1/2} \\ \mp b^*\Delta^{1/2} & \pm a^*\Delta^{1/2}\end{pmatrix}.$$

Redefining $\pm a\,\Delta^{-1/2}\to a$ and $\pm b\,\Delta^{-1/2}\to b$ we obtain the desired result. □

---

[1] A homomorphism is a map from one space to another in which the algebraic structure is conserved.

3.6.4 Infinitesimal rotation operators - the case $s = 1/2$

Let us have a closer look at the unitary transformation (3.139) corresponding to a passive rotation over the angle α about the z axis,

$$U_z(\alpha) = \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix}. \tag{3.146}$$

Using the well-known relation $e^{\pm i\varphi} = \cos\varphi \pm i\sin\varphi$ we can write $U_z(\alpha)$ in the form

$$U_z(\alpha) = \mathbb{1}\cos(\alpha/2) + i\sigma_z \sin(\alpha/2). \tag{3.147}$$

Note the appearance of the Pauli matrix σ_z. The change of $U_z(\alpha)$ by an *infinitesimal* passive rotation about the z axis is given by the partial derivative with respect to α and evaluates to

$$\frac{\partial U_z(\alpha)}{\partial\alpha} = (i\sigma_z/2)U_z(\alpha). \tag{3.148}$$

Since $U_z(0) = \mathbb{1}$ the above expression shows that σ_z satisfies the relation

$$i\sigma_z/2 = \left.\frac{\partial U_z(\alpha)}{\partial\alpha}\right|_{\alpha=0}. \tag{3.149}$$

Hence, the Pauli matrix σ_z can be obtained from the unitary transformation $U_z(\alpha)$ in the limit $\alpha \to 0$. For arbitrary α the differential equation (3.148) is solved by the exponential operator

$$U_z(\alpha) = e^{i\alpha\sigma_z/2}. \tag{3.150}$$

In other words, to determine $U_z(\alpha)$ for an arbitrary angle α all we need to know is the operator $i\sigma_z/2$. For this reason $i\sigma_z/2$ is called the *generator* of rotation about the z axis. In the language of group theory $i\sigma_z/2$ is one of the generators of the group SU(2) and $U_z(\alpha)$ is a representation of one of the elements of this group.

Since our choice of quantization axis was arbitrary Eq. (3.150) suggests the generalization

$$U_r(\varphi) = e^{i\varphi\sigma_r/2}, \tag{3.151}$$

where $U_r(\varphi)$ is the unitary transformation for a passive rotation (of the $s = \frac{1}{2}$ system) over the angle φ about the direction $\hat{\mathbf{r}}$, with $\sigma_r \equiv \hat{\mathbf{r}} \cdot \boldsymbol{\sigma}$. The correctness of this generalization follows in a few steps by expansion of the exponential operator - see Problem 3.10,

$$U_r(\varphi) = e^{i(\varphi/2)\sigma_r} = \mathbb{1}\cos(\varphi/2) + i\sigma_r \sin(\varphi/2). \tag{3.152}$$

Note that by specializing (3.152) to $\hat{\mathbf{r}} \to \hat{\mathbf{z}}$ and $\hat{\mathbf{r}} \to \hat{\mathbf{y}}$ we immediately regain the unitary matrices (3.139) and (3.141), respectively. For other directions the matrices are more complicated because σ_r only has a simple form along the x, y and z axes. In any case, the half-angle notation nicely reminds us of the domain doubling of the Euler angles.

Problem 3.10. Show that the unitary operator $U_r(\varphi)$ for the passive rotation of a $s = \frac{1}{2}$ system over the angle α about the direction $\hat{\mathbf{r}}$ can be written in the form

$$e^{i(\alpha/2)\sigma_r} = \mathbb{1}\cos(\alpha/2) + i\sigma_r \sin(\alpha/2).$$

Solution. For $s = \frac{1}{2}$ we have

$$U_r(\alpha) = e^{i\alpha S_r/\hbar} = e^{i(\alpha/2)\sigma_r} = \sum_n \frac{i^n (\alpha/2)^n \sigma_r^n}{n!}$$

Since $\sigma_r^2 = \mathbb{1}$, see Eq. (3.50), we have $\sigma_r^{2n} = 1$ and $\sigma_r^{2n+1} = \sigma_r$ and the expansion can be separated into its even and odd terms,

$$e^{i(\alpha/2)\sigma_r} = \mathbb{1}\sum_n \frac{(-1)^n (\alpha/2)^{2n}}{(2n)!} + i\sigma_r \sum_n \frac{(-1)^n (\alpha/2)^{2n+1}}{(2n+1)!}.$$

Recognizing the expansions for the sine and the cosine, this expression takes the desired form. □

3.6.5 Infinitesimal rotation operators - generalization

Let us return to Eq. (3.152). This expression holds for unitary transformations corresponding to passive rotations over a *finite* angle. Its validity is restricted to the case $s = \frac{1}{2}$ because the sine/cosine decomposition relies on the property $\sigma_r^2 = \mathbb{1}$ of the Pauli matrix (see Problem 3.10). Interestingly, this restriction does not hold for *infinitesimal* rotations. As we shall see below, an infinite product of infinitesimal rotations also leads to the exponential relation (3.151), even if the condition $\sigma_r^2 = \mathbb{1}$ is *not* imposed.

To demonstrate this generalization, we analyze $U_r(\varphi)$ as the product of n subsequent rotations over the angle φ/n (about the direction $\hat{\mathbf{r}}$). For $n \to \infty$ this product becomes an infinite product of infinitesimal rotations,

$$U_r(\varphi) = \lim_{n\to\infty} \left[U_r(\varphi/n)\right]^n . \tag{3.153}$$

For infinitesimal angles Eq. (3.152) reduces to

$$U_r(\varphi/n)|_{n\to\infty} = \mathbb{1} + (i\varphi/2n)|_{n\to\infty}\sigma_r. \tag{3.154}$$

Recalling Eq. (3.44) we substitute $i\sigma_r/2 = iS_r/\hbar$. Evaluating the infinite product (3.153) we find

$$U_r(\varphi) = \lim_{n\to\infty} \left(\mathbb{1} + i\frac{\varphi}{n}S_r/\hbar\right)^n = e^{i\varphi S_r/\hbar}. \tag{3.155}$$

This result is obtained without imposing the condition $\sigma_r^2 = \mathbb{1}$ (see Problem 3.11). Rearranging Eq. (3.154) we find that any operator $iS_r/\hbar$ that satisfies the relation (3.154) can be interpreted as the operator for an infinitesimal small rotation about the direction $\hat{\mathbf{r}}$,

$$iS_r/\hbar = \lim_{n\to\infty} \frac{U_r(\varphi/n) - \mathbb{1}}{\varphi/n} = \left.\frac{\partial U_r(\varphi)}{\partial\varphi}\right|_{\varphi=0} . \tag{3.156}$$

Problem 3.11. Show that the unitary operator $U_r(\varphi)$, corresponding to a passive rotation over the angle φ about the direction $\hat{\mathbf{r}}$, can be written as an exponential operator of $S_r \equiv \hat{\mathbf{r}} \cdot \mathbf{S}$

$$U_r(\varphi) = \lim_{n\to\infty} \left(\mathbb{1} + i\frac{\varphi}{n}S_r/\hbar\right)^n = e^{i\varphi S_r/\hbar},$$

irrespective of the value of the quantum number s.

Solution. The unitary operator for an infinitesimally small (but nonzero) passive rotation over the angle $(\varphi/n)|_{n\to\infty}$ about the direction $\hat{\mathbf{r}}$, is given by

$$U_r(\varphi/n)|_{n\to\infty} = \mathbb{1} + (i\frac{\varphi}{n}S_r/\hbar)|_{n\to\infty}.$$

Substituting this expression into (3.153) we obtain

$$U_r(\varphi) = \lim_{n\to\infty} \left(\mathbb{1} + i\frac{\varphi}{n}S_r/\hbar\right)^n = \lim_{n\to\infty} \sum_{k=0}^{n} \binom{n}{k} (i\varphi S_r/\hbar)^k \text{ with } \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Substituting the values for the binomial coefficients we obtain

$$U_r(\varphi) = \lim_{n\to\infty} \left[\left(1 + i\varphi S_r/\hbar + \frac{n(n-1)}{2!}\frac{1}{n^2}(i\varphi S_r/\hbar)^2 + \frac{n(n-1)(n-2)}{3!}\frac{1}{n^3}(i\varphi S_r/\hbar)^3 + \cdots\right.\right.$$
$$\left.\cdots + n\frac{1}{n^{n-1}}(i\varphi S_r/\hbar)^{n-1} + \frac{1}{n^n}(i\varphi S_r/\hbar)^n\right).$$

Rewriting this expression as an expansion in powers of $(1/n)$ it becomes

$$\begin{aligned} U_r(\varphi) &= \lim_{n\to\infty} \left(1 + i\varphi S_r/\hbar + \frac{1}{2!}(1-\frac{1}{n})(i\varphi S_r/\hbar)^2 + \frac{1}{3!}(1-\frac{3}{n}+\frac{2}{n^2})(i\varphi S_r/\hbar)^3 + \cdots\right) \\ &= 1 + i\varphi S_r/\hbar + \frac{1}{2!}(i\varphi S_r/\hbar)^2 + \frac{1}{3!}(i\varphi S_r/\hbar)^3 + \cdots , \end{aligned}$$

where all terms depending on n have vanished. In the last line we recognize the expansion of the exponential operator - see Eq. (L.31). As no presumptions are made with respect to the actual value of the spin, this result holds for any value of s. □

3.7 Angular momentum

3.7.1 Introduction

In Section 3.1 we established that a vector operator $\mathbf{J}$ caries the properties of angular momentum (as introduced in Chapter 1) if its cartesian components J_x, J_y and J_z are hermitian operators that satisfy the commutation relations (3.1). In the present section we shall make a fresh start and define angular momentum as an infinitesimal rotation imposed on a physical system. As we shall see, this definition leads in a few steps to conditions of Section 3.1.

To set the stage, we consider a *stationary state* of the Schrödinger Hamiltonian, $\psi_{nlm}(\mathbf{r})$, where $\mathbf{r} = (r, \theta, \phi)$ is the position with respect to the center of rotation (in spherical coordinates). We introduce an operator $U_z(\Delta\phi)$ to impose on ψ_{nlm} a passive rotation over the angle $\Delta\phi$ about the z axis. At this point $U_z(\Delta\phi)$ is unknown but has to be unitary in order to conserve the normalization of the state under rotation. Then, the change of ψ_{nlm} by an *infinitesimal* passive rotation about the z axis can be expressed in terms of $U_z(\Delta\phi)$ by evaluating the partial derivative of ψ_{nlm} with respect to ϕ,

$$\begin{aligned}\frac{\partial\psi_{nlm}(r,\theta,\phi)}{\partial\phi} &\equiv \lim_{\Delta\phi\to 0}\frac{\psi_{nlm}(r,\theta,\phi+\Delta\phi)-\psi_{nlm}(r,\theta,\phi)}{\Delta\phi}\\ &= \lim_{\Delta\phi\to 0}\frac{U_z(\Delta\phi)\,\psi_{nlm}(r,\theta,\phi)-\psi_{nlm}(r,\theta,\phi)}{\Delta\phi}\\ &= \lim_{\Delta\phi\to 0}\frac{U_z(\Delta\phi)-1}{\Delta\phi}\psi_{nlm}(r,\theta,\phi)\,. \end{aligned} \tag{3.157}$$

As this expression holds for arbitrary $\psi_{nlm}(\mathbf{r})$ the partial derivative can be expressed in the form

$$\frac{\partial}{\partial\phi} = \lim_{n\to\infty}\frac{U_z(\phi/n)-1}{\phi/n} = \left.\frac{\partial U_z(\phi)}{\partial\phi}\right|_{\phi=0}. \tag{3.158}$$

Here we defined $\Delta\phi \equiv \phi/n$, with integer n. In this notation, the (unknown) unitary operator $U_z(\phi/n)\,|_{n\to\infty}$ corresponds to a passive rotation over the *infinitesimal* angle $\delta\phi = (\phi/n)\,|_{n\to\infty}$ about the z axis.

At this point we leave the mathematical convention of passive rotations to give preference to physical rotations of the system. In this way we conform ourselves to the convention of Chapter 1 in which the orbital angular momentum, $\mathbf{L} = \mathbf{r}\times\mathbf{p}$, is defined as a right-handed rotation of the physical system. This change of convention is simple to implement because a *physical* rotation over the angle α about the arbitrary direction $\hat{\mathbf{a}}$ is equivalent to a *passive* rotation over the angle $-\alpha$ about the same axis. Then, the unitary operator $P_a(\alpha)$ corresponding to a *physical* rotation over the angle α about the direction $\hat{\mathbf{a}}$ is defined as

$$P_a(\alpha) \equiv U_a(-\alpha)\,. \tag{3.159}$$

As an aside we mention that by inverting the direction of $\hat{\mathbf{a}}$ we have

$$P_{-a}(\alpha) = P_a(-\alpha) = U_a(\alpha)\,. \tag{3.160}$$

Recalling Eq. (1.25) we find that the operator for *orbital angular momentum* about the z axis is given by

$$L_z = -i\hbar\frac{\partial}{\partial\phi} = i\hbar\lim_{n\to\infty}\frac{P_z(\phi/n)-1}{\phi/n} = i\hbar\left.\frac{\partial P_z(\phi)}{\partial\phi}\right|_{\phi=0}, \tag{3.161}$$

where an explicit expression for $P_z(\phi)$ remains to be obtained. Note that L_z is hermitian.

3.7.2 Differential operators - formal definition of angular momentum operators

In the previous section we found that orbital angular momentum can be represented b with $\mu, \nu \in \{-1, 0, +1\}$, wherey a unitary operator which imposes an infinitesimal rotation on the wavefunction of a physical system. The expressions (3.156) and (3.161) suggest to define any type of angular momentum in terms of a differential operator,

$$-iJ_a/\hbar \equiv \lim_{n\to\infty} \frac{P_a(\varphi/n) - \mathbb{1}}{\varphi/n} = \left.\frac{\partial P_a(\varphi)}{\partial\varphi}\right|_{\varphi=0}. \tag{3.162}$$

Here $P_a\left(\varphi/n\right)|_{n\to\infty}$ is the unitary operator corresponding to the *physical* rotation over the infinitesimal angle $\delta\varphi = \varphi/n|_{n\to\infty}$ about the direction $\hat{\mathbf{a}}$ and $\mathbb{1}$ is the unit operator. Inverting Eq. (3.162) we obtain for $P_a(\delta\varphi)$ the operator identity

$$P_a\left(\varphi/n\right)|_{n\to\infty} = \mathbb{1} - (i\varphi/n)|_{n\to\infty} J_a/\hbar. \tag{3.163}$$

Note that this expression is unitary ($P_a^\dagger = P_a^{-1}$) provided the operator

$$J_a \equiv \hat{\mathbf{a}} \cdot \mathbf{J} \tag{3.164}$$

is hermitian ($J_a^\dagger = J_a$). The unitary operator for a rotation over the *finite* angle φ about the direction $\hat{\mathbf{a}}$ is given by the following infinite product of infinitesimal operators (see Problem 3.11),

$$P_a(\varphi) \equiv \lim_{n\to\infty}\left(\mathbb{1} - i\frac{\varphi}{n}J_a/\hbar\right)^n = e^{-i\varphi J_a/\hbar}. \tag{3.165}$$

The significance of Eqs. (3.162)-(3.165) can hardly be over-emphasized because the properties of angular momentum follow in a few lines from the definition (3.162). To convince ourselves, all we need to do is show that the criteria for angular momentum operators (as introduced in Section 3.1) are satisfied; i.e., J_a has to be hermitian (as we already established above) and the cartesian components J_x, J_y and J_z have to satisfy the commutation relations (3.1). As we will make no assumption about the type of system under rotation such as the dimension of the subspace in which $P_a(\varphi)$ operates, our definition holds for *any* type of angular momentum.

To obtain the commutation relations (3.1) we use a geometrical argument by noting that infinitesimal rotations do not commute. This is illustrated in Fig. 3.6. In the upper part of the figure ($a1$) we first rotate a ball over the angle $\delta_y = \varphi_y/n$ about the y axis and ($a2$) subsequently $\delta_x = \varphi_x/n$ about the x axis. In the lower part ($b1$) we first rotate over the angle δ_x about the x axis and ($b2$) subsequently δ_y about the y axis. Comparing ($a2$) and ($b2$) we find that in the latter case it takes a small additional rotation $\delta_y\delta_x = \varphi_x\varphi_y/n^2$ to realize the orientation of ($a2$). In the limit $n \to \infty$ this additional rotation is about the z axis as indicated in ($b3$),

$$\lim_{n\to\infty}\left[P_x(\varphi_x/n)P_y(\varphi_y/n) - P_z(\varphi_x\varphi_y/n^2)P_y(\varphi_y/n)P_x(\varphi_x/n)\right] = 0. \tag{3.166}$$

Expanding the exponential operators to lowest non-vanishing order in $1/n$ we obtain

$$\left(\varphi_x\varphi_y/n^2\right)(J_yJ_x - J_xJ_y)/\hbar^2 + i\left(\varphi_x\varphi_y/n^2\right)J_z/\hbar = 0 \quad\Leftrightarrow\quad [J_x, J_y] = i\hbar J_z \tag{3.167}$$

and by cyclic permutation we find also the other commutation relations of the set (3.1). From this point on all properties of angular momentum follow from the algebra developed in Section 3.1.

Example 1 - hydrogenic wavefunctions

At this point we are in the position to calculate actual rotations. First, we demonstrate this for a rotation of the orbital wavefunction $\psi_{nlm}(r,\theta,\phi)$ about the z axis. Specializing to orbital angular momentum we calculate

$$P_z(\varphi)\psi_{nlm}(r,\theta,\phi) = e^{-i\varphi L_z/\hbar}\psi_{nlm}(r,\theta,\phi) = e^{-im\varphi/\hbar}\psi_{nlm}(r,\theta,\phi), \tag{3.168}$$

where we used $L_z\psi_{nlm} = -i\hbar\partial_\phi\psi_{nlm} = m\hbar\psi_{nlm}$ (see Section 1.1.3). In particular, we have $P_z(2\pi)\psi_{nlm}(r,\theta,\phi) = \psi_{nlm}(r,\theta,\phi)$.

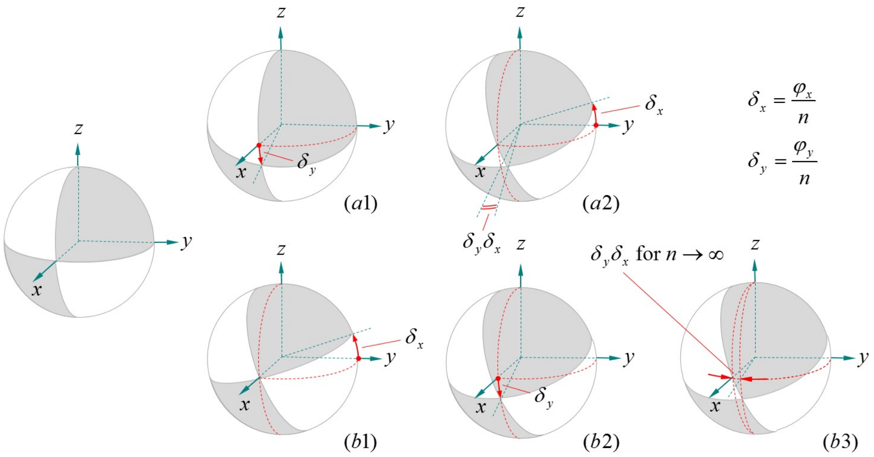


Figure 3.6: Infinitesimal rotations do not commute. Upper part: ($a1$) rotation of a ball about the y axis; ($a2$) rotation about the x axis. Lower part: ($b1$) rotation of the ball about the x axis; ($b2$) rotation about y axis; ($b3$) additional rotation required to obtain the same orientation as shown in ($a2$). All figures show the orientation of the ball *after* rotation over the angle indicated by the arrow.

Example 2 - angular momentum states in Dirac notation

Turning to the general case, we consider the eigenstates $\{|j,m\rangle\}$ of the angular momentum operators $\mathbf{J}^2$ and J_z. In this case the effect of an arbitrary rotation about the direction $\hat{\mathbf{a}}$ is given by

$$P_a(\varphi)|j,m\rangle = \sum_{m'} |j,m'\rangle\langle j,m'|e^{-i\varphi J_a/\hbar}|j,m\rangle. \tag{3.169}$$

The unitary matrix

$$[P_a(\varphi)]_{m'm} \equiv \langle j,m'|e^{-i\varphi J_a/\hbar}|j,m\rangle, \tag{3.170}$$

is called a *linear representation* of the rotation operator, $P_a(\varphi)$. In principle, the operators $P_a(\varphi)$, $\mathbb{1}$ and J_a operate on vectors of the full Hilbert space of $\mathbf{J}^2$ and J_z but as the $P_a(\varphi)$ do not mix states differing in j we can restrict ourselves - for given j - to the corresponding $d = 2j+1$ dimensional subspace V^d of Hilbert space. This block-diagonal form of the matrix (3.170) is called the *irreducible* form of the angular momentum representation. In particular, the eigenstates $|j,m\rangle$ transform in accordance with *one* of the irreducible representations of the full rotation group (the one corresponding to the value j). Recalling the identity (3.159), we note that by specializing the matrix (3.170) to the case $j = \frac{1}{2}$ we regain the Eqs. (3.139) and (3.141).

For a given basis $\{|j,m\rangle\}$ the unitary transformations $P_a(\varphi)$ (corresponding to proper rotations about the direction $\hat{\mathbf{a}}$ in real space) can be generated with the aid of Eq. (3.170). The set of all $P_a(\varphi)$ constitutes a *group*, the group SU(2). The general element of this group is $P_a(\varphi)$, $P_a(-\varphi)$ is its inverse and for $\varphi = 0$ we obtain the unit element. Moreover, the set is closed under the group operation. The operators J_a are called the *generators* of the group. The rotation matrices $[P_a(\varphi)]_{m'm}$ are called *representations* of the group. Depending on the dimension $d = 2j+1$ of the basis $\{|j,m\rangle\}$ the operators $P_z(\varphi)$ generate $d = 1,2,3,\cdots$ dimensional *irreducible* representations of the group SU(2). Note that in Section 3.6.3 the above was already established for the special case $s = \frac{1}{2}$ without introducing differential operators. This case is called the *fundamental* representation of the group SU(2).

3.7.3 Integral versus half-integral angular momentum

We are now equipped to analyze the effect of a physical rotation on a system of arbitrary angular momentum. This will lead us to an important insight in the difference between integer and half-integer angular momenta. We consider for this purpose a system of angular momentum j, with $\{|j,m\rangle\}$ being the basis defined by the operators $\mathbf{J}^2$ and J_z. An arbitrary state $|\chi\rangle$ of this system can be decomposed in the form

$$|\chi\rangle = \sum_{m=-j}^{j} |j,m\rangle\langle j,m|\chi\rangle. \tag{3.171}$$

Now we apply a physical rotation of this system over 2π about the z axis. Setting $\hat{\mathbf{a}} \to \hat{\mathbf{z}}$ and $\varphi \to 2\pi$ in Eq. (3.165) we find for the corresponding unitary transformation

$$P_z(2\pi)|\chi\rangle = e^{-i2\pi J_z/\hbar}|\chi\rangle = \sum_{m=-j}^{j} e^{-i2\pi m}|j,m\rangle\langle j,m|\chi\rangle. \tag{3.172}$$

Note that each term contains the same phase factor, $+1$ for integral j and -1 for half-integral j. Thus we obtain

$$P_z(2\pi)|\chi\rangle = \begin{cases} \;\;|\chi\rangle \text{ for integral } j \\ -|\chi\rangle \text{ for half-integral } j. \end{cases} \tag{3.173}$$

Note that for integral j the rotation properties are regular; i.e., all eigenfunctions of the decomposition are in phase after rotation of the physical system over 2π. This is readily verified for the spherical harmonics derived from the Schrödinger equation in Section 1.1.3. In contrast, for half-integral angular momentum we find $P_z(2\pi+\varphi) = -P_z(\varphi)$ and we have to *rotate over an additional* 2π (4π in total) to recover the original state $|\chi\rangle$. This was first demonstrated in 1975 in famous neutron interferometry experiments [90, 119]. Obviously, in the real space of the laboratory the range of angles $0 \le \varphi < 2\pi$ is not distinguishable from the range $2\pi \le \varphi < 4\pi$. Therefore, the unitary transformations of the group SU(2) capture a property of half-integral angular momenta that is absent in the real space rotation matrices of the group SO(3) - *spin differs from classical rotation.* For a given physical rotation over the angle φ (about the z axis) we can equally well choose $P_z(\varphi)$ or $-P_z(\varphi)$ to describe the corresponding unitary transformation. Recall that we arrived at same conclusion in Section 3.6.3. Apparently, the "wavefunctions" of half-integral spin systems are double-valued functions as was first noticed by Wolfgang Pauli [80]. These wavefunctions are called spinors to distinguish their rotation properties from those of the (single-valued) states of integer angular momenta (the spherical harmonics). Experimentally this double valuedness is of no consequence because the global phase of the state does not affect the expectation values of the angular momentum operators. In Section 3.8 we discuss how to generate matrix representations for unitary transformations of arbitrary angular momenta and illustrate this for the examples $s = \frac{1}{2}$ and $s = 1$.

Importantly, the double covering is of no consequence for unitary transformations of the *operators* because the transformation (3.119) involves U and $U^\dagger$ symmetrically. This causes the sign of U to drop out of the transformation; i.e., it does not affect expectation values - as expected for observables. In contrast, the double covering has important consequences for the *states* as these become double valued as expressed by Eqs. (3.140) and (3.142). So, whereas the operators transform like classical angular momentum operators, the transformation properties of the states have no classical analogue.

3.7.4 Physical rotation of angular momentum systems - general case

Our next task is to generalize the discussion to include rotations about arbitrary axes. According to the Euler rotation theorem, any *passive* rotation R can be written as three subsequent coordinate rotations over the Euler angles α, β and γ: $R(\alpha,\beta,\gamma) = R_{z''}(\gamma)R_{y'}(\beta)R_z(\alpha)$. To obtain the corresponding *physical* rotation $R_P(\alpha,\beta,\gamma)$ we invert the subsequent Euler rotations,

$$R_P(\alpha,\beta,\gamma) = R_{z''}(-\gamma)R_{y'}(-\beta)R_z(-\alpha) = R(-\alpha,-\beta,-\gamma). \tag{3.174}$$

Note that $R_P(\alpha,\beta,\gamma) \neq R^{-1}(\alpha,\beta,\gamma) = R(-\gamma,-\beta,-\alpha)$. Next, the physical rotation is expressed in terms of rotations about laboratory-fixed axes with the aid of Eq. (3.133),

$$R_P(\alpha,\beta,\gamma) = R(-\alpha,-\beta,-\gamma) = R_z(-\alpha)R_y(-\beta)R_z(-\gamma). \tag{3.175}$$

Changing to the corresponding unitary transformation (see Sections 3.6.4 and 3.7.1) we arrive at

$$P_R \equiv P(\alpha,\beta,\gamma) = U(-\alpha,-\beta,-\gamma) = P_z(\alpha)P_y(\beta)P_z(\gamma) = e^{-i\alpha J_z/\hbar}e^{-i\beta J_y/\hbar}e^{-i\gamma J_z/\hbar}. \tag{3.176}$$

The rotation matrices $\mathscr{D}^j_{m'm}$

Knowing the unitary matrices corresponding to the Euler rotations we can write the unitary transformations corresponding to an arbitrary physical rotation as

$$|\chi\rangle' = P_R|\chi\rangle = \sum_{m,m'=-j}^{+j} |j,m'\rangle\langle j,m'|P_R|j,m\rangle\langle j,m|\chi\rangle. \tag{3.177}$$

The rotation matrix element is commonly denoted by

$$\mathscr{D}^j_{m'm}(\alpha,\beta,\gamma) \equiv \langle j,m'|P_R(\alpha,\beta,\gamma)|j,m\rangle. \tag{3.178}$$

In view of the diagonality of P_z the rotation matrix simplifies to

$$\mathscr{D}^j_{m'm}(\alpha,\beta,\gamma) = \langle j,m'|e^{-i\alpha J_z/\hbar}P_y(\beta)e^{-i\gamma J_z/\hbar}|j,m\rangle = e^{im'\alpha}d^j_{m'm}(\beta)e^{-im\gamma}, \tag{3.179}$$

where the matrix $d^j_{m'm}(\beta)$ is given by

$$d^j_{m'm}(\beta) \equiv \langle j,m'|e^{-i\beta J_y/\hbar}|j,m\rangle. \tag{3.180}$$

It is straightforward to evaluate this matrix for any integral or half-integral value of j (see Problems 3.13 and 3.14 for the case $j=\frac{1}{2}$). A general formula for the matrix elements $d^j_{m'm}(\beta)$ was derived by Wigner [120]

$$d^j_{m'm}(\beta) = \sum_\kappa \frac{(-1)^\kappa\sqrt{(j+m')!(j-m')!(j+m)!(j-m)!}}{(j+m'-\kappa)!(j-m-\kappa)!\kappa!(\kappa-m'+m)!} \times [\cos(\beta/2)]^{2j+m'-m-2\kappa}\,[\sin(\beta/2)]^{2\kappa with \mu,\nu\in\{-1,0,+1\}, where -m'+m}, \tag{3.181}$$

where the summation runs over all values of κ for which the factorials are meaningful. Note that the $d^j_{m'm}(\beta)$ are *real*, which explains, in hindsight, the preference for the $z-y-z$ rotation sequence - See Fig. 3.4. The Wigner formula is easily embedded in computer algebra to manipulate angular momenta of arbitrary size on their generalized Bloch sphere.

Problem 3.12. Show that pure rotations about the x, y and z directions are given by

$$P_z(\alpha) = P(\alpha,0,0),\;\; P_y(\beta) = P(0,\beta,0),\;\; P_x(\beta) = P(-\tfrac{1}{2}\pi,\beta,\tfrac{1}{2}\pi).$$

3.7.5 Spherical tensor operators - irreducible tensor operators

Substituting an eigenstate into Eq. (3.177), $|\chi\rangle \to |k,q\rangle$, we obtain an expression for the transformation of eigenstates under rotation

$$|k,q\rangle' = \sum_{q'} |k,q'\rangle \mathscr{D}^k_{q'q}(\alpha,\beta,\gamma). \tag{3.182}$$

Turning to the position representation, $\langle \hat{\mathbf{r}}|k,q\rangle = Y^q_k(\hat{\mathbf{r}})$ and $\langle \hat{\mathbf{r}}|k,q\rangle' = Y^q_k(\hat{\mathbf{r}}')$, we obtain the transformation properties of the $Y^q_k(\hat{\mathbf{r}})$,

$$Y^q_k(\hat{\mathbf{r}}') = \sum_{q'} Y^{q'}_k(\hat{\mathbf{r}}) \mathscr{D}^k_{q'q}(\alpha,\beta,\gamma). \tag{3.183}$$

In particular, for $\hat{\mathbf{r}} = \hat{\mathbf{z}}$ only the term $q' = 0$ contributes to the summation - see Eq. L.55 - and we find the following relation between the spherical harmonics and the rotation matrices:

$$Y^q_k(\hat{\mathbf{r}}') = \sqrt{\frac{2k+1}{4\pi}} \mathscr{D}^k_{0,q}(\alpha,\beta,\gamma), \tag{3.184}$$

where $\hat{\mathbf{r}}' = R(\alpha,\beta,\gamma)\hat{\mathbf{z}}$. In Section 3.1.1 we established that the standard components of $\mathbf{J}$ transform like the $Y^q_1(\hat{\mathbf{r}})$. So, using the correspondence (3.31) we find for the transformation properties of the standard components of $\mathbf{J}$

$$J'_q = \sum_{q'=-1}^{+1} J_{q'} \mathscr{D}^1_{q'q}(\alpha,\beta,\gamma). \tag{3.185}$$

Combining this expression with Eq. (3.119) we arrive at Eq. (3.124) as announced at the end of Section 3.6.1,

$$J'_q = P_R(\alpha,\beta,\gamma) J_q P^\dagger_R(\alpha,\beta,\gamma) = \sum_{q'=-1}^{+1} J_{q'} \mathscr{D}^1_{q'q}(\alpha,\beta,\gamma). \tag{3.186}$$

The above procedure can be generalized by introducing *spherical tensor operators* $\mathrm{T}^{(k)}$ as operators with standard components that transform like $Y^q_k(\hat{\mathbf{r}})$ [93]. This restricts the rank of the tensor to *integer* values, $k \in \{0,1,\cdots\}$. Replacing in Eq. (3.183) the $Y^q_k(\hat{\mathbf{r}})$ by the standard components $T_{k,q} \in \{T_{k,-k},\cdots,T_{k,k}\}$ we obtain

$$T'_{kq} = P_R T_{kq} P^\dagger_R = \sum_{q'=-k}^{+k} T_{kq'} \mathscr{D}^k_{q'q}. \tag{3.187}$$

Importantly, as the $\mathscr{D}^k_{q'q}(\alpha,\beta,\gamma)$ are also defined for half-integer k the procedure can be further generalized to hold for *any irreducible tensor operator* T_{kq}. As was demonstrated by Giulio Racah [86] these operators satisfy the following commutation relations:

$$[J_z, T_{k,q}] = q\,\hbar\,T_{k,q} \tag{3.188a}$$

$$[J_\pm, T_{k,q}] = \sqrt{k\,(k+1) - q\,(q\pm 1)}\,\hbar\,T_{k,q\pm 1}. \tag{3.188b}$$

This property is demonstrated in Problem K.1. Irreducible tensor operators act within the invariant subspace of a *pure* angular momentum state (for given j, the subspace spanned by the basis $\{|j,m\rangle\}$, with $-j \le m \le j$ - see Section 3.1). Angular momentum operators are irreducible tensor operators of rank 1 ($T_{1,q} \to J_q$). For these operators the commutation relations (3.188) reduce to those of angular momentum (cf. Appendix K.1.2). The simplest class of irreducible tensor operators are the spherical tensor operators of rank 0. These are known as *scalar invariants*. They have a single component, T_{00}, which is invariant under rotation ($\mathscr{D}_{00}^{0} = 1$),

$$T'_{00} = P_R T_{00} P_R^{\dagger} = T_{00}. \tag{3.189}$$

Note that the Hamiltonian of systems of identical particles *cannot* involve irreducible tensor operators of *half-integral* rank since these would give rise to transitions between bosonic and fermionic states, which contradicts the experimental observation that the statistical nature of particles is rigorously conserved (cf. Section 3.4.3).

3.8 Generating angular momentum representations

To demonstrate the procedure for generating representations we discuss a few examples. The unitary matrix $P_z(\varphi)$ for a rotation over an angle φ about the $\hat{\mathbf{z}}$ direction is generated by the operator $e^{-i\varphi J_z/\hbar}$. The matrix representation follows with Eq. (3.170). For the diagonal representation this takes a minimal effort,

$$[P_z(\varphi)]_{m',m} \equiv \langle j, m'|e^{-i\varphi J_z/\hbar}|j, m\rangle = e^{-i\varphi m}\delta_{m',m} \tag{3.190}$$

and using the definition (3.162),

$$J_z = i\hbar \left.\frac{\partial P_z(\varphi)}{\partial \varphi}\right|_{\varphi=0}, \tag{3.191}$$

we obtain

$$[J_z]_{m',m} = m\hbar\delta_{m',m}e^{-i\varphi m}|_{\varphi=0} = m\hbar\delta_{m',m}. \tag{3.192}$$

This could have been written down immediately since

$$[J_z]_{m',m} \equiv \langle j, m'|J_z|j, m\rangle = m\hbar\delta_{m',m}. \tag{3.193}$$

The added value of the formalism becomes evident when asking for the matrix representation of non-diagonal angular momentum operators. For instance, the unitary operator $P_y(\theta)$ corresponds to a physical rotation over an angle θ about the $\hat{\mathbf{y}}$ direction. In this case the matrix representation follows with the Wigner formula,

$$[P_y(\theta)]_{m',m} = \langle j, m'|e^{-i\theta J_y/\hbar}|j, m\rangle = d^j_{m'm}(\theta). \tag{3.194}$$

and

$$\langle j, m'|J_y|j, m\rangle = i\hbar \left.\frac{\partial d^j_{m'm}(\theta)}{\partial \theta}\right|_{\theta=0}. \tag{3.195}$$

3.8.1 Example - the case $j = 1/2$

To demonstrate the procedure for generating representations we first consider the example of angular momentum $j = \frac{1}{2}$ and rederive the results of Section 3.2.3. The unitary matrix $P_z(\varphi)$ for a rotation over an angle φ about the $\hat{\mathbf{z}}$ direction follows with Eq. (3.190). For $j = \frac{1}{2}$ we find

$$P_z(\varphi) = \begin{pmatrix} e^{-i\varphi/2} & 0 \\ 0 & e^{i\varphi/2} \end{pmatrix}. \tag{3.196}$$

To determine J_z we turn to the definition (3.162). First, we calculate the derivative of $P_z(\varphi)$,

$$\left.\frac{\partial P_z(\varphi)}{\partial \varphi}\right|_{\varphi=0} = \begin{pmatrix} -(i/2)e^{-i\varphi/2}|_{\varphi=0} & 0 \\ 0 & (i/2)e^{i\varphi/2}|_{\varphi=0} \end{pmatrix} = -i\begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}. \tag{3.197}$$

Substituting this result into Eq. (3.162) we obtain for the angular momentum operator

$$J_z = i\hbar \left.\frac{\partial P_z(\varphi)}{\partial\varphi}\right|_{\varphi=0} = \tfrac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3.198}$$

The corresponding eigenvectors are

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.199}$$

There are various equivalent ways to proceed. First of all we can use again the method demonstrated for J_z. To determine J_y we start from the unitary operator $P_y(\theta)$, representing a physical rotation over an angle θ about the $\hat{\mathbf{y}}$ direction. For $j = \frac{1}{2}$ the operator $P_y(\theta)$ follows with Eq. (3.194)

$$P_y(\theta) = d^{1/2}_{m'm}(\theta) = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix}. \tag{3.200}$$

Using the definition (3.162) we calculate the derivative of $P_z(\varphi)$,

$$\left.\frac{\partial P_y(\theta)}{\partial\theta}\right|_{\theta=0} = \tfrac{1}{2} \begin{pmatrix} -\sin(\theta/2)\,|_{\theta=0} & -\cos(\theta/2)\,|_{\theta=0} \\ \cos(\theta/2)\,|_{\theta=0} & -\sin(\theta/2)\,|_{\theta=0} \end{pmatrix} = \tfrac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{3.201}$$

Substituting this result into Eq. (3.162) we obtain for the angular momentum operator

$$J_y = i\hbar \left.\frac{\partial P_y(\varphi)}{\partial\varphi}\right|_{\varphi=0} = \tfrac{1}{2}\hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{3.202}$$

3.8.1.1 Transformation of the cartesian components of J

An alternative approach to obtain J_x and J_y is by unitary transformation of J_z and J_x,

$$J_x = P_y(\pi/2) J_z P_y^\dagger(\pi/2) \tag{3.203a}$$

$$J_y = P_z(\pi/2) J_x P_z^\dagger(\pi/2). \tag{3.203b}$$

The unitary operator $P_z(\varphi)$ corresponds to a physical rotation over an angle φ about the $\hat{\mathbf{z}}$ direction - see Eq. (3.190); for $j = \frac{1}{2}$ it is given by Eq. (3.196). Likewise, the unitary operator $P_y(\theta)$ corresponds to a physical rotation over an angle θ about the $\hat{\mathbf{y}}$ direction - see Eq. (3.194); for $j = \frac{1}{2}$ it is given by Eq. (3.200). Using $\varphi = \pi/2$ in $P_z(\varphi)$ and $\theta = \pi/2$ in $P_y(\theta)$ the unitary operators become

$$P_z(\pi/2) = \sqrt{\tfrac{1}{2}} \begin{pmatrix} 1-i & 0 \\ 0 & 1+i \end{pmatrix}, \quad P_y(\pi/2) = \sqrt{\tfrac{1}{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \tag{3.204}$$

Substituting these expressions into Eqs. (3.203) we find

$$J_x = \tfrac{1}{2}\hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad J_y = \tfrac{1}{2}\hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{3.205}$$

The shift operators follow from the definitions (3.4),

$$J_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad J_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{3.206}$$

With these expressions we generated the results of Section 3.2.3. The commutation relations (3.1) and (3.7) as well as the shift relations (3.24) are satisfied. Using Eq. (3.9) we calculate $\mathbf{J}^2 = \frac{3}{4}\hbar^2\mathbb{1}$, where $\mathbb{1}$ is the unit matrix. This not only shows that Eq. (3.22a) is satisfied but also that $\mathbf{J}^2$ is invariant under rotation - like any scalar operator (cf. Appendix K). This is not surprising because $\sqrt{\langle\mathbf{J}^2\rangle}$ is a measure for the magnitude of $\mathbf{J}$. Once the matrices for J_y and J_z are known the matrix expressions for $P_z(\varphi) = U_z(-\varphi) = e^{-i\varphi J_z/\hbar}$ and $P_y(\varphi) = U_y(-\varphi) = e^{-i\varphi J_y/\hbar}$ can also be derived by expansion of the exponential operators. This is demonstrated in Problems 3.13 and 3.14.

3.8.1.2 Transformation of the angular momentum eigenstates

As an example we derive some expressions for the eigenstates of the $s = \frac{1}{2}$ system with respect an arbitrary quantization axis starting from the states $|\uparrow\rangle$ and $|\downarrow\rangle$ defined in Eqs. (3.42). A change of the quantization axis from $\hat{\mathbf{z}} = (0, 0)$ to $\hat{\mathbf{r}} = (\theta, \phi)$ is realized by a rotation of the coordinate system over the Euler angles $\gamma = 0$, $\beta = -\theta$ and $\alpha = -\phi$. The corresponding unitary transformation is given by

$$\mathscr{D}^{1/2}(\alpha, \beta, \gamma) = P_z(\phi) P_y(\theta) P_z(0). \tag{3.207}$$

Substituting Eqs. (3.196) and (3.200) we obtain

$$\mathscr{D}^{1/2}(\phi, \theta, 0) = \begin{pmatrix} e^{-i\phi/2}\cos(\theta/2) & -e^{-i\phi/2}\sin(\theta/2) \\ e^{i\phi/2}\sin(\theta/2) & e^{i\phi/2}\cos(\theta/2) \end{pmatrix}. \tag{3.208}$$

Thus, the spin-up and spin-down states with respect to an arbitrary quantization axis in the $\hat{\mathbf{r}}$ direction are given by

$$|\uparrow\rangle_r = \mathscr{D}^{1/2}(\phi, \theta, 0)\,|\uparrow\rangle = e^{-i\phi/2}\begin{pmatrix} \cos(\theta/2) \\ e^{i\phi}\sin(\theta/2) \end{pmatrix} \tag{3.209a}$$

$$|\downarrow\rangle_r = \mathscr{D}^{1/2}(\phi, \theta, 0)\,|\downarrow\rangle = e^{-i\phi/2}\begin{pmatrix} -\sin(\theta/2) \\ e^{i\phi}\cos(\theta/2) \end{pmatrix}. \tag{3.209b}$$

Note that with Eq. (3.209a) we regained Eq. (3.62).

Problem 3.13. Show by expansion of the exponential operator $e^{i\alpha S_z/\hbar}$ that the two-dimensional unitary transformation corresponding to a rotation over an angle α about the z axis is given by

$$U_z(\alpha) = \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix}.$$

Solution. We start with the expansion of the exponential operator,

$$U_z(\alpha) = e^{i\alpha S_z/\hbar} = 1 + i\alpha S_z/\hbar + \frac{1}{2!}(i\alpha S_z/\hbar)^2 + \frac{1}{3!}(i\alpha S_z/\hbar)^3 + \cdots.$$

The even and the odd terms can be factorized in a form containing a common matrix,

$$(i\alpha S_z/\hbar)^{2n} = (i\alpha/2)^{2n}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \qquad (i\alpha S_z/\hbar)^{2n+1} = (i\alpha/2)^{2n+1}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Next we recall the expansions of the sine and the cosine,

$$\cos\varphi = \sum_{n=0}^{\infty}\frac{(-1)^n}{2n!}\varphi^{2n} = \sum_{n=0}^{\infty}\frac{(i\varphi)^{2n}}{2n!}$$

$$i\sin\varphi = \sum_{n=0}^{\infty}\frac{i\,(-1)^n}{(2n+1)!}\varphi^{2n+1} = \sum_{n=0}^{\infty}\frac{(i\varphi)^{2n+1}}{(2n+1)!}.$$

Summing the even and odd terms separately we obtain

$$U_z(\alpha) = \cos(\alpha/2)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + i\sin(\alpha/2)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Using the relation $e^{i\varphi} = \cos\varphi + i\sin\varphi$ we arrive at the desired expression. □

Problem 3.14. Show by expansion of the exponential operator $e^{i\beta S_y/\hbar}$ that the two-dimensional unitary transformation corresponding to a rotation over an angle β about the y axis is given by

$$U_y(\beta) = \begin{pmatrix} \cos(\beta/2) & \sin(\beta/2) \\ -\sin(\beta/2) & \cos(\beta/2) \end{pmatrix}.$$

Solution. The proof goes along the same lines as in Problem 3.10. We start with the expansion of the exponential operator. In the present case the even and the odd terms factorize differently in the following way,

$$(i\beta S_y/\hbar)^{2n} = (i\beta/2)^{2n}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \qquad (i\beta S_y/\hbar)^{2n+1} = (i\beta/2)^{2n+1}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

Recalling the expansions of the sine and cosine we obtain after summing over the even and odd terms

$$U_y(\beta) = \cos\beta/2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin\beta/2\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Adding the two contributions we arrive at the desired expression. □

3.8.2 Example: $l = 1$ orbital angular momentum

To demonstrate the procedure for generating representations we consider the example of $l = 1$ orbital angular momentum and rederive Eq. (3.33). The unitary matrix $P_z(\varphi)$ is generated by the operator $e^{-i\varphi L_z/\hbar}$. The matrix elements follow with Eq. (3.170),

$$[P_z(\varphi)]_{m',m} = \langle l, m'|e^{-i\varphi L_z/\hbar}|l, m\rangle = \begin{cases} e^{-i\varphi m/\hbar} & \text{for } m' = m \\ 0 & \text{for } m' \neq m. \end{cases} \tag{3.210}$$

For $l = 1$ we obtain

$$P_z(\varphi) = \begin{pmatrix} e^{-i\varphi} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\varphi} \end{pmatrix}. \tag{3.211}$$

To determine L_z we turn to the definition (3.162). First, we calculate the derivative of $P_z(\varphi)$,

$$\left.\frac{\partial P_z(\varphi)}{\partial \varphi}\right|_{\varphi=0} = \begin{pmatrix} -ie^{-i\varphi}|_{\varphi=0} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & ie^{i\varphi}|_{\varphi=0} \end{pmatrix} = -i\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{3.212}$$

Substituting this result into Eq. (3.162) we obtain for the angular momentum operator

$$L_z = i\hbar\left.\frac{\partial P_z(\varphi)}{\partial \varphi}\right|_{\varphi=0} = \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{3.213}$$

This is indeed the expression given by Eq. (3.33).

The unitary matrix $P_y(\varphi)$ is generated by the operator $e^{-i\varphi L_y/\hbar}$. The matrix elements follow with Eq. (3.170),

$$[P_y(\theta)]_{m',m} = \langle j, m'|e^{-i\theta J_y/\hbar}|j, m\rangle = d^j_{m'm}(\theta). \tag{3.214}$$

For $l = 1$ we obtain

$$P_y(\theta) = \begin{pmatrix} \frac{1}{2}(1+\cos\theta) & -\sqrt{\frac{1}{2}}\sin\theta & \frac{1}{2}(1-\cos\theta) \\ \sqrt{\frac{1}{2}}\sin\theta & \cos\theta & -\sqrt{\frac{1}{2}}\sin\theta \\ \frac{1}{2}(1-\cos\theta) & \sqrt{\frac{1}{2}}\sin\theta & \frac{1}{2}(1+\cos\theta) \end{pmatrix}. \tag{3.215}$$

To determine L_y we use the definition (3.162). First, we calculate the derivative of $P_y(\theta)$,

$$\left.\frac{\partial P_y(\theta)}{\partial \theta}\right|_{\theta=0} = \begin{pmatrix} -\frac{1}{2}\sin\theta|_{\theta=0} & -\sqrt{\frac{1}{2}}\cos\theta|_{\theta=0} & \frac{1}{2}\sin\theta|_{\theta=0} \\ \sqrt{\frac{1}{2}}\cos\theta|_{\theta=0} & -\sin\theta|_{\theta=0} & -\sqrt{\frac{1}{2}}\cos\theta|_{\theta=0} \\ \frac{1}{2}\sin\theta|_{\theta=0} & \sqrt{\frac{1}{2}}\cos\theta|_{\theta=0} & -\frac{1}{2}\sin\theta|_{\theta=0} \end{pmatrix} = \sqrt{\tfrac{1}{2}}\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{3.216}$$

Substituting this result into Eq. (3.162) we obtain for the angular momentum operator

$$L_y = i\hbar \left. \frac{\partial P_y(\theta)}{\partial \theta} \right|_{\theta=0} = \sqrt{\tfrac{1}{2}}\hbar \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \tag{3.217}$$

Note that this is the expression given in Eq. (3.35).

3.8.2.1 Transformation of the angular momentum eigenstates

As an example we derive some expressions for the eigenstates of the $l = 1$ system with respect an arbitrary quantization axis starting from the state $|1,1\rangle \equiv |1,1\rangle_z$ defined in Eqs. (3.32). A change of the quantization axis from $\hat{\mathbf{z}} = (0,0)$ to $\hat{\mathbf{r}} = (\theta,\phi)$ is realized by a rotation of the coordinate system over the Euler angles $\gamma = 0$, $\beta = -\theta$ and $\alpha = -\phi$. The corresponding unitary transformation is given by

$$\mathscr{D}^1(\alpha,\beta,\gamma) = P_z(\phi)P_y(\theta)P_z(0). \tag{3.218}$$

Substituting Eqs. (3.211) and (3.215) we obtain

$$\mathscr{D}^1(\phi,\theta,0) = \begin{pmatrix} \frac{1}{2}(1+\cos\theta)e^{-i\varphi} & -\sqrt{\frac{1}{2}}\sin\theta e^{-i\varphi} & \frac{1}{2}(1-\cos\theta)e^{-i\varphi} \\ \sqrt{\frac{1}{2}}\sin\theta & \cos\theta & -\sqrt{\frac{1}{2}}\sin\theta \\ \frac{1}{2}(1-\cos\theta)e^{i\varphi} & \sqrt{\frac{1}{2}}\sin\theta e^{i\varphi} & \frac{1}{2}(1+\cos\theta)e^{i\varphi} \end{pmatrix}. \tag{3.219}$$

Hence, by unitary transformation we find $|1,1\rangle_x$ and $|1,1\rangle_y$

$$|1,1\rangle_x = \mathscr{D}^1(0,\pi/2,0\,|1,1\rangle = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{2} & 1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} \tag{3.220a}$$

$$|1,1\rangle_y = \mathscr{D}^1(\pi/2,\pi/2,0)\,|1,1\rangle = \frac{1}{2}\begin{pmatrix} -i & i\sqrt{2} & -i \\ \sqrt{2} & 0 & -\sqrt{2} \\ i & i\sqrt{2} & i \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -i \\ \sqrt{2} \\ i \end{pmatrix}. \tag{3.220b}$$

The corresponding density matrices are $\rho_{|1,1\rangle_x}$and $\rho_{|1,1\rangle_y}$ follow directly from these expressions

$$\rho_{|1,1\rangle_x} = \frac{1}{4}\begin{pmatrix} 1 & \sqrt{2} & 1 \\ \sqrt{2} & 2 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix} \quad \text{and} \quad \rho_{|1,1\rangle_y} = \frac{1}{4}\begin{pmatrix} 1 & -i\sqrt{2} & -1 \\ i\sqrt{2} & 2 & -i\sqrt{2} \\ -1 & i\sqrt{2} & 1 \end{pmatrix}. \tag{3.221}$$

Note that these density matrices are idempotent and have unit trace as required for pure states.

3.8.2.2 Transformation of the cartesian components of J

An alternative approach to obtain J_x and J_y is by unitary transformation of J_z and J_x,

$$J_x = P_y(\pi/2)J_zP_y^{\dagger}(\pi/2) \tag{3.222a}$$

$$J_y = P_z(\pi/2)J_xP_z^{\dagger}(\pi/2). \tag{3.222b}$$

The unitary operator $P_z(\varphi)$ corresponds to a physical rotation over an angle φ about the $\hat{\mathbf{z}}$ direction - see Eq. (3.190); for $j = 1$ it is given by Eq. (3.211). Likewise, the unitary operator $P_y(\theta)$ corresponds to a physical rotation over an angle θ about the $\hat{\mathbf{y}}$ direction - see Eq. (3.214); for $j = 1$ it is given by Eq. (3.215). Using $\varphi = \pi/2$ in $P_z(\varphi)$ and $\theta = \pi/2$ in $P_y(\theta)$ the unitary operators become

$$P_z(\pi/2) = \begin{pmatrix} -i & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & i \end{pmatrix}, \quad P_y(\pi/2) = \frac{1}{2}\begin{pmatrix} 1 & -\sqrt{2} & 1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix}. \tag{3.223}$$

Substituting these expressions into Eqs. (3.222) we find

$$J_x = \sqrt{\tfrac{1}{2}}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_y = \sqrt{\tfrac{1}{2}}\hbar \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \tag{3.224}$$

The shift operators follow from the definitions (3.4),

$$J_+ = \sqrt{2}\hbar \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad J_- = \sqrt{2}\hbar \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{3.225}$$

With these expressions we generated the results of Section 3.2.1.

4

Fine Structure

In Chapter 2 we introduced the term *fine structure* for small deviations from the principal atomic structure. The deviation from the Bohr formula can be a small shift or involve the splitting of an energy level. A common feature is that the shift or splitting is much smaller than the separation from adjacent Bohr levels. Interestingly, fine-structure splittings were already recorded a century before the Bohr formula was derived. In 1814 Joseph von Fraunhofer studied the solar spectrum and noted the famous doublet splitting of the D lines of *sodium* (Na) at 589.5 nm [45]. We shall find that in the alkali elements this splitting arises from slight differences in screening of the nuclear charge by the core electrons (see Section 4.5).

The fine structure of atomic hydrogen has a completely different origin. Revealing this origin has been one of the great puzzles of modern physics - competing phenomena had to be unraveled and important new physics had to be discovered. The first step was made by Arnold Sommerfeld, who showed in 1916 that relativistic correction of the kinetic energy gives rise to a doublet splitting of the correct order of magnitude [106]. A rigorous relativistic description was possible only after Paul Dirac formulated his relativistic quantum theory for the electron in 1928 [35]. In this theory the appearance of spin and the associated magnetic moment were identified as intrinsic relativistic phenomena. Aside from *relativistic* corrections also *radiative* corrections are essential to explain the detailed fine structure of the hydrogen atom. In 1947 Willis Lamb and Robert Retherford measured a small shift between the $2s$ and $2p$ level in hydrogen that could not be explained by the Dirac theory [68]. In the same year Hans Bethe demonstrated that this so-called *Lamb shift* arises from the zero-point fluctuations of the quantized electromagnetic field [13]. This has been one of the key ingredients that led to the formulation of modern *quantum electrodynamics* (QED) in which relativistic quantum mechanics and electrodynamics are unified in a single theory.

Unfortunately, both the Dirac theory and QED fall outside the scope of this introductory course. We adopt the phenomenological point of view in which the consequences of the Dirac theory (like spin) are postulated. This has the advantage that the elements of atomic spectroscopy can be introduced without facing the full complexity of the atom at once. Following the historical sequence of events we introduce the relativistic corrections to the kinetic and potential energy of the electrons and discuss the properties of spin as well as the origin of spin-orbit coupling. This phenomenological approach can be justified as being a mapping of the Dirac theory onto an effective Hamiltonian which can be evaluated with perturbation theory.

Atoms consist of orbiting and spinning charges. Therefore, it is not surprising that these ingredients affect the atomic structure at the fundamental level. They manifest themselves as classical and non-classical forms of magnetism. The best known magnetic phenomenon is the splitting of spectral lines by an applied magnetic field. This was discovered by Pieter Zeeman in 1897 and known as the *Zeeman effect* [123]. In zero field, the electronic magnetism manifests itself in spin-orbit splitting. The importance of magnetic interactions reaches far beyond atomic physics into condensed matter physics and chemistry. Magnetic interactions between atoms are crucial for the understanding of molecular binding and cohesion in solids and find important applications, for instance in medical imaging. Interestingly, although decisive for major applications, in atomic physics magnetic interactions appear as weak perturbations of the principal atomic structure.

In the coming sections we introduce the atomic fine structure of *one-electron atoms.*. We start with hydrogenic atoms and identify four contributions to the level shift (see Fig. 4.1). We distinguish between orbital- and spin-related coupling to magnetic fields (Sections 4.1-4.2). The spin and orbital dynamics are combined in spin-orbit coupling (Section 4.5). This is a velocity-induced phenomenon and embodies the central issue of the chapter. We discuss how, in hydrogenic atoms, the various shifts enter on equal footing (Section 4.5), whereas in alkali-like atoms the spin-orbit shift dominates the physics (Section 4.6). This is explained by the role of core electrons which effectively screen the valence electron from the nuclear charge. The last section of the chapter deals with the competition between spin-orbit and Zeeman coupling (Section 4.7).

4.1 Relativistic and radiative shifts

4.1.1 Relativistic mass correction

Close to the nucleus the electrons are accelerated to relativistic velocities. This gives rise to lifting of the degeneracy of energy levels of different l but equal n. The relativistic shifts are small but comparable to the Zeeman splitting. The *relativistic Hamiltonian* of a spinless particle of rest mass m moving in a central field $\mathcal{V}(r)$ is given by

$$H = \sqrt{c^2\mathbf{p}^2 + m^2c^4} + \mathcal{V}(r). \tag{4.1}$$

Subtracting the rest energy mc^2 and the potential energy $\mathcal{V}(r)$ we obtain the kinetic energy T, which can be expanded in even powers of v/c,

$$\begin{aligned} T &= \sqrt{c^2\mathbf{p}^2 + m^2c^4} - mc^2 = mc^2\sqrt{1 + (\frac{\mathbf{v}}{c})^2} - mc^2 \\ &= \left[1 + \tfrac{1}{2}(\frac{\mathbf{v}}{c})^2 - \tfrac{1}{8}(\frac{\mathbf{v}}{c})^4 + \cdots - 1\right] mc^2 \\ &= \frac{\mathbf{p}^2}{2m} - \frac{1}{2mc^2}\left(\frac{\mathbf{p}^2}{2m}\right)^2 + \cdots . \end{aligned} \tag{4.2}$$

The second line shows how the mass of the moving body depends on v/c. For weakly relativistic velocities the expansion (4.2) can be used to calculate the *relativistic correction* to the kinetic energy by perturbation theory. In particular, using the reduced mass m_r of the electron in a hydrogenic atom we can write the atomic Hamiltonian in the form (H.1)

$$\mathcal{H} - m_rc^2 = T + \mathcal{V}(r) = \frac{\mathbf{p}^2}{2m_r} + \mathcal{V}(r) + \mathcal{H}_r^{\text{mass}} = \mathcal{H}_0 + \mathcal{H}_r^{\text{mass}}, \tag{4.3}$$

where $\mathcal{H}_0$ is the (non-relativistic) Schrödinger Hamiltonian (2.1) with eigenstates $R_{nl}(r)$ given by Eq. (2.37). The perturbation $\mathcal{H}_r^{\text{mass}}$ can be expressed in terms of $\mathcal{H}_0$ and $\mathcal{V}(r)$

$$\mathcal{H}_r^{\text{mass}} = -\frac{1}{2m_rc^2}\left(\frac{\mathbf{p}^2}{2m_r}\right)^2 = -\frac{1}{2m_rc^2}\left(\frac{p_r^2}{2m_r} + \frac{\mathbf{L}^2}{2m_rr^2}\right)^2 = -\frac{1}{2m_rc^2}\left[\mathcal{H}_0 - \mathcal{V}(r)\right]^2. \tag{4.4}$$

We now calculate the splitting of the Bohr level n by the relativistic mass correction. Since $\mathcal{H}_r^{\text{mass}}$ commutes with $\mathbf{L}^2$ and L_z, the relativistic mass correction is *diagonal* in the rotational basis $\{|lm_l\rangle\}$; i.e., l and m_l are "good" quantum numbers. This makes explicit diagonalization of the perturbation matrix superfluous and reveals the rotational basis as the *optimal basis* to evaluate the perturbation

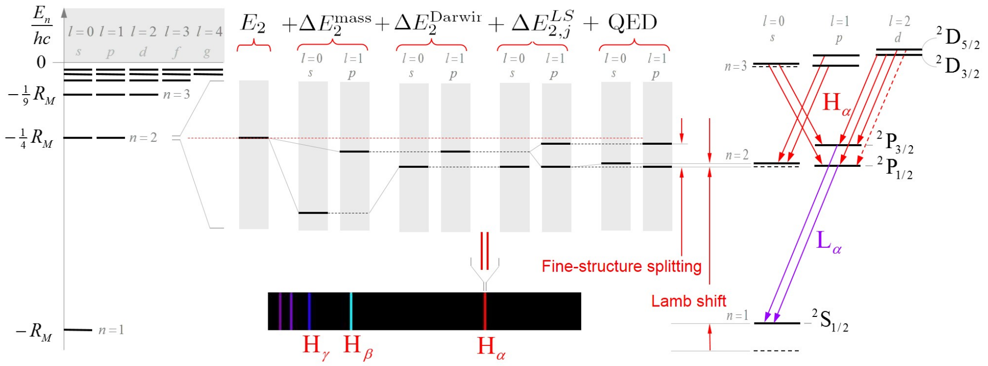


Figure 4.1: Fine-structure effects in hydrogenic atoms. From left to right are shown the unperturbed level at $E_2 = -\frac{1}{4}R_M$ and the cumulative effects of subsequently the relativistic mass correction shift ΔE_2^{mass}, the Darwin shift ΔE_2^{D}, the spin-orbit shifts $\Delta E_{2,j}^{LS}$ for $j = l \pm \frac{1}{2}$ and the Lamb shift. The latter cannot be understood with classical fields but requires quantum electrodynamics (QED). On the far right the Balmer transitions are shown. The dashed line is forbidden as a single photon electric-dipole transition.

(see Appendix H.3). Although states of different n are coupled by the perturbation ($\mathcal{H}_r^{\text{mass}}$ does *not* commute with $\mathcal{H}_0$) this is of no consequence as long as the second order correction is negligible (n is a "good quantum number to first order in perturbation theory"). The latter is satisfied as long as the first-order shifts are small in comparison to the level separation of the principal structure. Under these conditions the level shifts can be calculated in hydrogenic basis, $\{|nlm_l\rangle\}$, using the well-known expression from ordinary first-order perturbation theory - see Appendix H.3.1

$$\Delta E_{nlm}^{\text{mass}} = \langle nlm_l|\mathcal{H}_r^{\text{mass}}|nlm_l\rangle. \tag{4.5}$$

Using the hermiticity of $\mathcal{H}_0$ as well as the eigenvalues of the Schrödinger Hamiltonian we obtain

$$\begin{aligned}\Delta E_{nlm_l}^{\text{mass}} &= -\frac{1}{2m_rc^2}\langle nlm_l|\left[\mathcal{H}_0 - \mathcal{V}(r)\right]^2|nlm_l\rangle &(4.6)\\ &= -\frac{1}{2m_rc^2}[E_n^2 - 2E_n\langle nlm_l|\mathcal{V}(r)|nlm_l\rangle + \langle nlm_l|\mathcal{V}^2(r)|nlm_l\rangle]. &(4.7)\end{aligned}$$

Let us evaluate the relativistic shift $\Delta E_{nlm_l}^{\text{mass}}$ for hydrogenic atoms. Using the radial wavefunctions given in Section 2.3 we find that the shift is independent of the quantum number m_l. After integration over the angles θ and ϕ it takes the form

$$\Delta E_{nl}^{\text{mass}} = -\frac{1}{2m_rc^2}\left[E_n^2 - 2E_n\int_0^\infty \mathcal{V}(r)R_{nl}^2(r)r^2dr + \int_0^\infty \mathcal{V}^2(r)R_{nl}^2(r)r^2dr\right]. \tag{4.8}$$

Before evaluating the integral we note that the electrostatic interaction can be expressed in terms of the *fine-structure constant* $\alpha \approx 1/137$, $\mathcal{V}(r) = -(e^2/4\pi\varepsilon_0)Z/r = -\alpha^2 m_rc^2 Za/r$. Then, using Eq. (2.33) for E_n and Eqs. (2.49) for $\langle 1/\rho\rangle$ and $\langle 1/\rho^2\rangle$ we obtain

$$\begin{aligned}\Delta E_n^{\text{mass}} &= -E_n\frac{E_n}{2m_rc^2}[1 + 2\alpha^2 m_rc^2 Z/E_n\,\langle 1/\rho\rangle + \left(\alpha^2 m_rc^2 Z/E_n\right)^2\langle 1/\rho^2\rangle]\\ &= E_n\frac{\alpha^2Z^2}{4n^2}\left[1 - 2\frac{2n^2}{Z}\frac{Z}{n^2} + \left(\frac{2n^2}{Z}\right)^2\frac{Z^2}{n^3}\frac{1}{l+1/2}\right]\\ &= -E_n\frac{\alpha^2Z^2}{n^2}\left(\frac{3}{4} - \frac{n}{l+1/2}\right). &(4.9)\end{aligned}$$

This is a celebrated result - obtained by Arnold Sommerfeld in 1916 [106]. Note that it validates our use of first-order perturbation theory because the shifts are small compared to the level spacing of the principal structure, $\Delta E_n^{\text{mass}}/E_n \approx \alpha^2$. More interestingly, it shows that the relativistic corrections depend on the orbital angular momentum quantum number l and are biggest for states with low l. This l dependence lifts the accidental degeneracy of states with different angular momentum (see Fig. 4.1). Physically this is to be expected: the smaller the value of l, the closer the electron approaches the nucleus, the faster the motion and the larger the relativistic shift.

4.1.2 Darwin term

The relativistic velocity of the electron close to the nucleus also affects the Coulomb interaction. It turns out that under weakly relativistic conditions the Dirac equation can be reduced to a Schrödinger-type equation in which the effective Coulomb interaction is weaker. This is expressed by a correction term in the effective Hamiltonian known as the *Darwin term*,

$$\mathcal{H}_r^{\text{Darwin}} = \frac{e\hbar^2}{8m_e^2c^2}\text{div}\mathbf{E}. \tag{4.10}$$

The Darwin term has no classical analogue but it can be made plausible. The Dirac theory has the property that relativistic electrons exhibit a rapid quiver motion called *Zitterbewegung* [101, 51] over distances of the order of the (reduced) *Compton wavelength,* $\lambda_C = \hbar/m_ec = \alpha a_0 \simeq 3.86 \times 10^{-13}$ m. This length is a factor α smaller than the Bohr radius, $a_0 \simeq 5.29 \times 10^{-11}$ m, but much larger than the proton rms charge radius $r_p \simeq 8.78 \times 10^{-16}$ m. If we postulate this Zitterbewegung we can estimate the magnitude of the Darwin term by heuristically delocalizing the electron presuming a centrally symmetric charge distribution $\varrho(r')$ of size of the Compton wavelength and normalized to the total charge, $\int \varrho(r')d\mathbf{r}' = -e$. In this model an effective electrostatic energy is obtained by averaging over the charge distribution. For a nucleus of charge Ze at position $\mathbf{r}$ the average is given by

$$\mathcal{V}(\mathbf{r}) = Ze\int \varrho(r')\varphi(\mathbf{r}-\mathbf{r}')d\mathbf{r}'. \tag{4.11}$$

Here

$$\varphi(\mathbf{r}-\mathbf{r}') = \frac{1}{4\pi\varepsilon_0}\frac{1}{|\mathbf{r}-\mathbf{r}'|} \tag{4.12}$$

is a *nonlocal* scalar potential, nonlocal because the electron is delocalized as is accounted for by integrating over the dummy variable $\mathbf{r}'$. Obviously, the model cannot provide more than an order of magnitude estimate. So, we simply model the electron as a homogeneously charged sphere of radius R, with the charge density given by

$$\varrho(r') = \varrho_0 = -\frac{3e}{4\pi R^3}. \tag{4.13}$$

To evaluate the integral we choose $\mathbf{r}$ and $\mathbf{r}'$ relative to the center of the charge distribution as illustrated in Fig. 4.2. Using the multipole expansion (L.41) we find with the aid of the orthogonality relation for Legendre polynomials (L.45) that for a spherical charge distribution *only the $l = 0$ contribution is nonvanishing* and obtain for the electrostatic potential (see Problem 4.1)[1]

$$\varphi(r) = -\frac{1}{4\pi\varepsilon_0}\frac{e}{r} - \frac{1}{4\pi\varepsilon_0}\frac{e}{2R}\left[3-\left(\frac{r}{R}\right)^2 - 2\frac{R}{r}\right]_{r<R}. \tag{4.14}$$

To obtain the Darwin term we rewrite the potential energy in a form introducing a short-range correction to the pure Coulomb law

$$\mathcal{V}(r) = Ze\varphi(\mathbf{r}) = -\frac{Ze^2}{4\pi\varepsilon_0}\frac{1}{r} + \mathcal{V}'(r). \tag{4.15}$$

---

[1] Nonspherical nuclei give rise to the *electric* hyperfine interaction.

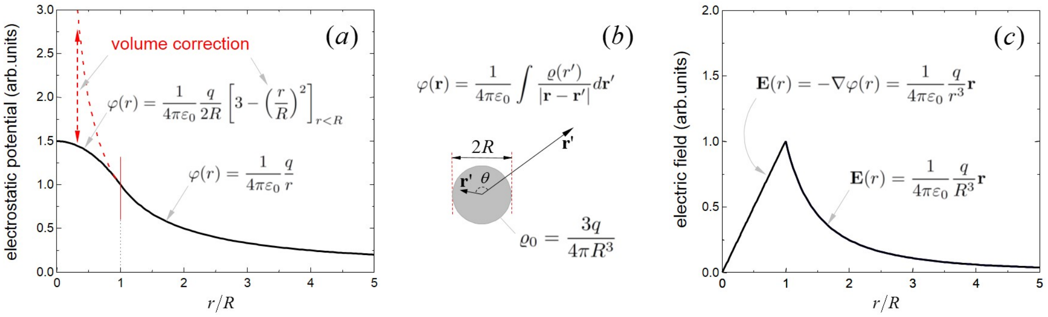


Figure 4.2: Electrostatic scalar potential (see a) and electric field (see c) of a homogeneously charged solid sphere of total charge q and radius R (see c). The dashed vertical arrow shows how the electrostatic potential can be obtained by correcting the Coulomb law for distances $r \leq R$.

Comparing Eqs. (4.14) and (4.15) we find that the correction is given by

$$\mathcal{V}'(r) = -\frac{Ze^2}{4\pi\varepsilon_0}\frac{1}{2R}\left(3-\frac{r^2}{R^2}-\frac{2R}{r}\right) \qquad \text{for } r \leq R. \tag{4.16}$$

We calculate the Darwin correction of the Bohr level n in the same way as the relativistic mass correction. We use ordinary first-order perturbation theory in the hydrogenic basis $\{|nlm_l\rangle\}$,

$$\Delta E_n^{\text{Darwin}} = \langle nlm_l|\mathcal{V}'(r)|nlm_l\rangle. \tag{4.17}$$

Since $\mathcal{V}'(r)$ commutes with $\mathbf{L}^2$ and L_z this perturbation is *diagonal* in the rotational basis $\{|lm_l\rangle\}$; i.e., l and m_l are "good" quantum numbers. Since $\mathcal{V}'(r)$ does *not* commute with $\mathcal{H}_0$, states of different n are coupled by the perturbation but this is of no consequence as long as the second-order shift is negligible (n is a "good quantum number to first order in perturbation theory" - see Appendix H.3). In the position representation we find

$$\Delta E^{\text{Darwin}} \simeq -\frac{Ze^2}{4\pi\varepsilon_0}\frac{1}{2R}\int\left(3-\frac{r^2}{R^2}-\frac{2R}{r}\right)R_{nl}^2(r)|Y_l^{m_l}(\hat{\mathbf{r}})|^2 d\mathbf{r} \qquad \text{with } r \leq R. \tag{4.18}$$

Since $\mathcal{V}'(r)$ has central symmetry the angular part evaluates to unity. Since $r \leq R \ll a_0$ the radial wavefunction can be replaced by its value in the origin and the Darwin correction is given by

$$\Delta E^{\text{Darwin}} \simeq -\frac{Ze^2}{4\pi\varepsilon_0}\frac{1}{2R}R_{nl}^2(0)\int_0^R\left(3-\frac{r^2}{R^2}-\frac{2R}{r}\right)r^2dr = \frac{Ze^2}{4\pi\varepsilon_0}\frac{R^2}{10}R_{nl}^2(0). \tag{4.19}$$

Importantly, this expression shows that the Darwin shift *only affects s states* because for nonzero angular momentum the radial wavefunction vanishes in the origin (see Fig. 2.2). The result (4.19) can also be obtained by using the concept of a *contact interaction*; i.e., an interaction for which the strength scales with the probability density of the electronic wavefunction at the position of the nucleus. Note that by including the following contact interaction directly into the Hamiltonian

$$\mathcal{H}_r^{\text{Darwin}} \simeq \frac{Ze^2}{\varepsilon_0}\frac{R^2}{10}\delta(\mathbf{r}) \tag{4.20}$$

we regain the Darwin shift (4.19) after averaging over the electronic wavefunction.[1] With insider knowledge of Eq. (4.10) we choose $R^2/10 = \lambda_C^2/8$ and obtain

$$\mathcal{H}_r^{\text{Darwin}} = \frac{Ze^2}{\varepsilon_0}\frac{\hbar^2}{8m_e^2c^2}\delta(\mathbf{r}). \tag{4.21}$$

---

[1] Note the relation $\delta(r) = 4\pi\delta(\mathbf{r})r^2$.

This results also follows directly from Eq. (4.10). In the Dirac theory the electron is point like and its Coulomb potential in the field of the nuclear charge given by $\varphi(r) = Ze/4\pi\varepsilon_0 r$. Since $\mathbf{E} = -\nabla\varphi$ and $\Delta 1/r = -4\pi\delta(\mathbf{r})$ the Darwin term (4.10) takes the form (4.21).

To conclude the section we calculate the total relativistic correction. For the Darwin correction we use $\hbar/m_e c = \alpha a_0$ and recalling Eq. (2.39) for the radial wavefunction in the origin we obtain

$$\Delta E^{\text{Darwin}} = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{a_0}\frac{Z^2}{2n^2}\frac{\alpha^2 Z^2}{n^2} n = -E_n \frac{\alpha^2 Z^2}{n^2} n. \tag{4.22}$$

Combining the two relativistic corrections

$$\mathcal{H}_r = \mathcal{H}_r^{\text{mass}} + \mathcal{H}_r^{\text{Darwin}}, \tag{4.23}$$

we obtain the total relativistic shift

$$\Delta E_n^{\text{rel}} = \Delta E_n^{\text{mass}} + \Delta E_n^{\text{Darwin}} = \begin{cases} -E_n \dfrac{\alpha^2 Z^2}{n^2}\left(\dfrac{3}{4} - \dfrac{n}{l + {}^1\!/_2}\right) & \text{for } l > 0 \\ -E_n \dfrac{\alpha^2 Z^2}{n^2}\left(\dfrac{3}{4} - n\right) & \text{for } l = 0. \end{cases} \tag{4.24}$$

The relativistic splitting between the nl and ns levels is given by

$$\Delta E_{nl-ns}^{\text{rel}} = \frac{\alpha^2 Z^2}{n^3}\left(\frac{2l-1}{2l+1}\right)\text{Ry} \tag{4.25}$$

In atomic hydrogen the relativistic splitting between the $2p$ and $2s$ levels is $\Delta E_{2p-2s}^{\text{rel}} = 7.296\,\text{GHz}$. As we shall see in the coming sections the relativistic corrections discussed here are of the same order of magnitude as typical magnetic corrections.

Problem 4.1. Show that for a homogeneously charged sphere (radius R, total charge q) the electrostatic potential can be written as the sum of a pure Coulomb potential plus a correction term for distances $r < R$,

$$\varphi(r) = \frac{1}{4\pi\varepsilon_0}\frac{q}{r} + \frac{1}{4\pi\varepsilon_0}\frac{q}{2R}\left[3 - \left(\frac{r}{R}\right)^2 - 2\frac{R}{r}\right]_{r<R}.$$

Solution. The electrostatic potential is given by

$$\varphi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0}\int \frac{\varrho(r')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'.$$

For a homogeneously charged sphere of total charge q the charge density is $\varrho_0 = 3q/4\pi R^3$. As the charge distribution is spherical we can evaluate the integral by restricting ourselves to the first term of the multipole expansion (L.41). Thus, we obtain

$$\int \frac{\varrho(r')}{r_>} d\mathbf{r}' = \begin{cases} \dfrac{\varrho_0}{r}\displaystyle\int_0^r 4\pi r'^2 dr' + \varrho_0 4\pi r' dr' = \dfrac{q}{2R}\left[3 - \left(\dfrac{r}{R}\right)^2\right] & r < R \\ \dfrac{\varrho_0}{r}\displaystyle\int_0^R 4\pi r'^2 dr' = \dfrac{q}{r} & r \geq R, \end{cases}$$

where $r_> = \max\{r, r'\}$. Adding and subtracting q/r for distances $r < R$ we obtain after putting the prefactor the desired expression (see Fig. 4.2). □

4.1.3 Lamb shift

Quantum electrodynamics (QED) is outside the scope of this introductory text but a flavor of its origin can be obtained from Welton's picture for the Lamb shift. Welton argued that the zero-point fluctuations of the quantized electromagnetic field will give rise to quiver motion of the electrons relative to the nucleus. This quiver motion will reduce the Coulomb interaction of the electrons with the nucleus because, like the Zitterbewegung in the case of the Darwin correction, the zero-point motion will tend to delocalize the electron. Thus, also in the present case we shall represent the electron by a centrally symmetric charge distribution $\varrho(r')$ normalized to the total charge, $\int \varrho(r')d\mathbf{r}' = -e$. Therefore, we simply presume the charge to be homogeneously distributed over a sphere of radius R_L. Hence, the Lamb shift can be estimated in the same way as the Darwin shift, which leads to the result

$$\Delta E^{\mathrm{Lamb}} = \frac{Ze^2}{4\pi\varepsilon_0}\frac{R_L^2}{10}R_{nl}^2(0). \tag{4.26}$$

This means that the Lamb shift only affects the s states. Without further explanation we state that the radius can be estimated to be $R_L \simeq \sqrt{3\alpha}\lambda_C$, which means that the Lamb shift is a few percent of the Darwin shift.

4.2 Hamiltonian for electronic motion in magnetic fields

In the presence of an externally applied magnetic field $\mathbf{B}$ the motion of a charged particle is affected by the *Lorentz* force, $\mathbf{F} = q(\mathbf{v}\times\mathbf{B})$, with q the charge of the particle and $\mathbf{v} = \dot{\mathbf{r}}$ its velocity. Unlike the electrical force $\mathbf{F} = q\mathbf{E}$, the Lorentz force is a *velocity dependent force.* It is an example of a *generalized force* for which a *generalized potential function* may be defined such that the canonical description of classical mechanics remains valid (see Problem 4.2). The generalized potential function for a particle of charge q is given by

$$U(\mathbf{r},\mathbf{v},t) = q(\varphi - \mathbf{v}\cdot\mathbf{A}). \tag{4.27}$$

Here $\varphi(\mathbf{r},t)$ is the scalar potential and $\mathbf{A}(\mathbf{r},t)$ the vector potential of the electromagnetic field, with $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$ and $\mathbf{E} = -\boldsymbol{\nabla}\varphi - \partial\mathbf{A}/\partial t$ expressions for the magnetic and electric fields, respectively.[1]

Like the force, also the momentum of charged particles is affected by the $\mathbf{B}$ field. The *generalized momentum* is given by the canonical expression $p_j = \partial\mathcal{L}/\partial v_j$, where $\mathcal{L} = T - U$ is the Lagrangian, with T the kinetic energy and U the generalized potential function of a system. Substituting $T = \frac{1}{2}m\mathbf{v}^2$ for the kinetic energy and Eq. (4.27) for the generalized potential energy we obtain for the canonical momentum

$$\mathbf{p} = m\mathbf{v} + q\mathbf{A}, \tag{4.28}$$

which shows that a particle of mass m and charge q does not carry only *kinetic momentum* $m\mathbf{v}$ but also *electromagnetic momentum* $q\mathbf{A}$.

To calculate the energy eigenstates of the atom in the presence of magnetic fields we have to quantize the classical Hamiltonian

$$H = \mathbf{p}\cdot\mathbf{v} - \mathcal{L}. \tag{4.29}$$

Substituting the Lagrangian we obtain

$$H = m\mathbf{v}^2 + q\mathbf{A}\cdot\mathbf{v} - \frac{1}{2}m\mathbf{v}^2 + q(\varphi - \mathbf{v}\cdot\mathbf{A}) = \frac{1}{2}m\mathbf{v}^2 + q\,\varphi(r), \tag{4.30}$$

which is seen to coincide with the familiar expression (1.1). To prepare for quantization we rewrite the classical Hamiltonian in terms of the canonical momentum

$$H = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2 + q\,\varphi(r). \tag{4.31}$$

---

[1] Note that the expressions for the fields indeed satisfy the Maxwell equations $\boldsymbol{\nabla}\cdot\mathbf{B} = 0$ and $\nabla\times\mathbf{E} = -\partial\mathbf{B}/\partial t$ because $\nabla\cdot(\nabla\times\mathbf{A}) = 0$ for any vector $\mathbf{A}$ and $\nabla\times\boldsymbol{\nabla}\varphi = 0$ for any scalar φ.

With the identification $\mathbf{p} \to -i\hbar\boldsymbol{\nabla}$ we obtain the quantum mechanical Hamiltonian for a charged particle in a classical electromagnetic field

$$\mathcal{H} = \frac{1}{2m}(-i\hbar\boldsymbol{\nabla} - q\mathbf{A})^2 + q\,\varphi(r). \tag{4.32}$$

To obtain the well-known Zeeman Hamiltonian we write Eq. (4.32) in the form[1]

$$\mathcal{H} = -\frac{\hbar^2}{2m}\Delta + \frac{i\hbar q}{2m}[(\boldsymbol{\nabla}\cdot\mathbf{A}) + (\mathbf{A}\cdot\boldsymbol{\nabla})] + \frac{q^2}{2m}\mathbf{A}^2 + q\,\varphi(r). \tag{4.33}$$

This expression may be simplified using the *gauge freedom*: $\mathbf{A}$ is not uniquely defined. Because we have $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\chi = 0$ for any choice of scalar potential χ, any vector potential $\mathbf{A} = \mathbf{A}'+\boldsymbol{\nabla}\chi$ results with $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$ in the same magnetic field. Hence, by choosing $\Delta\chi = -\boldsymbol{\nabla}\cdot\mathbf{A}'$ the vector potential can be made divergence-free

$$\boldsymbol{\nabla}\cdot\mathbf{A} = 0. \tag{4.34}$$

This choice for the vector potential is called the *Coulomb gauge*, also known under the name *radiation gauge*. It is the preferred gauge for atomic physics [27]. In the Coulomb gauge, the Hamiltonian (4.33) takes the form

$$\mathcal{H} = -\frac{\hbar^2}{2m}\Delta + \mathcal{V}(r) + \frac{i\hbar q}{m}\mathbf{A}\cdot\boldsymbol{\nabla} + \frac{q^2}{2m}\mathbf{A}^2, \tag{4.35}$$

where $\mathcal{V}(r) = q\,\varphi(r)$ is the Coulomb energy of the charge q.

Specializing further to a *homogeneous* magnetic field the vector potential can be written as[2]

$$\mathbf{A} = \tfrac{1}{2}\mathbf{B}\times\mathbf{r}. \tag{4.36}$$

This relation is readily verified by direct substitution in $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$. Substituting Eq. (4.36) into Eq. (4.35) we obtain[3]

$$\mathcal{H} = \mathcal{H}_0 - \frac{q}{2m}\mathbf{L}\cdot\mathbf{B} + \frac{q^2}{8m}r_\perp^2\mathbf{B}^2, \tag{4.37}$$

where $\mathcal{H}_0$ is the Schrödinger Hamiltonian, the second term stands for the orbital Zeeman coupling (cf. Section 4.3.3) and the third term for the atomic diamagnetism (cf. Section 4.3.2). The operator $r_\perp = r\sin\theta$ represents the component of the radius vector perpendicular to the direction of $\mathbf{B}$.

Problem 4.2. Show that the force $\mathbf{F} = q(\mathbf{E} + \mathbf{v}\times\mathbf{B})$ acting on a charge q, moving at velocity $\mathbf{v}$ in an electromagnetic field given by $\mathbf{E}(t)$ and $\mathbf{B}(t)$, may be described as arising from a generalized potential function $U(\mathbf{r},\mathbf{v},t) = q\,(\varphi - \mathbf{v}\cdot\mathbf{A})$, where $\varphi(\mathbf{r})$ is the scalar potential (Coulomb potential) and $\mathbf{A}(\mathbf{r},t)$ the vector potential of the electromagnetic field.

Solution. In order to satisfy the laws of classical mechanics a system should satisfy the Lagrange equations,

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}_i} - \frac{\partial\mathcal{L}}{\partial q_i} = 0, \tag{4.38}$$

where $\mathcal{L}(q_i,\dot{q}_i,t) \equiv T(q_i,\dot{q}_i,t) - U(q_i,\dot{q}_i,t)$ is the Lagrangian and $\{q_i,\dot{q}_i,t\}$ the generalized coordinates, with $i \in \{1,2,\cdots,l\}$ the coordinate index and l the number of degrees of freedom of the system. Substituting the Lagrangian into Eq. (4.38) the Lagrange equations take the form

$$\frac{d}{dt}\frac{\partial T}{\partial\dot{q}_i} - \frac{\partial T}{\partial q_i} = -\frac{\partial U}{\partial q_i} + \frac{d}{dt}\left(\frac{\partial U}{\partial\dot{q}_i}\right) \equiv Q_i(q_i,\dot{q}_i,t). \tag{4.39}$$

---

[1] Note that, *in general*, ∇ and $\mathbf{A}$ do *not* commute, $(\nabla\cdot\mathbf{A})\,\phi = \phi\,(\nabla\cdot\mathbf{A}) + (\mathbf{A}\cdot\nabla)\,\phi$.

[2] Note that $\nabla\cdot\mathbf{r} = 3$ and $(\mathbf{B}\cdot\nabla)\mathbf{r} = \mathbf{B}$.

[3] With Eq.(1.22) we obtain $2i\hbar\mathbf{A}\cdot\nabla = i\hbar(\mathbf{B}\times\mathbf{r})\cdot\nabla = \mathbf{B}\cdot(\mathbf{r}\times i\hbar\nabla) = -\mathbf{B}\cdot\mathbf{L}$.
Note that $(\mathbf{r}\times\mathbf{B})^2 = r^2\sin^2\theta\,B^2 = r_\perp^2\,\mathbf{B}\cdot\mathbf{B}$.

To solve our problem we turn to the case of a single body of mass m. In cartesian coordinates, the kinetic energy of the body is given by $T = \frac{1}{2}m\dot{r}_i^2$ where r_i represents its position, with $i \in \{x, y, z\}$, and $\dot{r}_i$ its velocity. Evaluating the l.h.s. of Eq. (4.39) the Lagrange equations for the motion become $m\ddot{r}_i = Q_i(r_i, \dot{r}_i, t)$, which is just the Newton equation of motion if Q_i is interpreted as a generalized force. Thus our task is reduced to demonstrating that the Lorentz force is obtained by evaluating the generalized force

$$Q_i(r_i, v_i, t) = -\frac{\partial U}{\partial r_i} + \frac{d}{dt}\left(\frac{\partial U}{\partial v_i}\right), \tag{4.40}$$

starting from the generalized potential function $U(\mathbf{r}, \mathbf{v}, t) = q\,(\varphi - \mathbf{v}\cdot\mathbf{A})$, which with the summation convention takes the form $U = q\,[\varphi\,(\mathbf{r}) - v_j A_j\,(\mathbf{r}, t)]$. The first term of Eq. (4.40) yields

$$-\partial_i U = q\,[-\partial_i\varphi + v_j\partial_i A_j\,(\mathbf{r}, t)]\,. \tag{4.41}$$

and the second term can be written as

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial U}{\partial v_i}\right) &= q\frac{d}{dt}\left(\frac{\partial\varphi\,(\mathbf{r}, t)}{\partial v_i} - \frac{\partial v_j A_j\,(\mathbf{r}, t)}{\partial v_i}\right) = -q\frac{d}{dt}A_i\,(\mathbf{r}, t)\\ &= -q\left(\frac{\partial A_i\,(\mathbf{r}, t)}{\partial t} + \frac{\partial A_i\,(\mathbf{r}, t)}{\partial r_j}\frac{dr_j}{dt}\right) = q\,[-\partial_t A_i\,(\mathbf{r}, t) - \dot{r}_j\partial_j A_i\,(\mathbf{r}, t)]\,.\end{aligned} \tag{4.42}$$

Combining Eqs. (4.41) and (4.42) we obtain for the generalized force

$$Q_i = q\,[-\partial_i\varphi - \partial_t A_i + v_j\partial_i A_j - v_j\partial_j A_i]\,. \tag{4.43}$$

Recalling the expression for the $\mathbf{E}$ field in terms of the scalar and vector potentials, $\mathbf{E} = -\boldsymbol{\nabla}\varphi - \partial\mathbf{A}/\partial t$, we recognize in the first two terms within the brackets of Eq. (4.43) the E_i-component of the $\mathbf{E}$ field,

$$E_i = -\partial_i\varphi - \partial_t A_i. \tag{4.44}$$

Likewise, using $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$, we recognize in the last two terms the component $(\mathbf{v}\times\mathbf{B})_i$ of the $\mathbf{v}\times\mathbf{B}$ field,

$$(\mathbf{v}\times\mathbf{B})_i = \varepsilon_{ijk}v_j(\boldsymbol{\nabla}\times\mathbf{A})_k = \varepsilon_{ijk}\varepsilon_{klm}v_j\partial_l A_m = v_j\partial_i A_j - v_j\partial_j A_i. \tag{4.45}$$

Hence, Eq. (4.43) can be rewritten as $Q_i = q\left[E_i + (\mathbf{v}\times\mathbf{B})_i\right]$, which is indeed the expression for the Lorentz force as had to be proved. □

4.3 Hydrogen-like atom in an external magnetic field

4.3.1 Effective magnetic moment

Let us consider a hydrogen-like atom at position $\mathbf{r}$ passing through an *inhomogeneous* magnetic field $\mathbf{B}(\mathbf{r})$ as is sketched for a cloud of atoms in Fig. 4.3. Suppose that the atom is in eigenstate $|nlm\rangle$ with respect to the quantization axis defined by the direction of $\mathbf{B}$. If this direction does not change along the trajectory of the atom (or changes adiabatically; i.e., sufficiently slowly), the atom stays in the same eigenstate and its energy can be regarded as a state-dependent effective potential,

$$\mathcal{U}(\mathbf{r}) = \langle\mathcal{H}_0\rangle_{nlm} + \frac{e}{2m_r}\langle\mathbf{L}\cdot\mathbf{B}(\mathbf{r})\rangle_{nlm} + \frac{e^2}{8m_r}\langle r_\perp^2\rangle_{nlm}\mathbf{B}^2(\mathbf{r}). \tag{4.46}$$

To obtain this expression we substituted the charge $q = -e$ and the reduced mass m_r of the electron into the Hamiltonian (4.37). The use of ordinary first-order perturbation theory in this context is justified in Section 4.3.3. The presence of the field gradient results in a force on the atom,

$$\mathbf{F} = -\operatorname{grad}\mathcal{U}(\mathbf{r}) = -\frac{\partial\mathcal{U}}{\partial B}\operatorname{grad}B = \mu_{\text{eff}}\operatorname{grad}B. \tag{4.47}$$

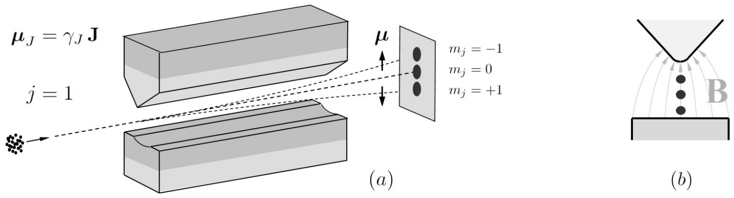


Figure 4.3: Stern-Gerlach apparatus for determining the magnetic moment of atoms by measuring the magnetic force induced by the gradient of $|\mathbf{B}(\mathbf{r})|$: (a) for atoms with unit total electronic angular momentum $J = 1$ three spots are observed. Note that $\gamma_J < 0$; i.e., atoms with magnetic moment polarized parallel to the magnetic field are deflected towards high field (*high-field seekers*); those with magnetic moment anti-parallel towards low field (*low-field seekers*); in state $m_J = 0$ the atoms are not deflected; (b) magnet configuration with a constant magnetic field gradient along the vertical direction in the symmetry plane.

Hence, the force depends on the gradient of the *absolute value* $B(\mathbf{r}) = |\mathbf{B}(\mathbf{r})|$ and can be used to measure μ_{eff}, the *effective magnetic moment.* This method is called the *Stern-Gerlach method* [107]. The effective magnetic moment is given by the first derivative of the effective potential

$$\mu_{\text{eff}} = -\frac{\partial \mathcal{U}}{\partial B} = -(e\hbar/2m_r)m_l - \left(e^2/4m_r\right)\langle r_\perp^2\rangle B. \tag{4.48}$$

The second derivative

$$\alpha_M = \frac{\partial^2 \mathcal{U}}{\partial B^2} = \left(e^2/4m_r\right)\langle r_\perp^2\rangle \tag{4.49}$$

is known as the *magnetic polarizability (magnetizability).* To obtain these results we approximated the field as homogeneous over the size of the atom.

Eqs. (4.46) and (4.48) suggests to rewrite the Hamiltonian (4.37) in the form of an effective Hamiltonian for a classical magnetic dipole in a magnetic field,

$$\mathcal{H} = \mathcal{H}_0 - \boldsymbol{\mu}_L \cdot \mathbf{B} - \tfrac{1}{2}\boldsymbol{\mu}_{dia} \cdot \mathbf{B}, \tag{4.50}$$

where

$$\boldsymbol{\mu}_L = -(e/2m_r)\mathbf{L} \tag{4.51}$$

represents the operator of a *permanent* magnetic moment and is called the *orbital magnetic moment*, and

$$\boldsymbol{\mu}_{dia} = -\left(e^2/4m_r\right) r^2 \sin^2\theta\, \mathbf{B} \equiv -\alpha_M \mathbf{B}, \tag{4.52}$$

is the operator for an *induced* magnetic moment known as the *diamagnetic moment.* As expected for an induced moment, $\boldsymbol{\mu}_{dia}$ opposes the direction of the applied magnetic field $\mathbf{B}$; also $\boldsymbol{\mu}_L$ and $\mathbf{L}$ point in opposite directions, in this case caused by the negative electron charge.

The relation between the magnetic moment $\boldsymbol{\mu}_L$ and the angular momentum $\mathbf{L}$ will be written in the form [1]

$$\boldsymbol{\mu}_L = \gamma_L \mathbf{L} \equiv -g_L \mu_B \left(\mathbf{L}/\hbar\right), \tag{4.53}$$

where $\mathbf{L}/\hbar$ is the dimensionless angular momentum operator ($\mathbf{L}$ in units of $\hbar$). The coefficient

$$\gamma_L = -e/2m_r \equiv -g_L \mu_B/\hbar, \tag{4.54}$$

relating the magnetic moment $\boldsymbol{\mu}_L$ to the angular momentum $\mathbf{L}$ in Eq. (4.51), is called the *gyromagnetic ratio* of the orbiting electron. Its value coincides with that of an orbiting classical body of charge $-e$ and mass m_r. The quantity

$$\mu_B = e\hbar/2m_e \tag{4.55}$$

---

[1] In these lectures we define gyromagnetic ratios always in the same way: positive if the magnetic moment is parallel to the angular momentum. The electron g factors (g_L and g_e) are (traditionally) defined as positive numbers. In contrast, the nuclear g factors will be defined to have the same sign as the nuclear gyromagnetic ratio.

is the *Bohr magneton* ($\mu_B \simeq 9.27 \times 10^{-24}$ J/T). The prefactor g_L is called the orbital g factor. Note that g_L is defined as a positive dimensionless number. Specification of the g factor is the conventional way to compare small differences in orbital magnetic moment between the elements. The orbital g factor can be expressed in terms of the electron-to-nucleus mass ratio

$$g_L = m_e/m_r = (1 + m_e/M) \simeq 1. \tag{4.56}$$

4.3.2 Diamagnetic coupling

Atomic *diamagnetism* finds its origin in the third term of the Hamiltonian (4.50),

$$\mathcal{H}_{dia} = -\tfrac{1}{2}\boldsymbol{\mu}_{dia} \cdot \mathbf{B} = \tfrac{1}{2}\alpha_M B^2. \tag{4.57}$$

For hydrogenic atoms in s states, $|nlm_l\rangle = |n00\rangle \equiv |ns\rangle$, which are eigenstates of the Schrödinger Hamiltonian $\mathcal{H}_0$, the magnetic polarizability can be calculated with the expression for ordinary first-order perturbation theory

$$\alpha_M^{(ns)} = \langle ns|\alpha_M|ns\rangle = (e^2/4m_r)\langle 00|\sin^2\theta|00\rangle\langle r^2\rangle_{ns}. \tag{4.58}$$

The angular average evaluates to $\langle 00|\sin^2\theta|00\rangle = 2/3$ (see Problem 2.5); the reduced matrix element for the radial average follows with Eq. (2.49). For $l = 0$ we obtain

$$\alpha_M^{(1s)} = \frac{n^2}{Z^2}\frac{e^2a^2}{12m_r}(5n^2+1). \tag{4.59}$$

Note that the diamagnetism of the hydrogen ground state is *extremely small*; for a $1\,T$ applied field the induced moment is only $\mu_{dia} = (ea^2/\hbar)g_l\mu_B B \simeq 4 \times 10^{-6}\mu_B$. However, it is important for high-n Rydberg atoms ($n \gtrsim 50$) because the polarizability scales with n^4.

For materials, the magnetic polarizability is related to the *diamagnetic susceptibility* χ_m^{dia}, which is a dimensionless number determining the *diamagnetic magnetization* (magnetic moment per unit volume) for a given applied field, $\mathbf{M}_{dia} = \chi_m^{dia}\mathbf{H}$. For dilute systems (in which the mutual interaction between the atoms may be neglected) the magnetization at position $\mathbf{r}$ is simply the product of the diamagnetic moment and the atomic (number) density $n(\mathbf{r})$,

$$\mathbf{M}_{dia} = n(\mathbf{r})\boldsymbol{\mu}_{dia}. \tag{4.60}$$

Using Eq. (4.52) and $\mathbf{B} = \mu_0\mathbf{H}$ we find for the diamagnetic susceptibility of dilute systems

$$\chi_m^{dia} = -\mu_0\alpha_M\, n(\mathbf{r}). \tag{4.61}$$

4.3.3 Orbital Zeeman coupling

The *orbital Zeeman effect* finds its origin in the second term of the effective Hamiltonian (4.50) which is known as the orbital Zeeman Hamiltonian

$$\mathcal{H}_Z = -\boldsymbol{\mu}_L \cdot \mathbf{B}, \tag{4.62}$$

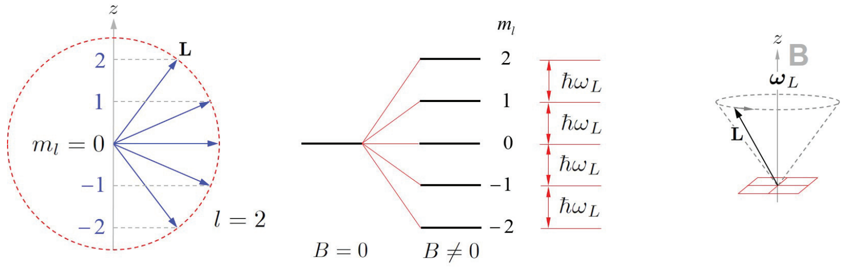
where $\boldsymbol{\mu}_L$ is the orbital magnetic moment operator (4.53). Note that Eq. (4.62) coincides with the expression for the energy of a classical magnetic dipole moment in a magnetic field.

Let us analyze how the degeneracy of the magnetic sublevels is lifted by applying a magnetic field.[1] Since $\mathbf{L}^2$ and L_z commute with $\mathcal{H}_0$ the orbital Zeeman Hamiltonian is *diagonal* in the hydrogenic basis, $\{|nlm_l\rangle\}$; i.e., the orbital Zeeman shift is given by

$$\Delta E_Z = \langle nlm_l|\mathcal{H}_Z|nlm_l\rangle = \langle lm_l|\mathcal{H}_Z|lm_l\rangle. \tag{4.63}$$

---

[1] Pure orbital Zeeman splitting is not observed in atomic systems. It always comes in combination with spin-orbit interaction (see Section 4.5.1).



(*a*) (*b*) (*c*)

Figure 4.4: (*a*) Semi-classical vector model of angular momentum for the case $l = 2$. The diagram shows the $2l+1$ projections of the semi-classical vector **L** along the direction of the magnetic field (z-direction); (*b*) The $(2l+1)$-fold degeneracy of the energy-levels is lifted when applying a magnetic field (Zeeman effect-in the absence of electronic spin). For 1 Bohr magneton the splitting amounts 1.4 MHz/Gauss (67 μK/Gauss); (c) The projections stay constant in time, also in the presence of Larmor precession which has the same angular frequency ($\boldsymbol{\omega}_L$) for all m_l values.

Hence, l and m_l are good quantum numbers and, for given l, the Zeeman operator acts in a $(2l+1)$-dimensional sector of Hilbert space. The relativistic corrections do not affect the magnetic shifts and diamagnetic corrections may be neglected as long as the magnetic field is sufficiently small.

Substituting Eq. (4.53) into (4.62) the Zeeman Hamiltonian becomes

$$\mathcal{H}_Z = -\gamma_L \mathbf{L} \cdot \mathbf{B} = g_L \mu_B B \left(L_z/\hbar \right), \tag{4.64}$$

where the z direction is again chosen along the direction of the **B** field. This Hamiltonian expresses the coupling of **L** to the **B** field (*orbital Zeeman coupling*). In this notation the Zeeman shifts of the $2l+1$ magnetic sublevels are given by

$$\Delta E_Z = -\gamma_L B \langle lm_l | L_z | lm_l \rangle = g_L \mu_B B m_l. \tag{4.65}$$

Comparing Eqs. (4.67) and (4.65) we note that, classically spoken, $\boldsymbol{\mu}_L$ is never parallel to **B** (see Fig. 4.4*a*). The quantity referred to as the *magnitude* of the orbital magnetic moment (short: *the* orbital moment) is defined as

$$\mu_L \equiv g_L (e/2m_e) \langle ll | L_z | ll \rangle = g_L \mu_B l. \tag{4.66}$$

It corresponds to the *maximum observable* magnetic moment (for given l) and should not be confused with the operator $\boldsymbol{\mu}_L$ or with the *semi-classical magnitude* of the orbital moment,

$$\sqrt{\langle \boldsymbol{\mu}_L^2 \rangle} = \sqrt{\langle lm_l | \boldsymbol{\mu}_L \cdot \boldsymbol{\mu}_L | lm_l \rangle} = g_L (e/2m_e) \sqrt{\langle lm_l | \mathbf{L}^2 | lm_l \rangle} = g_L \mu_B \sqrt{l(l+1)}. \tag{4.67}$$

As mentioned in Section 4.3.1 the *effective magnetic moment* of a given magnetic sublevel is defined as the derivative of the Zeeman energy with respect to magnetic field. With Eq. (4.65) we find

$$\mu_{\text{eff}} = -\partial[\Delta E_Z(B)]/\partial B = -g_L \mu_B m_l. \tag{4.68}$$

For the purely orbital Zeeman coupling μ_{eff} is a constant for given value of m_l but in more general cases of magnetic coupling it will depend on magnetic field (cf. Section 4.7). For an effective magnetic moment of 1 Bohr magneton, $\mu_{\text{eff}} = \mu_B \simeq 9.27 \times 10^{-24}$ J/T, the level shift corresponds to 1.4 MHz/Gauss (67 μK/Gauss). With Eq. (4.65) we have established that in a magnetic field the $(2l+1)$-fold orbital degeneracy is lifted, with the lowest energy corresponding to the lowest value of the magnetic quantum number ($m_l = -l$). This is illustrated in Fig. 4.4*b*.

4.3.4 Larmor precession

In a semi-classical picture, the Zeeman Hamiltonian (4.62) gives rise to a torque,

$$d\mathbf{L}(t)/dt = \boldsymbol{\mu}_L \times \mathbf{B}, \tag{4.69}$$

tending to line up the magnetic moment with the magnetic field giving rise to precession about $\mathbf{B}$ because hamiltonian evolution is dissipation free. This is known as *Larmor precession*. Substituting Eq. (4.51) into Eq. (4.69) the expression for the torque describes the precession of the angular momentum vector $\mathbf{L}$ about the field direction,

$$d\mathbf{L}(t)/dt = \gamma_L \mathbf{L} \times \mathbf{B}. \tag{4.70}$$

Rewriting this expression in the form

$$d\mathbf{L}/dt = \boldsymbol{\omega}_L \times \mathbf{L} \tag{4.71}$$

we find that the angular momentum $\mathbf{L}$ precesses about $\mathbf{B}$ at angular frequency

$$\boldsymbol{\omega}_L = -\gamma_L \mathbf{B}. \tag{4.72}$$

This is illustrated in Fig. 4.4*c*. The precession frequency is known as the *Larmor* frequency,

$$\omega_L \equiv |\boldsymbol{\omega}_L| = g_L \mu_B B/\hbar. \tag{4.73}$$

In terms of the Larmor frequency the Zeeman energy can be written as

$$\Delta E_Z = -\hbar\gamma_L B m_l = g_L \mu_B B m_l = \hbar\omega_L m_l. \tag{4.74}$$

Quantum mechanically, the precession follows from the Heisenberg equations of motion. In this approach all time-dependence is contained in the Heisenberg operator $\mathbf{L}(t) \equiv U^\dagger(t,t_0)\mathbf{L}U(t,t_0)$, where $U(t,t_0) = \exp[-i(t-t_0)\mathcal{H}_Z/\hbar]$ is the evolution operator (see Appendix F.2). As the Zeeman Hamiltonian is time independent, the precession of $\mathbf{L}(t)$ is described by

$$d\mathbf{L}(t)/dt = -(i/\hbar)\,[\mathbf{L}(t), \mathcal{H}_Z]. \tag{4.75}$$

Writing the Zeeman Hamiltonian in the form $\mathcal{H}_Z = -\gamma_L\, \mathbf{L}\cdot\mathbf{B}$ the equation of motion becomes

$$d\mathbf{L}(t)/dt = (i/\hbar)\,\gamma_L[\mathbf{L}(t), \mathbf{L}(t)\cdot\mathbf{B}]. \tag{4.76}$$

To evaluate this expression we consider the time evolution of the angular momentum components $L_i(t)$. With the aid of the commutation relations (1.29) and using the Einstein summation convention we have $[L_i, L_k] = -\varepsilon_{ijk} i\hbar L_j$ and Eq. (4.76) becomes

$$dL_i/dt = (i/\hbar)\,\gamma_L[L_i, L_k B_k] = (i/\hbar)\,\gamma_L[L_i, L_k]B_k = \gamma_L \varepsilon_{ijk} L_j B_k. \tag{4.77}$$

Returning to the vector notation this becomes

$$d\mathbf{L}(t)/dt = \gamma_L\,[\mathbf{L}(t)\times\mathbf{B}]\,. \tag{4.78}$$

This expression may be rewritten in two equivalent forms

$$d\mathbf{L}/dt = \boldsymbol{\mu}_L \times \mathbf{B} \text{ and } \quad d\mathbf{L}/dt = \boldsymbol{\omega}_L \times \mathbf{L}. \tag{4.79}$$

Here the *operator* $\mathbf{L}$ is found to be described by the well-known equation of motion (4.71) from classical mechanics describing the precession at angular frequency $\boldsymbol{\omega}_L$ of a spinning top of angular momentum $\mathbf{L}$ under the influence of the torque $\boldsymbol{\mu}_L \times \mathbf{B}$.

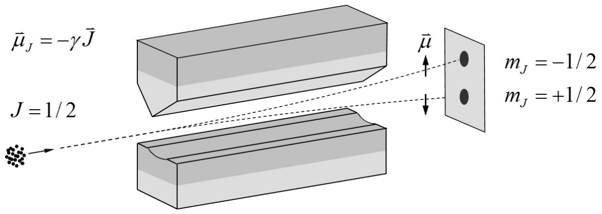


Figure 4.5: Observation of half-integral angular momentum atoms by the Stern-Gerlach method.

4.3.5 Spin

Spin was postulated by Goudsmit and Uhlenbeck in 1925 to provide a *physical interpretation* for the appearance of half-integral quantum numbers in the analysis of the atomic spectra [113, 114]. At the time of the proposal spectroscopic evidence for the existence of half integral quantum numbers was available and particularly elegantly demonstrated in the atomic beam experiment of Stern and Gerlach with silver atoms - see Fig. 4.5 [50]. Building on the spectroscopic evidence Pauli had already established the double-valuedness of the *electronic* wavefunction (see Appendix F.1.4) which points to the presence of an additional quantum number for the electron (a fourth degree of freedom) but he refrained from giving an interpretation. In his words the double-valuedness was "not describable classically" [80]. In their postulate Uhlenbeck and Goudsmit attribute this degree of freedom to self-rotation of the electron with angular momentum 1/2 in units of $\hbar$ and a *gyromagnetic ratio* twice as large as the orbital value $-e/2m_e$.

Superficially, the presence of electron spin may not come as a surprise because it is well known from classical mechanics that an orbiting satellite will generally show solid body rotation. However, in 1925 the modeling of the electron as some rotating cloud of charge had a two decade history of disappointment. It was not clear what was holding the charge together and to generate the required magnetic moment special relativity had to be violated; i.e., the cloud had to rotate so fast that locally the speed of light was exceeded. A difficulty of all mechanical models was that they even failed to provide a *self-consistent* description for the *gyromagnetic ratio.* The postulate of Goudsmit and Uhlenbeck was no exception in this respect: on the one hand the g factor of the electron had to have the value 2 (to describe the Zeeman effect in high fields - see Section 4.7.4); on the other hand using this value to estimate the zero-field splitting the calculated value was a factor of 2 too large (by itself this would imply a g factor of 4). Therefore, it is not surprising that the postulate only started to gain acceptance after it was demonstrated by Thomas that the proper relativistic analysis yields the value 2 in both cases (see Section 4.4.2) [108, 110]. However, it remained to be explained why spin represents a *fixed* quantity of angular momentum (called the *intrinsic angular momentum*) rather than variable amounts as observed for classical bodies.

In 1928 the mentioned properties were captured in the famous Dirac theory for the electron [35]. In this theory the spinning electron emerges as a manifestly non-classical object. It represents the first example of an *elementary particle* on which the *standard model* is based. In this respect the postulate marks the beginning of *particle physics.* The electron is an elementary particle of mass m_e, elementary charge $q = -e$ and spin $s = \frac{1}{2}$. Within experimental error the particle is of *zero size* [46] and *without* electric dipole moment [58]. The particle spin, together with its magnetic quantum number $m_s = \pm\frac{1}{2}$ defines the spin state of the electron. A difficulty of the Dirac theory is that the energy contained in the electric field of a point charge diverges. It took until 1948 before this problem was resolved with the development of *quantum electrodynamics* (QED) by Tomonaga, Schwinger and Feynman. Close to the singularity the electromagnetic energy density becomes so large that *virtual* excitation (see Appendix H.2) of electron-positron pairs becomes substantial. These virtual pairs constitute electric dipoles that live only briefly but long enough to be polarized by the electric field. This is known as *vacuum polarization.* The induced electric field counteracts the Coulomb field of the electron in such a way that the divergence can be eliminated and the experimental values of the electron charge and mass are obtained.

In view of the above it is evident that we should not think simplistically about the electron. This being said, the essential atomic physics can be explained without entering in the origin of spin and so we shall do; after all, the spin was postulated to successfully represent the properties of the atom. An exceptionally lucent account of the early developments that led to our current notion of spin is given in the book *The Story of Spin* by Tomonaga [112].

4.3.6 Spin Zeeman coupling

The spin manifests itself by its magnetic moment

$$\boldsymbol{\mu}_s = \gamma_e \mathbf{S} \equiv -g_e \mu_B \left(\mathbf{S}/\hbar\right), \tag{4.80}$$

where $\mathbf{S}$ is the electron-spin operator and

$$\gamma_e \equiv -g_e \mu_B/\hbar \tag{4.81}$$

the electron gyromagnetic ratio, $\gamma_e/2\pi \simeq -2.8025$ MHz/Gauss, with μ_B the Bohr magneton and g_e the electron g factor defined as a *positive* number. Hence, like $\boldsymbol{\mu}_L$ and $\mathbf{L}$ also $\boldsymbol{\mu}_s$ and $\mathbf{S}$ point in opposite directions. In analogy to the orbital case (4.66) the *magnitude* of the intrinsic magnetic moment (short: *the* intrinsic moment) is defined as the maximum observable moment,

$$\mu_e \equiv \tfrac{1}{2} g_e \mu_B. \tag{4.82}$$

Advanced analysis based on the Dirac theory for the electron predicts $g_e = 2$. Experiment shows g_e to be slightly larger, $g_e \simeq 2.002\,319$. Therefore, the electron is said to have an *anomalous magnetic moment* [67],

$$g_e = 2(1 + a_e). \tag{4.83}$$

The deviation is known as the *electron anomaly*, $a_e \equiv (g_e - 2)/2 \simeq \alpha/2\pi$, with numerical value $a_e \simeq 1.159\,65 \times 10^{-3}$ [77]. The electron anomaly finds its origin in the zero-point fluctuations of the electromagnetic field. The free-electron value differs slightly from that of electrons bound in atoms. It continues to be subject of intensive experimental and theoretical research as it provides an important test of quantum electrodynamics (QED). In contemporary experiments the spin state of a single electron is measured without destroying the state by observing spin-flip events of a trapped electron in real time [77, 46, 54]. This is an example of a Quantum Non Demolition (QND) experiment. It has demolished old wisdom [82] that it is fundamentally impossible to observe the magnetic moment of an electron by a Stern-Gerlach-like experiment [31].

Knowing the magnetic moment, the spin Zeeman Hamiltonian takes the generic form

$$\mathcal{H}_Z = -\boldsymbol{\mu}_s \cdot \mathbf{B}. \tag{4.84}$$

Substituting Eq. (4.80) we obtain

$$\mathcal{H}_Z = (g_e \mu_B/\hbar)\, \mathbf{S} \cdot \mathbf{B} = g_e \mu_B B \left(S_z/\hbar\right), \tag{4.85}$$

where the z direction is again chosen along the direction of the $\mathbf{B}$ field. This Hamiltonian expresses the coupling of $\mathbf{S}$ to the $\mathbf{B}$ field (*spin Zeeman coupling*). The Zeeman Hamiltonian (4.84) gives rise to a torque, which can be written in two equivalent forms

$$d\mathbf{S}/dt = \boldsymbol{\mu}_s \times \mathbf{B} \text{ and } \quad d\mathbf{S}/dt = \boldsymbol{\omega}_e \times \mathbf{S}, \tag{4.86}$$

where $\boldsymbol{\omega}_e = -\gamma_e \mathbf{B}$ is the angular frequency vector. The torque tends to line up the magnetic moment with the magnetic field but (in view of the absence of dissipation) this gives rise to precession about $\mathbf{B}$ at frequency

$$\omega_e = |\boldsymbol{\omega}_e| = g_e \mu_B B/\hbar = (g_e/g_L)\omega_L, \tag{4.87}$$

where ω_L is the Larmor frequency (4.73) and g_L the orbital angular momentum g factor.

We analyze the level splitting by the spin Zeeman coupling in the same way as we did for the orbital Zeeman coupling.[1] Since $\mathbf{S}^2$ and S_z commute with $\mathcal{H}_0$, $\mathbf{L}^2$ and L_z, the spin Zeeman operator is *diagonal* in the hydrogenic basis extended by a spin sector $\{|nlm_l\rangle\} \to \{|nlm_l s m_s\rangle\}$,

$$\Delta E_Z = \langle nlm_l s m_s|\mathcal{H}_Z|nlm_l s m_s\rangle = \langle s m_s|\mathcal{H}_Z|s m_s\rangle. \tag{4.88}$$

Hence, s and m_s are good quantum numbers, with $s = \frac{1}{2}$ for one-electron atoms. In general, the spin Zeeman operator acts in a $(2S+1)$-dimensional sector of Hilbert space. Substituting Eq. (4.85 for $\mathcal{H}_Z$ the spin Zeeman shift becomes

$$\Delta E_Z = -\gamma_e B\langle s m_s|S_z|s m_s\rangle = g_e \mu_B m_s B = (g_e/g_L) m_s \hbar\omega_L. \tag{4.89}$$

Like the orbital Zeeman energy, also the spin Zeeman energy can be expressed in a form containing either the Bohr magneton μ_B or the Larmor frequency ω_L. For a spin-up electron ($g_e \simeq 2$ and $m_s = \frac{1}{2}$) the effective magnetic moment equals 1 Bohr magneton, just like in the case of the orbital Zeeman coupling ($\mu_L \simeq \mu_e$). Thus, in this case the Zeeman shift is 1.4 MHz/Gauss.[2] To observe the characteristic two spot Stern-Gerlach image of a spin-$\frac{1}{2}$ system (see Fig. 4.5) the effective magnetic moment must be close to 1 Bohr magneton along the semi-classical trajectory of the atoms (in Section 5.4.2 we show that this condition is satisfied for $B > |B_{\mathrm{hfs}}|$, where $|B_{\mathrm{hfs}}|$ is the *hyperfine crossover field*).

4.3.7 Zeeman Hamiltonian for the electron

At this point we can formulate the Zeeman Hamiltonian for the electron, which describes the interaction of the electron with an externally applied static magnetic field. Combining Eqs. (4.84), (4.62) and (4.57) we find,

$$\mathcal{H}_Z = -(\boldsymbol{\mu}_s + \boldsymbol{\mu}_L)\cdot\mathbf{B} + \tfrac{1}{2}\alpha_M B^2. \tag{4.90}$$

Later, in Chapter 5, a nuclear contribution will be added to obtain the Zeeman Hamiltonian for the complete atom. The Zeeman Hamiltonian is an effective hamiltonian. It can be used with perturbation theory as long as the electronic wavefunctions are not substantially distorted by the applied field. This is the case for a large class of experiments in atomic physics laboratories. It breaks down under extreme conditions. Near neutron stars the magnetic field can be millions of times stronger than the strongest fields, 100-1000 T, that can be achieved in laboratories on earth. In these lectures we restrict ourselves to perturbative conditions.

4.4 Fine-structure Hamiltonian

4.4.1 Addition of orbital angular momentum and spin

As long as we may neglect the interaction with the environment the atom represents a closed mechanical system and the total angular momentum is a conserved quantity. In general, this total angular momentum is the result of several contributions. In the description of the atomic fine structure we restrict ourselves to orbital angular momentum ($\mathbf{L}$) and spin ($\mathbf{S}$). The total angular momentum is given by the vector sum

$$\mathbf{J} = \mathbf{L} + \mathbf{S}. \tag{4.91}$$

---

[1] Pure electron spin Zeeman splitting is observed in isotopes of *chromium*, Cr, and *molybdenum*, Mo (see Section 9.2.2). In one-electron atoms spin Zeeman coupling does not occur without hyperfine interaction (see Chapter 5).

[2] Note that we are dealing here with *level* shifts, not to be confused with the Zeeman shift of spin-flip *transition frequencies*, which are twice as big (for spin 1/2).

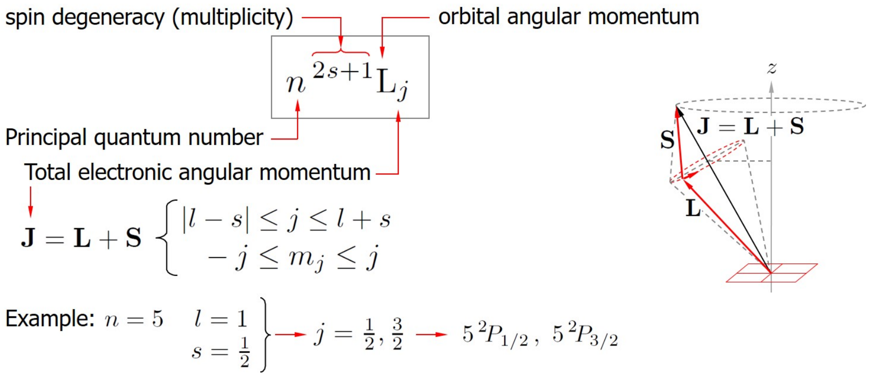


Figure 4.6: Addition of angular momenta allowing for $\mathbf{L}\cdot\mathbf{S}$ coupling. The states (so-called LS Terms), are written in the spectroscopic notation (Term notation) as $n\,^{(2s+1)}\mathrm{L}_j$, where n is the principal quantum number, s the quantum number for the electronic spin, and j the quantum number for the total electronic angular momentum $\mathbf{J}$ (which is a conserved quantity in the absence of coupling to the environment). The symbol L stands for the total orbital angular momentum $\mathrm{L}=S,P,D,F\cdots$ for $l=0,1,2,\cdots$, respectively. Note that the angular momentum projections L_z and S_z are proportional to J_z. This provides the geometric foundation of the Wigner-Eckart theorem as applied to vector operators.

Starting from the commutation relations for $\mathbf{L}$ and $\mathbf{S}$ separately it is straightforward to show that also $\mathbf{J}$ obeys commutation relations of the type (1.29),

$$[J_x,J_y]=i\hbar J_z,\ [J_y,J_z]=i\hbar J_x \text{ and } [J_z,J_x]=i\hbar J_y. \tag{4.92}$$

Recalling Section 3.1 this implies that $\mathbf{J}$ satisfies all properties of quantized angular momenta,

$$\mathbf{J}^2\,|lsj,m_j\rangle=j(j+1)\hbar^2\,|lsj,m_j\rangle \tag{4.93}$$

$$J_z\,|lsj,m_j\rangle=m_j\hbar\,|lsj,m_j\rangle\,, \tag{4.94}$$

with m_j restricted to the interval

$$-j\le m_j\le j. \tag{4.95}$$

In many cases the values of l and s are irrelevant or obvious from the context. In such cases these quantum numbers are usually omitted from the notation, writing simply $|j,m_j\rangle$ rather than $|lsj,m_j\rangle$. As shown in Section 3.7 the value of j is restricted by the values of l and s; j can take all values differing by one unit of angular momentum in the interval as expressed by the *triangle inequality*

$$|l-s|\le j\le l+s. \tag{4.96}$$

Note that j is positive definite for half-integral spin (the vector sum can only become zero if the two angular momenta are equal in magnitude and opposite in sense). The action of the shift operators $J_\pm$ are given by (see Eq. 3.20)

$$J_\pm\,|j,m_j\rangle=\sqrt{j\,(j+1)-m_j(m_j\pm 1)}\hbar\,|j,m_j\pm 1\rangle\,. \tag{4.97}$$

Since the components of $\mathbf{J}$ commute with the Schrödinger Hamiltonian, the total electronic angular momentum is a conserved quantity. Also, when extending the Hamiltonian by a so-called *internal coupling term* of the type $\mathcal{H}'\propto\mathbf{L}\cdot\mathbf{S}$ the total angular momentum $\mathbf{J}$ remains conserved (just as we expect for a closed mechanical system). In the latter case, $\mathbf{L}$ and $\mathbf{S}$ exert a torque on each other, which results (given the absence of dissipation) in a joint precession of $\mathbf{L}$ and $\mathbf{S}$ about $\mathbf{J}$ as is illustrated in Fig. 4.6. This is expressed by the commutation relations demonstrated in Problems 4.3-4.6 (see also Section 3.5).

For given value of j the operator $\mathbf{J}$ acts in a $(2j+1)$-dimensional subspace of the $(2l+1)\,(2s+1)$-dimensional product space defined by the so-called *uncoupled basis* $\{|lm_l sm_s\rangle \equiv |lm_l\rangle \otimes |sm_s\rangle\}$. The sum of the $(2j+1)$-dimensional subspaces of all values of j allowed by the triangular inequality can also be represented by the so-called *coupled basis* $\{|lsjm_j\rangle\}$. The coupled and the uncoupled basis are related by a unitary transformation (see Problem J.2).

Problem 4.3. Prove the following commutation relations $[\mathbf{L}^2, \mathbf{L}\cdot\mathbf{S}] = [\mathbf{S}^2, \mathbf{L}\cdot\mathbf{S}] = 0$.

Solution. Decomposing $\mathbf{L}\cdot\mathbf{S}$ we find $[\mathbf{L}^2, \mathbf{L}\cdot\mathbf{S}] = [\mathbf{L}^2, L_xS_x] + [\mathbf{L}^2, L_yS_y] + [\mathbf{L}^2, L_zS_z] = 0$ because $[\mathbf{L}^2, L_iS_i] = L_i[\mathbf{L}^2, S_i] + [\mathbf{L}^2, L_i]S_i = 0$ for $i \in \{x, y, z\}$. $[\mathbf{S}^2, \mathbf{L}\cdot\mathbf{S}] = 0$: Idem replacing $\mathbf{L}^2$ by $\mathbf{S}^2$. □

Problem 4.4. Prove the commutation relation $[J_z, \mathbf{L}\cdot\mathbf{S}] = 0$.

Solution. Decomposing $\mathbf{L}\cdot\mathbf{S}$ we find $[J_z, \mathbf{L}\cdot\mathbf{S}] = \frac{1}{2}[J_z, \mathbf{J}^2] - \frac{1}{2}[J_z, \mathbf{L}^2] - \frac{1}{2}[J_z, \mathbf{S}^2] = 0$ because $[J_z, \mathbf{J}^2] = 0$, $[J_z, \mathbf{L}^2] = [L_z + S_z, \mathbf{L}^2] = 0$ and $[J_z, \mathbf{S}^2] = [L_z + S_z, \mathbf{S}^2] = 0$. □

Problem 4.5. Prove the commutation relation $[\mathbf{J}^2, \mathbf{L}\cdot\mathbf{S}] = 0$.

Solution. Decomposing $\mathbf{J}^2$ we find $[\mathbf{J}^2, \mathbf{L}\cdot\mathbf{S}] = [\mathbf{L}^2, \mathbf{L}\cdot\mathbf{S}] + [\mathbf{S}^2, \mathbf{L}\cdot\mathbf{S}] + 2[\mathbf{L}\cdot\mathbf{S}, \mathbf{L}\cdot\mathbf{S}] = 0$ because $[\mathbf{L}^2, \mathbf{L}\cdot\mathbf{S}] = [\mathbf{S}^2, \mathbf{L}\cdot\mathbf{S}] = 0$ (see Problem 4.3) and $[\mathbf{L}\cdot\mathbf{S}, \mathbf{L}\cdot\mathbf{S}] = (\mathbf{L}\cdot\mathbf{S})^2 - (\mathbf{L}\cdot\mathbf{S})^2 = 0$. □

Problem 4.6. Prove the commutation relations $[\mathbf{J}^2, \mathbf{L}^2] = [\mathbf{J}^2, \mathbf{S}^2] = 0$, where $\mathbf{J} = \mathbf{L} + \mathbf{S}$.

Solution. Decomposing $\mathbf{J}^2$ we find $[\mathbf{J}^2, \mathbf{L}^2] = [\mathbf{L}^2, \mathbf{L}^2] + [\mathbf{S}^2, \mathbf{L}^2] + 2[\mathbf{L}\cdot\mathbf{S}, \mathbf{L}^2] = 0$ since $[\mathbf{L}\cdot\mathbf{S}, \mathbf{L}^2] = 0$ (see Problem 4.3); further $[\mathbf{L}^2, \mathbf{L}^2] = 0$ (any scalar operator commutes with itself) and $[\mathbf{S}^2, \mathbf{L}^2] = 0$ because $\mathbf{S}$ and $\mathbf{L}$ operate in orthogonal vector spaces. $[\mathbf{J}^2, \mathbf{S}^2] = 0$: Idem replacing $\mathbf{L}^2$ by $\mathbf{S}^2$. □

Problem 4.7. Show that $[L_z, \mathbf{L}\cdot\mathbf{S}] \neq 0$ and $[S_z, \mathbf{L}\cdot\mathbf{S}] \neq 0$ although $[J_z, \mathbf{L}\cdot\mathbf{S}] = 0$.

Solution. Decomposing $\mathbf{L}\cdot\mathbf{S}$ we find (using the summation convention) $[L_z, \mathbf{L}\cdot\mathbf{S}] = [L_z, L_\nu S_\nu] = L_\nu[L_z, S_\nu] + [L_z, L_\nu]S_\nu = [L_z, L_\nu]S_\nu$. Because $[L_z, L_\nu] = 0$ for $\nu = z$ we have $[L_z, \mathbf{L}\cdot\mathbf{S}] = [L_z, L_x]S_x + [L_z, L_y]S_y = i\hbar(L_yS_x - L_xS_y) \neq 0$. Likewise, we find $[S_z, \mathbf{L}\cdot\mathbf{S}] = i\hbar(L_xS_y - L_yS_x) \neq 0$. □

4.4.2 Velocity-induced magnetic field

As is well-known from classical electrodynamics, a particle moving at velocity $\mathbf{v}$ in the presence of an electric field $\mathbf{E}$ experiences a velocity-induced magnetic field, which is given by

$$\mathbf{B} = (\mathbf{E} \times \mathbf{v})/c^2, \tag{4.98}$$

where c is the speed of light. This follows from a special-relativistic analysis in the limit where Lorentz contraction can be neglected (cf. Appendix D). Nonrelativistically, this result can also be obtained with the aid of the Biot and Savart law by considering the current resulting from the relative motion of the nucleus with respect to the electron. Taking the relativistic point of view, an electron orbiting at velocity $\mathbf{v}$ in the electric field $\mathbf{E}$ of a nucleus will experience the velocity-induced field and its spin magnetic moment will couple to it. This coupling is similar to the spin Zeeman coupling and is known as *spin-orbit coupling.* To estimate this coupling we note that the electric field of the nucleus at the position of the electron is given by

$$\mathbf{E} = -\boldsymbol{\nabla}\varphi(r) = -\frac{\mathbf{r}}{r}\frac{d\varphi(r)}{dr}, \tag{4.99}$$

where $\varphi(r)$ is the electromagnetic scalar potential of the nucleus. Substituting $\mathbf{L} = m_r(\mathbf{r} \times \mathbf{v})$ for the orbital angular momentum, the velocity-induced magnetic field can be written as

$$\mathbf{B}_L = -\frac{1}{m_r c^2}\frac{1}{r}\frac{d\varphi(r)}{dr}\mathbf{L}. \tag{4.100}$$

This expression holds for all *hydrogen-like atoms.* For *hydrogenic* atoms $\varphi(r)$ is a simple Coulomb potential. For *alkali-like* atoms the nuclear charge is screened by a spherical shell of *core electrons* (see Section 4.6 for a phenomenological introduction of screening phenomena). A way to account for the screening is to replace the Coulomb potential by a *screened Coulomb potential*[1]

$$\varphi(r) = \frac{e}{4\pi\varepsilon_0}\frac{Z_{nl}(r)}{r}, \tag{4.101}$$

where $e\,Z_{nl}(r)$ is called the *effective nuclear charge* for an electron (the *spectator electron*) in the nl shell at distance r from the nucleus. Close to the nucleus screening is absent, at large distance screening by the core electrons is maximal,

$$\lim_{r\to 0} Z_{nl}(r) = Z \quad \text{and} \quad \lim_{r\to\infty} Z_{nl}(r) = Z_c \equiv 1 + Z_{ion}. \tag{4.102}$$

Here $e\,Z_c$ is called the *Rydberg charge* of the atom/ion of ionic charge $e\,Z_{ion}$. The Rydberg charge is the effective nuclear charge in the limit of perfect screening by the core electrons and is used for the description of atoms/ions excited to Rydberg states. For neutral atoms $Z_c = 1$, for a singly charged ion $Z_c = 2$, *etc.*. In the hydrogenic case ($Z_{ion} = Z - 1$) screening is absent at all distances; i.e., $Z_{nl}(r) \equiv Z = Z_c$. Since the derivative of the scalar potential determines $\mathbf{B}_L$ it is convenient to introduce the *spin-orbit screening function* $\mathcal{Z}_{nl}(r)$,

$$\frac{d\varphi(r)}{dr} = \frac{e}{4\pi\varepsilon_0}\frac{1}{r^2}\left(r\frac{dZ_{nl}(r)}{dr} - Z_{nl}(r)\right) \equiv -\frac{e}{4\pi\varepsilon_0}\frac{\mathcal{Z}_{nl}(r)}{r^2}. \tag{4.103}$$

As long as $Z_{nl}(r)$ *decreases monotonically* with increasing r we have

$$\mathcal{Z}_{nl}(r) \equiv Z_{nl}(r) - r\,dZ_{nl}(r)/dr > 0, \tag{4.104}$$

with $\mathcal{Z}_{nl}(r) \equiv Z$ in the hydrogenic case (for any state $|\kappa\rangle$ of the electron).

Treating the nucleus heuristically as a homogeneously charged sphere of radius r_n and total charge Ze, the electric field $\mathbf{E}(r)$ corresponds to a Coulomb field, for $r \le r_n$ attenuated proportionally to r (see Fig. 4.2). Hence, the velocity-induced magnetic field can be written as

$$\mathbf{B}_L = \begin{cases} \mathcal{Z}_{nl}(r)\dfrac{\mu_0}{4\pi}\dfrac{e}{m_r}\dfrac{\mathbf{L}}{r^3} & \text{for } r > r_n \\[2ex] Z\dfrac{\mu_0}{4\pi}\dfrac{e}{m_r}\dfrac{\mathbf{L}}{r_n^3} & \text{for } r \le r_n, \end{cases} \tag{4.105}$$

where μ_0 is the *magnetic permeability of vacuum.*. Note that the finite core size prevents the divergence of the $\mathbf{E}$ and $\mathbf{B}$ fields for $r \to 0$. Classically, the field experienced by the electron as a result of its orbital motion at nuclear distance $r = a_0$ is *large*, as follows from the prefactor

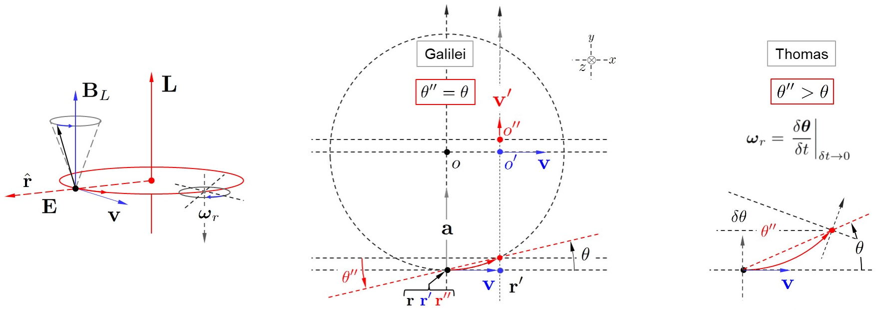
$$\frac{\mu_0}{4\pi}\frac{e}{m_e}\frac{\hbar}{a_0^3} \simeq 12.517\ \text{T}. \tag{4.106}$$

Quantum mechanically we have to average over the orbital wavefunction, which results in a reduction of this field. Calculating the modulus of the velocity-induced magnetic field using first-order perturbation theory we obtain for an electron in the state $|\kappa\rangle = |nlm_l\rangle$

$$B_L = \sqrt{\langle nlm_l|\mathbf{B}_L^2|nlm_l\rangle} = \frac{\mu_0}{4\pi}\frac{e\hbar}{m_r}\langle\mathcal{Z}(r)/r^3\rangle_{nl}\sqrt{\langle lm_l|\,(\mathbf{L}/\hbar)^2\,|lm_l\rangle}. \tag{4.107}$$

---

[1] We restrict ourselves to centrally symmetric screening potentials.



(a) (b) (c)

Figure 4.7: (*a*) Vector diagram showing the precession cone of the electron spin **S** in the velocity induced field $\mathbf{B}_L$. Note that the spin precesses in the same direction as the orbital motion **L**. The so-called Thomas rotation of the electron rest frame proceeds in the opposite direction at rate $\boldsymbol{\omega}_r$; (*b*) Orbital motion in three inertial frames analyzed under galilean relativity ($\theta'' = \theta$); (c) idem under special relativity; note that the Thomas rotation is clockwise in the xy plane of the orbit, $\delta\theta = \theta'' - \theta > 0$.

Turning to dimensionless quantities $\mathbf{L}/\hbar$ and $\rho = r/a$, with $a = (m_e/m_r)\, a_0$, and neglecting the finite size of the nucleus we obtain by substitution of Eqs. (2.49c) and (1.52) for the *hydrogenic* case

$$B_L = \begin{cases} \dfrac{\mu_0}{4\pi} \dfrac{e}{m_r} \dfrac{\hbar}{a^3} \dfrac{Z^4}{n^3} \dfrac{\sqrt{l(l+1)}}{(l+1)\,(l+{}^1\!/\!_2)\,l} & \text{for } l > 0 \\ 0 & \text{for } l = 0. \end{cases} \tag{4.108}$$

For $l > 0$ Eq. (4.108) is well-behaved. For instance, for the $2P$ level of hydrogen ($Z = 1,\ n = 2,\ l = 1$) we calculate $B_L \simeq 0.74$ T. For $l = 0$ the internal field tends to diverge but taking into account the nonzero nuclear size this divergence turns out to be an artifact of the pure Coulomb field. Using Eq. (4.105) (for $r < r_n$) the divergence is avoided and B_L turns out to be zero as a result of the angular averaging, $\langle 00|\,(\mathbf{L}/\hbar)^2\,|00\rangle = 0$.

4.4.3 Spin-orbit coupling and Thomas precession

The velocity-induced field $\mathbf{B}_L$ exerts a torque on the magnetic moment of the orbiting electron. This is called *spin-orbit* coupling. Just as for the *spin-Zeeman* coupling the torque gives rise to precession of the electron spin about the direction of the magnetic field,

$$(d\mathbf{S}/dt)_{rest} = \boldsymbol{\mu}_s \times \mathbf{B}_L = -g_e(e/2m_e)\mathbf{S} \times \mathbf{B}_L. \tag{4.109}$$

In view of Eq. (4.100) $\mathbf{B}_L$ points in the same direction as the angular momentum **L** (see Fig. 4.7*a*). Eq. (4.109) holds for the electron rest frame[1] (in which the nucleus orbits about the electron). For convenience the nucleus is taken to be at rest in the laboratory.

To quote Llewellyn Thomas: "*There is, however, an error in the above reasoning*" [108]. We overlooked the centripetal acceleration $\mathbf{a} = d\mathbf{v}/dt$, which is essential to keep the electron in its orbit (see Fig. 4.7*b*). The repair requires some special relativity. As shown in Problem 4.8 the curvature of the orbit gives rise to a *pure rotation* of the moving electron rest frame with respect to the laboratory frame at the rate $\boldsymbol{\omega}_r = -\boldsymbol{\omega}_T$ (for the configuration of Fig. 4.7*b*), where

$$\boldsymbol{\omega}_T \equiv \frac{\mathbf{v} \times \mathbf{a}}{2c^2}. \tag{4.110}$$

This is a purely *kinematical* effect[1]; i.e., it arises independently from the origin of the curvature (so it also exists for $\mathbf{B}_L = 0$). Note that in the absence of a magnetic field also spin precession must be

---

[1] see Appendix D.1

absent in the electron rest frame, $B = 0 \Leftrightarrow d\mathbf{S}/dt = 0$. This is a consequence of the conservation of angular momentum in an inertial system. However, absence of precession in the electron rest frame implies the presence of precession in the laboratory frame. This is called *Thomas precession* [108, 110]. In hindsight this precession is not surprising for an observer in the laboratory because the force that curves the orbit implies a torque on the spinning electron.

For $\mathbf{B}_L \neq 0$ we have two independent mechanisms resulting in precession about the same axis *but in opposite direction* (for the configuration of Fig. 4.7*b*). The net precession rate is given by

$$(d\mathbf{S}/dt)_{lab} = (d\mathbf{S}/dt)_{rest} - \boldsymbol{\omega}_T \times \mathbf{S}. \tag{4.111}$$

Substituting Eq. (4.109) we can write the precession rate in the form

$$(d\mathbf{S}/dt)_{lab} = -g_e(e/2m_e)\mathbf{S} \times \mathbf{B}_L + \mathbf{S} \times \boldsymbol{\omega}_T = -g_e(e/2m_e)\mathbf{S} \times \mathbf{B}_{LS}, \tag{4.112}$$

where

$$\mathbf{B}_{LS} = \mathbf{B}_L - \frac{\boldsymbol{\omega}_T}{g_e(e/2m_e)} \tag{4.113}$$

is the *effective internal field, the spin-orbit field* as experienced by the electron spin.

In the atom the acceleration is in radial direction,

$$\mathbf{a} = -\frac{1}{m_r}\boldsymbol{\nabla}\mathcal{V}(r) = \frac{e}{m_r}\frac{\mathbf{r}}{r}\frac{d\varphi(r)}{dr}. \tag{4.114}$$

Hence, the Thomas precession is given by

$$\boldsymbol{\omega}_T = -\frac{e}{2m_r^2c^2}\frac{1}{r}\frac{d\varphi(r)}{dr}\mathbf{L}. \tag{4.115}$$

For the spin-orbit field we find using Eqs. (4.113), (4.100), and (4.115)

$$\mathbf{B}_{LS} = -\left(1 - \frac{1}{g_e}\frac{m_e}{m_r}\right)\frac{1}{m_rc^2}\frac{1}{r}\frac{d\varphi(r)}{dr}\mathbf{L}. \tag{4.116}$$

Approximating $g_e m_r/m_e \simeq 2$ we notice that the Thomas precession reduces the spin-orbit field by $\sim 50\%$.

Problem 4.8. Show for the configuration of Fig. 4.7*b* that *in special relativity* the electron rest frame rotates with respect to the nuclear rest frame at the angular frequency

$$\boldsymbol{\omega}_r = -\frac{\mathbf{v} \times \mathbf{a}}{2c^2}.$$

Solution. To describe the motion of an electron orbiting about a nucleus we introduce three inertial frames, S, S' and S'' , in Fig. 4.7*b* labeled by their origins o, o', and o'', respectively. The electron position in the three frames is denoted by $\mathbf{r}, \mathbf{r}', \mathbf{r}''$, where $\mathbf{r} = (x, y, z)$. Note that the position of the origins in their own frame is $(0, 0, 0)$, which corresponds to the zero vector of the individual frames,

$$\mathbf{r}_o = \mathbf{r}'_{o'} = \mathbf{r}''_{o''} = 0.$$

The z axes are taken to be normal to the plane of the orbit ($z = z' = z''$). Let us suppose that at $t = 0$ the electron is at position $\mathbf{r}$ on the negative y axis of the inertial system S (as sketched in Fig. 4.7*b*). We place our self at this position (the laboratory fixed frame for a nucleus at rest) and follow the motion of the electron at rest in S''. We consider the situation in which S' moves with respect to S at velocity v along the x direction and S'' moves with respect to S' at velocity v' along the y direction as illustrated in Fig. 4.7*b*. The inertial frames S and S' are related by the Lorentz transformation

$$\begin{aligned} x' &= \gamma'(x - vt) && \Leftrightarrow x = \gamma'(x' + vt') \\ y' &= y && \Leftrightarrow y = y' \\ t' &= \gamma'(t - vx/c^2) && \Leftrightarrow t = \gamma'(t' + vx'/c^2), \end{aligned}$$

where $\gamma' = (1 - v^2/c^2)^{-\frac{1}{2}}$. Likewise, the inertial frames S' and S'' are related by the Lorentz transformation

$$\begin{aligned} x'' &= x' &&\Leftrightarrow x' = x'' \\ y'' &= \gamma''(y' - v't') &&\Leftrightarrow y' = \gamma''(y'' + v't'') \\ t'' &= \gamma''(t' - v'y'/c^2) &&\Leftrightarrow t' = \gamma''(t'' + v'y''/c^2), \end{aligned}$$

where $\gamma'' = (1 - v'^2/c^2)^{-\frac{1}{2}}$. The angle θ (with respect to the x direction) by which the electron is observed from the origin of the laboratory-fixed frame is given by

$$\tan\theta = \frac{y}{x} = \frac{y_{o''}}{x_{o''}} = \frac{y'_{o''}}{\gamma'[x'_{o''} + vt']} = \frac{\gamma''(y''_{o''} + v't'')}{\gamma'[x''_{o''} + v\gamma''(t'' + v'y''_{o''}/c^2)]} = \frac{\gamma''v't''}{\gamma'v\gamma''t''} = \frac{1}{\gamma'}\frac{v'}{v}.$$

Here we used $y/x = y_{o''}/x_{o''}$ because inertial systems are defined up to a translation (homogeneity of free space - see Appendix D.1). Furthermore, we used the origin property $(x''_{o''}, y''_{o''}, z''_{o''}) = (0,0,0)$. On the other hand, the angle θ'' (respect to the x'' direction) by which the origin o is observed from the position of the electron is given by

$$\tan\theta'' = \frac{-y''}{-x''} = \frac{y''_o}{x''_o} = \frac{\gamma''(y'_o - v't')}{x'_o} = \frac{\gamma''[y_o - v'\gamma'(t - vx_o/c^2)]}{\gamma'(x_o - vt)} = \frac{-v'\gamma'\gamma''t}{-\gamma'vt} = \gamma''\frac{v'}{v}.$$

Here we used $y''/x'' = y''_o/x''_o$ and $(x_o, y_o, z_o) = (0,0,0)$. Note that in the non-relativistic limit $\theta = \theta''$ because $\gamma', \gamma'' \to 1$ for $v, v' \to 0$, just as expected for galilean relativity. For $v'_y \ll v_x \ll c$ we have $\tan\theta \simeq \theta$ and $\tan\theta'' \simeq \theta''$ and obtain in special relativity

$$\Delta\theta = \theta'' - \theta = \left(\gamma'' - \frac{1}{\gamma'}\right)\frac{v'}{v} = \left(1 + \frac{v'^2}{v^2}\right)\frac{v^2}{2c^2}\frac{v'}{v} \simeq \frac{vv'}{2c^2}.$$

For the motion along the orbit we have $\mathbf{a} = d\mathbf{v}/dt$. Since, for $t \to 0$, the acceleration $\mathbf{a}$ points along the y direction (towards the center of curvature) we may approximate $v' = \delta v = a\delta t$ and write

$$\delta\theta = \frac{v\,a}{2c^2}\delta t.$$

Note that the direction of rotation is clockwise in the xy plane. Furthermore, since $\mathbf{v}$ and $\mathbf{a}$ are orthogonal at $t = 0$ and the result must independent of the initial position (homogeneity of space and time) we obtain for the rotation rate

$$\boldsymbol{\omega}_r = \frac{\delta\boldsymbol{\theta}}{\delta t} = -\frac{\mathbf{v} \times \mathbf{a}}{2c^2}. \qquad \square$$

4.4.4 Spin-orbit interaction

Substituting the expression for the spin-orbit field (4.116) into the spin-Zeeman Hamiltonian (4.84) we obtain the Hamiltonian describing the *spin-orbit interaction*

$$\mathcal{H}_{LS} = -\boldsymbol{\mu}_s \cdot \mathbf{B}_{LS} = g_e(e/2m_e)\mathbf{S} \cdot \mathbf{B}_{LS} = \xi(r)\mathbf{L} \cdot \mathbf{S}. \tag{4.117}$$

This Hamiltonian expresses the coupling between $\mathbf{L}$ and $\mathbf{S}$ (*spin-orbit coupling*). The function $\xi(r)$ is the *coupling strength*,

$$\xi(r) = -\frac{(g'_e - 1)}{2m_r^2c^2}\frac{e}{r}\frac{d\varphi(r)}{dr} = \begin{cases} \xi_0 \dfrac{\mathcal{Z}_{nl}(r)}{r^3} & \text{for } r > r_n \\ \xi_0 \dfrac{Z}{r_n^3} & \text{for } r \le r_n, \end{cases} \tag{4.118}$$

where the prefactor can be written in various equivalent forms,

$$\xi_0 \simeq \frac{1}{2m_r^2c^2}\frac{e^2}{4\pi\varepsilon_0} = \frac{\mu_0}{4\pi}2\gamma_L^2 = \frac{1}{2}\alpha^4 m_r c^2 \frac{a^3}{\hbar^2}, \tag{4.119}$$

as follows by approximating $g'_e = g_e m_r/m_e \simeq 2$ and the use of Eqs. (4.54), (2.11) and (2.13). Importantly, we note that $\xi(r) > 0$ if $\mathcal{Z}_{nl}(r) > 0$. Furthermore, note the appearance of the *fine-structure constant* ($\alpha \approx 1/137$) in the energy scale $\alpha^4 m_r c^2$.

4.4.5 Fine structure Hamiltonian for hydrogen-like atoms

We are now in a position to write down the Hamiltonian for the magnetic fine structure of hydrogen-like atoms; i.e., for an electron, of charge $q = -e$ and intrinsic magnetic moment $\boldsymbol{\mu}_s$, moving in the central field of the atomic nucleus and in the presence (or absence) of an externally applied magnetic field,

$$\mathcal{H} = -\frac{\hbar^2}{2m_r}\Delta + \mathcal{V}(r) + \mathcal{H}_r - \frac{i\hbar e}{m_r}(\mathbf{A}\cdot\boldsymbol{\nabla}) + \frac{e^2}{2m_r}\mathbf{A}^2 - \boldsymbol{\mu}_s\cdot(\mathbf{B}+\mathbf{B}_{LS}). \tag{4.120}$$

This expression holds in the Coulomb gauge, $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$. From left to right we recognize first the Schrödinger Hamiltonian $\mathcal{H}_0$ and its relativistic correction $\mathcal{H}_r$, the $\mathbf{A}\cdot\boldsymbol{\nabla}$ term describes the motion of the electron in the vector potential of an externally applied magnetic field $\mathbf{B}$ (for a homogeneous field $\mathbf{A} = \frac{1}{2}\mathbf{B}\times\mathbf{r}$). The last term is the spin Zeeman coupling, involving both the external field $\mathbf{B}$ and the spin-orbit field $\mathbf{B}_{LS}$. We start by considering the atom in the absence of an external field ($\mathbf{B} = 0$); i.e., for a vanishing vector potential ($\mathbf{A} = 0$). In Section 4.7 we discuss fine structure in the presence of an external field ($\mathbf{B} \neq 0$).

4.5 Fine structure in zero field

4.5.1 Effective Hamiltonian and spin-orbit splitting

In the absence of external fields the fine-structure Hamiltonian (4.120) reduces to the form

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_r + \xi(r)\mathbf{L}\cdot\mathbf{S}, \tag{4.121}$$

where $\mathcal{H}_0$ is the Schrödinger Hamiltonian and $\mathcal{H}_r$ stands for the relativistic correction. The spin-orbit interaction causes the n^2-fold degeneracy of the principal atomic structure to be lifted in zero field. This is called *spin-orbit splitting* and is illustrated for ^2P terms in Fig. 4.8. Since $\xi(r)$ does not commute with $\mathcal{H}_0$ and $\mathcal{H}_r$, the spin-orbit interaction couples states of different principal quantum number. Hence, to calculate the splitting, we have to turn to perturbation theory (of a degenerate level). Fortunately, both $\xi(r)$ and $\mathbf{L}\cdot\mathbf{S}$ commute with $\mathbf{L}^2$, $\mathbf{S}^2$, $\mathbf{J}^2$ and J_z (see Problems 4.3, 4.4, 4.5 and 4.6). This makes the spin-orbit interaction *diagonal* in the coupled angular momentum basis, $\{|lsjm_j\rangle\}$; i.e., l, s, j and m_j are good quantum numbers. Since l remains a good quantum number, the *parity* of the angular momentum state is not affected by the spin-orbit interaction. For a manifolds of given n, the spin-orbit shift can be calculated using the well-known expression (H.56) from ordinary first-order perturbation theory,[1]

$$\Delta E_{n,j}^{LS} = \langle nlsjm_j|\mathcal{H}_{LS}|nlsjm_j\rangle = (\zeta_{nl}/\hbar^2)\langle lsjm_j|\mathbf{L}\cdot\mathbf{S}|lsjm_j\rangle, \tag{4.122}$$

where ζ_{nl} is a radial integral known as the *spin-orbit coupling constant*,

$$\zeta_{nl} = \langle\xi(r)\rangle_{nl}\hbar^2 = \begin{cases} \frac{1}{2}\alpha^4 m_r c^2\langle\mathcal{Z}(r)/\rho^3\rangle_{nl} & \text{for } l > 0 \\ \text{finite} & \text{for } l = 0. \end{cases} \tag{4.123}$$

The divergence expected for the $l = 0$ radial integral is avoided (ζ_{nl} = finite) provided we account for the nonzero nuclear size. The coupling constant has the properties of a so-called reduced matrix element: it depends on the quantum numbers n and l but not on m_l (see Problem 2.2). Its numerical value is *positive* ($\zeta_{nl} > 0$) in hydrogen-like atoms. In *many-electron* atoms the coupling constant can also become negative (see Section 9.5.5). The radial integral $\langle\mathcal{Z}(r)/r^3\rangle_{nl}$ can be determined numerically but generally not without a substantial effort because the determination of $\mathcal{Z}_{nl}(r)$

---

[1] The coupling to different n levels enters only to second or higher order in perturbation theory. In first order only the states within the manifold *of given* n contribute (n is a good quantum number to first order in perturbation theory - see Appendix H.3).

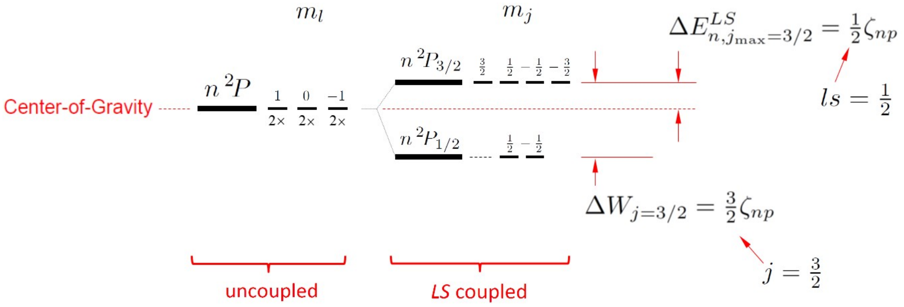


Figure 4.8: Fine-structure splitting for a $n\,^2P$ Term ($l = 1$, $s = \frac{1}{2}$). Shown are the unperturbed level as well as the spin-orbit splitting into the $n\,^2P_{3/2}$ (*stretched*) and $^2P_{1/2}$ manifolds. The splitting is drawn for the *regular* (hydrogen-like) case ($\zeta_{np} > 0$). Note that the *center-of-gravity* of the levels is conserved in the splitting. The magnitude of the splitting, ΔW_j, satisfies with the *Landé interval rule.*

requires knowledge of the effective charge distribution and this in turn depends on the screening. On the other hand, using the spin-orbit shift rules (see Section 4.5.2 below), it is straightforward to determine the integral experimentally from a measurement of the spin-orbit splitting in zero magnetic field.

In view of the discussion above it is convenient to write the atomic Hamiltonian in the following approximate form,

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_r + (\zeta_{nl}/\hbar^2)\mathbf{L}\cdot\mathbf{S}, \tag{4.124}$$

where the numerical constant ζ_{nl} is defined by Eq. (4.123). Eq. (4.124) is an *effective fine-structure Hamiltonian* that holds for given values of the quantum numbers n and l. It has the advantage over Eq. (4.121) that it offers the possibility to determine ζ_{nl} experimentally as a phenomenological constant. Importantly, for the last term of the Hamiltonian (4.124) not only l but also n is a good quantum number. This is an approximation: calculations based on Eq. (4.124) are equivalent with calculations based on the actual Hamiltonian (4.121) as long as the principal structure of the atom is accounted for only to first order in perturbation theory. This is a good approximation as long the fine-structure splitting of a level of given n is small compared to the principal-level splitting, which is the case for $\zeta_{nl} \ll E_{n+1} - E_n$. As this condition is satisfied in many practical cases, we shall use the effective Hamiltonian (4.124) to describe the fine structure of hydrogen-like atoms.

4.5.2 Shift rules for spin-orbit coupling in zero field

Using the operator identity

$$\mathbf{L}\cdot\mathbf{S} = \tfrac{1}{2}(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) \tag{4.125}$$

the *spin-orbit shift* of the *LS* manifold with quantum number j evaluates to

$$\Delta E^{LS}_{n,j} = \tfrac{1}{2}\zeta_{nl}\left[j(j+1) - l(l+1) - s(s+1)\right]. \tag{4.126}$$

Note that for *positive* coupling constant ($\zeta_{nl} > 0$) the energy of the *LS* manifolds increases with j. This is called *regular* (i.e., hydrogen-like) ordering of the j levels; for $\zeta_{nl} < 0$ the order of the j levels is called *inverted.* The manifold of highest multiplicity ($j = j_{\max} = l + s$) corresponds to the so called *stretched state* of the *LS* coupling scheme. From Eq. (4.126) we derive *three shift rules* for the spin-orbit manifolds of given j, l and s (see Fig. 4.8):

- the shift of the *stretched state* ($j = j_{\max} = l + s$) is given by

$$\Delta E^{LS}_{n,j_{\max}} = \zeta_{nl}ls. \tag{4.127}$$

- the relative shift of two adjacent j manifolds follows the *Landé interval rule*:

$$\Delta W_j = \Delta E^{LS}_{n,j} - \Delta E^{LS}_{n,j-1} = \zeta_{nl} j_>, \tag{4.128}$$

where $j_>$ is the quantum number of the manifold with the highest multiplicity of *adjacent pairs*, $j_> = max(j, j-1) = j$.

- the weighted mean of the energies of *all j manifolds* of a given LS coupling scheme coincides with the energy of the $(2l+1)\,(2s+1)$-fold degenerate unperturbed level (see Problem 4.9),

$$\frac{1}{(2l+1)\,(2s+1)} \sum_{j=|l-s|}^{l+s} (2j+1)\,\Delta E^{LS}_{n,j} = 0. \tag{4.129}$$

This is called the *center-of-gravity rule.*

The fine structure of all *hydrogen-like* atoms is determined by the angular momentum properties of a single electron. Hence, the electronic states are *electron spin doublets* ($s = \frac{1}{2}$) of varying orbital angular momentum. To facilitate verbal discussion the manifolds are usually referred to in the *Term* notation; e.g., $1\,^2S$, $2\,^2P$, $3\,^2D$, etc..[1] In doublet terms the coupling of the spin to the orbital angular momentum can result in only two values of total angular momentum: $j = l + \frac{1}{2}$ (parallel coupling; e.g., $n\,^2P_{3/2}$) and $j' = l - \frac{1}{2}$ (anti-parallel coupling; e.g., $n\,^2P_{1/2}$). For these two cases we have

$$\Delta E^{LS}_{n,j} = \begin{cases} \left.\begin{array}{ll} +\frac{1}{2}\zeta_{nl} l & \left(j = l + \frac{1}{2}\right) \\ -\frac{1}{2}\zeta_{nl}(l+1) & \left(j = l - \frac{1}{2}\right) \end{array}\right\} & \text{for } l > 0 \\ 0 & \text{for } l = 0. \end{cases} \tag{4.130}$$

Note in this example that for hydrogen-like atoms ($\zeta_{nl} > 0$) the state with the highest multiplicity has indeed the highest energy (as mentioned above). Note further that the interval rule holds,

$$\Delta W_{l+1/2} = \Delta E^{LS}_{n,l+1/2} - \Delta E^{LS}_{n,l-1/2} = \zeta_{nl}(l + \tfrac{1}{2}). \tag{4.131}$$

The sum of the level shifts is given by

$$\Delta E^{LS}_{n,l+1/2} + \Delta E^{LS}_{n,l-1/2} = -\zeta_{nl}/2 \tag{4.132}$$

and the center of gravity of the levels is indeed zero,

$$(2l+2)\,\zeta_{nl} l/2 - 2l\zeta_{nl}\,(l+1)\,/2 = 0. \tag{4.133}$$

Problem 4.9. Show that the center of gravity of the fine-structure manifold coincides with the energy of the unperturbed level.

Solution. Note that the trace is independent of the choice of representation,

$$\operatorname{tr}(\mathcal{H}'_{LS}) = \operatorname{tr}\left(U^{-1}\mathcal{H}_{LS}U\right) = \operatorname{tr}\left(UU^{-1}\mathcal{H}_{LS}\right) = \operatorname{tr}\left(\mathcal{H}_{LS}\right),$$

where U is a unitary transformation between the two bases and we used (in the second step) the matrix property (M.26) that the trace of a product of operators is invariant under cyclic permutation of the operators. Rather than working in the coupled basis $\{|nlsjm_j\rangle\}$ we choose the uncoupled basis $\{|nlm_lsm_s\rangle\}$ to calculate the trace (4.129),

$$\operatorname{tr}\left(\mathcal{H}_{LS}\right) = \sum_{m_s=-s}^{s} \sum_{m_l=-l}^{l} (\zeta_{nl}/\hbar^2)\langle lm_lsm_s|\mathbf{L}\cdot\mathbf{S}|lm_lsm_s\rangle.$$

---

[1] In the most general context a *Term* is defined as a manifold of states spanning the angular momentum subspace of given L and S; for hydrogen-like atoms $L = l$ and $S = s = \frac{1}{2}$.

Using the inner product rule (3.109b) and noting that only the L_zS_z terms are diagonal we find

$$\operatorname{tr}\,(\mathcal{H}_{LS}) = \sum_{m_s=-s}^{s}\sum_{m_l=-l}^{l}\zeta_{nl}m_l m_s = 0.$$

For the center of gravity we have to normalize on the $(2l+1)\,(2s+1)$-fold degeneracy of the unperturbed level but obviously this does not change the result if the center of gravity is zero. □

4.5.3 Fine structure of hydrogenic atoms

For the special case of *hydrogenic* atoms the orbital integrals can be solved analytically. Setting $\mathcal{Z}_{nl}(r) \equiv Z$ in Eq. (4.118) we obtain for the coupling strength

$$\xi(r) = \tfrac{1}{2}\alpha^4 m_r c^2 \frac{a^3}{r^3}\frac{Z}{\hbar^2}. \tag{4.134}$$

Hence, the coupling constant follows from Eq. (4.123)

$$\zeta_{nl} = \begin{cases} \tfrac{1}{2}\alpha^4 m_r c^2 Z\,\langle\rho^{-3}\rangle_{nl} & \text{for } l>0 \\ \text{finite} & \text{for } l=0. \end{cases} \tag{4.135}$$

Evaluating the radial integral, see Eqs. (2.49), this becomes

$$\zeta_{nl} = \begin{cases} \tfrac{1}{2}\alpha^4 m_r c^2 \dfrac{Z^4}{n^3}\dfrac{1}{l(l+{}^1\!/_2)(l+1)} = -\dfrac{\alpha^2 Z^2}{n}E_n\dfrac{1}{l(l+{}^1\!/_2)(l+1)} & \text{for } l>0 \\ \text{finite} & \text{for } l=0, \end{cases} \tag{4.136}$$

where E_n is the zero-order energy given by Eq. (2.33); for $l = 0$ we took into account the finite nuclear size. Thus, using Eq. (4.131) we find for the *fine structure splitting*

$$\Delta W_{l+1/2} = -\frac{\alpha^2 Z^2}{n}E_n\frac{1}{l(l+1)} = \frac{\alpha^2 Z^4}{2n^3}\frac{1}{l(l+1)}\ \text{Hartree} \quad \text{for } l>0. \tag{4.137}$$

The numerical value of this splitting is 10.949 GHz (experimental value: 10.969 GHz). Using Eqs. (4.130) we find for the level shifts

$$\Delta E^{LS}_{n,j} = \begin{cases} \left.\begin{aligned} &-\frac{\alpha^2 Z^2}{n^2}E_n\frac{n}{(l+{}^1\!/_2)(l+1)}\frac{1}{2} && (j=l+\tfrac{1}{2}) \\ &+\frac{\alpha^2 Z^2}{n^2}E_n\frac{n}{l(l+{}^1\!/_2)}\frac{1}{2} && (j=l-\tfrac{1}{2}) \end{aligned}\right\} & \text{for } l>0 \\ 0 & \text{for } l=0. \end{cases} \tag{4.138}$$

Remarkably, adding Eqs. (4.24) and (4.138) brings us a *single* expression for the *fine-structure shift* which depends on only two quantum numbers (n, j) rather that three (n, l, s),

$$\Delta E_{n,j} = \Delta E^{\text{rel}}_n + \Delta E^{LS}_n = -E_n\frac{\alpha^2 Z^2}{n^2}\left(\frac{3}{4} - \frac{n}{j+{}^1\!/_2}\right). \tag{4.139}$$

For $j = \frac{1}{2}$ this expression corresponds to the $l = 0$ relativistic shift of Eq. (4.24) whereas for $j = l \pm \frac{1}{2}$ it gives the fine-structure splitting for $l > 0$. Interestingly, the value $j = \frac{1}{2}$ is obtained for $l = 0$ but also for $l = 1$ (using $j = l - \frac{1}{2}$). This results in an "accidental" degeneracy of the ${}^2P_{1/2}$ and ${}^2S_{1/2}$ levels in hydrogenic atoms but is absent in the hydrogen-like atoms at large. Furthermore, note that the energy splittings of the *fine structure* are typically a factor α^2 smaller than the energy splitting of the *principal structure* ($\alpha \approx 1/137$ is the *fine structure constant*).

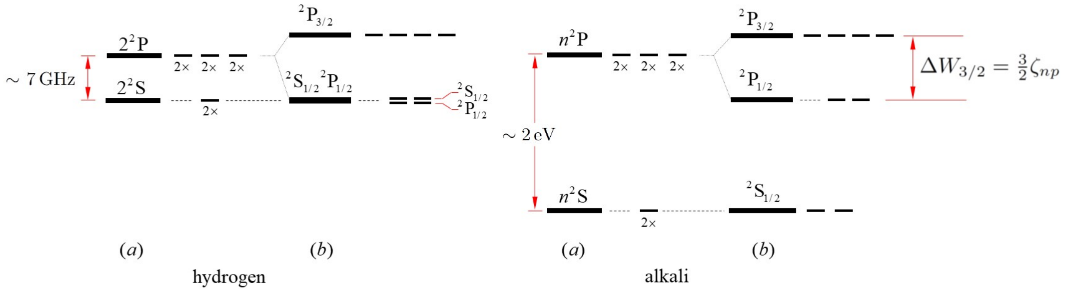


Figure 4.9: Term diagram showing the difference in fine structure between hydrogen and alkali-like atoms: (a) without spin-orbit splitting, (b) with spin-orbit splitting. The hydrogen fine structure is characterized by two levels with a splitting of only $\sim 7\,\mathrm{GHz}$, whereas in the alkali case one observes three levels and the splitting is much larger ($\sim 2\,\mathrm{eV}$). These differences originate in the absence or presence of screening of the nuclear charge by core electrons (which is less effective for s electrons than for p electrons).

4.6 Fine structure of alkali-like atoms in zero field

4.6.1 Introduction

Historically, the accidental degeneracy of the $2^2P_{1/2}$ and $2^2S_{1/2}$ levels in hydrogen played an important role in the development of quantum mechanics. In atomic spectroscopy this degeneracy manifests itself as a fine-structure splitting into *two* levels ($^2P_{3/2}$ and $^2P_{1/2}/^2S_{1/2}$, not considering the small Lamb shift). In alkali atoms this degeneracy is absent and *three* levels are observed ($^2P_{3/2}$, $^2P_{1/2}$ and $^2S_{1/2}$). Furthermore, as illustrated in Fig. 4.9, the $2\,^2S$-$2\,^2P$ splitting in hydrogen is only $\sim 7\,\mathrm{GHz}$ (see Section 4.1.2) whereas in the alkalies it is $\sim 2\,\mathrm{eV}$, more than four orders of magnitude larger. In hindsight, these differences had to be understood before the electron spin could be postulated [112]. It slowed down progress and became a great scientific puzzle: why is it that two quantum numbers suffice to describe hydrogen (n, j) but a third one is required for the alkalies (n and l and s)?

To understand the alkali atoms we have to understand the physics of *screening by core electrons*. Core electrons are present in the alkalies but absent in the hydrogenic case. Our approach will be mostly phenomenological. Starting from the experimental values in the NIST Atomic Level Database [66] we develop a qualitative picture of the atomic structure in which *electron shells* provide the underpinning of the *periodic system of the elements*. This means that we shall temporarily jump ahead, using features from the physics of many-electron atoms for which the theory will be presented in Chapters 6 through 9. After discovering the enormous impact of the core on the principal structure of the atom (Section 4.6.2) we turn in section 4.6.3 to the consequences for spin-orbit coupling. Throughout the discussion of screening Hartree atomic units will be used.

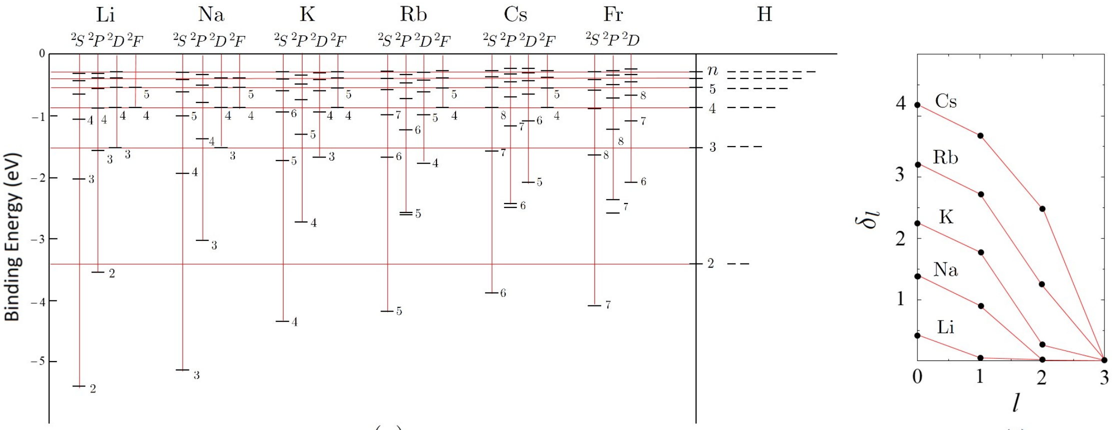
4.6.2 Screening by core electrons - effect on principal structure

Alkali and alkali-like atoms differ from hydrogenic atoms by the presence of *core electrons*. As will be shown in Section 9.2 the electron core is spherical, so it does not affect the central symmetry of the atom; i.e., the rotational structure is not affected. In contrast, the radial structure changes dramatically as the core electrons seriously modify the charge distribution within the atom. It may speak for itself that this affects the radial distribution function of the orbitals with immediate consequences for the binding energies of the electrons.

To introduce the phenomenology of screening we show in Fig. 4.10a and Fig. 4.11a the energy levels of the valence electron in alkali atoms next to those of hydrogen.[1] The levels are labeled by

---

[1] Beware that we focus on the similarities between the behavior of the *valence* electron in the alkalies and the



(*a*) (*b*)

Figure 4.10: (*a*) Energy levels of the alkali atoms for different states of the *valence electron* next to those of hydrogen. The data are taken from the NIST Atomic Spectra Database Levels Data [66]. Note that the binding energies in the alkali atoms remain close to those of the hydrogenic electron in spite of enormous differences in nuclear charge. In particular, for $l \to \infty$ and/or $n \to \infty$ the levels converge to those of H. This similarity is the result of screening of the nuclear charge by the core electrons. In the data for Rb, Cs and Fr the fine-structure splitting of the $5\,^2P$, $6\,^2P$ and $7\,^2P$ terms becomes resolved on the eV energy scale of the diagram. (*b*) Quantum defect plot of the same data [28]. Note that for $l \geq 3$ the alkali atoms behave hydrogenic; i.e., the screening by the core electrons is perfect. Furthermore, for a given atomic species the quantum defect is independent of the principal quantum number, $\delta_{nl} \simeq \delta_l$.

the principal quantum number (n) for given value of the angular momentum (s, , p, d, $\cdots$). Note that the alkali levels cover roughly the same range of energies as those of hydrogen in spite of the enormous difference in nuclear charge: $Z = 1$ in *hydrogen* (H) and $Z = 87$ in *francium* (Fr). This points to efficient *screening* of the nucleus by the core electrons. The key idea is illustrated for *potassium* (K) in Fig. 4.11*b*. The electron core acts as a cloud of negative charge in-between the nucleus and the valence electron. In classical electrodynamics such a charge distribution is called *space charge.* In quantum mechanics the space charge arises as a *mean field* of negative charge distributed according to the probability density of the electronic orbitals (see Fig. 4.12). This mean field lowers the electrostatic potential around the nucleus, thus giving rise a screened Coulomb potential with (in atomic units) an *effective nuclear charge* $Z_{nl}(r)$ as introduced in Section 4.4.2.

4.6.2.1 Screening constant and efficiency versus quantum defect

In the simplest description of screening, for each orbital nl the amount of screening is specified by a single number, the *screening constant* σ_{nl}. For an atom of nuclear charge Z, the screening constant is related to the *effective nuclear charge,*

$$Z_{nl} \equiv Z - \sigma_{nl}. \tag{4.140}$$

As a *rule of thumb,* σ_{nl} equals the number of core electrons involved in the screening. So, we expect: $Z_{1s} \simeq Z$ (no core) $Z_{2s} \simeq Z - 2$ ($1s^2$ core), $Z_{2p} \simeq Z - 4$ ($1s^2 2s^2$ core), $Z_{3s} \simeq Z - 10$ ($1s^2 2s^2 2p^6$ core), $Z_{3p} \simeq Z - 12$ ($1s^2 2s^2 2p^6 3s^2$ core), etc.. In terms of Z_{nl} the energies of the principal structure are given by

$$\varepsilon_{nl} = -\frac{Z_{nl}^2}{2n^2}. \tag{4.141}$$

---

hydrogenic electron of the Bohr atom. Core electron excitations are *not included* in this diagram.

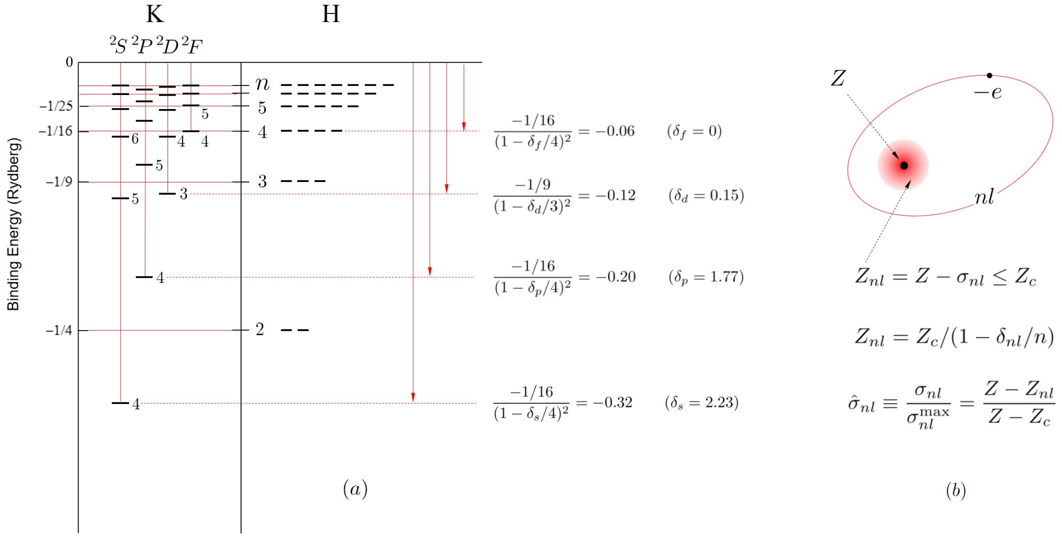


Figure 4.11: (a) Energy levels of *potassium* (K) compared to those of hydrogen. The *deviations from the hydrogenic values* are quantified by the quantum defects δ_s, δ_p, δ_d and δ_f; (b) Quantum defects result from incomplete screening of the nucleus by core electrons. This manifests itself as an *effective nuclear charge*, $Z_{nl} = Z - \sigma_{nl}$ (the *nuclear charge* Z lowered by the *screening constant* σ_{nl}). In the Rydberg limit the screening is maximal ($Z_{nl} = Z_c$) and the *screening efficiency* $\hat{\sigma}_{nl} \equiv \sigma_{nl}/\sigma_{nl}^{\max}$ becomes unity. The relation between δ_{nl} and Z_{nl} (and σ_{nl}) follows by equating Eqs. (4.143) and (4.141) for the same energy level.

Note that binding energy of the valence electron scales quadratically with Z_{nl}/n.

Another way of characterizing screening is by specification of the *quantum defect* δ_{nl} (also known as *Rydberg correction*). This quantity turns the principal quantum number, the integer n, into an *effective principal quantum number*, the positive real number

$$n_l^* = n - \delta_{nl}. \tag{4.142}$$

The quantum defect is a measure for the deviation from the hydrogenic distribution of levels. A list of quantum defects for *valence electron levels* in alkali atoms is given in Table 4.1. Note that, for a given alkali species, δ_{nl} shows little dependence on the principal quantum number, $\delta_{nl} \simeq \delta_l > 0$. On the other hand δ_l decreases rapidly with l as is illustrated in Fig. 4.10*b*. Hence, by measuring a few quantum defects (δ_s, δ_p, δ_d, $\cdots$) the whole level diagram can be reconstructed using the Bohr formula for the binding energies but with n_l^* replacing n,

$$\varepsilon_{nl} = -\frac{Z_c^2}{2n_l^{*2}} = -\frac{Z_c^2}{2n^2}\frac{1}{(1-\delta_l/n)^2}. \tag{4.143}$$

Here Z_c stands for the *Rydberg charge* introduced in Section 4.4.2; for neutral atoms $Z_c = 1$, for singly charged ions $Z_c = 2$, *etc.*.

Interestingly, for $n \gg 1$ the ratio of the level shift (with respect to the hydrogenic value) over the level separation value goes to a constant. The limiting value is called the *Rydberg limit*. This limit can be determined by expanding Eq. (4.143) to leading order in powers of δ_{nl}/n. For the *level shift* with respect to the hydrogenic value we find

$$\delta\varepsilon_{nl} = \frac{Z_c^2}{n^3}\delta_{nl} + \cdots \qquad (n \gg 1). \tag{4.144}$$

Table 4.1: Fine-structure-avaraged quantum defects for selected atomic energy levels. The table is based on the NIST Atomic Spectra Database Levels Data [66]. Metastable helium (He*) will be discussed in Section 6.4.4.

El.	n_s	n_s^*	δ_s	Z_{ns}	$\hat{\sigma}_s$	n_p	$\bar{n}_p^*$	$\bar{\delta}_p$	Z_{np}	$\hat{\sigma}_p$	n_d	$\bar{n}_d^*$	$\bar{\delta}_d$	Z_{nd}	$\hat{\sigma}_d$
H	1	1.000	0.000	1	1	–	–	–	–	–	–	–	–	–	–
He*	2	1.689	0.311	1.18	0.82	2	1.938	0.062	1.03	0.97	–	–	–	–	–
Li	2	1.589	0.411	1.26	0.87	2	1.959	0.041	1.02	0.99	–	–	–	–	–
	3	2.596	0.404	1.16	0.92	3	2.956	0.044	1.02	0.99	3	2.999	0.001	1.000	1
Na	3	1.627	1.373	1.84	0.92	3	2.117	0.883	1.42	0.96	3	2.990	0.010	1.003	1
	4	2.643	1.357	1.51	0.95	4	3.133	0.867	1.28	0.97	4	3.988	0.012	1.003	1
K	4	1.770	2.230	2.26	0.93	4	2.234	1.766	1.79	0.96	3	2.854	0.146	1.051	1
	5	2.801	2.199	1.79	0.96	5	3.265	1.735	1.53	0.97	4	3.797	0.203	1.054	1
Rb	5	1.805	3.195	2.77	0.95	5	2.288	2.712	2.19	0.97	4	2.767	1.233	1.446	0.99
	6	2.845	3.155	2.11	0.97	6	3.325	2.675	1.80	0.98	5	3.706	1.294	1.349	0.99
Cs	6	1.869	4.131	3.21	0.96	6	2.351	3.649	2.55	0.97	5	2.552	2.448	1.959	0.98
	7	2.920	4.080	2.40	0.97	7	3.395	3.605	2.06	0.98	6	3.534	2.466	1.698	0.99

Since $\delta_{nl} \simeq \delta_l$ for $n \gg 1$, the *level separation*, $\Delta\varepsilon_{nl} \equiv \varepsilon_{(n+1)l} - \varepsilon_{nl}$, is given by

$$\Delta\varepsilon_{nl} = \frac{Z_c^2}{2(n-\delta_l)^2} - \frac{Z_c^2}{2(n+1-\delta_l)^2} = \frac{Z_c^2}{n^3} + \cdots \quad (n \gg 1). \tag{4.145}$$

In the *Rydberg limit* the ratio $\delta\varepsilon_{nl}/\Delta\varepsilon_{nl}$ converges to the quantum defect,

$$\lim_{n\to\infty} \frac{\delta\varepsilon_{nl}}{\Delta\varepsilon_{nl}} = \delta_l. \tag{4.146}$$

The screening constant is a good absolute measure for screening but often a relative measure is desired. As maximal screening is obtained for valence electrons in the Rydberg limit,

$$\sigma_{nl} = Z - Z_{nl} \leq Z - Z_c \equiv \sigma_{nl}^{\max},$$

a good relative measure for the screening is the *screening efficiency*, defined as

$$\hat{\sigma}_{nl} \equiv \sigma_{nl}/\sigma_{nl}^{\max} = (Z - Z_{nl})/(Z - Z_c). \tag{4.147}$$

This quantity varies from 0 in the absence of screening to 1 for perfect screening. The screening efficiencies for the valence electron of the alkali atoms are included in Table 4.1. For the ground state of *lithium* (Li), with only two screening electrons, it is 87%. For all other alkali states the screening is over 90% and increases with increasing values of n and l. The latter does not come as a surprise because for large n and l the valence electron resides at larger distances from the nucleus and the charge overlap with the electron core will be small.

The relation between the effective nuclear charge, Z_{nl}, and the quantum defect, δ_{nl}, follows by equating Eqs. (4.141) and (4.143),

$$Z_{nl} = Z_c/(1 - \delta_{nl}/n). \tag{4.148}$$

Since ε_{nl} scales quadratically with Z_{nl}/n the levels become hydrogenic (with respect to Z_c) for $n \gg 1$. Furthermore, since δ_l decreases for increasing l the binding energy also becomes hydrogenic for $l \gg 1$. This means that, for large l and/or n, the screening of the nucleus by the core electrons becomes close to perfect and the energy levels can be calculated by replacing Z with Z_c in the Bohr formula. A simple screened Coulomb potential may be constructed showing the absence of a dependence on n (for given Z) as well as the decline of screening with increasing l [9]. At the advanced level, the quantum defect is a key element in the theory of *Rydberg atoms* [47].

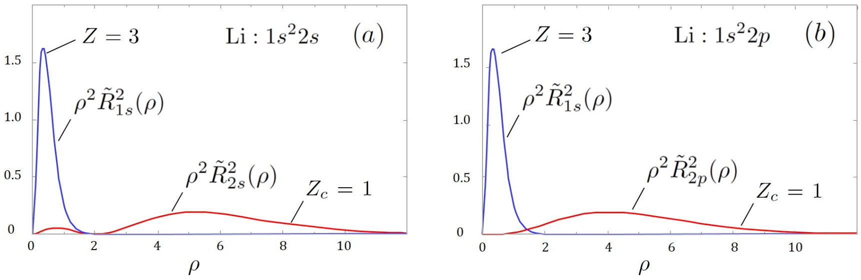


Figure 4.12: Radial distribution functions of hydrogenic wavefunctions give an impression of the electronic charge distribution. In the *lithium* atom (Li) the $1s$ orbital is a core orbital which is drawn for the unscreened nuclear charge ($Z=3$). We compare two valence electrons, both drawn for $Z=Z_c=1$: (a) the $2s$ electron penetrates deeply into the core which points to poor screening from the nucleus; (b) the $2p$ electron hardly penetrates into the core. This points to close to perfect screening. Note that screening corrections giving rise to deviations from the hydrogenic shape are not included in this plot.

In terms of the quantum defect the screening efficiency is given by

$$\hat{\sigma}_{nl} = 1 - \frac{Z_c}{Z-Z_c}\frac{\delta_{nl}/n}{1-\delta_{nl}/n}. \tag{4.149}$$

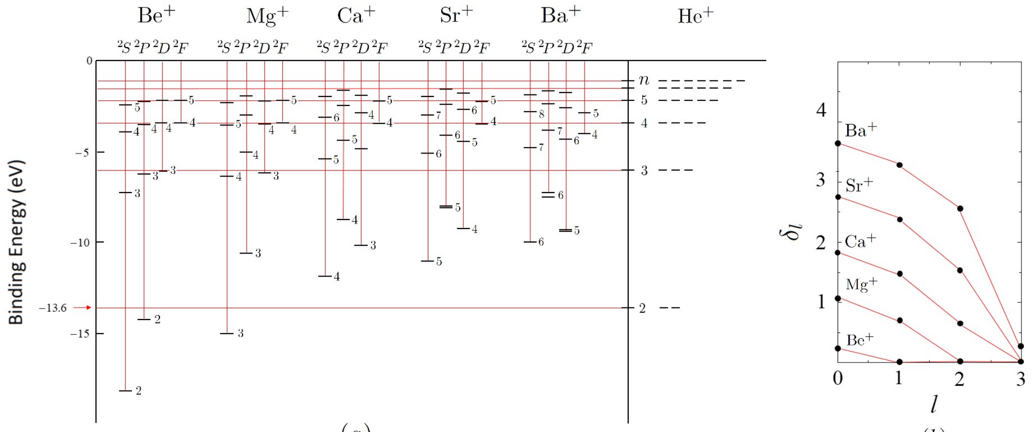
This expression shows that, for given Z, Z_c and n, a smaller quantum defect corresponds to better screening efficiency, with the screening becoming perfect for $l \gg 1$ and/or $n \gg 1$. When comparing configurations differing in Z and Z_c and/or n, the correspondence is not so intuitive. For the neutral *lithium* (Li) atom and the *calcium ion* (Ca$^+$) in the same electron configuration, the same screening efficiency is obtained for a smaller quantum defect in the ion (see Problem 4.10).

4.6.2.2 The phenomenology of screening

Equipped with the tools for characterizing the screening we are prepared to analyze the level diagrams of Fig. 4.10. We start with the example of *lithium* (Li). In this case we have $Z=3$ and the valence electron is screened from the nucleus by the $1s^2$ core ($\sigma_{nl} \simeq 2$, for $n \geq 2$). Hence, for perfect screening we expect for the effective nuclear charge $Z_{nl} = Z - \sigma_{nl} \simeq Z_c = 1$. The experimental values are found in Table 4.1. Apparently, the screening is close to perfect for the $2p$, $3p$ and $3d$ orbitals ($Z_{nl} \simeq 1$) but incomplete for the $2s$ and $3s$ orbitals. This is consistent with the mean field picture because s orbitals extend all the way to the nucleus (where screening is negligible), whereas orbitals with higher angular momenta (p, d, $f, \cdots$) fall off towards the characteristic central node (see Section 2.3). As illustrated in Fig. 4.12, the radial distribution of the $2p$ electron surrounds the $1s$ core like a halo whereas the $2s$ electron electron has a sizable overlap with the core. This points to incomplete screening for the $2s$ electron which explains the stronger binding in Fig. 4.10. As visible in Fig. 4.9 the corresponding difference in binding energy is typically 2 eV for the alkalies.

This *preferential binding of electrons in s orbitals* (or better: low-l orbitals) becomes manifest when studying the level ordering for the entire group of alkalies in Fig. 4.10. In hydrogenic atoms the binding energy decreases with increasing principal quantum number but this ordering is lost in the heavier alkalies. In the case of Na the binding of the $4s$ orbital is stronger than that of the $3d$ orbital ($\varepsilon_{4s} < \varepsilon_{3d}$), for K this becomes $\varepsilon_{4s}, \varepsilon_{4p} < \varepsilon_{3d}$ and even $\varepsilon_{5s} < \varepsilon_{3d}$. These *deviations from the hydrogenic ordering* have important consequences for the ground state configurations of the alkali atoms and, more in general, for the structure of the periodic system (cf. Chapter 9).

What would happen to the screening if we could increase the nuclear charge? It may speak for itself that the atom would contract but is the ordering of the energy levels affected? Questions like this can be answered by studying the level diagram of the singly ionized alkaline-earth ions



(a) (b)

Figure 4.13: (a) Energy levels of the alkaline-earth ions for different states of the *valence electron* next to those of hydrogenic helium. The data are taken from the NIST Atomic Spectra Database Levels Data [66]. Note that the binding energies remain close to that of the hydrogenic electron in He$^+$ as the result of screening by the core electrons. In particular, for $l \to \infty$ and/or $n \to \infty$ the levels converge to those of He$^+$; (b) Quantum defect plot of the same data. For $l \geq 3$ the ions behave hydrogenic; i.e., the screening by the core electrons is maximal. Note the close similarity with Fig. 4.10; i.e., the alkaline-earth ions behave alkali like. A qualitative difference with the alkali atoms is the appearance of the *metastable* ^2D levels in Ca$^+$, Sr$^+$ and Ba$^+$ which can be used to create optical qubits for quantum information processing [99].

shown in Fig. 4.13. These ions are *isoelectronic* (i.e., have the same number of electrons) with the adjacent alkali atom in the periodic system but the nuclear charge has increased from Z in the atom to $Z^+ = Z + 1$ in the ion. In the simplest model for screening also the effective nuclear charge increases by one: $Z^+_{nl} \simeq Z_{nl} + 1$. The corresponding *relative* increase, $Z^+_{nl}/Z_{nl} = (Z_{nl} + 1)/Z_{nl}$, is largest for the outer orbitals because Z_{nl} *falls off* towards Z_c in the Rydberg limit,

$$Z^+_c/Z_c > Z^+_{nl}/Z_{nl} > (Z + 1)/Z. \tag{4.150}$$

The ratio $Z^+_{nl}/Z_{nl} > 1$ implies a contraction of the electronic orbitals of the ion relative to that of the atom (in the same electron configuration). Recalling Eq. (2.49b) the inequalities 4.150 suggest that the contraction increases *progressively* (i.e., more than linearly) with growing distance from the nucleus, the largest contraction being expected for the valence orbitals. The emerging physical picture is one in which the valence electron is pulled towards/into the electron core (if present).

To further explore the phenomenology of screening we have a closer look at the singly ionized alkaline-earth ions. For these ions we have $Z^+_c = 2$, like in hydrogenic helium, He$^+$. This means that the Rydberg charge has doubled with respect to the neutral atom; i.e., to first approximation the binding energy of the valence electron has to increase by a factor 4. This is confirmed by Fig. 4.13. For comparison the levels of the He$^+$ ion are included in the figure. Apart from the larger binding energies, the level diagram is quite similar to that of the alkali atoms. For the l dependence this becomes evident by comparing Figs. 4.10*b* and 4.13*b*. The hydrogenic-helium limiting behavior is implicit in Eq. (4.143) because it depends on Z^+_c (not on Z) : $\delta_{nl} \to 0$ for $l \gg 1$ and $\delta_{nl}/n \to 0$ for $n \gg 1$.

A numerical comparison between atoms and ions (see Table 4.2) shows that in all cases the screening efficiency, as defined by Eq. (4.149), is smaller in the ions. This is in accordance with the physical picture of progressive contraction in which the valence electron is pulled into the core. Closer inspection of Table 4.2 reveals some interesting differences between atoms and ions. For the lowest d levels the quantum defect increases (i.e., the ion is less hydrogenic than the atom) but for the lowest s and p levels the opposite occurs. So, aside from the tendency of pulling the valence

Table 4.2: Fine-structure-averaged quantum defects of the *valence* electrons in isoelectronic pairs of alkali-like atoms. The table is based on the NIST Atomic Spectra Database Levels Data [66].

El.	n_s	n_s^*	δ_s	Z_{ns}/n_s	$\hat{\sigma}_s$	n_p	$\bar{n}_p^*$	$\bar{\delta}_p$	Z_{np}/n_p	$\hat{\sigma}_p$	n_d	$\bar{n}_d^*$	$\bar{\delta}_d$	Z_{nd}/n_d	$\hat{\sigma}_d$
Li	2	1.589	0.411	0.63	0.87	2	1.959	0.041	0.51	0.99	–	–	–	–	–
Be^+	2	1.729	0.271	1.16	0.84	2	1.954	0.046	1.02	0.98	–	–	–	–	–
B^{2+}	2	1.797	0.203	1.67	0.83	2	1.958	0.042	1.53	0.97	–	–	–	–	–
Na	3	1.627	1.373	0.61	0.92	3	2.117	0.883	0.47	0.96	3	2.990	0.010	0.33	0.999
Mg^+	3	1.903	1.097	1.05	0.88	3	2.265	0.735	0.88	0.94	3	2.970	0.030	0.67	0.998
Al^{2+}	3	2.075	0.925	1.45	0.87	3	2.372	0.628	1.27	0.92	3	2.950	0.050	1.02	0.995
K	4	1.770	2.230	0.56	0.93	4	2.234	1.766	0.45	0.96	3	2.854	0.146	0.35	0.997
Ca^+	4	2.141	1.859	0.93	0.90	4	2.497	1.503	0.80	0.93	3	2.312	0.687	0.86	0.97
Sc^{2+}	4	2.382	1.618	1.26	0.89	4	2.682	1.318	1.12	0.92	3	2.225	0.775	1.35	0.94
Ti^{3+}	4	2.557	1.443	1.56	0.87	4	2.822	1.178	1.42	0.91	3	2.244	0.756	1.78	0.93
V^{4+}	4	2.693	1.307	1.86	0.87	4	2.931	1.069	1.77	0.90	3	2.284	0.716	2.19	0.91
Cr^{5+}	4	2.802	1.198	2.14	0.86	4	3.020	0.980	1.99	0.89	3	2.326	0.674	2.58	0.90
Rb	5	1.805	3.195	0.55	0.95	5	2.288	2.712	0.44	0.97	4	2.767	1.233	0.36	0.99
Sr^+	5	2.221	2.779	0.90	0.93	5	2.604	2.396	0.77	0.95	4	2.432	1.569	0.82	0.96
Cs	6	1.869	4.131	0.53	0.96	6	2.351	3.649	0.43	0.97	5	2.552	2.448	0.39	0.98
Ba^+	6	2.332	3.668	0.86	0.94	6	2.721	3.279	0.74	0.96	5	2.415	2.585	0.83	0.96

electron into the electron core another physical mechanism comes into play. This mechanism has to be substantial because for the heavy ions *calcium* (Ca^+), *strontium* (Sr^+) and *barium* (Ba^+) the energy of the lowest d level drops *below* that of the lowest p level ($Z_{nd}/n_d > Z_{np}/n_p$). For the light ions, *beryllium* (Be^+) and *magnesium* (Mg^+), this is not the case. As we shall see this points to differences in screening properties between the inner core ($1s^2, 2s^2, 2p^6$) and outer core electrons ($3s^2, 3p^6$, *etc.*) which has the net effect of expelling valence charge away from the nuclear region. We illustrate this phenomenon with a few examples:

- In the case of the *lithium* isoelectronic sequence Li, Be^+ and B^{2+} only the $1s^2$ core is present. Table 4.2 reveals that the screening efficiency of the $2s$ valence orbital is modest. It decreases from 87% to 83% while the orbital becomes more hydrogenic when increasing the nuclear charge (δ_{2s} decreases by 50% from 0.411 to 0.203). The preferential binding decreases from $(Z_{ns}/Z_{np})^2 = 1.53$ in Li to 1.19 in B^{2+}. The increase in nuclear charge does not affect the $2p$ orbital which remains close to hydrogenic, also in the ions ($\delta_{2p} \simeq 0.041$). All this is plausible because the radial distribution function of the $1s$ core has little overlap with that of the $2p$ orbital as shown in Fig. 4.12*b*.

- Also for the *sodium* isoelectronic sequence Na, Mg^+ and Al^{2+} the outer core is absent. The inner core consists of closed $1s^2$, $2s^2$ and $2p^6$ shells. Table 4.2 shows that the screening efficiency of the $3s$ valence orbital (92% in Na) is better than that of the $2s$ orbital. Not surprisingly, because we have 10 rather than 2 core electrons involved in the screening. However, the $3s$ orbital is clearly less hydrogenic although this improves ionic charge (δ_{3s} decreases by 33% from 1.373 to 0.925). The preferential binding decreases from $(Z_{ns}/Z_{np})^2 = 1.68$ in Na to 1.30 in Al^{2+}. To a lesser extent this trend is also present for the $3p$ orbital (δ_{3p} decreases with 39% from 0.883 to 0.628). The screening of the valence electron in the $3d$ orbital remains close to perfect (> 99%). Although the quantum defect remains small it becomes *less* hydrogenic (δ_{3d} *increases* from 0.01 to 0.05). All this is plausible when comparing the radial distribution function (2.46) for the $2p$ orbital with $Z_c = Z - 4 = 8$ with that of the $3s$, $3p$ and $3d$ orbitals for which $Z_c = 2$. For the $3s$ and $3p$ orbitals the overlap is substantial, for the $3d$ orbital it is small.

- The perfect screening of the $3d$ orbital in Na, Mg$^+$ and Al^{2+} is lost when turning to the *potassium* isoelectronic sequence K, Ca$^+, \cdots$ Cr^{5+}. This has to do with the presence of the *outer* core ($3s^2$ and $3p^6$ shells). The outer core contracts with respect to the inner core because the increase in effective nuclear charge is largest for the outer shells. Increasing Z from 19 in K to 20 in Ca$^+$ the $1s^2$ core contraction is small because $Z^+_{1s}/Z_{1s} \simeq 1.05$. Also for the $2s^2$ and $2p$ the core contraction is small: $Z^+_{2s}/Z_{2s} \simeq 1.06$ and $Z^+_{2p}/Z_{2p} \simeq 1.07$. These estimates are made for maximal screening by the inner-lying shells: $Z^+_{nl}/Z_{nl} = (Z - \sigma_{nl} + 1)/(Z - \sigma_{nl})$ with $\sigma_{1s} = 0$, $\sigma_{2s} = 2$ and $\sigma_{2p} = 4$. For the *outer* core electrons the core contraction is larger, $Z^+_{3s}/Z_{3s} \simeq 1.11$ ($\sigma_{3p} = 10$) and $Z^+_{3p}/Z_{3p} \simeq 1.14$ ($\sigma_{3p} = 12$). Table 4.2 shows that the $4s$ valence orbital becomes more hydrogenic (δ_{4s} decreases by 1.032). To a lesser extent this is also the case for the $4p$ orbital (δ_{4p} decreases by 0.786). As the $3d$ orbital has much less overlap with the core electrons the increase in Z_c leads to a *reduction* in screening, the orbital becoming less hydrogenic (δ_{3d} *increases* by 0.528). Hence, the binding of the $3d$ electron grows relative to that of the $4s$ and $4p$ electrons. Further increasing Z_c by turning to doubly ionized *scandium* (Sc^{2+}) the $3d$ orbital contracts even further and its energy even drops below that of the $4s$ level, thus restoring the hydrogenic ordering of the levels. The contraction of the $3d$ orbital with growing nuclear charge is sometimes referred to as the *collapse of the* $3d$ *shell.* We return to this feature in Chapter 9.

In summary: we identified two opposing mechanisms which affect the screening of the valence electron. The increase in Z pulls the valence electron towards the nucleus, which *always* results in increased binding (Z_c doubles). In the presence of an outer core, the valence electron is pulled into the core electron cloud which further *increases* the binding by a *decrease* of screening. On the other hand, the same increase in Z results in a contraction of the electron core. This *decreases* the binding of those valence electrons that penetrate deeply into the core, the valence electrons in s orbitals and (to a lesser extent) p orbitals.

Problem 4.10. Consider a hydrogen-like isoelectronic pair of an atom and a singly charged ion in states with the same quantum numbers n and l. Show that the screening efficiency of the valence electron in atom and ion are equal if the following condition is satisfied:

$$\delta^+_{nl} = \tfrac{1}{2}\delta_{nl}\frac{1 - \delta^+_{nl}/n}{1 - \delta_{nl}/n},$$

where δ_{nl} and δ^+_{nl} are the quantum defects of the atom and ion, respectively, and $\delta_{nl}/n \ll 1$. The relation shows that in a general comparison between two isoelectronic states with equal n and l, a smaller quantum defect is not always equivalent with better screening.

Solution. For $\delta_{nl}/n \ll 1$ the screening efficiency is given by Eq. (4.149). For the atom the nuclear charge number is Z and the Rydberg charge number $Z_c = 1$; for the isoelectronic singly charge ion these quantities increase by one, $Z^{+1} = Z + 1$ and $Z^+_c = Z_c + 1$. Substituting this into Eq. (4.149) for atom and ion we find for equal screening:

$$\hat{\sigma}^+_{nl} = 1 - \frac{Z^+_c}{Z^+ - Z^+_c}\frac{\delta^+_{nl}/n}{1 - \delta^+_{nl}/n} = 1 - \frac{Z_c + 1}{Z - Z_c}\frac{\delta^+_{nl}/n}{1 - \delta^+_{nl}/n} = 1 - \frac{Z_c}{Z - Z_c}\frac{\delta_{nl}/n}{1 - \delta_{nl}/n} = \hat{\sigma}_{nl}.$$

Since $Z_c = 1$ we have $Z_c + 1 = 2Z_c$ and this leads to the desired condition. □

4.6.3 Screening by core electrons - effect on spin-orbit splitting

An alternative view on screening can be obtained by analyzing the influence of core electrons on the magnitude of the spin-orbit splitting. To introduce this alternative we present in Table 4.3 the

Table 4.3: Fine structure splitting of hydrogen and alkali atoms. The hydrogen and deuterium data are taken from [61]; those of lithium from [24, 98]; the rest of the table is based on the NIST Atomic Spectra Database Levels Data [66].

		$E(n\,^2S)$		$E(n\,^2P)$		$\lambda_{vac}(\mathrm{D}_2)$	$\lambda_{vac}(\mathrm{D}_1)$	$\Delta W_{3/2}$			
	n	n_s^*	(Ry)	$\bar{n}_p^*$	(Ry)	(nm)	(nm)	(GHz)	(Ry)	Z	Z_i
H	2	2.000	-0.25	2.000	-0.25	121.567	121.567	10.969 0	3.3×10^{-6}	1	1
D	2	2.000	-0.25	2.000	-0.25	121.534	121.534	10.972 0	3.3×10^{-6}	1	1
^{6}Li	2	1.590	-0.40	1.961	-0.26	670.977	670.992	10.052 8	3.1×10^{-6}	3	1
^{7}Li	2	1.589	-0.40	1.959	-0.26	670.962	670.977	10.053 4	3.1×10^{-6}	3	1
Na	3	1.627	-0.38	2.117	-0.22	589.158	589.756	515.521	1.6×10^{-4}	11	7
K	4	1.770	-0.32	2.234	-0.20	766.700	770.108	1 730.11	5.3×10^{-4}	19	15
Rb	5	1.805	-0.31	2.293	-0.19	780.241	794.979	7 122.92	2.3×10^{-3}	37	31
Cs	6	1.869	-0.29	2.362	-0.18	852.347	894.593	16 609.67	5.1×10^{-3}	55	49
Fr	7	1.828	-0.30	2.373	-0.18	718.185	817.166	50 562.67	15.4×10^{-3}	87	84

$\mathrm{D}_1 : n\,^2P_{1/2} - n\,^2S_{1/2}$; $\mathrm{D}_2 : n\,^2P_{3/2} - n\,^2S_{1/2}$; $\Delta W_{3/2} = E(n\,^2P_{3/2}) - E(n\,^2P_{1/2})$

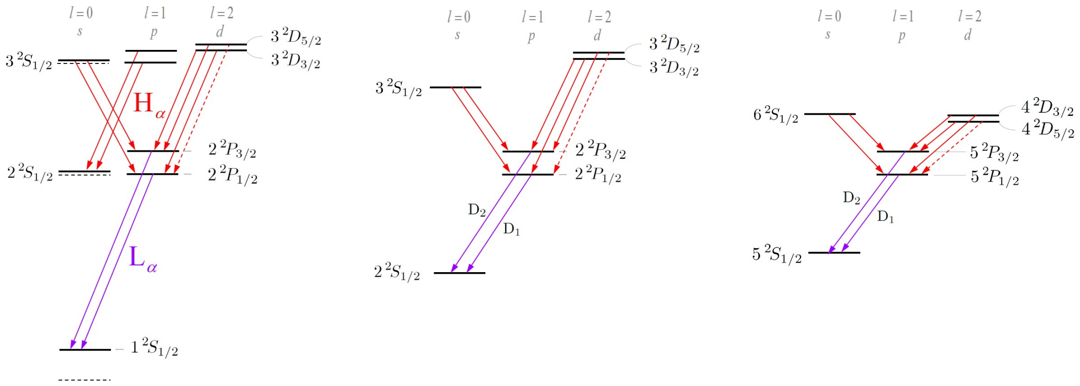
spin-orbit splitting of the (lowest-n) ^2P term, $\Delta W_{3/2}$, as observed in the alkali atoms, together with the Lyman α splitting in *hydrogen* (H) and *deuterium* (D). This ^2P splitting is usually referred to as the *doublet splitting* in hydrogen-like atoms. Included in the table are further the binding energies of the (lowest-n) ^2S and ^2P levels (represented by their effective principal quantum number as well as in Ry atomic units) along with the corresponding wavelength of the optical transitions to the $^2S_{1/2}$ ground states. For the alkalies these transitions are referred to as the D_1 and D_2 lines, respectively (see Fig. 4.14). The D_1, D_2 doublet shows up as a characteristic feature in the optical window of the alkali spectra.

Note that the magnitude of the doublet splitting grows by orders of magnitude when the atomic number increases from $Z = 1$ in *hydrogen* to $Z = 87$ in *francium* (Fr) but in all cases this splitting remains much smaller than the $n\,^2S - n\,^2P$ splitting (which is roughly 2 eV - see Fig. 4.9). The increase of the doublet splitting with Z may be large in absolute terms, it does not match the enormous increase predicted by the Z^4 scaling of the hydrogenic formula (4.137). As we are aware of the screening by core electrons this does not come as a surprise. Heuristically, the core electrons reduce the effective nuclear charge, which leads to a reduction of the velocity-induced magnetic field and the spin-orbit coupling accordingly. An impression of the influence of screening on the spin-orbit interaction can be obtained using a semi-empirical expression discovered by Landé [70]. This expression follows by substituting the observed binding energies, as given by Eq. (4.143), into Eq. (4.137) for the doublet splitting,

$$\Delta W_{l+1/2} = \alpha^2 \frac{Z_c^2 Z_i^2}{2\bar{n}_l^{*3} l(l+1)} \text{ Hartree.} \tag{4.151}$$

Here Z_c is the Rydberg charge, Z_i the *internal charge* and $\bar{n}_l^*$ is the doublet-averaged effective principal quantum number of the ^2P doublet. In hydrogen $Z_i = Z$ and $\bar{n}_l^* = n$. For the ^2P doublets in the alkalies, Z_i can be calculated with Eq. (4.151) using the experimental values for l, n_l^* and $\Delta W_{l+1/2}$; the results are listed in Table 4.3. Note that $(Z - Z_i)$ increases only modestly from ~ 2 in Li to ~ 6 in Cs with Z increasing from 3 to 58. This suggests that only the innermost core electrons are effective in suppressing the spin-orbit field. Roughly speaking, we have $Z_i \simeq Z - 4$ for p electrons. Francium differs because relativistic effects dominate. We return to these issues when discussing many-electron atoms.

Note that the D_1 and D_2 transition frequencies depend slightly on the isotope under consideration. In Table 4.3 this is demonstrated for the light isotopes, *hydrogen*, *deuterium* and *lithium*



hydrogen *lithium* *rubidium*

Figure 4.14: Fine structure and electric-dipole transitions in hydrogen-like atoms. *Left:* Lyman-α and Balmer-α transitions in atomic hydrogen; *Middle and Right*: The D_1, D_2 doublet structure observed in the spectra of alkali atoms is closely related to the rotational structure of the Lyman-α doublet in hydrogen. The fine structure of the $^2D_{3/2,5/2}$ levels is regular (hydrogen-like) in *lithium* but anomalous (inverted) in *rubidium*. Such anomalies arise from kinematic correlations between the electrons (see Section 7.1) but do not affect the selection rules.

(^{6}Li and ^{7}Li). For *sodium* (Na) and the heavier alkalies this *isotope shift* is too small to be visible at the wavelength resolution presented. The origin of the isotope shift is found in two isotope-dependent phenomena that affect the binding energy: (a) the *reduced mass correction*, introduced in Section 2.1; (b) the *nuclear volume correction*. For the hydrogen isotopes the nuclear volumes are small and the reduced mass correction large; i.e., the latter completely dominates the shift (see Problem 4.11). For the lithium isotopes the two corrections come in on equal footing. Knowledge of the mass related corrections are essential for the determination of the Lamb shift, crucial for precision tests of quantum electrodynamics. Note that, aside from the isotope effect on the binding energy of the $^2P_{1/2}$ and $^2P_{3/2}$ levels, there is also a small isotope effect of the doublet splitting itself. The latter has to be small because the mentioned corrections are predominantly common to the $^2P_{1/2}$ and $^2P_{3/2}$ levels. It is a challenge for theoreticians and experimentalists alike to attain the required precision for a consistent determination of the isotope effects on the D_1 and D_2 lines as well as that on the spin-orbit splitting [98].

Problem 4.11. The Lyman-α isotope shift between hydrogen and deuterium is $\Delta\lambda = 0.033$ nm (see Table 4.3). Show that this shift is (mostly) explained by the reduced mass correction.

Solution. Using the Rydberg formula (2.35) we find that the Lyman-α wavelength ratio is inversely proportional to the ratio of Rydberg constants

$$\frac{\lambda_H}{\lambda_D} = \frac{R_D}{R_H} = \frac{1+m_e/m_p}{1+m_e/m_d} \simeq 1.000272,$$

where $m_p = 1.007276466812$ amu is the proton mass and $m_d = 2.013553212712$ amu the deuteron mass. Hence, we calculate λ_H to be 0.0272% larger than that of deuterium, which corresponds to a reduced mass correction of 0.033057 nm and explains the observed isotope shift. □

4.6.4 Transition dipole moments in the presence of spin-orbit coupling

In the presence of spin-orbit coupling the m_l is no longer a good quantum number and this has consequences for the calculation of the transition-dipole elements of the electric-dipole operator. In this case the states are written as $|nlsjm_j\rangle$ and the transition dipole is of the form

$$\mathbf{D}_{eg} = D_{eg}\hat{\mathbf{u}}_q^* \tag{4.152}$$

where $e = |n'l's'j'm'\rangle$, $g = |nlsjm\rangle$ and

$$D_{eg} = -ea \,\langle n'l's'j'm'|\, \sqrt{4\pi/3}\rho Y_1^q(\hat{\mathbf{r}})\, |nlsjm\rangle \tag{4.153}$$

is the *transition dipole moment*, with $q = m' - m$. In this case we immediately turn to the canonical approach. Because the operators $\sqrt{4\pi/3}\, erY_1^q(\hat{\mathbf{r}})$ are irreducible vector operators we can apply the Wigner-Eckart theorem (K.24) and obtain

$$D_{eg} = -ea(-1)^{j'-m'} \langle n'l's'j'||\sqrt{4\pi/3}\rho Y_1(\hat{\mathbf{r}})||nlsj\rangle \begin{pmatrix} j' & 1 & j \\ -m' & q & m \end{pmatrix}. \tag{4.154}$$

In this case the reduced matrix element can be further reduced to the standard representation $\{\mathbf{L}^2, L_z, \mathbf{S}^2, S_z\}$. For this we use the repeated reduction formula (K.62),

$$\begin{aligned} \langle n'l's'j'||\sqrt{4\pi/3}\rho Y_1(\hat{\mathbf{r}})||nlsj\rangle &= (-1)^{j+l'+s+1}\delta_{s,s'} \\ &\times \langle n'l'||\sqrt{4\pi/3}\rho Y_1(\hat{\mathbf{r}})||nl\rangle \sqrt{(2j+1)\,(2j'+1)} \begin{Bmatrix} l' & j' & s \\ j & l & 1 \end{Bmatrix}. \end{aligned} \tag{4.155}$$

Substituting Eq. (2.79) this becomes

$$\begin{aligned} \langle n'l's'j'||\sqrt{4\pi/3}\rho Y_1(\hat{\mathbf{r}})||nlsj\rangle &= (-1)^{j+s+1+\max(l,l')}\delta_{s,s'}\delta_{l',l\pm1}\mathcal{R}_{n'l',nl}\times \\ &\times \sqrt{\max(l,l')}\sqrt{(2j+1)\,(2j'+1)} \begin{Bmatrix} l' & j' & s \\ j & l & 1 \end{Bmatrix}. \end{aligned} \tag{4.156}$$

Writing the transition dipole moment in the form

$$D_{n'l's'j'm'\leftrightarrow nlsjm} = -ea\mathcal{R}_{n'l',nl}\mathcal{A}^{l'ls}_{j'm'jm}\delta_{s,s'}\delta_{l',l\pm1}, \tag{4.157}$$

we find for the angular contribution (after reordering the $3j$ symbol)

$$\mathcal{A}^{l'ls}_{j'm'jm} = (-1)^{-m'+s+\max(l,l')}\sqrt{\max(l,l')}\sqrt{(2j+1)\,(2j'+1)} \begin{Bmatrix} l' & j' & s \\ j & l & 1 \end{Bmatrix} \begin{pmatrix} j & 1 & j' \\ m & q & -m' \end{pmatrix}. \tag{4.158}$$

The *selection rules* for fine-structure transitions (determined by the $3j$ symbol) are:

$$\Delta m = 0, \pm1 \quad \text{for } \pi, \sigma^{\pm} - \text{transitions} \tag{4.159a}$$

$$\Delta j = 0, \pm1 \quad (\text{excluding } \; j = 0 \leftrightarrow j' = 0). \tag{4.159b}$$

These selection rules come on top of the *electric-dipole parity selection rule* (2.73),

$$\Delta l = \pm1 \qquad \Delta s = 0\,. \tag{4.160}$$

In addition also the triangular inequalities $\Delta\,(l,s,j)$ and $\Delta\,(l'sj')$ must be satisfied. Note that the case $j = j' = 0$ is excluded because the triangular inequality $\Delta\,(j,j',1)$ cannot be satisfied. The rule $\Delta s = 0$ is trivial because in one-electron atoms the total spin is $s = \frac{1}{2}$, which is the intrinsic spin of the electron and cannot change. The allowed electric-dipole transitions between lower levels of hydrogen-like atoms are sketched in Fig. 4.14.

4.6.4.1 Example: $p \to s$ transitions in hydrogen-like atoms $(s = 1/2)$

The result for the angular matrix elements of $p \to s$ transitions in the presence of fine-structure coupling $\left(j' = \frac{1}{2}, \frac{3}{2} \to j = \frac{1}{2}\right)$ depends, for given j', on $m' \equiv m_{j'}$,

$$\mathcal{A}^{l'ls}_{\frac{3}{2}m',\frac{1}{2}\pm\frac{1}{2}} = (-)^{\frac{3}{2}-m'}\sqrt{8} \begin{Bmatrix} 1 & \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 \end{Bmatrix} \begin{pmatrix} \frac{1}{2} & 1 & \frac{3}{2} \\ \pm\frac{1}{2} & m' \mp \frac{1}{2} & -m' \end{pmatrix} = (-)^{\frac{3}{2}-m'}\tfrac{1}{3}\sqrt{\tfrac{3}{2} \pm m'}, \tag{4.161a}$$

$$\mathcal{A}^{l'ls}_{\frac{1}{2}m',\frac{1}{2}\pm\frac{1}{2}} = (-)^{\frac{3}{2}-m'}\sqrt{8} \begin{Bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 1 \end{Bmatrix} \begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \pm\frac{1}{2} & m' \mp \frac{1}{2} & -m' \end{pmatrix} = (-)^{\frac{3}{2}-3m'}\tfrac{1}{3}\sqrt{\tfrac{3}{2} \mp m'}. \tag{4.161b}$$

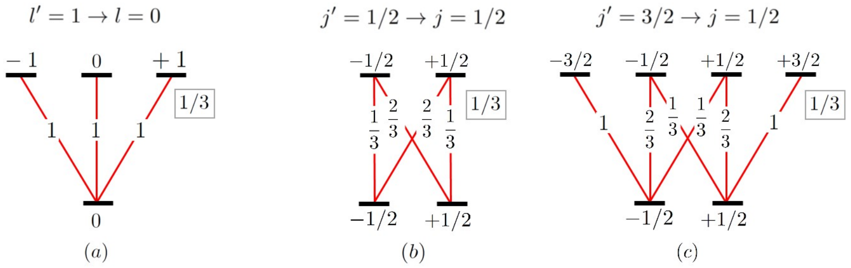


Figure 4.15: The angular contributions to the transition strength between two angular momentum levels (l' and l) is independent of the presence (or absence) of fine structure. This is illustrated for $p \to s$ transitions in hydrogen-like atoms with fine structure: (a) in principal structure (see also Fig. 2.4); (b) from $j' = \frac{1}{2}$ level; (c) from $j' = \frac{3}{2}$ level. Note that the sum of the probabilities starting from any of the $m_{j'}$ levels, always yields the *same* value (normalized to unity, not counting a common prefactor - here 1/3 in all three cases).

Summing over the magnetic quantum numbers of the $j = \frac{1}{2}$ level (which is the only s level) we obtain

$$\sum_{m=-1/2}^{1/2} \mathcal{A}^2_{j'm',\frac{1}{2}m} = \tfrac{1}{3}. \tag{4.162}$$

Note that (for given l', l, and s) this sum is *independent* of j' and m'; i.e., the sum is same for all values of the magnetic quantum number of the p level. This is illustrated in Fig. 4.15. This sum rule can be understood from the physical point of view when considering electric-dipole transitions. Since the electric-dipole operator does not couple to the spin (it acts in a different sector of Hilbert space), the spin cannot affect the total electric dipole transition rate. It only distributes the rate over the available final states in the form of σ_+, σ_- and π contributions.

4.7 Fine structure in an applied magnetic field

4.7.1 Introduction

In the presence of an externally applied magnetic field $\mathbf{B}$ the vector potential is nonzero, $\mathbf{A} = \frac{1}{2}\mathbf{B}\times\mathbf{r}$ (presuming the field to be homogeneous across the atom), and using the results of Sections 4.2 and 4.3, the Hamiltonian for the magnetic fine structure (4.120) can be written in the form of an *effective* Hamiltonian valid to first order in perturbation theory for given values of n and l - compare with Eq. (4.124)

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_r + (\zeta_{nl}/\hbar^2)\mathbf{L}\cdot\mathbf{S} - (\boldsymbol{\mu}_s + \boldsymbol{\mu}_L)\cdot\mathbf{B}. \tag{4.163}$$

Here $\mathcal{H}_0$ is the Schrödinger Hamiltonian (2.1) and $\mathcal{H}_r$ is the relativistic correction given by Eq. (4.23). Substituting Eqs. (4.53) and (4.80) for the magnetic moments the Hamiltonian takes the form

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_r + (\zeta_{nl}/\hbar^2)\mathbf{L}\cdot\mathbf{S} + (g_L\mu_B\mathbf{L} + g_e\mu_B\mathbf{S})\cdot\mathbf{B}/\hbar. \tag{4.164}$$

We write this as $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_r + \mathcal{H}'$, where

$$\mathcal{H}' = \mathcal{H}_{LS} + \mathcal{H}_Z \tag{4.165}$$

is the perturbation to be considered, with

$$\mathcal{H}_{LS} = (\zeta_{nl}/\hbar^2)\mathbf{L}\cdot\mathbf{S} \quad \text{and} \quad \mathcal{H}_Z = (g_L\mu_B L_z + g_e\mu_B S_z)B/\hbar. \tag{4.166}$$

representing the spin orbit coupling and the orbital and Zeeman terms, respectively, with the quantization axis (z axis) chosen along the magnetic field direction.

We calculate the splitting of a fine-structure level of given n and l using first-order perturbation theory for a degenerate level. Unlike the zero-field case, for finite fields the perturbation theory does not reduce to ordinary first-order perturbation theory. Although $\mathbf{L}\cdot\mathbf{S}$ and $\mathcal{H}_Z$ commute *separately* with $\mathcal{H}_0+\mathcal{H}_r$, they do *not* commute *mutually* (see Problem 4.7). Therefore, the operators do not share a basis. Physically, there is competition between the mutual coupling between $\mathbf{L}$ and $\mathbf{S}$ and the coupling of $\mathbf{L}$ and $\mathbf{S}$ individually with the external $\mathbf{B}$ field. In low fields the $\mathbf{L}\cdot\mathbf{S}$ coupling is dominant, whereas in high fields it is replaced by the $\mathbf{L}\cdot\mathbf{B}$ and $\mathbf{S}\cdot\mathbf{B}$ couplings. This crossover from low-field to high-field behavior is known as the *Paschen-Back effect* [78]. In intermediate fields the Hamiltonian is neither diagonal in the $\{|nlsjm_j\rangle\}$ (*coupled*) representation nor in the $\{|nlm_lsm_s\rangle\}$ (*uncoupled*) representation and to calculate the level shifts to first order we have to diagonalize the perturbation matrix. This is equivalent to solving the secular equation (I.24),

$$|\mathcal{H}'_{i,j}-\varepsilon\delta_{i,j}|=0, \tag{4.167}$$

where $\mathcal{H}'_{i,j}$ are the matrix elements of the perturbation $\mathcal{H}'$ in the representation of choice.

Also for *many-electron atoms,* with total orbital angular momentum L and total spin S, the above diagonalization procedure can be used to describe the fine structure. This is the case if the perturbation can be written in the form (4.165) and is known as LS coupling or *Russell-Saunders coupling* (cf. Section 9.5.4)

4.7.2 Matrix element in the uncoupled basis

To solve the secular equation in the *uncoupled* representation we use the inner product rule (3.109b) to write the perturbation in the form

$$\mathcal{H}'=(\zeta_{nl}/\hbar^2)\left[L_zS_z+\tfrac{1}{2}\left(L_+S_-+L_-S_+\right)\right]+(g_L\mu_BL_z+g_e\mu_BS_z)B/\hbar. \tag{4.168}$$

Note that this Hamiltonian conserves the total angular momentum along the z axis; i.e., m_l+m_s is a good quantum number (J_z commutes with the Hamiltonian) for all values of the magnetic field. This quantum number will be denoted by m_j. The diagonal matrix elements of the perturbation $\mathcal{H}'$ in the uncoupled basis $\{|nlm_lsm_s\rangle\}$ and for given values of $m_j=m_l+m_s$ are given by

$$\begin{aligned}\langle nlm_lsm_s|\mathcal{H}'|nlm_lsm_s\rangle &= \langle nl(m_j-m_s)sm_s|\mathcal{H}'|nl(m_j-m_s)sm_s\rangle\\ &=\zeta_{nl}(m_j-m_s)m_s+\left[g_Lm_j+(g_e-g_L)m_s\right]\mu_BB\\ &\equiv\langle m_jm_s|\mathcal{H}'|m_jm_s\rangle\equiv\mathcal{H}'_{m_jm_sm_s},\end{aligned} \tag{4.169}$$

with ζ_{nl} the spin-orbit coupling constant. Note that in these matrix elements the states $\{|nlm_lsm_s\rangle\}$ share fixed values for the quantum numbers n, l and s; only the values of m_l and m_s vary. Using the relation $m_l=m_j-m_s$ we gave preference to specification of m_j and m_s rather than m_l and m_s. Turning to the shorthand notation $|nlm_lsm_s\rangle\rightarrow|m_jm_s\rangle$ the matrix elements take the form $\langle m_jm'_s|\mathcal{H}'|m_jm_s\rangle$. Since m_j is conserved by $\mathcal{H}'$ the notation can be condensed to $\mathcal{H}'_{m_jm'_sm_s}$. The terms $\frac{1}{2}(\zeta_{nl}/\hbar^2)\left(L_+S_-+L_-S_+\right)$ give rise to off-diagonal elements and are calculated using the shift operators (1.59) and the conservation of m_j,

$$\begin{aligned}\langle nlm'_lsm'_s|\mathcal{H}'|nlm_lsm_s\rangle &= \langle nl(m_j-m_s\pm1)s(m_s\mp1)|\mathcal{H}'|nl(m_j-m_s)sm_s\rangle\\ &=\tfrac{1}{2}\zeta_{nl}\sqrt{l(l+1)-(m_j-m_s)(m_j-m_s\pm1)}\sqrt{s(s+1)-m_s(m_s\mp1)}\\ &\equiv\mathcal{H}'_{m_j(m_s\mp1)m_s}.\end{aligned} \tag{4.170}$$

Hence, the perturbation couples states differing in m_s such that $\Delta m_s=\pm1$ while $m_l+m_s=m_j$ remains conserved. For the coupled basis $\{|nlsjm_j\rangle\}$ this implies that the perturbation couples states differing in j such that $\Delta j=\pm1$ while m_j remains conserved. In the description of the field dependence preference will be given to expressing coupled states in terms of uncoupled states, rather than the other way around, because in the uncoupled representation the expressions for the limiting cases of low and high field remain intuitively transparent. In the coupled representation the latter is only the case for the low-field limit.

4.7.3 Diagonalization of the perturbation matrix for hydrogen-like atoms

For hydrogen-like atoms the solution of the secular equation (4.167) can be obtained analytically for arbitrary magnetic fields because the electronic states are *electron spin doublets* $\left(s = \frac{1}{2}\right)$. As shown in Section 4.5 the spin-orbit shifts in zero field can be expressed as

$$\Delta E_{n,j}^{LS} = \begin{cases} +\frac{1}{2}\zeta_{nl}l & \left(j = l + \frac{1}{2}\right) \\ -\frac{1}{2}\zeta_{nl}(l+1) & \left(j = l - \frac{1}{2}\right) \end{cases} \quad \text{for } l \neq 0, \tag{4.171}$$

where the manifold with the highest (lowest) multiplicity has the highest (lowest) energy. In the absence of orbital angular momentum ($l = 0$) this shift is absent. For $l > 0$ the perturbation matrix takes the form

$$\mathcal{H}' = \begin{pmatrix} \mathcal{H}'_{M\uparrow\uparrow} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \mathcal{H}'_{(M-1)\uparrow\uparrow} & \mathcal{H}'_{(M-1)\uparrow\downarrow} & \cdots & 0 & 0 & 0 \\ 0 & \mathcal{H}'_{(M-1)\downarrow\uparrow} & \mathcal{H}'_{(M-1)\downarrow\downarrow} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mathcal{H}'_{(-M+1)\uparrow\uparrow} & \mathcal{H}'_{(-M+1)\uparrow\downarrow} & 0 \\ 0 & 0 & 0 & \cdots & \mathcal{H}'_{(-M+1)\downarrow\uparrow} & \mathcal{H}'_{(-M+1)\downarrow\downarrow} & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \mathcal{H}'_{-M\downarrow\downarrow} \end{pmatrix}, \tag{4.172}$$

where $M \equiv l + \frac{1}{2}$ and ↑ is the usual short hand notation for $m_s = \frac{1}{2}$ (and ↓ for $m_s = -\frac{1}{2}$). The quantum number m_j varies within the range $-M \leq m_j \leq M$. Note that the matrix is a $d \times d$ matrix, where $d = (2l+1)(2s+1)$ is the dimension of the angular momentum subspace in which $\mathbf{L}$ and $\mathbf{S}$ operate. The characteristic equation is of the asymmetric type (cf. Appendix H.3) and factorizes into a product of two determinants of size (1×1) and $2l$ determinants of size (2×2), each characterized by its own value of $m_j = m_l + m_s$, with $-M \leq m_j \leq M$. With the aid of Eqs. (4.169) and (4.170) the matrix elements are found to be

$$\mathcal{H}'_{m_j\uparrow\uparrow} = +\tfrac{1}{2}\zeta_{nl}\left(m_j - \tfrac{1}{2}\right) + g_L\mu_B m_j B + \tfrac{1}{2}\mu'_B B \tag{4.173a}$$

$$\mathcal{H}'_{m_j\downarrow\downarrow} = -\tfrac{1}{2}\zeta_{nl}\left(m_j + \tfrac{1}{2}\right) + g_L\mu_B m_j B - \tfrac{1}{2}\mu'_B B \tag{4.173b}$$

$$\mathcal{H}'_{m_j\downarrow\uparrow} = +\tfrac{1}{2}\zeta_{nl}(l + \tfrac{1}{2})(1 - \tilde{m}_j^2)^{1/2} = \mathcal{H}'_{m_j\uparrow\downarrow}, \tag{4.173c}$$

where

$$\mu'_B \equiv (g_e - g_L)\,\mu_B \simeq \mu_B \tag{4.174a}$$

$$\tilde{m}_j \equiv m_j/(l + \tfrac{1}{2}) \quad \text{with} \quad -1 \leq \tilde{m}_j \leq 1. \tag{4.174b}$$

4.7.3.1 Pure states

Note that for $m_j = \pm(l + \frac{1}{2})$, the coupling term vanishes, $\mathcal{H}'_{m_j\downarrow\uparrow} = 0$. Therefore, the corresponding states are called *pure states*. They are also known under the name *stretched states* because the orbital and spin angular momentum spins are coupled "parallel" and have maximum projection along the quantization axis,

$$|j_{\max}, \pm(l + \tfrac{1}{2})\rangle = \left|l, \pm l; s, \pm\tfrac{1}{2}\right\rangle. \tag{4.175}$$

The corresponding energies are

$$\varepsilon_{\pm} = \tfrac{1}{2}\zeta_{nl} l \pm (\mu_L + \mu_e) B \quad \text{for} \quad m_j = \pm(l + \tfrac{1}{2}). \tag{4.176}$$

Here $\mu_L \equiv g_L \mu_B l$ and $\mu_e \equiv g_e \mu_B s$ are, respectively, the *orbital* and *intrinsic* magnetic moment of the electron (cf. Sections 4.3.3 and 4.3.5).

4.7.3.2 Mixed states

The other states, $m_j \neq \pm(l + \frac{1}{2})$, are called *mixed states* because they involve linear combinations of the uncoupled basis states,

$$|j_{\max}, m_j\rangle = +\cos\alpha |l, \left(m_j - \tfrac{1}{2}\right); s, \uparrow\rangle + \sin\alpha |l, \left(m_j + \tfrac{1}{2}\right); s, \downarrow\rangle \tag{4.177a}$$

$$|j_{\min}, m_j\rangle = -\sin\alpha |l, \left(m_j - \tfrac{1}{2}\right); s, \uparrow\rangle + \cos\alpha |l, \left(m_j + \tfrac{1}{2}\right); s, \downarrow\rangle. \tag{4.177b}$$

With this choice of coefficients the states $|j_{\max}, m_j\rangle$ and $|j_{\min}, m_j\rangle$ form an orthonormal pair. The states labeled $j_{\max}$ ($j_{\min}$) are taken to be part of the manifold with highest (lowest) multiplicity. In the limit $B \to 0$ the Zeeman energy vanishes and j becomes a good quantum number, $j_{\max} \to j = l + \frac{1}{2}$ and $j_{\min} \to j = l - \frac{1}{2}$. In this limit the Eqs. (4.177) take the form of the Clebsch-Gordan decomposition (J.49), where the phase factors satisfy the *Clebsch-Gordan phase convention* introduced in Section 4.4.1. We have chosen the phase factors *also for non-zero fields* in accordance with this convention. Because Eq. (4.177a) corresponds in the limit $B \to 0$ to a state of maximal j (parallel coupling) the phase factors are unity. The energies of the mixed states are given by solutions of the secular equation for given value of m_j,

$$W_{\pm}(m_j) = \tfrac{1}{2}\left(\mathcal{H}'_{m_j\uparrow\uparrow} + \mathcal{H}'_{m_j\downarrow\downarrow}\right) \pm \tfrac{1}{2}\sqrt{(\mathcal{H}'_{m_j\uparrow\uparrow} - \mathcal{H}'_{m_j\downarrow\downarrow})^2 + 4|\mathcal{H}'_{m_j\downarrow\uparrow}|^2}. \tag{4.178}$$

This equation can be written in the form of the *Breit-Rabi formula*,[1]

$$\varepsilon_{\pm}(m_j) = g_L \mu_B m_j B - \tfrac{1}{4}\zeta_{nl} \pm \tfrac{1}{2}\zeta_{nl}(l + \tfrac{1}{2})\sqrt{1 + 2\tilde{m}_j B/B_{\text{fs}} + (B/B_{\text{fs}})^2}. \tag{4.179}$$

The characteristic magnetic field,

$$B_{\text{fs}} \equiv \zeta_{nl}(l + \tfrac{1}{2})/\mu'_B, \tag{4.180}$$

with μ'_B defined by Eq. (4.174a), is called the *fine-structure field*. It represents an *effective internal magnetic field* (as if the external field cannot be completely switched off) and is a measure for the spin-orbit splitting in magnetic field units. Its absolute value,[2] $|B_{\text{fs}}|$, is the *fine-structure crossover* field - the intersection point of the low-field and high-field asymptotes of the magnetic field dependence of the energy levels. For the $2p$ level of atomic hydrogen $B_{\text{fs}} \simeq 0.8$ T (see Problem 4.12). With Eq. (4.179) we have obtained an expression for the magnetic field dependence of the spin-orbit shift of atomic states with given values of l and $m_j = m_l + m_s \neq \pm(l + \frac{1}{2})$. It describes the crossover from zero field to the high magnetic field limit. In Fig. 4.16 this is illustrated for ${}^2P_{3/2}$ and ${}^2P_{1/2}$ atomic states with the magnetic field given in units of the crossover field $|B_{\text{fs}}|$.

To obtain the eigenstates corresponding to the energies ε_+ and ε_- we have to determine the coefficients in Eqs. (4.177). This is done by the procedure described in Appendix H.3. For hydrogen-like atoms ($\zeta_{nl} > 0$) we have $\varepsilon_+ > \varepsilon_-$; i.e., ε_+ (ε_-) corresponds to the upper (lower) branch of the doublet. From the limit $B \to 0$ analyzed in Section 4.5 we know that ε_+ corresponds to $|j_{\max}, m_j\rangle$ and ε_- to $|j_{\min}, m_j\rangle$. The amplitudes of the coefficients, $A_+ \equiv \cos\alpha$ and $A_- \equiv \sin\alpha$, follow by substitution of the matrix elements (4.173a)-(4.173c) into Eq. (G.51)

$$A_{\pm}^2 = \frac{1 - \tilde{m}_j^2}{\left(\tilde{m}_j + B/B_{\text{fs}} \mp \sqrt{1 + 2\tilde{m}_j B/B_{\text{fs}} + (B/B_{\text{fs}})^2}\,\right)^2 + 1 - \tilde{m}_j^2}. \tag{4.181}$$

---

[1] The name refers to the work of Breit and Rabi in relation to the hyperfine structure of atomic hydrogen [22].

[2] Note that B_{fs} can become negative in many-electron atoms (see Section 9.5.5).

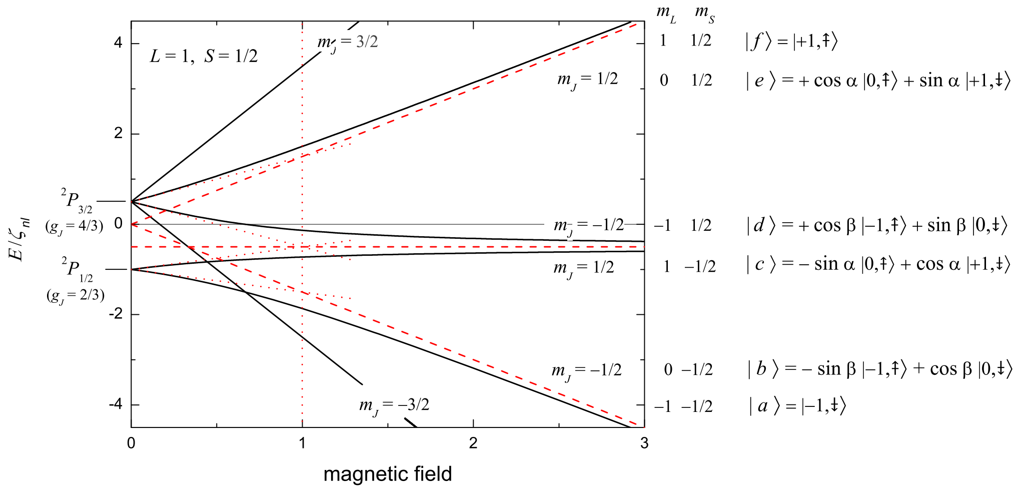


Figure 4.16: Fine structure diagram for the $^2P_{3/2}$ and $^2P_{1/2}$ electronic states and magnetic energy of the various magnetic sublevels from zero field to the Paschen-Back regime as calculated with Eqs. (4.179). The magnetic energy is given in units of the spin-orbit coupling constant ζ_{nl} and the magnetic field in units of the fine-structure crossover field, $|B_{\rm fs}|$. The dashed lines correspond to the high-field approximations - see Eq. (4.190). The dotted lines correspond to the low-field approximation - see Eq. (4.199). The states of equal m_J are LS coupled. The limiting behavior is: (a) $B \to \infty$: $\sin\alpha$, $\sin\beta \to 0$; (b) $B \to 0$: $\cos^2\alpha \to 2/3$, $\cos^2\beta \to 1/3$.

High-field limit ($B \gg |B_{\rm fs}|$)**:** In high magnetic fields we can expand the Breit-Rabi formula (4.179) in powers of $B_{\rm fs}/B$. Retaining only the terms up to first order in $B_{\rm fs}/B$ we obtain

$$\varepsilon_\pm(m_j) = g_L\mu_B m_j B - \tfrac{1}{4}\zeta_{nl} \pm \tfrac{1}{2}\mu'_B B(1 + \tilde{m}_j B_{\rm fs}/B). \tag{4.182}$$

The leading terms of this expression corresponds to the *high-field asymptotes* indicated by the dashed lines in Fig. 4.16. Using a similar expansion Eq. (4.181) takes the form

$$A_-^2 = \sin^2\alpha \simeq \tfrac{1}{4}(1 - \tilde{m}_j^2)(B_{\rm fs}/B)^2 \to 0. \tag{4.183}$$

This is an example of *weak coupling (strong asymmetry)* as defined in Appendix G.3.4.

Low-field limit ($B \ll |B_{\rm fs}|$)**:** For low magnetic fields we can expand Eq. (4.179) in powers of $B/B_{\rm fs}$. Retaining only the terms up to second order in $B/B_{\rm fs}$ we obtain

$$\varepsilon_\pm(m_j) = g_L\mu_B m_j B - \tfrac{1}{4}\zeta_{nl} \pm \tfrac{1}{2}\zeta_{nl}(l + \tfrac{1}{2})[1 + \tilde{m}_j B/B_{\rm fs} + \tfrac{1}{2}(1 - \tilde{m}_j^2)(B/B_{\rm fs})^2 + \cdots]. \tag{4.184}$$

The linear and quadratic contributions are referred to as *linear* and *quadratic* Zeeman effects. The linear terms correspond to the *low-field tangents* indicated by the dotted lines in Fig. 4.16. Expanding Eq. (4.181) to first order in $B/B_{\rm fs}$ we obtain

$$A_\pm^2 \simeq \tfrac{1}{2}(1 \pm \tilde{m}_j)\,[1 \pm (1 \mp \tilde{m}_j)B/B_{\rm fs} + \cdots]\,. \tag{4.185}$$

Note that this is *not* an example of the strong coupling limit (weak asymmetry) as defined in Appendix H.3. This would be the case for $m_j = 0$ but this value does not occur in hydrogen-like atoms.

Crossover field ($B = |B_{\rm fs}|$)**:** Equating Eqs. (4.182) and (4.184) for the high- and low-field tangents we obtain the following expression for the intersection point,

$$\mu_B' B(1 + \tilde{m}_j B_{\rm fs}/B) = \zeta_{nl}(l + \tfrac{1}{2})(1 + \tilde{m}_j B/B_{\rm fs}). \tag{4.186}$$

Solving for B we find for the intersection point $B = |B_{\rm fs}| = |\zeta_{nl}|(l + \frac{1}{2})/\mu_B'$. This point marks the crossover between the low-field and high-field regions. Note that the crossover field does *not* depend on m_j.

Problem 4.12. Calculate the fine-structure field $B_{\rm fs}$ for the $2p$ level of hydrogen.

Solution. Using Eqs. (4.180), (4.136) and (2.33) we obtain the following expression for the fine-structure field in hydrogenic atoms

$$B_{\rm fs} \simeq \alpha^2 \frac{Z^4}{n^4} \frac{hcR_M}{\mu_B} \frac{(l + {}^1\!/_2)n}{l(l + {}^1\!/_2)(l + 1)}.$$

For the $2p$ state of hydrogen ($n = 2$, $l = 1$ and $Z = 1$) this implies $\frac{1}{16}\alpha^2 hcR_M/\mu_B \simeq 0.8$ T. □

4.7.4 High-field limit - Paschen-Back effect

For magnetic fields much higher than the fine-structure field ($B \gg |B_{\rm fs}|$) the spin-orbit coupling is weak as compared to the Zeeman interaction. This causes the Zeeman coupling of **L** and **S** individually to the external field to dominate over the spin-orbit coupling between **L** and **S** - see Fig. 4.17. For increasing magnetic field, this manifests itself as a crossover from spin-orbit coupling of **L** and **S** ($B \ll |B_{\rm fs}|$) to Zeeman coupling of **L** and **S** individually to the magnetic field ($B \gg |B_{\rm fs}|$). This crossover is known as the *Paschen-Back effect.* Therefore, in the high-field limit the atomic Hamiltonian

$$\mathcal{H} = (\mathcal{H}_0 + \mathcal{H}_r + \mathcal{H}_Z) + \mathcal{H}_{LS} \tag{4.187}$$

is best analyzed in the *uncoupled* representation $\{|nlm_l sm_s\rangle\}$, where $\mathcal{H}_Z$ is diagonal and $\mathcal{H}_{LS}$ acts as the perturbation. In this basis and for given value of n, l, s the perturbation matrix is given by

$$\mathcal{H}^{LS}_{m_l' m_s'; m_l m_s} = (\zeta_{nl}/\hbar^2)\langle lm_l' sm_s'|\mathbf{L} \cdot \mathbf{S}|lm_l sm_s\rangle, \tag{4.188}$$

which can be written in the form

$$\mathcal{H}^{LS}_{m_l' m_s'; m_l m_s} = (\zeta_{nl}/\hbar^2)\langle lm_l' sm_s'|L_z S_z + \tfrac{1}{2}\left(L_+ S_- + L_- S_+\right)|lm_l sm_s\rangle. \tag{4.189}$$

As only the *diagonal matrix elements* of a weak perturbation *contribute to first order* in perturbation theory (see Appendix G.3.4 - weak coupling), in high fields the spin-orbit shift is given by $\Delta E^{LS}_{n;m_l m_s} = \zeta_{nl} m_l m_s$ and the *asymptotic field dependence* (dashed lines in Fig. 4.16) can be expressed as

$$\Delta E_{n;m_l m_s} = \left(g_L m_l + g_e m_s\right)\mu_B B + \zeta_{nl} m_l m_s. \tag{4.190}$$

Note that the derivation of this expression is valid for *arbitrary* values of l and s. Moreover, Eq. (4.190) may be generalized to the case of *many-electron atoms* under conditions of *Russell-Saunders coupling* using the arguments presented in Section 4.7.1 (cf. Section 9.5.4).

4.7.4.1 Hydrogen-like atoms

For hydrogen-like atoms the high-field limit is illustrated in Fig. 4.16 for the case $l = 1$, $s = \frac{1}{2}$, assuming $g_L = 1$ and $g_e = 2$. The stretched cases $m_l = 1$, $m_s = \frac{1}{2}$ and $m_l = -1$, $m_s = -\frac{1}{2}$ correspond to

$$\Delta E_{n;\pm m_l \pm m_s} = \tfrac{1}{2}\zeta_{nl} \pm 2\mu_B B \tag{4.191}$$

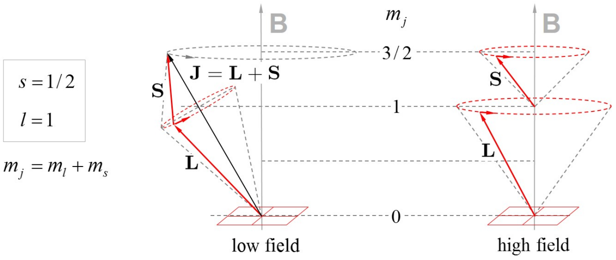


Figure 4.17: Precession of angular momenta in an externally applied magnetic field for the example $l = 1$, $s = \frac{1}{2}$. In *low* fields $\mathbf{L}$ and $\mathbf{S}$ are strongly spin-orbit coupled (by $\mathcal{H}_{LS}$) to form the total electronic angular momentum $\mathbf{J} = \mathbf{L}+\mathbf{S}$, which precesses slowly about the field direction as a result of a weak Zeeman coupling, $\mathcal{H}_Z = -\boldsymbol{\mu}_J \cdot \mathbf{B}$; this is most conveniently described in the *coupled* basis $\{|j, m_j\rangle\}$. In *high* fields $\mathbf{L}$ and $\mathbf{S}$ are predominantly Zeeman coupled to the external field, $\mathcal{H}_Z = -(\boldsymbol{\mu}_L + \boldsymbol{\mu}_S)\cdot\mathbf{B}$, and precess independently about the field direction; this is best described in the *uncoupled* basis $\{|l, m_l, s, m_s\rangle\}$. Note that $m_j = m_l + m_s$ is a good quantum number in both limits.

and coincide with the straight solid lines in Fig. 4.16. The $m_l = 0$ cases $m_s = \pm\frac{1}{2}$ correspond to

$$\Delta E_{n;\pm m_l \pm m_s} = \pm \mu_B B \tag{4.192}$$

and are represented by the inclined dashed lines in Fig. 4.16. The antiparallel cases $m_l = -1$, $m_s = \frac{1}{2}$ and $m_l = 1$, $m_s = -\frac{1}{2}$ correspond to

$$\Delta E_{n;\pm m_l \pm m_s} = -\tfrac{1}{2}\zeta_{nl} \tag{4.193}$$

and are shown as the horizontal dashed line in Fig. 4.16.

4.7.5 Low-field limit - Landé factor g_J

For magnetic fields much lower than the fine-structure field ($B \ll |B_{\mathrm{fs}}|$) the angular momenta $\mathbf{L}$ and $\mathbf{S}$ are spin-orbit coupled (by $\mathcal{H}_{LS}$) to form the total electronic angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ (see Fig. 4.17). Under these conditions the atomic Hamiltonian

$$\mathcal{H} = (\mathcal{H}_0 + \mathcal{H}_r + \mathcal{H}_{LS}) + \mathcal{H}_Z \tag{4.194}$$

is best analyzed in the *coupled* representation, $\{|nlsjm_j\rangle\}$, where $\mathbf{L}\cdot\mathbf{S}$ is diagonal and $\mathcal{H}_Z$ acts as the perturbation. In this basis and for given values of n, l, s, j the perturbation matrix is given by

$$\mathcal{H}^Z_{j'm_{j'};jm_j} = \langle j'm_{j'}|g_L L_z + g_e S_z\,|jm_j\rangle\,\mu_B B/\hbar. \tag{4.195}$$

Limiting ourselves to the diagonal terms, we obtain the following expression for the energy shift (Zeeman shift)

$$\Delta E^Z_{j,m_j} = \langle jm_j|g_L L_z + g_e S_z\,|jm_j\rangle\,\mu_B B/\hbar. \tag{4.196}$$

This expression is valid to only first order in perturbation theory because $\mathcal{H}_Z$ is in *not* diagonal in the coupled representation: only the *diagonal matrix elements* of a weak perturbation *contribute to first order* (see Appendix G.3.4 - weak coupling). Interestingly, the operators L_z and S_z cannot change m_l and m_s and, therefore, have to conserve m_j. This is a consequence of the m_j *selection rule* (3.90), $m_j = m_l + m_s$. However, L_z and S_z do *not* conserve j (see Problem 4.13). Only in *low* fields the perturbation can be made sufficiently weak to neglect the off-diagonal terms; i.e., to regard both j and m_j as good quantum numbers of the atomic Hamiltonian (4.194).. Furthermore, using the arguments given in Section 4.7.1, also the analysis of the low-field limit may be generalized to the case of *many-electron atoms* under conditions of *Russell-Saunders coupling* (cf. Section 9.5.4).

Intuitively the situation is also clear: for sufficiently low fields the magnetic moment of the atom will scale with the total electronic angular momentum **J**,

$$\boldsymbol{\mu}_J \equiv \gamma_J \mathbf{J} \equiv -g_J \mu_B \left(\mathbf{J}/\hbar\right) \quad \text{for } B \to 0. \tag{4.197}$$

Note that γ_J and g_J are defined with opposite sign, just as we did for γ_L and γ_e. The value of g_J will turn out to be positive. In terms of $\boldsymbol{\mu}_J$ the Zeeman energy can be described by the effective Hamiltonian

$$\mathcal{H}_Z = -(\boldsymbol{\mu}_L + \boldsymbol{\mu}_s) \cdot \mathbf{B} = -\boldsymbol{\mu}_J \cdot \mathbf{B} \quad \text{for } B \to 0. \tag{4.198}$$

This Hamiltonian yields a Zeeman shift on top of the zero-field spin-orbit shift ΔE_{nj}^{LS} (for given values of n, l, s, j):

$$\Delta E_{j,m_j}^{Z} = g_J \mu_B B \langle j m_j | (J_z/\hbar) | j m_j \rangle = g_J \mu_B m_j B \quad \text{for } B \to 0. \tag{4.199}$$

This expression corresponds to the *low-field tangents* indicated by the dotted lines in Fig. 4.16. As was to be expected, also the effective Hamiltonian (4.198) is diagonal in the $\{|nlsjm_j\rangle\}$ basis if the quantization axis is chosen along the direction of the **B** field. Our task is of course to determine the value of g_J.

Problem 4.13. Show that the fine-structure levels of a one-electron atom are coupled by the Zeeman interaction.

Solution. We show this for the orbital Zeeman interaction. Since L_z conserves $m_j = m_l + m_s$, we have to show $\langle j'm|L_z|jm\rangle \neq 0$ for $j' \neq j$, where $m = m_{j'} = m_j$. First we do a Clebsch-Gordan decomposition to the uncoupled representation using the triangle inequality $l - \frac{1}{2} \leq j \leq l + \frac{1}{2}$ as well as the projection conservation rule (J.5),

$$|jm\rangle = |l(m-\tfrac{1}{2}); \tfrac{1}{2}\tfrac{1}{2}\rangle\langle l(m-\tfrac{1}{2}); \tfrac{1}{2}\tfrac{1}{2}|jm\rangle + |l(m+\tfrac{1}{2}); \tfrac{1}{2} - \tfrac{1}{2}\rangle\langle l(m+\tfrac{1}{2}); \tfrac{1}{2} - \tfrac{1}{2}|jm\rangle.$$

Using this decomposition we obtain for the matrix element

$$\begin{aligned}\langle j'm|L_z|jm\rangle = {} & (m-\tfrac{1}{2})\langle j'm|l(m-\tfrac{1}{2}); \tfrac{1}{2}\tfrac{1}{2}\rangle\langle l(m-\tfrac{1}{2}); \tfrac{1}{2}\tfrac{1}{2}|jm\rangle \\ & + (m+\tfrac{1}{2})\langle j'm|l(m+\tfrac{1}{2}); \tfrac{1}{2} - \tfrac{1}{2}\rangle\langle l(m+\tfrac{1}{2}); \tfrac{1}{2} - \tfrac{1}{2}|jm\rangle,\end{aligned}$$

which is nonzero if $j' = l \pm \frac{1}{2}$ for $j = l \mp \frac{1}{2}$, with $m = \pm j$ or $m = \mp j'$. □

4.7.5.1 Wigner-Eckart theorem

To understand the relation between Eqs. (4.196) and (4.199) we recall the vector diagram in Fig. 4.6. The slow precession of $\mathbf{J} = \mathbf{L} + \mathbf{S}$ about the direction of a weak probe field (just serving to define a quantization axis) does not affect the projections of the angular momentum operators. As L_z and S_z commute with J_z these operators are invariant under rotation about the z axis; i.e., diagonal in the $\{|nlsjm_j\rangle\}$ basis. The matrix elements of L_z and S_z can be written in the form

$$\langle nlsjm_j|\, L_z \,|nlsjm_j\rangle = \langle lsj \,\|L\|\, lsj\rangle \,\langle jm_j|\, J_z \,|jm_j\rangle \tag{4.200a}$$

$$\langle nlsjm_j|\, S_z \,|nlsjm_j\rangle = \langle lsj \,\|S\|\, lsj\rangle \,\langle jm_j|\, J_z \,|jm_j\rangle\,, \tag{4.200b}$$

where the proportionality constants $\langle lsj \,\|L\|\, lsj\rangle$ and $\langle lsj \,\|S\|\, lsj\rangle$ are known as *reduced matrix elements*. They deserve special attention because they turn out to be *independent of* m_j. Since $J_z = L_z + S_z$ we find by adding Eqs. (4.200a) and (4.200b) that the reduced matrix elements add up to unity,

$$\langle lsj \,\|L\|\, lsj\rangle + \langle lsj \,\|S\|\, lsj\rangle = 1 = \langle lsj \,\|J\|\, lsj\rangle. \tag{4.201}$$

The above is intuitively clear from the vector diagram in Fig. 4.6. The projections of **L** and **S** along the total angular momentum vector **J**, i.e. $\mathbf{L} \cdot \mathbf{J}$ and $\mathbf{S} \cdot \mathbf{J}$, are conserved while **J** precesses about **B**,

$$\langle nlsjm_j|\, L_z \,|nlsjm_j\rangle = \frac{\langle lsjm_j|\, (\mathbf{L} \cdot \mathbf{J})\, J_z \,|lsjm_j\rangle}{\langle lsjm_j|\, \mathbf{J}^2 \,|lsjm_j\rangle}. \tag{4.202}$$

Since $\mathbf{L}\cdot\mathbf{J}$ and J_z are both diagonal in the low-field basis (see Problem 4.14) we can factorize the matrix element into the form (4.200a), where the m_j-independent part is given by

$$\langle lsj\,\|L\|\,lsj\rangle = \frac{\langle lsjm_j|\,\mathbf{L}\cdot\mathbf{J}\,|lsjm_j\rangle}{\langle lsjm_j|\,\mathbf{J}^2\,|lsjm_j\rangle} = \frac{j(j+1)+l(l+1)-s(s+1)}{2j(j+1)}. \tag{4.203}$$

An alternative derivation of this expression (not relying on the geometric argument presented here) is given in Problem 4.15. Eqs. (4.200) and similar expressions for the other vector components (including the shift operators) follow from the Wigner-Eckart theorem of group theory for the case of vector operators (cf. Appendix K.3.4). Importantly, the three components of the vector operators share the same reduced matrix element,

$$\langle nlsjm_j|\,L_z\,|nlsjm_j\rangle = \langle lsj\,\|L\|\,lsj\rangle\,\langle jm_j|\,J_z\,|jm_j\rangle \tag{4.204a}$$

$$\langle nlsj(m_j\pm 1)|\,L_\pm\,|nlsjm_j\rangle = \langle lsj\,\|L\|\,lsj\rangle\,\langle j(m_j\pm 1)|\,J_\pm\,|jm_j\rangle\,. \tag{4.204b}$$

These expressions embody the *vector projection rule* of coupled angular momenta.

4.7.5.2 Landé factor g_J

In search for g_J we return to Eqs. (4.196) and (4.199). With the aid of the Wigner-Eckart theorem (4.200) as well as the relation (4.201) we obtain the expression

$$\begin{aligned} g_J &= g_L\langle lsj\,\|L\|\,lsj\rangle + g_e\langle lsj\,\|S\|\,lsj\rangle \\ &= \tfrac{1}{2}\,(g_L+g_e) + \tfrac{1}{2}\,(g_L-g_e)\,[\langle lsj\,\|L\|\,lsj\rangle - \langle lsj\,\|S\|\,lsj\rangle]\,. \end{aligned} \tag{4.205}$$

Substituting the expressions for the reduced matrix elements (see above or Problem 4.15)

$$\langle lsj\,\|L\|\,lsj\rangle = \frac{j(j+1)+l(l+1)-s(s+1)}{2j(j+1)} \tag{4.206a}$$

$$\langle lsj\,\|S\|\,lsj\rangle = \frac{j(j+1)+s(s+1)-l(l+1)}{2j(j+1)} \tag{4.206b}$$

we obtain for the g_J factor,

$$g_J = \tfrac{1}{2}\,(g_L+g_e) + \tfrac{1}{2}\,(g_L-g_e)\,\frac{l(l+1)-s(s+1)}{j(j+1)}. \tag{4.207}$$

Approximating $g_L = 1$ and $g_e = 2$ we find that g_J takes the form

$$g_J = 1 + \frac{j(j+1)+s(s+1)-l(l+1)}{2j(j+1)}. \tag{4.208}$$

This expression is called the *Landé factor*. Note that just like the expressions for the high-field limit also the derivation of the expressions for the low-field limit is valid for *arbitrary* values of l and s. For the hydrogen-like atoms we have $s = \frac{1}{2}$ and the Landé factor reduces to the compact form

$$g_J = \begin{cases} 1 \pm \dfrac{1}{2l+1} & \left(j = l\pm\tfrac{1}{2}\right) \quad \text{for } l > 0 \\ 2 & \qquad\qquad\quad\;\; \text{for } l = 0. \end{cases} \tag{4.209}$$

Note that in this important case the g_J factor is always *positive* ($g_J > 0$) and largest for the state with the largest multiplicity (highest j).

In the important case of electronic s states Eq. (4.207) reduces to $g_J = g_e$ as expected for the absence of an orbital moment. However, due to confinement of the electron inside the atom, g_J

Table 4.4: Relative deviation from the free-electron g-value for hydrogen-like atoms; differences between isotopes of a given atom are at the ppb level [11, 4].

	H	Li	Na	K	Rb	Cs	units
$g_J/g_e - 1$	-17.4	$-20.9(1)$	$-23.4(1)$	$-24.3(1)$	$5.9(1)$	$98.5(1)$	ppm

differs slightly from the free-electron g_e value. These deviations are summarized for hydrogen-like atoms in Table 4.4. For the hydrogen atom $g_J(H)/g_e - 1 = -17.4(1.0) \times 10^{-6}$, which means that in hydrogen the electron g factor is reduced by 17.4 ppm [7]. The value of g_J cannot be determined without consideration of the nuclear properties because of coupling of the electronic and nuclear magnetic moments (see Chapter 5). Therefore, precision measurement of g_J requires analysis of the atomic hyperfine structure [11].

Problem 4.14. Derive the operator identity

$$\mathbf{L} \cdot \mathbf{J} = \tfrac{1}{2}[\mathbf{J}^2 + \mathbf{L}^2 - \mathbf{S}^2].$$

Solution. Since $\mathbf{J} = \mathbf{L} + \mathbf{S}$ we find

$$\mathbf{L} \cdot \mathbf{J} = \mathbf{L}^2 + \mathbf{L} \cdot \mathbf{S} = \mathbf{L}^2 + \tfrac{1}{2}[\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2] = \tfrac{1}{2}[\mathbf{J}^2 + \mathbf{L}^2 - \mathbf{S}^2]. \qquad \square$$

Problem 4.15. Derive the relation

$$\langle lsj \,\|L\|\, lsj\rangle = \frac{j(j+1) + l(l+1) - s(s+1)}{2j(j+1)}.$$

Solution. To determine $\langle lsj \,\|L\|\, lsj\rangle$ we evaluate the operator $\mathbf{S}^2 = (\mathbf{J} - \mathbf{L})^2 = \mathbf{J}^2 + \mathbf{L}^2 - 2\mathbf{J} \cdot \mathbf{L}$ (note that $\mathbf{J} \cdot \mathbf{L} = \mathbf{L} \cdot \mathbf{J}$) in the coupled representation,

$$\langle lsjm_j| \, \mathbf{S}^2 \, |lsjm_j\rangle = \langle lsjm_j| \, \mathbf{J}^2 + \mathbf{L}^2 - 2\mathbf{L} \cdot \mathbf{J} \, |lsjm_j\rangle \,.$$

Comparing the left-hand with the right-hand side we obtain

$$s(s+1) = j(j+1) + l(l+1) - 2\,\langle lsjm_j| \, \mathbf{L} \cdot \mathbf{J} \, |lsjm_j\rangle \,/\hbar^2.$$

With the aid of the inner-product rule (3.109b) and using Wigner-Eckart theorem (4.204) this relation becomes

$$\begin{aligned} s(s+1) &= j(j+1) + l(l+1) - 2\,\langle lsjm_j| \, L_z J_z + \tfrac{1}{2}[L_+ J_- + L_- J_+] \, |lsjm_j\rangle \,/\hbar^2 \\ &= j(j+1) + l(l+1) - 2\langle lsj \,\|L\|\, lsj\rangle \langle lsjm_j| J_z^2 + \tfrac{1}{2}[J_+ J_- + J_- J_+] \, |lsjm_j\rangle \,/\hbar^2 \\ &= j(j+1) + l(l+1) - 2\langle lsj \,\|L\|\, lsj\rangle \langle lsjm_j| \mathbf{J}^2 \, |lsjm_j\rangle \,/\hbar^2 \\ &= j(j+1) + l(l+1) - 2\langle lsj \,\|L\|\, lsj\rangle j(j+1). \end{aligned}$$

Solving for $\langle lsj \,\|L\|\, lsj\rangle$ we obtain the desired relation. $\square$

4.7.6 Example for hydrogen-like atoms

In Fig. 4.16 the low-field limit can be recognized for the case $l = 1$, $s = \frac{1}{2}$, assuming $g_L = 1$ and $g_e = 2$. For $j = l - s = \frac{1}{2}$ we calculate the Landé result $g_J = 2/3$; for $j = l + s = \frac{3}{2}$ we have $g_J = 4/3$. As shown by the solid lines in Fig. 4.16, in low field ($B \to 0$) we note that

- the 4-fold degeneracy of the ${}^2P_{3/2}$ level is lifted in accordance with

$$\Delta E_{n,j=3/2,m_j} \simeq \Delta E_{n,3/2} + \tfrac{4}{3}\mu_B m_j B, \qquad (4.210)$$

where $m_j = \pm\frac{1}{2}, \pm\frac{3}{2}$. Decomposing $|j, m_j\rangle = \sum_{m_l m_s} |lm_l; sm_s\rangle\langle lm_l; sm_s|jm_j\rangle$ the corresponding eigenstates can be written as

$$|\tfrac{3}{2}, \pm\tfrac{3}{2}\rangle = |1, \pm 1; \tfrac{1}{2}, \pm\tfrac{1}{2}\rangle \tag{4.211}$$

$$|\tfrac{3}{2}, \pm\tfrac{1}{2}\rangle = \sqrt{2/3}|1, 0; \tfrac{1}{2}, \pm\tfrac{1}{2}\rangle + \sqrt{1/3}|1, \pm 1; \tfrac{1}{2}, \mp\tfrac{1}{2}\rangle, \tag{4.212}$$

where the Clebsch-Gordan coefficients $\langle lm_l; sm_s|jm_j\rangle$ are taken from Appendix J.5.1 for the case of $(1 \times 1/2)$ coupling. Here the phase factors are unity because this is a case of maximal j. Note that the correct amplitudes of the coefficients are also obtained using Eq. (4.185) in the limit $B \to 0$. Hence, as we now have also established the sign, the full field dependence of the coefficients is known.

- the two-fold degeneracy of the $^2P_{1/2}$ level is lifted like

$$\Delta E_{n,j=1/2,m_j} \simeq \Delta E_{n,1/2} + \tfrac{2}{3}\mu_B m_j B, \tag{4.213}$$

with $m_j = \pm\frac{1}{2}$. In this case the eigenstates are

$$|\tfrac{1}{2}, \pm\tfrac{1}{2}\rangle = \mp\sqrt{1/3}|1, 0; \tfrac{1}{2}, \pm\tfrac{1}{2}\rangle \pm \sqrt{2/3}|1, \pm 1; \tfrac{1}{2}, \mp\tfrac{1}{2}\rangle, \tag{4.214}$$

where the Clebsch-Gordan coefficients $\langle lm_l sm_s|jm_j\rangle$ are again taken from Appendix J.5.1 for the case of $(1 \times 1/2)$ coupling. Also in this case the amplitude of the coefficients are obtained using Eq. (4.185) in the limit $B \to 0$.

5

Magnetic hyperfine structure

The term *hyperfine structure* is reserved for features of the atomic structure that arise from properties of the nucleus. From Chapter 1 we recall the *effective mass* correction, which accounts for the influence of the nuclear mass on the size of the electronic orbits. Likewise, the *nuclear volume* correction corrects for the non-zero nuclear volume (see Problem 4.1). The most prominent hyperfine features result from the magnetic dipole moment and electric quadrupole moment of the nucleus. The magnetic dipole moment, μ_I, is related to the *nuclear spin*, I, and gives rise to the *magnetic* hyperfine structure. The electric quadrupole moment, Q, arises from the *nuclear shape* and contributes to the *electric* hyperfine structure.

The existence of nuclear spin was conjectured by Wolfgang Pauli in 1924 in relation to observations of hyperfine structure in optical spectra [79]. The related phenomenology was demonstrated by Back and Goudsmit in 1928 [6]. However, as hyperfine shifts and splittings can be as much as a factor 1000 smaller than those of the fine structure, spectral resolution becomes an issue for precision measurements. Therefore, it took the invention of radio-frequency spectroscopy in atomic beams by Isodor Rabi in 1939 [84] before the hyperfine structure could be fully explored. The Rabi method was turned into a modern precision method by Norman Ramsey in 1949 [88, 87]. For the alkali atoms the experimental status was reviewed in 1977 [4]. Since the invention of optical frequency combs by Theodore Hänsch [55] and John Hall [53], hyperfine structure is studied with high-resolution optical spectroscopy (see e.g. [98]).

Although the nuclear properties are determined by the constituent particles (*nucleons* - protons and neutrons) and their orbits inside the nucleus, remarkably little detailed knowledge of nuclear and high-energy physics is required to understand a lot of the atomic hyperfine structure. For atomic spectroscopy it is important to know that the forces between the nucleons *conserve parity* in almost all practical cases. In exceptional cases the weak interaction (which does not conserve parity) plays a role. In any case, we do not have to deal with many other challenging issues of the nucleus. For instance, the structure of the nucleons is highly non-trivial because the interactions between the quarks (the building blocks of the nucleons) cannot be treated perturbatively. In this respect, the name *proton spin puzzle* of quantum chromodynamics (QCD) speaks for itself [2]: why is the spin of the proton 1/2 ? The quark spins (each spin-$\frac{1}{2}$) turn out to account for only some 30% of its value, and the largest contribution is attributed to a combination of the gluon spin (spin-1) and the orbital motion of quarks and gluons.

In this chapter we introduce the magnetic hyperfine structure of *one-electron atoms.* As we did for the fine structure we focus on atomic systems in which a *single* electron dominates the electronic angular momentum properties. Hence, the electronic states remain *electron spin doublets* ($s = \frac{1}{2}$) of varying orbital angular momentum. However, also non-hydrogen-like features arise. Aside from the spin of the valence electron, also the *spin of the core electrons* (core polarization - see Section 9.3.2) as well as the nuclear *spin* (magnetic hyperfine) and *shape* (electric hyperfine) play a role.

In view of the above it is appropriate to introduce the atomic hyperfine structure in two rounds.

In the present chapter we focus on the *magnetic* hyperfine structure as observed in the hydrogen atom and in the ground states of the alkali elements. In this way we can introduce the principal features of the magnetic hyperfine structure, including the well-known Zeeman effect. We emphasize the formal similarity with the phenomenology of the fine structure. In the hydrogenic case the core does not have to be addressed because core electrons are simply absent. In the alkali-atom *ground* states the core affects the results quantitatively but qualitatively the hyperfine structure remains essentially hydrogen like. In *excited* states, evident non-hydrogen-like behavior can be observed. We illustrate this for the $l > 0$ excited states (^2P, ^2D terms) of the alkali elements.

The physics of the *electric* hyperfine structure is quite different. In hydrogen ($I = \frac{1}{2}$) it is completely absent because the nucleus is spherical. For $I > \frac{1}{2}$ (as is the case for many alkali atoms) the nuclei are nonspherical. This affects the hyperfine structure but is of no consequence for the electronic ground states of the alkali atoms because the electronic charge distribution is spherical (^2S terms) and the coupling to electric multipole moments averages to zero. So, one may argue that the *electric* hyperfine structure is less prominent than other features of the atomic structure and can be left for a more advanced course. This leaves the magnetic hyperfine structure as the back bone of the non-relativistic theory for the hyperfine structure of free atoms. This being said we should be aware that interesting anomalies immediately arise once we deviate from the elementary picture of a non-relativistic one-electron atom with a spherical nucleus. A comprehensive introduction can be found in the book by Lloyd Armstrong [5].

5.1 Nuclear magnetic coupling

5.1.1 Nuclear Zeeman coupling

Any charged particle with angular momentum gives rise to a magnetic moment and the atomic nucleus is no exception to this rule. Therefore, just as we associated the electronic spin Zeeman energy to the electron spin, the nuclear spin gives rise to a nuclear spin Zeeman contribution. Writing the Hamiltonian for the magnetic energy in the generic form of a Zeeman interaction we have

$$\mathcal{H}_Z = -\boldsymbol{\mu}_I \cdot \mathbf{B}, \tag{5.1}$$

where

$$\boldsymbol{\mu}_I = \gamma_I \mathbf{I} \equiv g_I \mu_N (\mathbf{I}/\hbar) \tag{5.2}$$

is the *nuclear magnetic moment*, with $\mathbf{I}$ the nuclear spin, γ_I the *gyromagnetic ratio* of the nucleus, g_I the *nuclear g factor* and

$$\mu_N \equiv e\hbar/2m_p = (m_e/m_p)\mu_B \tag{5.3}$$

the *nuclear magneton*, $\mu_N \simeq 5 \times 10^{-27}$ J/T. The dimensionless number $m_e/m_p \approx 1/1836$ is referred to as the *electron-to-proton mass ratio.* Note that $e\hbar/2m_p$ is the gyromagnetic ratio of a classical rotating body of charge e and mass equal to the proton mass m_p. The g_I can be positive or negative, which is an important indicator that the nucleus is more than a simple rotating cloud of positive charge; aside from the rotational motion, also the intrinsic spins of the nucleons contribute to the moment. The quantum numbers I and m_I define the nuclear spin state. In analogy to the orbital case (4.66) the *magnitude* of the nuclear magnetic moment (short: *the* nuclear moment) is defined as the maximum observable nuclear moment (i.e., the value obtained for $m_I = I$),

$$\mu_I \equiv g_I \mu_N I. \tag{5.4}$$

For the nuclear spin quantum number I we shall use capitals rather than the lower-case notation used for the angular momentum quantum numbers of a single electron (s, l and j). The g factor is related to the gyromagnetic ratio of the nucleus,

$$\gamma_I \equiv g_I \mu_N/\hbar. \tag{5.5}$$

Table 5.1: Nuclear spin, nuclear moments and ground-state hyperfine constant and crossover field for selected isotopes [4] (Ca data from [12]). In the literature authors often avoid the nuclear magneton and use a different sign convention for the nuclear g factor: $g_I' = -g_I(m_e/m_p)$.

	I	g_I	μ_I (μ_N)	$\gamma_I/2\pi$ (MHz/T)	Q (barn)	a_{hfs} (MHz)	g_J	B_{hfs} (Gauss)
H	1/2	5.5857	2.79285	42.578	-	1420.405 752	2.002284	507
$^{2}\mathrm{H}$	1	0.8574	0.85744	6.536	0.0029	327.384 353	2.002284	175
$^{3}\mathrm{He}$	1/2	−4.2550	−2.12750	−32.434	-	-		-
$^{4}\mathrm{He}$	0	-	-	-	-	-		-
$^{6}\mathrm{Li}$	1	0.8226	0.82257	6.270	−0.0008	152.136 841	2.002301	82
$^{7}\mathrm{Li}$	3/2	2.1710	3.25646	16.548	−0.041	401.752 043	2.002301	287
$^{9}\mathrm{Be}$	3/2	−0.7852	−1.1778	−5.9853	0.053	-		-
$^{20,22}\mathrm{Ne}$	0	-	-	-	-	-		-
$^{21}\mathrm{Ne}$	3/2	−0.4412	−0.66180	−3.3631	0.103	-		-
$^{23}\mathrm{Na}$	3/2	1.4784	2.21766	11.270	0.101	885.813 064	2.002296	633
$^{24,26}\mathrm{Mg}$	0	-	-	-	-	-		-
$^{25}\mathrm{Mg}$	5/2	−0.3422	−0.85545	−2.6083	0.201	-		-
$^{36,38,40}\mathrm{Ar}$	0	-	-	-	-	-		-
$^{39}\mathrm{K}$	3/2	0.2601	0.39147	1.989	0.049	230.859 860	2.002294	165
$^{40}\mathrm{K}$	4	−0.3245	−1.29810	−2.474	−0.061	−285.730 800	2.002294	−459
$^{41}\mathrm{K}$	3/2	0.1433	0.21489	1.092	0.060	127.006 935	2.002294	90
$^{40,\cdots,48}_{\mathrm{even}}\mathrm{Ca}$	0	-	-	-	-	-		-
$^{43}\mathrm{Ca}$	7/2	−0.3765	−1.31764	−2.8697	−0.049	−806.402 072	2.002257	−1151
$^{78,\cdots,86}_{\mathrm{even}}\mathrm{Kr}$	0	-	-	-	-	-		-
$^{83}\mathrm{Kr}$	9/2	−0.215 7	−0.97067	−1.6442	0.253	-		-
$^{85}\mathrm{Rb}$	5/2	0.5413	1.35335	4.126	0.23	1011.910 813	2.002331	1084
$^{87}\mathrm{Rb}$	3/2	1.8345	2.75182	13.984	0.127	3417.341 306	2.002331	2441
$^{84,86,88}\mathrm{Sr}$	0	-	-	-	-	-		-
$^{87}\mathrm{Sr}$	9/2	−0.2430	−1.09360	−1.8524	0.335	-		-
$^{133}\mathrm{Cs}$	7/2	0.7380	2.58291	5.625	−0.004	2298.157 943	2.002540	3284
$^{130,\cdots,138}_{\mathrm{even}}\mathrm{Ba}$	0	-	-	-	-	-		-
$^{135}\mathrm{Ba}$	3/2	0.5586	0.837943	4.258	0.160	-		-
$^{137}\mathrm{Ba}$	3/2	0.6249	0.937365	4.634	0.245	-		-

As we do not enter into the internal structure of the nucleus we simply treat γ_I as an empirical value. Recalling the discussion of Section 4.3.4 we infer that the torque on the nuclear spin is given by

$$d\mathbf{I}(t)/dt = \boldsymbol{\mu}_I \times \mathbf{B}. \tag{5.6}$$

This torque gives rise to precession of $\mathbf{I}$ about $\mathbf{B}$ at angular frequency

$$\boldsymbol{\omega}_I = -\gamma_I \mathbf{B}. \tag{5.7}$$

For a free proton the gyromagnetic ratio is given by $\gamma_p = g_p(e/2m_p) = g_p\mu_N/\hbar$, with $\gamma_p/2\pi \simeq 4.26$ kHz/Gauss and $g_p \simeq 5.5857$ being the proton g factor. Likewise, for the deuteron $g_d \simeq 0.8574$ and we calculate with $\gamma_d = g_d\mu_N/\hbar$ for the gyromagnetic ratio $\gamma_d/2\pi \simeq 0.653\,59$ kHz/Gauss. For some selected isotopes the values are tabulated in Table 5.1. Note that the ratio of the *nuclear magneton* and *Bohr magneton* is given by the electron-to-proton mass ratio, $\mu_N/\mu_B = m_e/m_p \approx 1/1836$. Therefore, the nuclear-spin Zeeman coupling is typically three orders of magnitude weaker than the electronic-spin Zeeman coupling introduced in Section 4.3.5. As a result this coupling only weakly perturbs the atomic fine structure discussed in Chapter 4. Substituting Eq. (5.2) into the

Zeeman Hamiltonian (5.1) the nuclear Zeeman term takes the form

$$\mathcal{H}_Z = -\gamma_I \mathbf{I}\cdot\mathbf{B} = -g_I\mu_N B\,(I_z/\hbar)\,, \tag{5.8}$$

where the z direction is again chosen along the direction of the **B** field. This Hamiltonian expresses the coupling of **I** to the **B** field (*nuclear Zeeman coupling*).

We analyze the level splitting by the nuclear spin Zeeman coupling in the same way as we did for the electron spin Zeeman coupling using ordinary first-order perturbation theory.[1] For the nuclear spin Zeeman term this is done in the hydrogenic basis further extended by a spin sector $\{|nlm_lsm_s\rangle\}\rightarrow\{|nlm_lsm_sIm_I\rangle\}$, $\Delta E_Z = \langle nlm_lsm_sIm_I|\mathcal{H}_Z|nlm_lsm_sIm_I\rangle$. Since $\mathbf{I}^2$ and I_z commute with $\mathcal{H}_0$, $\mathbf{L}^2$, L_z, $\mathbf{S}^2$ and S_z the nuclear spin Zeeman operator acts in a $2I+1$ dimensional sector of Hilbert space and is *diagonal* in the spin basis $\{|Im_I\rangle\}$; i.e., I and m_I are good quantum numbers. The nuclear-Zeeman energy shift is given by

$$\Delta E_Z = -\gamma_I B\langle I,m_I|I_z|I,m_I\rangle = -g_I\mu_N m_I B. \tag{5.9}$$

5.1.2 Total atomic angular momentum

As long as we may neglect the interaction with the environment (i.e., in zero field) the atom represents a closed mechanical system and the *total angular momentum* is a conserved quantity. In the presence of hyperfine interaction this quantity is denoted by **F**. In hydrogen-like atoms it results from the contribution of three angular momenta,

$$\mathbf{F} = \mathbf{L}+\mathbf{S}+\mathbf{I}. \tag{5.10}$$

Since the addition of three angular momenta can be done in different ways it is an important simplifying factor that the coupling of the nuclear spin to the other angular momenta is weak, much weaker than the coupling between **L** and **S**. Therefore, we may (and shall) presume that the nuclear spin couples only to the total electronic angular momentum $\mathbf{J}=\mathbf{L}+\mathbf{S}$ rather than to **L** and **S** separately. In this way the hyperfine structure of the atom can be obtained by considering the coupling of only two angular momenta,[2]

$$\mathbf{F} = \mathbf{J}+\mathbf{I}. \tag{5.11}$$

Starting from this expression it is straightforward to show that **F** obeys commutation relations of the type (1.29),

$$[F_x,F_y] = i\hbar F_z,\ [F_y,F_z] = i\hbar F_x \text{ and } [F_z,F_x] = i\hbar F_y. \tag{5.12}$$

Therefore, **F** satisfies all properties of quantized angular momenta,

$$\mathbf{F}^2\,|jIF,m_F\rangle = F(F+1)\hbar^2\,|jIF,m_F\rangle \tag{5.13}$$

$$F_z\,|jIF,m_F\rangle = m_F\hbar\,|jIF,m_F\rangle\,, \tag{5.14}$$

with m_F restricted to the interval

$$-F\le m_F\le F. \tag{5.15}$$

In many cases the specification of j and I is irrelevant because their value is obvious from the context. In such cases these quantum numbers are usually omitted in the notation, writing simply $|F,m_F\rangle$ rather than $|jIF,m_F\rangle$. For given values of j and I the value of F can vary from $|j-I|$ to $j+I$ in steps of unit angular momentum in accordance with the *triangle inequality*

$$|j-I|\le F\le j+I. \tag{5.16}$$

---

[1] Pure nuclear spin Zeeman splitting is observed in isotopes of closed shell atoms (e.g., the inert gases). In one-electron atoms nuclear spin Zeeman coupling does not occur without hyperfine interaction (see later in this chapter).

[2] In heavy many-electron atoms the identity $\mathbf{J}=\mathbf{L}+\mathbf{S}$ is *not* satisfied (see jj coupling - Section 9.5.3). However, irrespective of the detailed coupling scheme, **J** remains well defined and Eq. (5.11) remains valid in zero field.

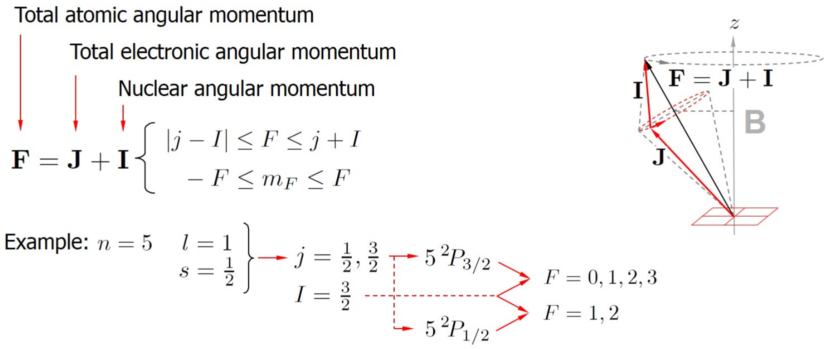


Figure 5.1: Addition of angular momenta in the presence of nuclear spin, presuming $\mathbf{L}\cdot\mathbf{S}$ coupling at the fine-structure level. The precession along the z axis occurs when the symmetry is broken by a non-zero magnetic field in this direction.

Note that F can become zero (unlike j in the case of the fine structure of hydrogen-like atoms). This happens for $j = I$. The actions of the shift operators $F_\pm$ are given by Eq. (3.24),

$$F_\pm \left|F, m_F\right\rangle = \sqrt{F\,(F+1) - m_F(m_F \pm 1)}\hbar\left|F, m_F \pm 1\right\rangle. \tag{5.17}$$

Since j, I and F are quantized the coupling of $\mathbf{I}$ and $\mathbf{J}$ cannot exist without their joint precession about $\mathbf{F}$ as is illustrated in Fig. 5.1. In view of this precession $\mathbf{I}$ and $\mathbf{J}$ are clearly not conserved. The conservation of $\mathbf{F}$ is lost when a nonzero magnetic field $\mathbf{B}$ breaks the symmetry. This gives rise to a precession of $\mathbf{F}$ about the field direction as is also indicated in Fig. 5.1.

The operator $\mathbf{F}$ acts in a $(2F+1)$-dimensional subspace of the full $(2j+1)\,(2I+1)$-dimensional product space defined by the so-called *uncoupled basis* $\{|jm_jIm_I\rangle \equiv |jm_j\rangle \otimes |Im_I\rangle\}$. Together the $(2F+1)$-dimensional subspaces of all possible values of F allowed by the triangular inequality provide the so-called *coupled basis* $\{|jIFm_F\rangle\}$ of the same $(2j+1)\,(2I+1)$-dimensional product space. Since F_z commutes with both J_z and I_z we know that these operators are diagonal in both bases and, since $F_z = J_z + I_z$, the magnetic quantum numbers are related by

$$m_F = m_j + m_I. \tag{5.18}$$

Note that this is the selection rule (3.90).

5.2 Magnetic hyperfine interaction

5.2.1 Introduction

The magnetic hyperfine interaction finds its origin in the coupling of the magnetic moment of the nucleus to the degrees of freedom of the electron and to the external magnetic field $\mathbf{B}$. Approximating the nucleus by a point dipole, $\boldsymbol{\mu}_I$, its contribution to the vector potential is given by

$$\mathbf{A}_I\left(\mathbf{r}\right) = \frac{\mu_0}{4\pi}\frac{\boldsymbol{\mu}_I \times \mathbf{r}}{r^3}, \tag{5.19}$$

where $\mathbf{r}$ is the usual operator for the position of the electron relative to the nucleus (see Fig. 5.2). The corresponding magnetic field is the nuclear magnetic dipole field (see Problem 5.1),

$$\mathbf{B}_I\left(\mathbf{r}\right) = \nabla \times \mathbf{A}_I\left(\mathbf{r}\right) = \frac{\mu_0}{4\pi}\frac{1}{r^3}\left[3\left(\boldsymbol{\mu}_I \cdot \hat{\mathbf{r}}\right)\hat{\mathbf{r}} - \boldsymbol{\mu}_I\right] + \tfrac{2}{3}\mu_0\boldsymbol{\mu}_I\delta\left(\mathbf{r}\right). \tag{5.20}$$

This field consists of two contributions: the first term is the ordinary nuclear dipole field, observable *outside* the nucleus; the second term accounts for the field *inside* the nucleus.

In the Coulomb gauge, $\nabla\cdot\mathbf{A}=0$, the atomic Hamiltonian, including hyperfine interaction, takes the form

$$\mathcal{H}=-\frac{\hbar^2}{2m_r}\Delta+\mathcal{V}(r)+\mathcal{H}_r-\frac{i\hbar e}{m_r}\mathbf{A}\cdot\boldsymbol{\nabla}+\frac{e^2}{2m_r}\mathbf{A}^2-\boldsymbol{\mu}_s\cdot(\mathbf{B}_{LS}+\mathbf{B}_I+\mathbf{B})-\boldsymbol{\mu}_I\cdot\mathbf{B}. \tag{5.21}$$

The latter two terms are related to the postulated spins of electron and nucleus. They represent the interaction of $\boldsymbol{\mu}_s$ with the spin-orbit field ($\mathbf{B}_{LS}$), the nuclear dipole field ($\mathbf{B}_I$) and the external field (spin Zeeman term), and further the interaction of $\boldsymbol{\mu}_I$ with the external field (nuclear Zeeman term). The $\mathbf{A}\cdot\boldsymbol{\nabla}$ term represents the sum of the orbital Zeeman coupling (coupling of the orbital motion of the electron to the external field $\mathbf{B}$ - see Section 4.2) and the nuclear-spin-orbit coupling (coupling of the orbital motion to the nuclear field $\mathbf{B}_I$). Separating the principal structure and fine-structure terms from the rest of the Hamiltonian, Eq. (5.21) can be expressed as the sum of 5 contributions

$$\mathcal{H}=\mathcal{H}_0+\mathcal{H}_r+\mathcal{H}_{LS}+\mathcal{H}_{\rm hfs}+\mathcal{H}_Z, \tag{5.22}$$

where $\mathcal{H}_0$ is the Schrödinger Hamiltonian, H_r the relativistic correction term, $\mathcal{H}_{LS}$ the spin-orbit Hamiltonian (4.117),

$$\mathcal{H}_Z=-(\boldsymbol{\mu}_L+\boldsymbol{\mu}_s+\boldsymbol{\mu}_I)\cdot\mathbf{B} \tag{5.23}$$

the Zeeman Hamiltonian and

$$\mathcal{H}_{\rm hfs}=-\frac{i\hbar e}{m_r}\mathbf{A}_I\cdot\boldsymbol{\nabla}+\frac{e^2}{2m_r}\mathbf{A}_I^2-\boldsymbol{\mu}_s\cdot\mathbf{B}_I \tag{5.24}$$

the magnetic hyperfine contribution. The $\mathbf{A}_I^2$ term may be neglected because its contribution is much smaller than the contribution of the terms linear in $\mathbf{A}_I$. A similar approximation was made in the case of the fine structure; at a distance of the Bohr radius from the nucleus the nuclear field in hydrogen is far too weak to induce any appreciable magnetic polarization, $\mathbf{B}_I(a_0)\simeq10^{-2}$ T.

Problem 5.1. The magnetic field of a point dipole is given by

$$\mathbf{B}_I=\frac{\mu_0}{4\pi}\frac{1}{r^3}[3(\boldsymbol{\mu}_I\cdot\hat{\mathbf{r}})\hat{\mathbf{r}}-\boldsymbol{\mu}_I]+\tfrac{2}{3}\mu_0\boldsymbol{\mu}_I\delta(\mathbf{r}).$$

Derive this expression starting from the expression for the vector potential of the dipole field outside a homogeneous magnetized sphere, $\mathbf{A}_I=(\mu_0/4\pi)(\boldsymbol{\mu}_I\times\mathbf{r})/r^3$, and the vector potential $\mathbf{A}=\frac{1}{2}(\mathbf{B}\times\mathbf{r})$ of a homogeneous field inside the sphere (see Fig.5.2).

Solution. (*a*) Using Eq. (N.9) the magnetic field $\mathbf{B}_I=\boldsymbol{\nabla}\times\mathbf{A}_I$ of the nuclear magnetic moment outside the nucleus can be written as

$$(4\pi/\mu_0)\,\mathbf{B}_I=\boldsymbol{\nabla}\times(\boldsymbol{\mu}_I\times\mathbf{r})/r^3=-(\boldsymbol{\mu}_I\cdot\boldsymbol{\nabla})\mathbf{r}/r^3+\boldsymbol{\mu}_I(\boldsymbol{\nabla}\cdot\mathbf{r}/r^3).$$

This can be reformulated in the form

$$(4\pi/\mu_0)\,\mathbf{B}_I=-r^{-3}(\boldsymbol{\mu}_I\cdot\boldsymbol{\nabla})\mathbf{r}-\mathbf{r}\left(\boldsymbol{\mu}_I\cdot\boldsymbol{\nabla}r^{-3}\right)+r^{-3}\boldsymbol{\mu}_I(\boldsymbol{\nabla}\cdot\mathbf{r})+\boldsymbol{\mu}_I\left(\mathbf{r}\cdot\boldsymbol{\nabla}r^{-3}\right).$$

Using Eqs. (N.13), (N.11) and (N.18) we obtain after collecting terms the first two terms of the expression for the dipole field.
(*b*) At the surface of a homogeneously magnetized sphere of radius R the vector potential becomes

$$\mathbf{A}|_{r=R}=\tfrac{1}{2}(\mathbf{B}\times\mathbf{r})|_{r=R}=(\mu_0/4\pi)(\boldsymbol{\mu}_I\times\mathbf{r})/r^3|_{r=R},$$

which implies $\mathbf{B}=(\mu_0/2\pi)\,\mu_0\boldsymbol{\mu}_I/R^3$. Integrating the field over the volume of the sphere we obtain

$$\int_0^{(4/3)\pi R^3}\mathbf{B}(\mathbf{r})d\mathbf{r}=\tfrac{2}{3}\mu_0\boldsymbol{\mu}_I\,.$$

The same result is obtained for a point dipole by choosing $\mathbf{B}=\frac{2}{3}\mu_0\boldsymbol{\mu}_I\delta(\mathbf{r})$. □

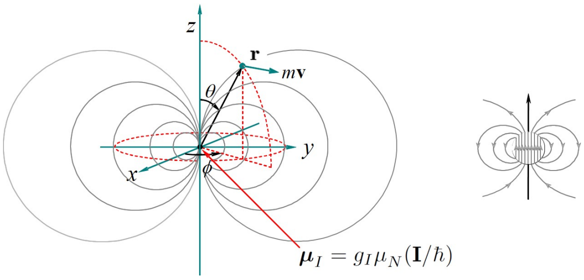


Figure 5.2: *Left*: Electron moving in the magnetic field of the nuclear magnetic dipole. *Right*: Not only the dipole field outside the nucleus but also the field inside the nucleus contributes to the hyperfine interaction. The latter contribution is estimated in the text by treating the nucleus as a homogeneously magnetized sphere of vanishing size (point dipole).

5.2.2 Three contributions to the magnetic hyperfine interaction in zero field

Substituting Eq. (5.19) into the Hamiltonian (5.24) we find that the magnetic hyperfine interaction can be written as the sum of three terms,

$$\mathcal{H}_{\mathrm{hfs}} = \mathcal{H}_{IL} + \mathcal{H}_{dd} + \mathcal{H}_{\mathrm{Fermi}}\,, \tag{5.25}$$

which are sufficient to calculate the level shifts in hydrogen-like atoms.

- the first term follows from $\mathcal{H}_{IL} = -(i\hbar e/m_r)\mathbf{A}_I \cdot \boldsymbol{\nabla}$ and represents the *nuclear-spin-orbit interaction* of the moving electron with the magnetic field of the nucleus. Substituting Eq. (5.19) and using Eqs. (N.1), (4.54), (4.81) and (5.2) this term takes the form [1]

  $$\mathcal{H}_{IL} = 2\frac{g_L}{g_e}\frac{\mu_0}{4\pi}\frac{\gamma_I|\gamma_e|}{r^3}\mathbf{I}\cdot\mathbf{L} \simeq \frac{\mu_0}{4\pi}\frac{\gamma_I|\gamma_e|}{r^3}\mathbf{I}\cdot\mathbf{L}, \tag{5.26}$$

  where we recalled that $2g_L/g_e \simeq 1$.

- the second and third term both follow from $-\boldsymbol{\mu}_s \cdot \mathbf{B}_I(\mathbf{r})$. Substituting Eq. (5.20) and using Eqs. (4.80) and (5.2) we obtain the *magnetic dipole-dipole interaction*

  $$\mathcal{H}_{dd} = \frac{\mu_0}{4\pi}\frac{\gamma_I|\gamma_e|}{r^3}\left[3\left(\mathbf{I}\cdot\hat{\mathbf{r}}\right)\left(\mathbf{S}\cdot\hat{\mathbf{r}}\right) - \mathbf{I}\cdot\mathbf{S}\right] \tag{5.27a}$$

- as well as the *Fermi contact interaction*

  $$\mathcal{H}_{\mathrm{Fermi}} = \tfrac{2}{3}\mu_0\gamma_I|\gamma_e|\delta\left(\mathbf{r}\right)\mathbf{I}\cdot\mathbf{S}. \tag{5.28}$$

5.2.3 Magnetic dipole-dipole interaction using spherical tensor operators

Note that the operators $\mathcal{H}_{dd}$ and $\mathcal{H}_{\mathrm{Fermi}}$ are symmetric in the role of the nuclear spin and the electron spin. In other words, we can equally well can say that the electron spin probes the magnetic field of the nucleus or that the nucleus probes the magnetic field of the electron spin. In view of the importance of the interaction between two magnetic dipoles we derive - as an intermezzo - an expression for this interaction in which the central symmetry of the atom is optimally exploited. For two magnetic dipoles, denoted by $\boldsymbol{\mu}_I = \gamma_I\mathbf{I}$ and $\boldsymbol{\mu}_S = \gamma_S\mathbf{S}$, the dipole-dipole interaction takes the following well-known form

$$\mathcal{H}_{dd} = \frac{\mu_0}{4\pi}\frac{1}{r^3}\left[3\left(\boldsymbol{\mu}_I\cdot\hat{\mathbf{r}}\right)\left(\boldsymbol{\mu}_S\cdot\hat{\mathbf{r}}\right) - \boldsymbol{\mu}_I\cdot\boldsymbol{\mu}_S\right]. \tag{5.29}$$

---

[1] Note that $-i\hbar\left(\boldsymbol{\mu}_I\times\mathbf{r}\right)\cdot\boldsymbol{\nabla} = -i\hbar\boldsymbol{\mu}_I\cdot\left(\mathbf{r}\times\boldsymbol{\nabla}\right) = \boldsymbol{\mu}_I\cdot\mathbf{L}$.

The difficulty in evaluating this operator is related to the inner products of the dipoles with the radial direction operator. These inner products complicate the separation of the position dependence from the dependence on angular momentum. For a system of central symmetry this separation is best achieved by decomposition of $\mathcal{H}_{dd}$ in terms of the standard components T_2^m of a rank 2 spherical tensor (see Problem 5.2),

$$\mathcal{H}_{dd} = \frac{\mu_0}{4\pi}\frac{\gamma_S\gamma_I}{r^3}\sqrt{\frac{4\pi}{5}}\sum_{m=-2}^{2}(-1)^m Y_2^{-m}(\hat{\mathbf{r}})T_2^m. \tag{5.30}$$

The operators T_2^m can be expressed in terms of the standard components $I_{\pm1}, I_0$ and $S_{\pm1}, S_0$ of the constituent rank 1 tensors (the vector operators $\mathbf{I}$ and $\mathbf{S}$). Actually, in order to facilitate evaluation of the matrix elements of $\mathcal{H}_{dd}$ in the $\{|sm_s, Im_I\rangle\}$ representation, we do not use the standard components $I_{\pm1}, I_0$ and $S_{\pm1}, S_0$ but rather I_z and S_z and the shift operators $I_\pm$ and $S_\pm$,

$$T_2^0 = 2I_zS_z - \tfrac{1}{2}\left(I_-S_+ + I_+S_-\right), \tag{5.31a}$$

$$T_2^{\pm1} = \mp\left(I_\pm S_z + I_zS_\pm\right)\sqrt{3/2}, \tag{5.31b}$$

$$T_2^{\pm2} = I_\pm S_\pm\sqrt{3/2}. \tag{5.31c}$$

Note that the T_2^m satisfy the commutation relations (K.3), as required for spherical tensor operators. The inclusion of the phase factor $(-1)^m$ in Eq. (5.30) assures that the T_2^m transform like the Y_2^m, hence satisfy the Condon and Shortley phase convention.

Problem 5.2. Verify the tensor decomposition (5.30).

Solution. We first we express the inner products $\mathbf{I}\cdot\hat{\mathbf{r}}$ and $\mathbf{S}\cdot\hat{\mathbf{r}}$ in spherical coordinates (θ and ϕ):

$$\mathbf{I}\cdot\hat{\mathbf{r}} = I_z\cos\theta + I_x\sin\theta\cos\phi + I_y\sin\theta\sin\phi = I_z\cos\theta + \tfrac{1}{2}I_-\sin\theta e^{i\phi} + \tfrac{1}{2}I_+\sin\theta e^{-i\phi}.$$

Replacing $\mathbf{I}$ by $\mathbf{S}$ we obtain $\mathbf{S}\cdot\hat{\mathbf{r}}$. For the product of these inner products we find

$$\begin{aligned}(\mathbf{I}\cdot\hat{\mathbf{r}})\,(\mathbf{S}\cdot\hat{\mathbf{r}}) = {} & I_zS_z\cos^2\theta + \tfrac{1}{4}(I_-S_+ + I_+S_-)\sin^2\theta \\ & + \tfrac{1}{2}(I_zS_- + I_-S_z)\sin\theta\cos\theta\, e^{i\phi} + \tfrac{1}{2}(I_zS_+ + I_+S_z)\sin\theta\cos\theta\, e^{-i\phi} \\ & + \tfrac{1}{4}I_-S_-\sin^2\theta e^{2i\phi} + \tfrac{1}{4}I_+S_+\sin^2\theta e^{-2i\phi}.\end{aligned}$$

Recalling the inner product rule (3.109b) we have $\mathbf{I}\cdot\mathbf{S} = I_zS_z + \frac{1}{2}\left(I_+S_- + I_-S_+\right)$ and recognizing the spherical harmonics $Y_2^{\pm1}(\hat{\mathbf{r}})$ and $Y_2^{\pm1}(\hat{\mathbf{r}})$ we obtain

$$\begin{aligned}3\,(\mathbf{I}\cdot\hat{\mathbf{r}})\,(\mathbf{S}\cdot\hat{\mathbf{r}}) - \mathbf{I}\cdot\mathbf{S} = {} & I_zS_z(3\cos^2\theta - 1) + \tfrac{1}{4}\left(I_+S_- + I_-S_+\right)(3\sin^2\theta - 2) \\ & + \sqrt{4\pi/5}\left[-\left(I_zS_- + I_-S_z\right)Y_2^1(\hat{\mathbf{r}}) + \left(I_zS_+ + I_+S_z\right)Y_2^{-1}(\hat{\mathbf{r}})\right]\sqrt{3/2} \\ & + \sqrt{4\pi/5}\left[I_-S_-Y_2^2(\hat{\mathbf{r}}) + I_+S_+Y_2^{-2}(\hat{\mathbf{r}})\right]\sqrt{3/2}.\end{aligned}$$

At this point we identify the $(-1)^qT_2^q$ with the prefactors of $Y_2^{-q}(\hat{\mathbf{r}})$ for $q = \pm1, \pm2$. Since $2 - 3\sin^2\theta = 3\cos^2\theta - 1 = 2\sqrt{4\pi/5}\,Y_2^0(\hat{\mathbf{r}})$ we finally identify T_2^0 with the prefactor of $Y_2^0(\hat{\mathbf{r}})$. □

5.3 Hyperfine interaction in zero field

5.3.1 Introduction - effective Hamiltonian

In the previous section we identified three mechanisms through which the angular momentum of the electron can couple to that of the nucleus. Before continuing with the derivation of the associated coupling coefficients and hyperfine shifts we summarize the main results to be obtained later in this chapter. We shall establish that the three contributions of the *magnetic* hyperfine interaction add up to a *single* effective Hamiltonian,

$$\mathcal{H}_{\text{hfs}} = (\mathcal{A}_{\text{hfs}}/\hbar^2)\,\mathbf{I}\cdot\mathbf{J}, \tag{5.32}$$

where $\mathcal{A}_{\text{hfs}}$ is called the *hyperfine coupling constant.* Note that this Hamiltonian is of the same general form as the spin-orbit Hamiltonian. Within the framework of first-order perturbation theory $\mathcal{H}_{\text{hfs}}$ acts in the $(2j+1)(2I+1)$-degenerate subspace of the fine-structure level $n\,^{2s+1}L_j$. Within this approximation j, I, F and m_F are good quantum numbers, with $-F \leq m_F \leq F$ (note that $\mathbf{J}^2$, $\mathbf{I}^2$, $\mathbf{F}^2$ and F_z commute with $\mathcal{H}_{\text{hfs}}$). For the electronic ground state the approximation is excellent because the hyperfine splitting is small as compared to the splitting of the principal structure. For the electronically excited states the approximation is only good as long as the fine-structure splitting is sufficiently large. For the time being we distinguish two cases:

- *Hydrogen-like atoms:* For zero orbital angular momentum ($l = 0$) only the *Fermi contact interaction* contributes. In this case we have $\mathbf{J} = \mathbf{S}$ and one uses the lower case notation for the coupling constant, $\mathcal{A}_{\text{hfs}} \to a_{\text{hfs}}$. In this notation Eq. (5.32) takes the form

  $$\mathcal{H}_{\text{hfs}} = (a_{\text{hfs}}/\hbar^2)\,\mathbf{I}\cdot\mathbf{S}, \tag{5.33}$$

  where the coupling constant for the $n\,^{2s+1}L_j$ fine-structure level is given by

  $$a_{\text{hfs}}\left(n\,^{2s+1}L_j\right) = \frac{\mu_0}{4\pi}\frac{2}{3}\gamma_I|\gamma_e|\frac{\hbar^2}{a^3}\tilde{R}^2_{ns}(0) \qquad (l=0)\,. \tag{5.34}$$

- *Hydrogenic atoms:* For $l > 0$ the Fermi contact interaction does not contribute and the form (5.32) follows from the addition of $\mathcal{H}_{IL}$ and $\mathcal{H}_{dd}$. In this case the coupling constant is written in the upper case notation. For the $n\,^{2s+1}L_j$ fine-structure level it is given by

  $$\mathcal{A}_{\text{hfs}}\left(n\,^{2s+1}L_j\right) = \frac{\mu_0}{4\pi}\gamma_I|\gamma_e|\frac{\hbar^2}{a^3}\langle nl||\rho^{-3}||nl\rangle\frac{l(l+1)}{j(j+1)} \qquad (l>0)\,, \tag{5.35}$$

  where $\rho = r/a$ is the reduced distance to the nucleus, $a = (m_e/m_r)a_0$ being the atomic unit of length and m_r the reduced mass of the electron.

Note that $\mathcal{H}_{\text{hfs}}$ is an *effective* Hamiltonian because the radial dependence of the hyperfine interaction is averaged over the orbital wavefunction, thus neglecting the coupling to other manifolds with different principal quantum number. Therefore, it is valid only to first-order in perturbation theory. For reasons of compactness of notation in most cases we do not specify the fine-structure Term but simply write $\mathcal{A}_{\text{hfs}}$ (for $l > 0$) or a_{hfs} (for $l = 0$) for the coupling constant. Importantly, the expressions for $\mathcal{A}_{\text{hfs}}$ and a_{hfs} are both proportional to γ_I. As all other quantities in the definition of the coupling constant are positive the hyperfine shift is *positive* for atoms with *positive* nuclear gyromagnetic ratio and *negative* for atoms where this ratio is *negative.*[1] For selected isotopes the observed ground state coupling constants a_{hfs} are listed in Table 5.1.

To finish this introduction we compare the hyperfine coupling with fine-structure coupling in hydrogenic atoms. Rewriting Eq. (4.135) with the aid of Eq. (4.119) the fine-structure coupling takes the form

$$\zeta_{nl} = \frac{\mu_0}{4\pi}2\gamma_L^2\frac{\hbar^2}{a^3}Z\,\langle nl||\rho^{-3}||nl\rangle. \tag{5.36}$$

---

[1] In many-electron atoms $\mathcal{A}_{\text{hfs}}$ can become negative for positive g_I (see Fig. 5.3). This arises as a result of polarization of the electron core (see Section 9.3.2). Core polarization also affects the value of the ground state hyperfine splitting of alkali atoms but is not strong enough to affect its sign. Hence, the *ground state* hyperfine splitting of the alkali atoms is hydrogen like. For this reason these ground states are discussed in the context of the present chapter.

Comparing this expression with Eq. (5.35) we find for the ratio of the coupling constants

$$\frac{\mathcal{A}_{\text{hfs}}\left(n\,^{2s+1}L_j\right)}{\zeta_{nl}} \simeq \frac{\gamma_I}{|\gamma_L|}\frac{l(l+1)}{j(j+1)}\frac{1}{Z} \qquad (l>0)\,. \tag{5.37}$$

Here we approximated $\gamma_e \simeq 2\gamma_L$. Since $\gamma_I/|\gamma_L| \simeq g_I m_e/m_p$ we find that the hyperfine coupling constant is typically a factor 1000 smaller than the corresponding constant for the fine structure.

5.3.2 Zero-field hyperfine shift in hydrogen-like atoms - the case $l=0$

We start the derivation of *magnetic* hyperfine shifts with hydrogen-like atoms of zero orbital angular momentum (s states) in the absence of external fields. This case is of considerable importance because all hydrogen-like atoms have $l=0$ ground states. Note that in the absence of orbital angular momentum also the spin-orbit coupling is absent, which means that only the electronic and nuclear spins contribute to the total angular momentum,

$$\mathbf{F} = \mathbf{S} + \mathbf{I}. \tag{5.38}$$

As the Schrödinger Hamiltonian, $\mathcal{H}_0$, commutes with $\mathbf{I}\cdot\mathbf{S}$ the orbital wavefunctions separate from the spin part, suggesting the basis $\{|nlm_l; sIFm_F\rangle\}$ to calculate the perturbation. Interestingly, as we shall see, to obtain Eq. (5.34) the Fermi contact term (5.28) suffices: $\Delta E_{\text{hfs}} = \Delta E_{\text{Fermi}}$. Factoring out the orbital matrix element $\langle nlm_l|\delta(\mathbf{r})|nlm_l\rangle = \langle n00|\delta(\mathbf{r})|n00\rangle$ and recalling the relation $\delta(r) = 4\pi\delta(\mathbf{r})r^2$, we obtain to first order in perturbation theory

$$\Delta E_{\text{hfs}}(n\,^{2s+1}S_{1/2}) = (a_{\text{hfs}}/\hbar^2)\langle sIFm_F|\mathbf{I}\cdot\mathbf{S}|sIFm_F\rangle, \tag{5.39}$$

where the hyperfine coupling constant is given by

$$\begin{aligned} a_{\text{hfs}}(n\,^{2s+1}S_{1/2}) &= \tfrac{2}{3}\mu_0\gamma_I|\gamma_e|\hbar^2\langle n00|\delta(\mathbf{r})|n00\rangle \\ &= \frac{\mu_0\gamma_I|\gamma_e|\hbar^2}{a^3}\frac{2}{3}\int\delta(\boldsymbol{\rho})\,\tilde{R}^2_{ns}(\rho)\left|Y^0_0(\hat{\boldsymbol{\rho}})\right|^2 d\boldsymbol{\rho} \\ &= \frac{\mu_0}{4\pi}\frac{\gamma_I|\gamma_e|\hbar^2}{a^3}\frac{2}{3}\tilde{R}^2_{ns}(0) \qquad (l=0) \end{aligned} \tag{5.40}$$

with $\rho = r/a$ being the radial distance in atomic units $a = (m_e/m_r)a_0$ (see Section 2.1). This completes the derivation of Eq. (5.34). A special feature of the delta function is that it does not couple states with different principal quantum number. Hence, with regard to contact interactions n is a good quantum number.

5.3.2.1 Fermi contact suffices for $l=0$

What about the other terms of the hyperfine Hamiltonian? They turn out to vanish. We shall demonstrate that for s states neither $\mathcal{H}_{IL}$ nor $\mathcal{H}_{dd}$ contribute to the hyperfine shift. For this purpose we choose the basis $\{|nlm_l; sIFm_F\rangle\}$ in which $\mathbf{L}$ is decoupled from $\mathbf{F}=\mathbf{S}+\mathbf{I}$. Note that this choice of basis is both possible and logical for subspaces in which orbital angular momentum is absent; i.e., for $l=m_l=0$. In this basis the absence of an energy shift is straightforward for the nuclear-spin-orbit interaction $\mathcal{H}_{IL}$,

$$\Delta E_{IL}(n\,^{2s+1}L_j) = 2\frac{g_L}{g_e}\frac{\mu_0}{4\pi}\gamma_I|\gamma_e|\langle n00; sIFm_F|\frac{\mathbf{I}\cdot\mathbf{L}}{r^3}|n00; sIFm_F\rangle = 0, \tag{5.41}$$

because all components of $\mathbf{L}$ yield zero when acting on a state of zero orbital angular momentum.

A bit more effort is required to show that the magnetic dipole-dipole operator ($\mathcal{H}_{dd}$) does not contribute to the $l = 0$ hyperfine shift. We shall find that its contribution averages to zero for *spherical* orbitals. For this purpose we consider the first-order shift in the $\{|n00; sIFm_F\rangle\}$ representation,

$$\Delta E_{dd}(n\,^{2s+1}L_j) = \frac{\mu_0}{4\pi}\gamma_I|\gamma_e|\langle n00; sIFm_F|\frac{1}{r^3}\left[3\left(\mathbf{I}\cdot\hat{\mathbf{r}}\right)\left(\mathbf{S}\cdot\hat{\mathbf{r}}\right) - \mathbf{I}\cdot\mathbf{S}\right]|n00; sIFm_F\rangle. \tag{5.42}$$

To separate the position dependence from the spin dependence we express the $\mathcal{H}_{dd}$ in terms of second-order spherical harmonics - see Eq. (5.30)

$$\mathcal{H}_{dd} = \frac{\mu_0}{4\pi}\frac{\gamma_e\gamma_I}{r^3}\sum_{m=-2}^{2} T_2^m Y_2^{-m}(\hat{\mathbf{r}}), \tag{5.43}$$

where the tensor components T_2^m are defined by Eqs. (5.31). Thus we obtain

$$\Delta E_{dd}(n\,^{2s+1}L_j) = \frac{\mu_0}{4\pi}\gamma_I\gamma_e\sum_{m=-2}^{2}\langle n00|Y_2^{-m}(\hat{\mathbf{r}})/r^3|n00\rangle\langle sIFm_F|T_2^m|sIFm_F\rangle. \tag{5.44}$$

Note that because the T_2^m are independent of $\mathbf{r}$ we could factor out the $\mathbf{r}$ dependent matrix element

$$\langle n00|Y_2^m(\hat{\mathbf{r}})/r^3|n00\rangle = \int\frac{\left[R_{ns}(r)\right]^2}{r^3}Y_2^m(\hat{\mathbf{r}})|Y_0^0(\hat{\mathbf{r}})|^2 d\mathbf{r}. \tag{5.45}$$

This integral vanishes as follows by evaluating the angular part. Rather than using direct integration, which is straightforward, we demonstrate this with the aid of the general formula (L.58),

$$\int Y_2^m(\hat{\mathbf{r}})|Y_0^0(\hat{\mathbf{r}})|^2 d\hat{\mathbf{r}} = \sqrt{\frac{5}{4\pi}}\begin{pmatrix}0\,2\,0\\0\,0\,0\end{pmatrix}\begin{pmatrix}0 & 2 & 0\\0 & m & 0\end{pmatrix} = 0. \tag{5.46}$$

Here the $3j$ symbols are zero because the triangle inequality is violated. Hence, there is no need for any further integration, all matrix elements of $\mathcal{H}_{dd}$ vanish in the $\{|n00; sIFm_F\rangle\}$ representation. Note that this holds not only for the diagonal matrix elements but also for the off-diagonal matrix elements of the T_2^m. Actually, along the same lines of reasoning also the coupling to s states with different principal quantum number vanishes. This shows that also in second order a dipole-dipole contribution to the hyperfine shift of s levels is absent. Although we set out to show that $\Delta E_{dd} = 0$ to first order in perturbation theory we found that this equation holds quite rigorously. Only when the spherical shape of the s orbitals is lost (e.g., in an external electric field) the dipole-dipole shift no longer vanishes.

5.3.3 Zero-field hyperfine shift in hydrogen-like atoms - the case $l > 0$

For electrons of non-zero orbital angular momentum the Fermi contact contribution is absent because any radial wavefunction with $l > 0$ is zero at the origin: $R_{nl}(0) = 0$ for $l > 0$. Unfortunately, this does not mean that we can forget about the contact interaction for all valence electrons in non-spherical orbital states. As it turns out, in many-electron atoms, polarization of the electron core by the valence electrons give rise to an induced Fermi contact interaction which can be so strong that it affects even the sign of $\mathcal{A}_{\mathrm{hfs}}$, in particular for alkali atoms in electronically excited states. The phenomenon of core polarization is discussed in Section 9.3.2. In hydrogen, core electrons are absent and the simple rule holds: for electrons with nonzero orbital angular momentum the contact interaction vanishes and $\mathcal{H}_{IL}$ and $\mathcal{H}_{dd}$ fully determine the hyperfine shift. For this reason, the results to be obtained in the current section apply to the hydrogenic case.

So let us focus on the contributions of $\mathcal{H}_{IL}$ and $\mathcal{H}_{dd}$ to the hyperfine shift. Rather than considering these contributions independently we calculate their sum,

$$\mathcal{H}_{IL} + \mathcal{H}_{dd} = \frac{\mu_0}{4\pi}\frac{\gamma_I|\gamma_e|}{r^3}[3\,(\mathbf{I}\cdot\hat{\mathbf{r}})\,(\mathbf{S}\cdot\hat{\mathbf{r}}) + \mathbf{I}\cdot(\mathbf{L}-\mathbf{S})]. \tag{5.47}$$

Actually, we approximated $2g_L/g_e \simeq 1$, which is a good approximation since $(g_e - 2)\,m_e/M \simeq 10^{-6}$. The shift follows with

$$\Delta E_{\mathrm{hfs}}(n\,^{2s+1}L_j) = \frac{\mu_0}{4\pi}\gamma_I|\gamma_e|\langle nlsjIFm_F|\frac{1}{r^3}[3\,(\mathbf{I}\cdot\hat{\mathbf{r}})\,(\mathbf{S}\cdot\hat{\mathbf{r}}) + \mathbf{I}\cdot(\mathbf{L}-\mathbf{S})]|nlsjIFm_F\rangle. \tag{5.48}$$

To calculate the matrix elements we rewrite the equation in the form

$$\Delta E_{\mathrm{hfs}}(n\,^{2s+1}L_j) = \gamma_I\langle nlsjIFm_F|\mathbf{I}\cdot\mathbf{B}_J|nlsjIFm_F\rangle, \tag{5.49}$$

where

$$\mathbf{B}_J = \frac{\mu_0}{4\pi}\frac{|\gamma_e|}{r^3}[3\hat{\mathbf{r}}\,(\mathbf{S}\cdot\hat{\mathbf{r}}) + (\mathbf{L}-\mathbf{S})] \tag{5.50}$$

may be interpreted as the effective magnetic field generated by the electron at the position of the nucleus. Since $\mathbf{B}_J$ is a vector operator acting in the space spanned by the spin-orbit basis $\{|nlsjm_j\rangle\}$, we can apply the Wigner-Eckart theorem to its components,

$$\langle nlsjm_j|B_{J_z}\,|nlsjm_j\rangle = \langle nlsj||B_J||nlsj\rangle\langle jm_j|J_z\,|jm_j\rangle \tag{5.51a}$$

$$\langle nlsj(m_j\pm1)|B_{J_\pm}\,|nlsjm_j\rangle = \langle nlsj||B_J||nlsj\rangle\langle j(m_j\pm1)|J_\pm\,|jm_j\rangle\,. \tag{5.51b}$$

Here we have written both the diagonal and the off-diagonal matrix elements of $\mathbf{B}_J$ to emphasize that the same reduced matrix element applies to both cases (reduced matrix elements do not depend on magnetic quantum numbers). This property gives us the freedom to determine $\langle nlsj||B_J||nlsj\rangle$ by considering only *diagonal* matrix elements (irrespective of the operator being diagonal or not). To determine $\langle nlsj||B_J||nlsj\rangle$ we choose to use the diagonal matrix elements of the operator $\mathbf{J}\cdot\mathbf{B}_J$,

$$\langle nlsjm_j|\mathbf{J}\cdot\mathbf{B}_J\,|nlsjm_j\rangle = \langle nlsj||B_J||nlsj\rangle\langle jm_j|\mathbf{J}^2\,|jm_j\rangle\,. \tag{5.52}$$

As shown in Problem 5.3 the evaluation of the matrix elements yields for hydrogen-like atoms ($s=\frac{1}{2}$)

$$\langle nlsj||B_J||nlsj\rangle = \frac{\mu_0}{4\pi}|\gamma_e|\langle nl||r^{-3}||nl\rangle\frac{l(l+1)}{j(j+1)}. \tag{5.53}$$

The expression for the hyperfine shift is obtained by returning to Eq. (5.48). We temporarily decompose the matrix elements for the hyperfine shift in the $\{|Im_Ijm_j\rangle\}$ basis. This enables us to apply Eqs. (5.51),

$$\begin{aligned}\Delta E_{\mathrm{hfs}} &= \gamma_I\sum_{m_I'm_j',m_Im_j}\langle Fm_F|jm_j'Im_I'\rangle\langle nlsjm_j'Im_I'|\mathbf{I}\cdot\mathbf{B}_J|nlsjm_jIm_I\rangle\langle jm_jIm_I|Fm_F\rangle\\ &= \gamma_I\langle nlsj||B_J||nlsj\rangle\sum_{m_I'm_j',m_Im_j}\langle Fm_F|jm_j'Im_I'\rangle\langle jm_j'Im_I'|\mathbf{I}\cdot\mathbf{J}|jm_jIm_I\rangle\langle jm_jIm_I|Fm_F\rangle\\ &= \gamma_I\langle nlsj||B_J||nlsj\rangle\langle nlsjIFm_F|\mathbf{I}\cdot\mathbf{J}|nlsjIFm_F\rangle.\end{aligned} \tag{5.54}$$

Substituting the reduced matrix element (5.53) we obtain

$$\Delta E_{\mathrm{hfs}}(n\,^{2s+1}L_j) = (\mathcal{A}_{\mathrm{hfs}}/\hbar^2)\langle jIFm_F|\mathbf{I}\cdot\mathbf{J}|jIFm_F\rangle, \tag{5.55}$$

where the coupling constant is given by

$$\mathcal{A}_{\rm hfs}(n\,^{2s+1}L_j) = \frac{\mu_0}{4\pi}\gamma_I|\gamma_e|\hbar^2\langle nl||r^{-3}||nl\rangle\frac{l(l+1)}{j(j+1)}. \tag{5.56}$$

Writing the radial matrix element in dimensionless form using $\rho = r/a$, with $a = (m_e/m_r)a_0$ being the atomic unit of length, we see that Eq. (5.56) coincides exactly with the expression for the hyperfine coupling constant announced in Eq. (5.35). Since $\mathbf{I}\cdot\mathbf{J}$ is diagonal in the $\{|jIFm_F\rangle\}$ basis the perturbation theory has been reduced to ordinary perturbation theory. This makes the analysis of the hyperfine structure into a straightforward exercise (as we demonstrate later in this chapter). This being said it should be remembered that the result is based on the assumption that the hyperfine coupling is much weaker than the spin-orbit interaction. If this is not satisfied the primary coupling of $\mathbf{L}$ and $\mathbf{S}$ into $\mathbf{J}$ cannot be justified. Furthermore, since the radius operator does *not* commute with the Schrödinger Hamiltonian we should be aware that states of different principal quantum number contribute in second order.

Problem 5.3. Show that for $s = \frac{1}{2}$ the reduced matrix element of the electronic field at the nucleus is given by

$$\langle nlsj||B_J||nlsj\rangle = \frac{\mu_0}{4\pi}|\gamma_e|\langle nl||r^{-3}||nl\rangle\frac{l(l+1)}{j(j+1)}.$$

Solution. Starting from Eq. (5.52) we have

$$\begin{aligned}\langle nlsj||B_J||nlsj\rangle &= \frac{\langle nlsjm_j|\mathbf{J}\cdot\mathbf{B}_J\,|nlsjm_j\rangle}{\langle jm_j|\mathbf{J}^2\,|jm_j\rangle}\\ &= \frac{\mu_0}{4\pi}\frac{|\gamma_e|}{j(j+1)\hbar^2}\langle nlsjm_j|\frac{1}{r^3}\{3\,(\mathbf{J}\cdot\hat{\mathbf{r}})\,(\mathbf{S}\cdot\hat{\mathbf{r}})+\mathbf{J}\cdot(\mathbf{L}-\mathbf{S})\}|nlsjm_j\rangle.\end{aligned}$$

Noting that $\mathbf{J}\cdot\hat{\mathbf{r}} = \mathbf{S}\cdot\hat{\mathbf{r}}$ because $\mathbf{L}\cdot\hat{\mathbf{r}} = (\mathbf{r}\times\mathbf{p})\cdot\hat{\mathbf{r}} = 0$ we obtain, using the tensor decomposition (5.30) and retaining only the terms conserving m_j,

$$\begin{aligned}\langle nlsj||B_J||nlsj\rangle &= \frac{\mu_0}{4\pi}\frac{|\gamma_e|}{j(j+1)\hbar^2}\langle nlsjm_j|\frac{1}{r^3}\{[3\,(\mathbf{S}\cdot\hat{\mathbf{r}})^2-\mathbf{S}^2]+\mathbf{L}^2\}|nlsjm_j\rangle\\ &= \frac{\mu_0}{4\pi}\frac{|\gamma_e|}{j(j+1)\hbar^2}\langle nlsjm_j|\frac{1}{r^3}\left\{\left[4S_z^2-(S_-S_++S_+S_-)\right]Y_2^0(\hat{\mathbf{r}})+\mathbf{L}^2\right\}|nlsjm_j\rangle\\ &= \frac{\mu_0}{4\pi}\frac{|\gamma_e|}{j(j+1)\hbar^2}\langle nlsjm_j|\frac{1}{r^3}\left\{(6S_z^2-2\mathbf{S}^2)Y_2^0(\hat{\mathbf{r}})+\mathbf{L}^2\right\}|nlsjm_j\rangle.\end{aligned}$$

Decomposing the states $|nlsjm_j\rangle$ to the uncoupled basis $\{|lm_lsm_s\rangle\}$ we find, using $s=\frac{1}{2}$, that the electron-spin contribution vanishes

$$\langle lsjm_j|6S_z^2-2\mathbf{S}^2|lsjm_j\rangle = \sum_{m_lm_s}\langle lm_lsm_s|6S_z^2-2\mathbf{S}^2|lm_lsm_s\rangle\langle lm_lsm_s|lsjm_j\rangle^2 = 0.$$

Hence, factoring out the angular-independent part we obtain

$$\langle nlsj||B_J||nlsj\rangle = \frac{\mu_0}{4\pi}\frac{|\gamma_e|}{j(j+1)\hbar^2}\langle nl||r^{-3}||nl\rangle\langle lsjm_j|\mathbf{L}^2|lsjm_j\rangle,$$

which results in the requested expression when substituting $\langle lsjm_j|\mathbf{L}^2|lsjm_j\rangle = l(l+1)\hbar^2$. □

5.3.4 Magnetic hyperfine structure of hydrogen-like atoms

At this point we formulate the effective spin Hamiltonian for hydrogen-like atoms including magnetic hyperfine interaction (5.32),

$$\mathcal{H} = \mathcal{H}_0+\mathcal{H}_r+(\zeta_{nl}/\hbar^2)\,\mathbf{L}\cdot\mathbf{S}+(\mathcal{A}_{\rm hfs}/\hbar^2)\,\mathbf{I}\cdot\mathbf{J}. \tag{5.57}$$

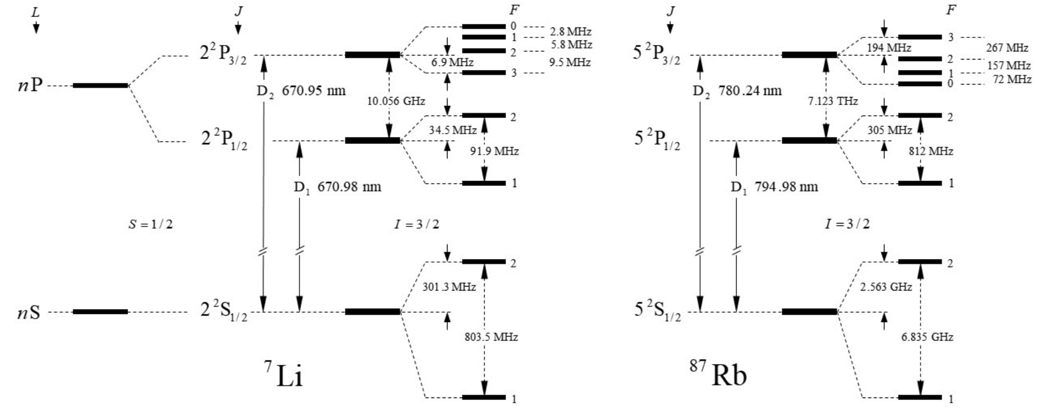


Figure 5.3: Zero-field hyperfine splittings of the lowest electronic states of ^{7}Li [11, 23] and ^{87}Rb [15, 122, 8]. As both species have the same nuclear spin $\left(I = \frac{3}{2}\right)$ the level diagrams are similar. Note that ^{87}Rb behaves like hydrogen whereas ^{7}Li does not (inversion of the hyperfine structure of the $^2P_{3/2}$ level; cf. Section 9.3.2).

This Hamiltonian suffices to calculate magnetic hyperfine shifts for hydrogen-like atoms of given n, l and j. For hydrogenic atoms the constants ζ_{nl} and $\mathcal{A}_{\mathrm{hfs}}$ can be calculated from first principles. For alkali-like atoms in their electronic ground state ζ_{nl} and $\mathcal{A}_{\mathrm{hfs}}$ can be regarded as empirical. For the excited states of alkali-like atoms they become non-hydrogen-like because the $\mathcal{A}_{\mathrm{hfs}}$ coefficients can become negative (cf. Section 9.3.2). As examples we show in Fig. 5.3 the hyperfine splitting of $n\,^2S_{1/2}$, $n\,^2P_{1/2}$ and $n\,^2P_{3/2}$ terms of ^{7}Li and ^{87}Rb.

To calculate the hyperfine splitting we use perturbation theory for degenerate levels. After splitting off the perturbation $\mathcal{H}' = \mathcal{H}_{\mathrm{hfs}}$ we calculate the energy shifts relative to the fine-structure levels $n\,^{2s+1}L_j$, characterized by n, l, s and j. As long as the hyperfine shifts remain small compared to the fine-structure splitting the quantum numbers n, l and j may be considered as "good quantum numbers to first-order in perturbation theory of a degenerate level" - see Eq. (H.56). The mathematical justification for this is that the second-order contribution to the energy shift (H.63) is negligibly small because the energy denominator of Eq. (H.63) is large. Physically, the quantum number j remains a good quantum number because the fine-structure coupling is much stronger than the hyperfine couplings. The fine-structure Hamiltonian (4.124) commutes with both $\mathbf{J}^2$ and $\mathbf{I}^2$ and because these operators also commute with $\mathbf{I}\cdot\mathbf{J}$ we know that j and I are good quantum number of the full Hamiltonian (5.57). Furthermore, as $\mathbf{F} = \mathbf{J} + \mathbf{I}$ and $\mathbf{F}^2 = \mathbf{J}^2 + \mathbf{I}^2 + 2\mathbf{I}\cdot\mathbf{J}$ we find that $\mathbf{F}^2$ and $F_z = J_z + I_z$ commute with both the fine-structure and the hyperfine Hamiltonian. Therefore, also F and m_F are good quantum numbers and $\mathcal{H}_{\mathrm{hfs}}$ is diagonal in the joint basis $\{|n(ls)jIFm_F\rangle\}$ of $\mathbf{L}\cdot\mathbf{S}$, $\mathbf{F}^2$ and F_z; i.e., for given values of n, l and j the quantum numbers F and m_F are good quantum numbers (s and I are always good quantum numbers in atomic physics because these represent intrinsic properties of the electron and the nucleus, respectively). The level shift caused by the hyperfine interaction depends on F and is then given by

$$\Delta E_F^{IJ}(n\,^{2s+1}L_j) = \left(\mathcal{A}_{\mathrm{hfs}}/\hbar^2\right)\langle jIFm_F|\mathbf{I}\cdot\mathbf{J}|jIFm_F\rangle, \tag{5.58}$$

where $\mathcal{A}_{\mathrm{hfs}}$ is the *hyperfine coupling constant* for given quantum numbers n, l and j.

5.3.5 Shift rules for the magnetic hyperfine coupling in zero field

Using the operator identity

$$\mathbf{I}\cdot\mathbf{J} = \tfrac{1}{2}(\mathbf{F}^2 - \mathbf{I}^2 - \mathbf{J}^2) \tag{5.59}$$

the *hyperfine shift* of the manifold with quantum number F takes the form

$$\Delta E_F^{IJ}(n\,^{2s+1}L_j) = \tfrac{1}{2}\mathcal{A}_{\mathrm{hfs}}\left[F(F+1) - j(j+1) - I(I+1)\right]. \tag{5.60}$$

Importantly, for negative $\mathcal{A}_{\mathrm{hfs}}$ the multiplet structure is *inverted.* As the form of Eq. (5.60) is identical to that of Eq. (4.126) derived for the fine-structure splitting this leads again to three shift rules of the type introduced in Section 4.5.2:

- the shift of the *stretched state* is given by

$$\Delta E_{F_{\max}}^{IJ}(n\,^{2s+1}L_j) = \mathcal{A}_{\mathrm{hfs}}\, jI. \tag{5.61}$$

  Recall that the stretched state corresponds to the F manifold with the largest multiplicity, $F = F_{\max} = j + I$.

- the relative shift of two adjacent F manifolds follows an *interval rule* of the *Landé type*:

$$\Delta W_F = \Delta E_F^{IJ}(n\,^{2s+1}L_j) - \Delta E_{F-1}^{IJ}(n\,^{2s+1}L_j) = \mathcal{A}_{\mathrm{hfs}} F_>, \tag{5.62}$$

  where $F_>$ is the quantum number of the manifold with the highest multiplicity of *adjacent pairs*, $F_> = max(F, F-1) = F$.

- the hyperfine manifolds satisfy the *center-of-gravity rule*:

$$\frac{1}{(2I+1)\,(2j+1)} \sum_{F=|j-I|}^{j+I} (2F+1)\, \Delta E_F^{IJ}(n\,^{2s+1}L_j) = 0, \tag{5.63}$$

  where $(2I+1)\,(2j+1)$ is the degeneracy of the unperturbed $n\,^{2s+1}L_j$ level.

5.3.6 Magnetic hyperfine structure of hydrogenic atoms in zero field

For *hydrogenic* atoms the orbital parts in Eqs. (5.35) and (5.34) can be expressed analytically; using Eqs. (2.37) and (2.49c) we obtain

$$a_{\mathrm{hfs}} = \frac{\mu_0}{4\pi}\frac{8}{3}\frac{\gamma_I|\gamma_e|\hbar^2}{[(1+m_e/m_p)a_0]^3}\frac{Z^3}{n^3} \qquad (l=0) \tag{5.64a}$$

$$\mathcal{A}_{\mathrm{hfs}} = \frac{\mu_0}{4\pi}\frac{\gamma_I|\gamma_e|\hbar^2}{[(1+m_e/m_p)a_0]^3}\frac{Z^3}{n^3}\frac{1}{j(j+1)(l+1/2)} \qquad (l>0). \tag{5.64b}$$

For all *hydrogen-like* atoms in s levels the hyperfine states are *spin doublets* $\left(j = \frac{1}{2}\right)$. In such systems coupling of the nuclear spin $\mathbf{I}$ to the total electronic angular momentum $\mathbf{J}$ can result in only two values of total angular momentum: $F = I + \frac{1}{2}$ (parallel coupling) and $F' = I - \frac{1}{2}$ (anti-parallel coupling). For these two cases we have

$$\Delta E_F^{IJ}(n\,^{2s+1}L_{1/2}) = \begin{cases} \left.\begin{array}{ll} +\frac{1}{2}a_{\mathrm{hfs}} I & \left(F = I + \frac{1}{2}\right) \\ -\frac{1}{2}a_{\mathrm{hfs}}(I+1) & \left(F = I - \frac{1}{2}\right) \end{array}\right\} & \text{for } F > 0 \\ 0 & \text{for } F = 0. \end{cases} \tag{5.65}$$

For $\gamma_I > 0$ ($\gamma_I < 0$) the state with the highest multiplicity has the highest (lowest) energy.

5.3.6.1 Hydrogen in its electronic ground state $1\,^2S_{1/2}$

The electronic ground state of atomic hydrogen is the $1\,^2S_{1/2}$ state. With nuclear spin $I = \frac{1}{2}$ the hyperfine levels correspond to $F = 0, 1$. Starting from Eq. (5.64) the hyperfine splitting may be rewritten in terms of fundamental constants,

$$\Delta\nu_{\mathrm{hfs}} = a_{\mathrm{hfs}}/h = \tfrac{4}{3} g_p (1 + a_e)\alpha^4 \frac{m_e}{m_p}\frac{m_e c^2/h}{(1 + m_e/m_p)^3}, \tag{5.66}$$

where $a_e = (g_e - 2)/2 = 1.159\,652\,180\,91(26) \times 10^{-3}$ is the electron anomaly (see Section 4.3.6), $g_p = 5.585\,694\,702(17)$ the proton g factor, $\alpha = 7.297\,352\,5664(17) \times 10^{-3}$ the fine-structure constant, $m_e = 9.109\,383\,56(11) \times 10^{-31}$ kg the electron rest mass $m_e/m_p = 5.446\,170\,213\,52(52) \times 10^{-4}$ the electron-proton mass ratio and $c = 299\,792\,458$ m s^{-1} the defined value for the speed of light in vacuum. With the quoted values for the fundamental constants we calculate $\Delta\nu_{\mathrm{hfs}} \simeq 1.420\,49$ GHz. The experimental value is [117, 32]

$$\Delta\nu_{\mathrm{hfs}} = 1\,420\,405\,751.770(3)\ \mathrm{Hz}. \tag{5.67}$$

Remarkably, the simple theory presented here already gives agreement at the level $1 : 10^4$. With the most advanced theories one, possibly two additional digits can be accounted for. When not accounting for the electron anomaly agreement is worse, about $1 : 10^3$. Historically, this discrepancy gave rise to the suggestion of the existence of the electron anomaly and its relevance in relation to quantum electrodynamics [21]. The experimental precision became possible with the development of the *hydrogen maser*, the first example of an *atomic clock* [89]. The *precision* of the hydrogen maser is better than the actual *accuracy*, which is limited by collisions of the hydrogen atoms in the clock. Beware that the hyperfine structure of *deuterium* (nuclear spin $I = 1$) is complicated by the presence of a nuclear electric quadrupole moment.

Hyperfine transitions in ground state hydrogen are also well-known from radio astronomy as they give rise to electromagnetic radiation at 1.420 GHz, (wavelength 21 cm). This wavelength is not absorbed in the earth atmosphere. The 21 cm radiation of cosmic origin is famous because it was used to reveal, for the first time, the spiral structure of our galaxy [116]. In relation to the properties of *antihydrogen*, the hyperfine transitions of the hydrogen ground state attracted renewed interest [32].

5.3.7 Transition dipole moments in the presence of hyperfine coupling

In the presence of hyperfine coupling m_j is no longer a good quantum number and this has consequences for the calculation of the transition-dipole moment of the electric-dipole operator. The eigenstates of the Hamiltonian are given by the basis $|nlsjIFM\rangle$ and the transition dipole is of the form

$$\mathbf{D}_{eg} = D_{eg}\hat{\mathbf{u}}_q^*, \tag{5.68}$$

where $e = |n'l's'j'I'F'M'\rangle$, $g = |nlsjIFM\rangle$ and

$$D_{eg} = -ea\,\langle n'l's'j'I'F'M'|\,\sqrt{4\pi/3}\rho Y_1^q(\hat{\mathbf{r}})\,|nlsjIFM\rangle \tag{5.69}$$

is the *transition dipole moment*, with $q = M' - M$. Starting again with the Wigner-Eckart theorem we have

$$D_{eg} = -ea(-1)^{F'-M'}\langle n'l's'j'I'F'||\sqrt{4\pi/3}\rho Y_1(\hat{\mathbf{r}})||nlsjIF\rangle \begin{pmatrix} F' & 1 & F \\ -M' & q & M \end{pmatrix}. \tag{5.70}$$

In this case the reduced matrix element $\langle n'l's'j'I'F'||\sqrt{4\pi/3}\rho Y_1(\hat{\mathbf{r}})||nlsjIF\rangle$ can be further reduced. Therefore, we turn to the reduction formula (K.62),

$$\begin{aligned}\langle n'l's'j'I'F'||\sqrt{4\pi/3}\rho Y_1(\hat{\mathbf{r}})||nlsjIF\rangle &= (-1)^{F+j'+I+1}\delta_{I,I'}\times \\ &\times \langle n'l's'j'||\sqrt{4\pi/3}\rho Y_1(\hat{\mathbf{r}})||nlsj\rangle\sqrt{(2F+1)(2F'+1)}\begin{Bmatrix} j' & F' & I \\ F & j & 1 \end{Bmatrix}.\end{aligned} \tag{5.71}$$

Substituting Eq. (4.156) this becomes

$$\begin{aligned}\langle n'l's'j'I'F'||\sqrt{4\pi/3}\rho Y_1(\hat{\mathbf{r}})||nlsjIF\rangle &= (-1)^{F+j'+I+j+s+\max(l,l')}\mathcal{R}_{n'l',nl}\delta_{l',l\pm1}\delta_{s,s'}\delta_{I,I'} \\ &\times \sqrt{\max(l,l')(2j+1)(2j'+1)(2F+1)(2F'+1)}\begin{Bmatrix} l' & j' & s \\ j & l & 1 \end{Bmatrix}\begin{Bmatrix} j' & F' & I \\ F & j & 1 \end{Bmatrix}.\end{aligned} \tag{5.72}$$

Writing the transition dipole moment in the form

$$D_{eg} = -ea\mathcal{R}_{n'l',nl}\mathcal{A}^{l'ls}_{j'IF'M',jIFM}\delta_{l',l\pm1}\delta_{s,s'}\delta_{I,I'}, \tag{5.73}$$

we find for the angular contribution (after reordering the $3j$ symbol)

$$\begin{aligned}\mathcal{A}^{l'ls}_{j'IF'M',jIFM} &= (-1)^{-M'+j'+j+I+s+1+\max(l,l')}\\ &\times\sqrt{\max(l,l')\,(2j+1)\,(2j'+1)\,(2F+1)\,(2F'+1)}\begin{Bmatrix} l' & j' & s\\ j & l & 1\end{Bmatrix}\begin{Bmatrix} j' & F' & I\\ F & j & 1\end{Bmatrix}\begin{pmatrix} F & 1 & F'\\ M & q & -M'\end{pmatrix}.\end{aligned} \tag{5.74}$$

The *selection rules* for hyperfine transitions (determined by the $3j$ symbol) are:

$$\Delta M = 0, \pm1 \quad \text{for } \pi, \sigma^{\pm} - \text{transitions} \tag{5.75a}$$

$$\Delta F = 0, \pm1 \quad (\text{excluding } F = 0 \nleftrightarrow F' = 0) \tag{5.75b}$$

These selection rules come on top of those of the fine structure

$$\Delta j = 0, \pm1 \quad (\text{excluding } j = 0 \nleftrightarrow j' = 0) \tag{5.76a}$$

$$\Delta l = \pm1 \qquad \Delta s = 0 \qquad \Delta I = 0. \tag{5.76b}$$

In addition also the triangular inequalities $\Delta\,(l,s,j)$, $\Delta\,(l'sj')$, $\Delta\,(I,j,F)$ and $\Delta\,(I,j',F')$ must be satisfied. Note that the cases $F = F' = 0$ and $j = j' = 0$ are excluded because the triangular inequalities $\Delta\,(F,F',1)$ and $\Delta\,(j,j',1)$ cannot be satisfied.

5.3.7.1 Example: $p \to s$ transitions in hydrogen-like atoms ($s = 1/2$) with $I = 3/2$

The result for the angular contributions to the strength of $p \to s$ transitions between two fine-structure levels ($j' \to j$) and in the presence of hyperfine coupling ($F' = 3,2,1,0 \to F = 2,1$) depends, for given I, F', on F and $M' \equiv M_{F'}$. Summing over the magnetic quantum numbers of the $F = 1,2$ levels (which are the only two s levels) we obtain

$$\sum_{F=1}^{2}\sum_{M=-F}^{F}\mathcal{A}^2_{\frac{3}{2}IF'M',\frac{1}{2}IFM} = \tfrac{1}{3}. \tag{5.77}$$

Note that (for given I) this sum is *independent* of F' and M'; i.e., the sum is same for all values of the magnetic quantum number of the p level. This is illustrated in Fig. 5.4. The sum rule can be understood from the physical point of view by considering electric-dipole transitions. Since the electric-dipole operator does not couple to spin degrees of freedom (it acts in a different sector of Hilbert space), the spins cannot affect the total electric dipole transition rate. It only distributes the rate over the available final states in the form of σ_+, σ_- and π contributions. For transitions from the $F' = 2$ level 50 % passes through the $F' = 2 \to F = 2$ channel and 50 % passes through the $F' = 2 \to F = 1$ channel. Likewise, for transitions from the $F' = 1$ level 1/6 passes through the $F' = 1 \to F = 2$ channel and 5/6 passes through the $F' = 1 \to F = 1$ channel.

5.4 Hyperfine structure in an applied magnetic field

Adding to the Hamiltonian (5.57) the Zeeman Hamiltonian (5.23) we obtain the effective Hamiltonian for the hyperfine-structure of hydrogen-like atoms in an externally applied magnetic field,

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_r + (\zeta_{nl}/\hbar^2)\,\mathbf{L}\cdot\mathbf{S} + (\mathcal{A}_{\text{hfs}}/\hbar^2)\,\mathbf{I}\cdot\mathbf{J} - (\boldsymbol{\mu}_L + \boldsymbol{\mu}_s + \boldsymbol{\mu}_I)\cdot\mathbf{B}, \tag{5.78}$$

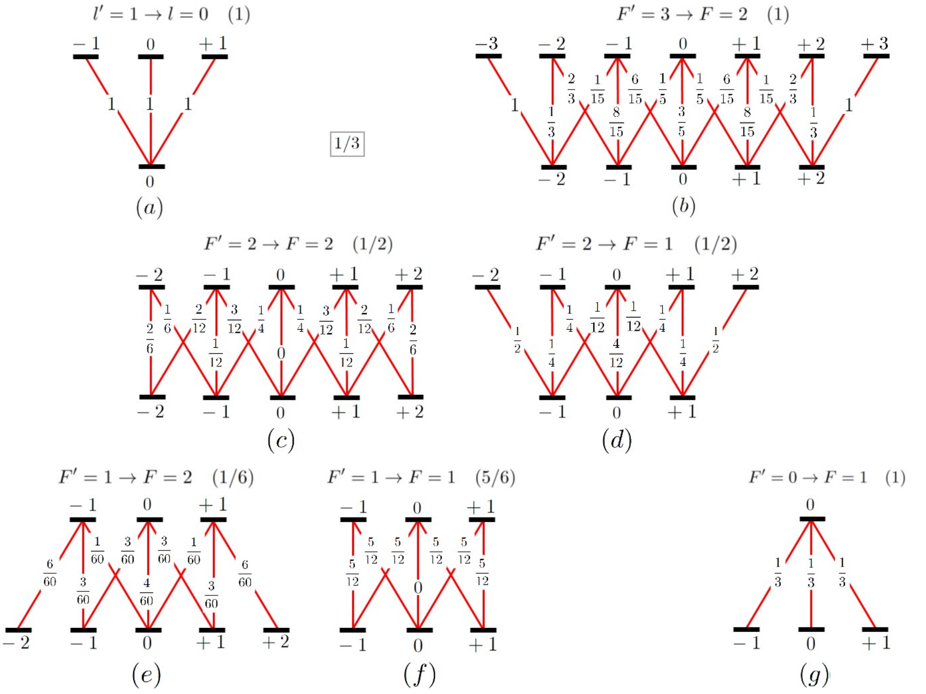


Figure 5.4: The angular contributions to the transition strength between two angular momentum levels (l' and l) is independent of the presence (or absence) of hyperfine structure. This is illustrated for $p \to s$ transitions in hydrogen-like atoms with both fine structure and hyperfine structure (the example of ^{87}Rb): (a) in principal structure (see also Figs. 2.4 and 4.15); (b) from $F' = 3$ level; (c) and (d) from $F' = 2$ level; (e) and (f) from $F' = 1$ level and (g) from $F' = 0$ level. Note that the sum of the probabilities starting from any of the $M_{F'}$ levels, always yields the same value (unity, not counting the common prefactor 1/3).

which can be written in the form

$$\mathcal{H} = \mathcal{H}_{\rm fs} + (\mathcal{A}_{\rm hfs}/\hbar^2)\,\mathbf{I}\cdot\mathbf{J} + [g_J(e/2m_e)\mathbf{J} - g_I(e/2m_p)\mathbf{I}]\cdot\mathbf{B}\ , \tag{5.79}$$

where $\mathcal{H}_{\rm fs}$ represents the zero field fine-structure Hamiltonian (see Section 4.5.1),

$$\mathcal{H}_{\rm fs} = \mathcal{H}_0 + \mathcal{H}_r + (\zeta_{nl}/\hbar^2)\,\mathbf{L}\cdot\mathbf{S}. \tag{5.80}$$

Eq. (5.79) is a suitable Hamiltonian for use in “first-order in perturbation theory of a degenerate level” as long as n, l and j are good quantum numbers; i.e., for magnetic fields much smaller than the fine-structure crossover field ($B \ll |B_{\rm fs}|$). Choosing the quantization axis (z axis) along the magnetic field direction Eq. (5.79) can be written in the form

$$\mathcal{H} = \mathcal{H}_{\rm fs} + (\mathcal{A}_{\rm hfs}/\hbar^2)\,\mathbf{I}\cdot\mathbf{J} + (g_J\mu_B J_z - g_I\mu_N I_z)B/\hbar. \tag{5.81}$$

In analogy with Eq. (4.164) we write the Hamiltonian in the form $\mathcal{H} = \mathcal{H}_{\rm fs} + \mathcal{H}'$, where

$$\mathcal{H}' = \mathcal{H}_{\rm hfs} + \mathcal{H}_Z, \tag{5.82}$$

is the perturbation to be considered, with

$$\mathcal{H}_{\rm hfs} = (\mathcal{A}_{\rm hfs}/\hbar^2)\,\mathbf{I}\cdot\mathbf{J} \quad \text{and} \quad \mathcal{H}_Z = (g_J\mu_B J_z - g_I\mu_N I_z)B/\hbar \tag{5.83}$$

representing the hyperfine coupling and the (electronic plus nuclear) Zeeman terms, respectively. Although both $\mathbf{I}\cdot\mathbf{J}$ and $\mathcal{H}_Z$ commute separately with $\mathcal{H}_{\rm fs}$, mutually they do *not* commute. Physically, there is competition between the mutual coupling of **J** and **I** and the coupling of **J** and **I** individually with the external **B** field. In low fields the hyperfine coupling is dominant, whereas in the high field limit it may be neglected (better: gives a field-independent contribution). This is called the *Back-Goudsmit effect* [6], the exact analogue of the *Paschen-Back effect* [78] in the atomic fine structure in which **L** and **S** decouple (see also [79]). As the fine-structure coupling is stronger than the hyperfine coupling, the Paschen-Back decoupling typically takes place at higher fields than the hyperfine decoupling. In intermediate fields the perturbation $\mathcal{H}'$ is neither diagonal in the $\{|nlsjIFm_F\rangle\}$ (*coupled*) representation nor in the $\{|nlsjm_jIm_I\rangle\}$ (*uncoupled*) representation and we have to solve a secular equation of the type (I.24) just as in the case of the atomic fine structure (Section 4.7).

5.4.1 Matrix elements in the uncoupled basis

To solve the secular equation in the *uncoupled* representation we use the inner product rule (3.109b) to write the perturbation in the form

$$\mathcal{H}' = (a_{\mathrm{hfs}}/\hbar^2)\left[I_zJ_z + \tfrac{1}{2}\left(I_+J_- + I_-J_+\right)\right] + (g_J\mu_BJ_z - g_I\mu_NI_z)B/\hbar. \tag{5.84}$$

Note that this Hamiltonian conserves the total angular momentum along the z axis; i.e., $m_j + m_I = m_F$ is a good quantum number (F_z commutes with the Hamiltonian) for all values of the magnetic field. The diagonal matrix elements are for a given value of m_F

$$\begin{aligned}\langle nIm_Ijm_j|\mathcal{H}'|nIm_Ijm_j\rangle &= \langle nI\left(m_F - m_j\right)jm_j|\mathcal{H}'|nI\left(m_F - m_j\right)jm_j\rangle \\ &= a_{\mathrm{hfs}}\left(m_F - m_j\right)m_j - g_I\mu_Nm_FB + \left(g_J\mu_B + g_I\mu_N\right)m_jB \\ &\equiv \langle m_Fm_j|\mathcal{H}'\,|m_Fm_j\rangle \equiv \mathcal{H}'_{m_Fm_jm_j},\end{aligned} \tag{5.85}$$

with $m_e/m_p = \mu_N/\mu_B$ being the electron-proton mass ratio. Note that in these matrix elements the states $\{|nlsjm_jIm_I\rangle\}$ share fixed values for the quantum numbers n, l, s, j and I; only the values of m_j and m_I vary. Using the relation $m_I = m_F - m_j$ we gave preference to specification of m_F and m_j rather than m_j and m_I. Turning to the shorthand notation $|nlsjm_jIm_I\rangle \to |m_Fm_j\rangle$ the matrix elements take the form $(m_Fm'_j|\mathcal{H}'|m_Fm_j\rangle$. Since m_F is conserved by $\mathcal{H}'$ the notation can be condensed to $\mathcal{H}'_{m_Fm_jm_j}$. The terms $\frac{1}{2}\left(a_{\mathrm{hfs}}/\hbar^2\right)\left(I_+J_- + I_-J_+\right)$ give rise to off-diagonal elements and are calculated using the shift operators (1.59) and the conservation of m_F,

$$\begin{aligned}\langle nIm'_Ijm'_j|\mathcal{H}'|nIm_Ijm_j\rangle &= \langle nI\left(m_F - m_j \pm 1\right)j\left(m_j \mp 1\right)|\mathcal{H}'|nI\left(m_F - m_j\right)jm_j\rangle \\ &= \tfrac{1}{2}a_{\mathrm{hfs}}\sqrt{I\left(I+1\right) - \left(m_F - m_j\right)\left(m_F - m_j \pm 1\right)}\sqrt{j\left(j+1\right) - m_j(m_j \mp 1)} \\ &\equiv \langle(m\left(m_j \mp 1\right)|\mathcal{H}'|m_Fm_j\rangle \equiv \mathcal{H}'_{m_F(m_j\mp1)m_j}.\end{aligned} \tag{5.86}$$

Hence, the perturbation couples states differing in m_j such that $\Delta m_j = \pm 1$ while $m_I + m_j = m_F$ remains conserved. For the coupled basis $\{|nIjFm_F\rangle\}$ this implies that the perturbation couples states differing in F such that $\Delta F = \pm 1$ while m_F remains conserved. In the description of the field dependence preference will be given to expressing coupled states in terms of uncoupled states, rather than the other way around, because in the uncoupled representation the expressions for the limiting cases of low and high field remain intuitively transparent. In the coupled representation the latter is only the case for the low-field limit.

5.4.2 Hydrogen-like atoms with $j = 1/2$

For hydrogen-like atoms in states with $j = \frac{1}{2}$ and $F = I \pm \frac{1}{2}$ the solution of the secular equation can be obtained analytically for arbitrary magnetic field. This case applies to the electronic ground states of hydrogen ($1\,^2S_{1/2}$) and all the alkali atoms ($n\,^2S_{1/2}$ with $n = 2 - 7$). As we are dealing with s orbitals the orbital magnetic moment is zero and g_J coincides with the free-electron g value,

$g_J = g_e$. The ${}^2S_{1/2}$ states behave like the *spin doublets* discussed in Section 4.7.3. The hyperfine shifts in zero field were obtained in Section 5.3.4,

$$\Delta E^{I1/2}_{I\pm1/2}(n\,{}^2S_{1/2}) = \begin{cases} +\frac{1}{2}a_{\mathrm{hfs}}I & \text{for } F = I+\frac{1}{2} \\ -\frac{1}{2}a_{\mathrm{hfs}}\,(I+1) & \text{for } F = I-\frac{1}{2}. \end{cases} \quad \text{for } F>0. \tag{5.87}$$

However, unlike the spin doublets discussed in Section 4.7.3 the state with the lowest multiplicity is *not necessarily* the state with the lowest energy but this depends on the sign of a_{hfs} and therefore on the sign of the nuclear gyromagnetic ratio γ_I. For $\gamma_I > 0$ ($\gamma_I < 0$) the state of lowest multiplicity has the lowest (highest) energy, as illustrated in Fig. 5.5. In the absence of nuclear spin ($I = 0$) this shift is absent. For $I > 0$ the perturbation matrix takes the form

$$\mathcal{H}' = \begin{pmatrix} \mathcal{H}'_{M\uparrow\uparrow} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \mathcal{H}'_{(M-1)\uparrow\uparrow} & \mathcal{H}'_{(M-1)\uparrow\downarrow} & \cdots & 0 & 0 & 0 \\ 0 & \mathcal{H}'_{(M-1)\downarrow\uparrow} & \mathcal{H}'_{(M-1)\downarrow\downarrow} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \mathcal{H}'_{(-M+1)\uparrow\uparrow} & \mathcal{H}'_{(-M+1)\uparrow\downarrow} & 0 \\ 0 & 0 & 0 & \cdots & \mathcal{H}'_{(-M+1)\downarrow\uparrow} & \mathcal{H}'_{(-M+1)\downarrow\downarrow} & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \mathcal{H}'_{-M\downarrow\downarrow} \end{pmatrix}, \tag{5.88}$$

where $M \equiv I + \frac{1}{2}$ and ↑ is the short hand notation for $m_j = \frac{1}{2}$ (and ↓ for $m_j = -\frac{1}{2}$). The quantum number m_F varies within the range $-M \le m_F \le M$. Note that the matrix is a $d \times d$ matrix, where $d = (2I+1)(2j+1)$ is the dimension of the angular momentum subspace in which $\mathbf{I}$ and $\mathbf{J}$ operate. The secular equation is of the asymmetric type (cf. Appendix H.3) and factorizes into a product of two (1×1) and $I+1$ (2×2) determinants, each characterized by its value of $m_F = m_j + m_I$. With the aid of Eqs. (5.85) and (5.86) the matrix elements are found to be

$$\mathcal{H}'_{m_F\uparrow\uparrow} = +\tfrac{1}{2}a_{\mathrm{hfs}}\left(m_F - \tfrac{1}{2}\right) - g_I\mu_N m_F B + \tfrac{1}{2}\mu_+ B \tag{5.89a}$$

$$\mathcal{H}'_{m_F\downarrow\downarrow} = -\tfrac{1}{2}a_{\mathrm{hfs}}\left(m_F + \tfrac{1}{2}\right) - g_I\mu_N m_F B - \tfrac{1}{2}\mu_+ B \tag{5.89b}$$

$$\mathcal{H}'_{m_F\downarrow\uparrow} = +\tfrac{1}{2}a_{\mathrm{hfs}}(I+\tfrac{1}{2})(1-\tilde{m}_F^2)^{1/2} = \mathcal{H}'_{m_F\uparrow\downarrow}, \tag{5.89c}$$

where

$$\mu_\pm \equiv g_J\mu_B \pm g_I\mu_N \tag{5.90a}$$

$$\tilde{m}_F \equiv m_F/(I+\tfrac{1}{2}) \quad \text{with} \quad -1 \le \tilde{m}_F \le 1. \tag{5.90b}$$

Note that the symbol μ_- is not relevant in the present context. It is included in the definition (5.90a) for future convenience.

5.4.2.1 Pure states

For $m_F = \pm(I+\frac{1}{2})$ the coupling term vanishes, $\mathcal{H}'_{m_F\downarrow\uparrow} = 0$. These cases correspond to the *stretched states*,

$$|F_{\max}, \pm(I+\tfrac{1}{2})\rangle = \left|s, \pm\tfrac{1}{2}; I, \pm I\right\rangle. \tag{5.91}$$

The corresponding energies are

$$\varepsilon_\pm = \tfrac{1}{2}a_{\mathrm{hfs}}I \mp g_I\mu_N(I+\tfrac{1}{2})B \pm \tfrac{1}{2}\mu_+ B \qquad \text{for } m_F = \pm(I+\tfrac{1}{2}). \tag{5.92}$$

Note that the last term dominates the field dependence.

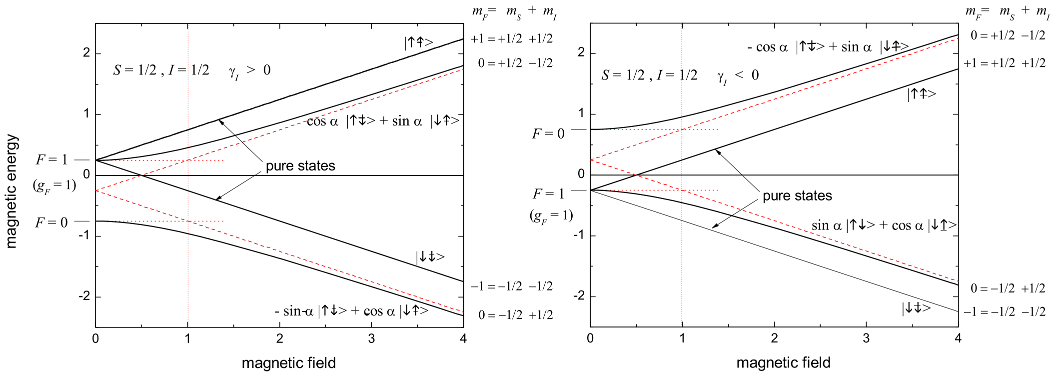


Figure 5.5: Magnetic energy versus magnetic field in units of the hyperfine field $|B_{\rm hfs}|$. Left: positive hyperfine shift. Right: negative hyperfine shift. We adopt the convention $\sin\alpha \to 0$ for $B \gg |B_{\rm hfs}|$. The dashed (dotted) lines correspond to the high-field (low-field) tangents of the energy curves.

5.4.2.2 Mixed states

The other states ($-1 < \tilde{m}_F < 1$) are *mixed states* because they involve linear combinations of the uncoupled basis states. Depending on the sign of γ_I we write this linear combination in the form (see Fig. 5.5)

$$\gamma_I > 0 \quad \begin{cases} |F_{\rm max}, m_F\rangle = +\cos\alpha|\uparrow, I, \left(m_F - \frac{1}{2}\right)\rangle + \sin\alpha|\downarrow, I, \left(m_F + \frac{1}{2}\right)\rangle \\ |F_{\rm min}\,, m_F\rangle = -\sin\alpha|\uparrow, I, \left(m_F - \frac{1}{2}\right)\rangle + \cos\alpha|\downarrow, I, \left(m_F + \frac{1}{2}\right)\rangle \end{cases} \tag{5.93a}$$

$$\gamma_I < 0 \quad \begin{cases} |F_{\rm min}\,, m_F\rangle = -\cos\alpha|\uparrow, I, \left(m_F - \frac{1}{2}\right)\rangle + \sin\alpha|\downarrow, I, \left(m_F + \frac{1}{2}\right)\rangle \\ |F_{\rm max}, m_F\rangle = +\sin\alpha|\uparrow, I, \left(m_F - \frac{1}{2}\right)\rangle + \cos\alpha|\downarrow, I, \left(m_F + \frac{1}{2}\right)\rangle. \end{cases} \tag{5.93b}$$

In both cases the states $|F_{\rm max}, m_F\rangle$ and $|F_{\rm min}, m_F\rangle$ are seen to form an orthonormal pair. The difference in notation between the two cases is introduced solely to improve the readability of the expressions (5.93a) and (5.93b); it implies that $\sin\alpha \simeq \alpha \to 0$ in the limit of high magnetic fields, irrespective of the value of γ_I. This notation offers the convenience of indicating which uncoupled state can be neglected in high magnetic fields. From the limit $B \to 0$ we know that for $\gamma_I > 0$ the *upper* branch of the spin doublets must correspond to the $F_{\rm max}$ manifold; likewise, for $\gamma_I < 0$ the *lower* branch of the spin doublets must correspond to the $F_{\rm min}$. In the $B \to 0$ limit the Zeeman energy vanishes and F becomes a good quantum number $\left(F_{\rm max} \to F = I + \frac{1}{2},\ F_{\rm min} \to F = I - \frac{1}{2}\right)$ and the Eqs. (5.93a) and (5.93b) take the form of a Clebsch-Gordan decomposition - compare with Eq. (J.49). The phase factors are chosen in accordance with the *Condon and Shortley phase convention*, where the phase factors are unity for the branch of maximal F and $m_F = F$ (parallel coupling).

The energies of the mixed states are given by solutions of the secular equation for given value of m_F,

$$W_\pm = \tfrac{1}{2}\left(\mathcal{H}'_{m_F\uparrow\uparrow} + \mathcal{H}'_{m_F\downarrow\downarrow}\right) \pm \tfrac{1}{2}\sqrt{(\mathcal{H}'_{m_F\uparrow\uparrow} - \mathcal{H}'_{m_F\downarrow\downarrow})^2 + 4|\mathcal{H}'_{m_F\downarrow\uparrow}|^2}. \tag{5.94}$$

This equation can be expressed in the form of the *Breit-Rabi formula* [22],

$$\varepsilon_\pm(m_F) = -g_I\mu_N m_F B - \tfrac{1}{4}a_{\rm hfs} \pm \tfrac{1}{2}|a_{\rm hfs}|(I + \tfrac{1}{2})\sqrt{1 + 2\tilde{m}_F B/B_{\rm hfs} + (B/B_{\rm hfs})^2}. \tag{5.95}$$

Since $|\tilde{m}_F| < 1$, the discriminant is always positive. The modulus $|a_{\rm hfs}|$ assures that ε_+ always represents the largest of the two roots ($\varepsilon_+ \geq \varepsilon_-$); i.e., corresponds the upper branch of the hyperfine doublet. Eq. (5.95) expresses the magnetic field dependence of the hyperfine shift for hydrogen-like atoms with $j = \frac{1}{2}$ and for given values of F and $m_F = m_I + m_j \neq \pm(I + \frac{1}{2})$. The characteristic field

$$B_{\rm hfs} = a_{\rm hfs}(I + \tfrac{1}{2})/\mu_+ \tag{5.96}$$

is called the *hyperfine field* with μ_+ defined by Eq. (5.90a); $B_{\rm hfs}$ is an effective *internal magnetic field* which can be positive or negative depending on the sign of $a_{\rm hfs}$. Its absolute value, $|B_{\rm hfs}|$, is the *hyperfine crossover field* and corresponds to the intersection point of the low-field and high-field asymptotic expressions for the hyperfine energy. For the ground state of atomic hydrogen the crossover field is $B_{\rm hfs} \simeq 50.7$ mT, much smaller than the fine-structure crossover field estimated in Section 4.7.3. Thus, in many cases of practical interest the high-field limit is reached for the hyperfine coupling when the atom is still in the low-field limit for the fine structure ($|B_{\rm hfs}| \ll |B_{\rm fs}|$). If this comfortable separation of energy scales is absent one has to diagonalize the combined secular equations for the fine structure and the hyperfine structure, which is already a task for which computers are best suited. Actually, some will say: why bother about analytic solutions at all when computer diagonalizations are swift and flawless even in complicated cases! In this respect it is good to be aware that the proper assignment of the basis vectors and a consistent phase convention remains an issue for which analytic understanding remains indispensable.

To obtain the eigenstates corresponding to the energies ε_+ and ε_- we have to determine the coefficients in Eqs. (5.93). This is done by the procedure described in Appendix H.3. The amplitudes of the coefficients, $A_+ \equiv \cos\alpha$ and $A_- \equiv \sin\alpha$, follow by substitution of the matrix elements (5.89a)-(5.89c) into Eq. (G.51),

$$\left.\begin{array}{l}\gamma_I > 0 : A_\pm^2 \\ \gamma_I < 0 : A_\mp^2\end{array}\right\} = \frac{1 - \tilde{m}_F^2}{\left(\tilde{m}_F + B/B_{\rm hfs} \mp \sqrt{1 + 2\tilde{m}_F B/B_{\rm hfs} + (B/B_{\rm hfs})^2}\right)^2 + 1 - \tilde{m}_F^2}. \tag{5.97}$$

Clock states ($m_F = 0$)**:** A special class of mixed hyperfine states are *clock states*, states with $m_F = 0$. Near zero field these states are insensitive for small variations in magnetic field which explains their importance for precision time and frequency measurements. For $m_F = 0$ the Breit-Rabi formula takes the form

$$\varepsilon_\pm(0) = -\tfrac{1}{4}a_{\rm hfs} \pm \tfrac{1}{2}|a_{\rm hfs}|(I + \tfrac{1}{2})\sqrt{1 + (B/B_{\rm hfs})^2}. \tag{5.98}$$

The amplitudes are given by

$$A_\pm^2 = \left(B/B_{\rm hfs} \mp \sqrt{1 + (B/B_{\rm hfs})^2}\right)^{-2}. \tag{5.99}$$

High-field limit ($B \gg |B_{\rm hfs}|$)**:** In high magnetic fields we can expand the Breit-Rabi formula (5.95) in powers of $B_{\rm hfs}/B$. Retaining only the terms up to first order in $B_{\rm hfs}/B$ we obtain

$$\varepsilon_\pm(m_F) = -g_I\mu_N m_F B - \tfrac{1}{4}a_{\rm hfs} \pm \tfrac{1}{2}\mu_+ B(1 + \tilde{m}_F B_{\rm hfs}/B). \tag{5.100}$$

This expression corresponds to the *high-field asymptotes* indicated by the dashed lines in Fig. 5.5. Using a similar expansion Eq. (5.97) takes the form

$$A_-^2 = \sin^2\alpha \simeq \tfrac{1}{4}(1 - \tilde{m}_F^2)(B/B_{\rm hfs})^2 \to 0. \tag{5.101}$$

This is an example of *weak coupling (strong asymmetry)* as defined in Appendix G.3.4.

Low-field limit ($B \ll |B_{\rm hfs}|$)**:** For low magnetic fields we can expand Eq. (5.95) in powers of $B/B_{\rm hfs}$. Retaining the terms up to second order in $B/B_{\rm hfs}$ we obtain

$$\begin{aligned}\varepsilon_\pm(m_F) = -g_I\mu_N m_F B &- \tfrac{1}{4}a_{\rm hfs} \\ &\pm \tfrac{1}{2}|a_{\rm hfs}|(I + \tfrac{1}{2})[1 + \tilde{m}_F B/B_{\rm hfs} + \tfrac{1}{2}(1 - \tilde{m}_F^2)(B/B_{\rm hfs})^2 + \cdots].\end{aligned} \tag{5.102}$$

The linear and quadratic contributions are referred to as *linear* and *quadratic* Zeeman effects. The linear terms correspond to the *low-field tangents* indicated by the dotted lines in Fig. 5.5. Note that the quadratic shift is strongest for small $|m_F|$, in particular for $m_F = 0$. The quadratic shift is always positive for the upper (+) and negative for the lower (−) manifold as is evident from Fig. 5.5. Restricting ourselves to the leading terms we distinguish two cases:

a.) $0 < |m_F| < I + \frac{1}{2}$: In this case the linear terms are dominant,

$$\varepsilon_\pm(m_F) = -g_I\mu_N m_F B - \tfrac{1}{4}a_{\text{hfs}} \pm \tfrac{1}{2}|a_{\text{hfs}}|(I + \tfrac{1}{2})(1 + \tilde{m}_F B/B_{\text{hfs}}) + \cdots . \tag{5.103}$$

This expression corresponds to the low-field tangents indicated by the dotted lines in Fig. 5.9. Expanding Eq. (5.97) to first order in B/B_{hfs}we find

$$A_\pm^2 \simeq \tfrac{1}{2}(1 \pm \tilde{m}_F)\,[1 \pm (B/B_{\text{hfs}})(1 \mp \tilde{m}_F) + \cdots] . \tag{5.104}$$

b.) $m_F = 0$: Clock states correspond to the case of *strong coupling (weak asymmetry)* as defined in Appendix H.3. The Breit-Rabi formula takes a quadratic field dependence

$$\varepsilon_\pm(0) = -\tfrac{1}{4}a_{\text{hfs}} \pm \tfrac{1}{2}|a_{\text{hfs}}|(I + \tfrac{1}{2}) \pm \tfrac{1}{4}|a_{\text{hfs}}|(I + \tfrac{1}{2})(B/B_{\text{hfs}})^2 + \cdots . \tag{5.105}$$

The absence of a linear field dependence in the limit $B \to 0$ makes these levels insensitive to small magnetic fields, which explains their importance for application in *atomic clocks.* Expanding Eq. (5.99) in powers of B/B_{hfs} we note that the linear term is absent and the amplitudes of the coefficients can be approximated by

$$A_\pm^2 \simeq \tfrac{1}{2}\left[1 \pm (B/B_{\text{hfs}})^2 + \cdots\right] . \tag{5.106}$$

Crossover field ($B = |B_{\text{hfs}}|$)**:** Equating the expressions (5.100) and (5.103) for the high- and low-field tangents we obtain an expression for the intersection point,

$$\mu_+ B(1 + \tilde{m}_F B_{\text{hfs}}/B) = |a_{\text{hfs}}|(I + \tfrac{1}{2})(1 + \tilde{m}_F B/B_{\text{hfs}}). \tag{5.107}$$

Solving for B we find for the intersection point $B = |B_{\text{hfs}}| = |a_{\text{hfs}}|(I + \frac{1}{2})/\mu_+$. This point marks the crossover between the low-field and high-field regions.

5.4.3 High-field limit

For magnetic fields much higher than the hyperfine field but much lower than the fine-structure field ($|B_{\text{hfs}}| \ll B \ll |B_{\text{fs}}|$) the hyperfine coupling is weak compared to the Zeeman interaction whereas the fine-structure coupling is still strong. For increasing magnetic field, this manifests itself as a crossover from the hyperfine coupling between **J** and **I** ($B \ll |B_{\text{hfs}}|$) to Zeeman coupling of **J** and **I** individually to the magnetic field ($B \gg |B_{\text{hfs}}|$) - see Fig. 5.7. This crossover is known as the *Back-Goudsmit effect* and is similar to the *Paschen-Back effect* in the fine structure. Therefore, in the high-field limit the atomic Hamiltonian

$$\mathcal{H} = (\mathcal{H}_{\text{fs}} + \mathcal{H}_Z) + \mathcal{H}_{IJ} \tag{5.108}$$

is best analyzed in the *uncoupled* hyperfine representation $\{|nlsjm_jIm_I\rangle\}$ in which $\mathcal{H}_Z$ is diagonal and $\mathcal{H}_{IJ}$ is the perturbation. In this basis and for given values of n, l, s, j the perturbation matrix is given by

$$\mathcal{H}^{IJ}_{m_j'm_I';m_jm_I} = (a_{\text{hfs}}/\hbar^2)\langle jm_j'Im_I'|\mathbf{I}\cdot\mathbf{J}|jm_jIm_I\rangle, \tag{5.109}$$

which can be written in the form

$$\mathcal{H}^{IJ}_{m_j'm_I';m_jm_I} = (a_{\text{hfs}}/\hbar^2)\langle jm_j'Im_I'|I_zJ_z + \tfrac{1}{2}\left(I_+J_- + I_-J_+\right)|jm_jIm_I\rangle. \tag{5.110}$$

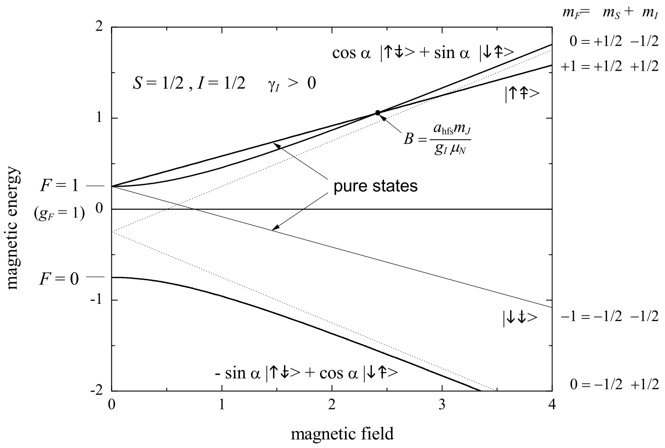


Figure 5.6: Crossing of the nuclear Zeeman levels at at $B \simeq 16.7$ T in hydrogen. This happens when the nuclear Zeeman shift equals the hyperfine spitting. Note that the figure is *not* to scale (the nuclear moment was chosen to be *artificially large* to show both the hyperfine splitting and the intersection point in one plot).

As only the *diagonal matrix elements* of a weak perturbation *contribute to first order* in perturbation theory (see Appendix G.3.4 - weak coupling), in high fields the hyperfine shift is given by $\Delta E^{IJ}_{m_j m_I}\left(n\,^{2s+1}L_j\right) = a_{\rm hfs} m_I m_j$ and the *asymptotic field dependence* (dashed lines in Fig. 5.5) can be expressed as

$$\Delta E_{m_j m_I}\left(n\,^{2s+1}L_j\right) = a_{\rm hfs} m_I m_j + (g_J \mu_B m_j - g_I \mu_N m_I)B. \tag{5.111}$$

Note that this expression is valid for arbitrary values of j and I.

5.4.3.1 Very-high-field limit ($B > a_{\rm hfs} m_j / g_I \mu_N$)

The levels in the hyperfine manifold with the largest multiplicity will intersect when the nuclear Zeeman shift becomes larger than the zero-field hyperfine splitting (see Fig. 5.6). This happens for $g_I \mu_N B > a_{\rm hfs} m_j > 0$ (presuming j to be a good quantum number). Beyond the intersection the ordering of the nuclear sublevels is the same in the lower and upper manifold, as is to be expected in sufficiently high field. Interestingly, at $B = a_{\rm hfs} m_j / g_I \mu_N$ the overall hyperfine shift is independent of m_I, $\Delta E_{m_j m_I}\left(n\,^{2s+1}L_j\right) = g_J \mu_B m_j B$. This means that all m_I levels of a manifold with $m_j > 0$ coincide, which is an example of *accidental degeneracy.* For the ground states of atomic hydrogen, deuterium and tritium this happens in the manifolds with $m_s = \frac{1}{2}$ at $B \simeq 16.7$ T.

5.4.4 Low-field limit - linear Zeeman shift and hyperfine g factor (g_F)

For magnetic fields much lower than the hyperfine field ($B \ll |B_{\rm hfs}|$) the angular momenta $\mathbf{J}$ and $\mathbf{I}$ are strongly hyperfine coupled (by $\mathcal{H}_{IJ}$) to form the total electronic angular momentum $\mathbf{F} = \mathbf{J} + \mathbf{I}$ (see Fig. 5.7). Under these conditions the atomic Hamiltonian

$$\mathcal{H} = (\mathcal{H}_{\rm fs} + \mathcal{H}_{IJ}) + \mathcal{H}_Z \tag{5.112}$$

is best analyzed in the *coupled* representation $\{|nlsjIFm_F\rangle\}$, where $\mathbf{L}\cdot\mathbf{S}$ and $\mathbf{J}\cdot\mathbf{I}$ are diagonal and $\mathcal{H}_Z$ is the perturbation. In this basis and for given values of n, l, s, j, I, F the perturbation matrix is given by

$$\mathcal{H}^Z_{F'm_{F'};Fm_F} = \langle F'm_{F'}|g_J \mu_B J_z - g_I \mu_N I_z\,|Fm_F\rangle\, B/\hbar. \tag{5.113}$$

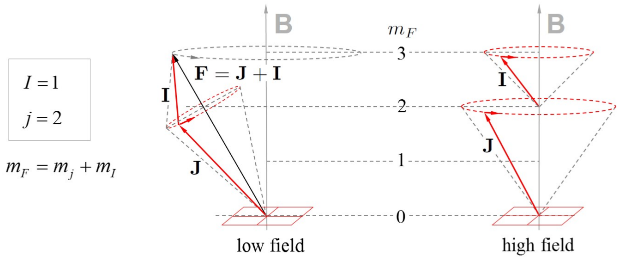


Figure 5.7: Precession of angular momenta in an externally applied magnetic field for the example $j = 2$, $I = 1$. In *low* fields **J** and **I** are hyperfine coupled to form the total atomic angular momentum $\mathbf{F} = \mathbf{J} + \mathbf{I}$, which precesses slowly about the field direction as a result of a weak Zeeman coupling, $\mathcal{H}_Z$, to the magnetic field; this is most conveniently described in the *coupled* basis $\{|F, m_F\rangle\}$. In *high* fields **J** and **I** are predominantly Zeeman coupled to the external field and precess mostly independently about the field direction; this is best described in the *uncoupled* basis $\{|j, m_j, I, m_I\rangle\}$. Note that $m_F = m_j + m_I$ is a good quantum number in both limits.

Expressing the nuclear magneton into the Bohr magneton, $\mu_N \equiv (m_e/m_p)\mu_B$, and limiting ourselves to the diagonal terms, we obtain the following expression for the energy shift (Zeeman shift)

$$\Delta E^Z_{F,m_F}\left(n\,^{2s+1}L_j\right) = \langle jIFm_F|g_J J_z - g_I I_z(m_e/m_p)\,|jIFm_F\rangle\,\mu_B B/\hbar, \tag{5.114}$$

This expression is valid to only first order in perturbation theory because $\mathcal{H}_Z$ is in *not* diagonal in the coupled representation: only the *diagonal matrix elements* of a weak perturbation *contribute to first order* (see Appendix G.3.4 - weak coupling). As the operators J_z and I_z cannot change m_j and m_I, the quantum number m_F has to be conserved. This is a consequence of the m_F *selection rule* (3.90), $m_F = m_j + m_I$. However, J_z and I_z do *not* conserve F (compare Problem 4.13). Only in *low* fields the perturbation can be made sufficiently weak to neglect the off-diagonal terms and treat both F and m_F as good quantum numbers of the atomic Hamiltonian (4.194).

The physics in low field is completely analogous to the low-field limit of the fine-structure: in sufficiently low fields the magnetic moment of the atom will scale proportionally to the total atomic angular momentum **F**,

$$\boldsymbol{\mu}_F \equiv \gamma_F \mathbf{F} \equiv -g_F \mu_B\,(\mathbf{F}/\hbar) \qquad \text{for } B \to 0 \text{ and } F > 0. \tag{5.115}$$

Note that γ_F and g_F are defined with opposite sign, just as we did for γ_L, γ_e and γ_J; as will be shown the sign of g_F can be positive or negative depending on F. We only have to address the case $F > 0$ because for $F = 0$ the magnetic moment must be absent. Hence, the Zeeman energy can be described by the effective Hamiltonian $\mathcal{H}_Z = -\boldsymbol{\mu}_F \cdot \mathbf{B}$. This Hamiltonian yields a Zeeman shift on top of the fine- and hyperfine structure shifts determined by the quantum numbers n, l, s, j, I, F :

$$\Delta E^Z_{F,m_F} = g_F \mu_B B \langle Fm_F|(F_z/\hbar)\,|Fm_F\rangle = g_F \mu_B m_F B \qquad \text{for } B \to 0\ . \tag{5.116}$$

This expression corresponds to the *low-field tangents* indicated by the dotted lines in Fig. 5.5. Note that the effective Hamiltonian is diagonal in the $\{|nlsjIFm_F\rangle\}$ basis.

In search for g_F we compare Eqs. (5.114) and (5.116) and obtain for $F > 0$

$$g_F = g_J \langle jIF||J||jIF\rangle - g_I(m_e/m_p)\langle jIF||I||jIF\rangle. \tag{5.117}$$

To arrive at this result we applied the Wigner-Eckart theorem. As J_z, I_z and F_z are diagonal in the $\{|nlsjIFm_F\rangle\}$ basis, the matrix elements of J_z and I_z become for $F > 0$

$$\langle nlsjIFm_F|\,J_z\,|nlsjIFm_F\rangle = \langle jIF\,\|J\|\,jIF\rangle\,\langle Fm_F|\,F_z\,|Fm_F\rangle \tag{5.118a}$$

$$\langle nlsjIFm_F|\,I_z|nlsjIFm_F\rangle = \langle jIF\,\|I\|\,jIF\rangle\,\langle Fm_F|\,F_z\,|Fm_F\rangle\,, \tag{5.118b}$$

where the proportionality constants $\langle jIF\,\|J\|\,jIF\rangle$ and $\langle jIF\,\|I\|\,jIF\rangle$ are *reduced matrix elements* and *independent of* m_F. Note that since $F_z = J_z + I_z$ we find by adding Eqs. (5.118a) and (5.118b) that the reduced matrix elements add up to unity,

$$\langle jIF||J||jIF\rangle + \langle jIF||I||jIF\rangle = 1 = \langle lsj\,\|F\|\,lsj\rangle \quad (\text{for } F > 0). \tag{5.119}$$

With the aid of this expression we obtain for $F > 0$

$$g_F = \begin{cases} -g_I(m_e/m_p) + [g_J + g_I(m_e/m_p)]\langle jIF||J||jIF\rangle \simeq g_J\langle jIF||J||jIF\rangle & \text{for } j \neq 0 \\ -g_I(m_e/m_p) & \text{for } j = 0. \end{cases} \tag{5.120}$$

Note that the derivation is valid for arbitrary values of **J** and **I**. For hydrogen-like atoms the condition $j \neq 0$ is always satisfied. The projections of **I** and **J** along the total angular momentum vector **F**, *i.e.* $\mathbf{I}\cdot\mathbf{F}$ and $\mathbf{J}\cdot\mathbf{F}$, are conserved while **F** precesses about **B**,

$$\langle nlsjIFm_F|\, J_z\, |nlsjIFm_F\rangle = \frac{\langle jIFm_F|\,(\mathbf{J}\cdot\mathbf{F})\,F_z\,|jIFm_F\rangle}{\langle jIFm_F|\,\mathbf{F}^2\,|jIFm_F\rangle}. \tag{5.121}$$

Since $\mathbf{J}\cdot\mathbf{F}$ is diagonal in the low-field basis,

$$\mathbf{J}\cdot\mathbf{F} = \mathbf{J}^2 + \mathbf{J}\cdot\mathbf{I} = \mathbf{J}^2 + \tfrac{1}{2}[\mathbf{F}^2 - \mathbf{J}^2 - \mathbf{I}^2] = \tfrac{1}{2}[\mathbf{F}^2 + \mathbf{J}^2 - \mathbf{I}^2],$$

we find by comparison with Eq. (5.118a)

$$\langle jIF||J||jIF\rangle = \frac{\langle jIFm_F|\,\mathbf{J}\cdot\mathbf{F}\,|jIFm_F\rangle}{\langle jIFm_F|\,\mathbf{F}^2\,|jIFm_F\rangle} = \frac{F(F+1) + j(j+1) - I(I+1)}{2F(F+1)} \quad \text{for } F > 0. \tag{5.122}$$

which is independent of m_F, as it should for a reduced matrix element. The same result can be obtained with the algebraic procedure demonstrated in Problem 4.15.

For the important special case $j = \frac{1}{2}$ the expression (5.120) reduces to the form

$$g_F \simeq \pm g_J \frac{1}{2I+1} \quad (F = I \pm \tfrac{1}{2};\ j = \tfrac{1}{2};\ F > 0). \tag{5.123}$$

For $F = 0$ also $m_F = 0$ and g_F has no meaning. Note that g_F is always *negative* for the state with the *smallest* multiplicity $\left(F = I - \frac{1}{2} > 0\right)$, *i.e.* for electronic and nuclear spins coupled in opposite directions. This is intuitively clear because for $I \geq 1$ the nuclear spin dominates the angular momentum of the $j = \frac{1}{2}$ atom whereas the electron determines the magnetic moment. Hence the atoms behave like a classically spinning ball of positive charge. For similar reasons, g_F is always *positive* for the state with the *highest* multiplicity $\left(F = I + \frac{1}{2}\right)$. The sign of the nuclear gyromagnetic ratio is irrelevant in this context because the contribution of the nuclear magnetic moment is negligibly small anyhow.

5.4.5 Low-field limit - quadratic Zeeman shift

In the expansion (5.102) we obtained the expression for the *quadratic Zeeman shift* for the special case $j = \frac{1}{2}$.[1] For arbitrary value of j the quadratic shift can be calculated with the aid of second-order perturbation theory,

$$\Delta E^{(2)Z}_{F,m_F}\left(n\,^{2s+1}L_j\right) = \sum_{F'} \frac{|\langle IjF'm_F|g_J\mu_B J_z - g_I\mu_N I_z|IjFm_F\rangle B/\hbar|^2}{E_{F,m_F} - E_{F',m_F}}. \tag{5.124}$$

---

[1] This quadratic shift is not to be confused with the diamagnetic contribution, which is much weaker but quadratic in B at the level of the Hamiltonian.

Note that this contribution indeed scales like B^2. As J_z and I_z conserve m_F the Zeeman operator couples states of different $F' \in \{|I-j|, \cdots, I+j\}$ for given value of m_F.

We demonstrate the procedure for the special case $j = \frac{1}{2}$, in which only two hyperfine levels are coupled and the result is already known from Eq. (5.102). The energy splitting may be approximated by its zero-field value; i.e., by the interval rule (5.62), $\Delta W_F = a_{\rm hfs} F = a_{\rm hfs}(I + \frac{1}{2})$; note that the energy splitting also has a field dependence but as this leads to a cubic shift it is not included here. The quadratic shifts of the upper ($F = I + \frac{1}{2}$) and lower ($F = I - \frac{1}{2}$) hyperfine levels (of given m_F) only differ in the sign, with the quadratic shifts of the upper manifold all positive and those of the lower manifold all negative. For the upper/lower manifold the quadratic shift becomes

$$\Delta E^{(2)Z}_{F,m_F}\left(n\,^{2s+1}L_j\right) = \pm \frac{|\langle IjF'm_F|g_J\mu_B J_z - g_I\mu_N I_z|IjFm_F\rangle B/\hbar|^2}{a_{\rm hfs}(I+\frac{1}{2})}. \tag{5.125}$$

where $F' = I + \frac{1}{2}$ refers to the upper and $F = I - \frac{1}{2}$ to the lower manifold. Adding and subtracting $g_I\mu_N J_z$ and using the orthogonality relation $\langle IjF'm_F|IjFm_F\rangle = 0$ as well as the definitions (5.90a) and (5.90b) the Eq. (5.125) can be reduced to

$$\Delta E^{(2)Z}_{F,m_F}\left(n\,^{2s+1}L_j\right) = \pm a_{\rm hfs}(I+\tfrac{1}{2})\,(B/B_{\rm hfs})^2\,|\langle IjF'm_F|J_z/\hbar|IjFm_F\rangle|^2. \tag{5.126}$$

The matrix element follows by Clebsch-Gordan decomposition

$$\langle IjF'm_F|J_z/\hbar|IjFm_F\rangle = \sum_{m_j,m_I} m_j \langle F'm_F|Im_I jm_j\rangle\langle Im_I jm_j|Fm_F\rangle. \tag{5.127}$$

Noting that $m_I = m_F - m_j$ and turning to $3j$ symbols (using the property that $2I+1+F+F'$ and $2I - 2m_F$ are always even) we find

$$\langle IjF'm_F|J_z/\hbar|IjFm_F\rangle = \sum_{m_j=\pm1/2} m_j\sqrt{2(I+1)2I}\begin{pmatrix} I & I+\frac{1}{2} & \frac{1}{2} \\ (m_F-m_j) & -m_F & m_j \end{pmatrix}\begin{pmatrix} I & I-\frac{1}{2} & \frac{1}{2} \\ (m_F-m_j) & -m_F & m_j \end{pmatrix}. \tag{5.128}$$

Using Eqs. (J.17) this becomes

$$\langle IjF'm_F|J_z/\hbar|IjFm_F\rangle = \sum_{m_j=\pm1/2} (-1)^{m_j-\frac{3}{2}}\frac{1}{2}\sqrt{\frac{I+1\pm(m_F-m_j)}{(2I+1)}}\sqrt{\frac{I\mp(m_F-m_j)}{(2I+1)}}. \tag{5.129}$$

Both the $m_j = -\frac{1}{2}$ and the $m_j = \frac{1}{2}$ contribution yield the same result and adding these we obtain

$$\langle IjF'm_F|J_z/\hbar|IjFm_F\rangle = \tfrac{1}{2}\sqrt{1-\tilde{m}_F^2}. \tag{5.130}$$

Substituting this expression into Eq. (5.126) we arrive at

$$\Delta E^{(2)Z}_{F,m_F}\left(n\,^{2s+1}L_j\right) = \pm\tfrac{1}{4}a_{\rm hfs}(I+\tfrac{1}{2})\,(B/B_{\rm hfs})^2\,(1-\tilde{m}_F^2). \tag{5.131}$$

Note that we indeed regained the quadratic term of Eq. (5.102). Since $\tilde{m}_F^2 \le 1$ we re-established that the quadratic Zeeman shift is *largest* for the state of *smallest* $|m_F|$. This is important in particular for phenomena which are insensitive for the linear Zeeman effect (see e.g. [60]).

5.5 Ground state hyperfine structure of hydrogen-like atoms

5.5.1 Hydrogen (^{1}H) in the electronic ground state $^2S_{1/2}$ ($I = 1/2$)

The electronic ground state of atomic hydrogen is the $1\,^2S_{1/2}$ state. In Fig. 5.8 the low-field limit can be recognized for the case of the ground state of atomic hydrogen ($j = s = I = \frac{1}{2}$; $l = 0$).

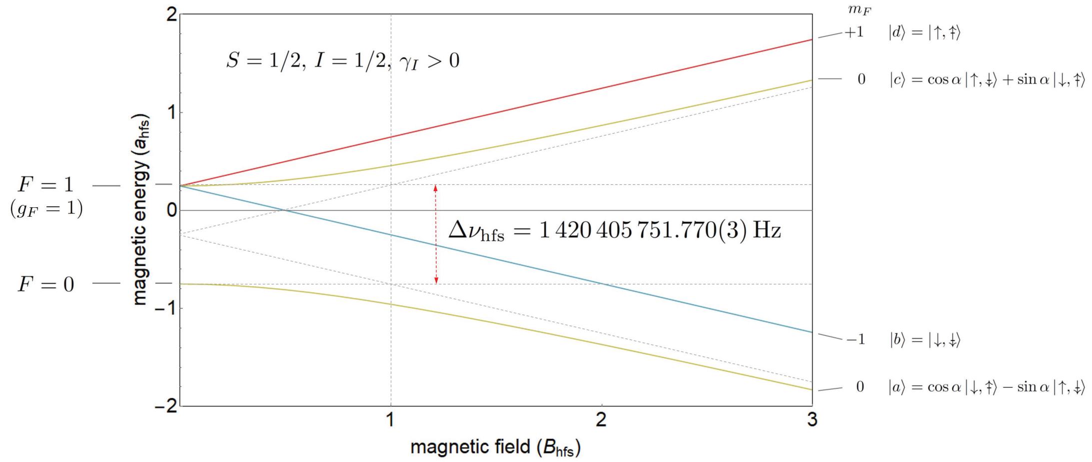


Figure 5.8: Breit-Rabi diagram for the hyperfine structure of atomic hydrogen in its ground state $1^2S_{1/2}$. Indicated is the clock transition ($\lambda = 21\,\text{cm}$) on which the hydrogen maser is based and famous for its astrophysical importance. The dotted lines correspond to the high-field asymptotes - see Eq. (5.111). The magnetic field is given in units of the hyperfine crossover field $B_{\text{hfs}} = |\mathbf{B}_{\text{hfs}}| = a_{\text{hfs}}/\mu_+ = 507$ Gauss. Note that the upper (lower) branches correspond in high fields with electron spin up (down). The zero-field hyperfine splitting is calculated to an accuracy of $1 : 10^4$ in Section 5.3.6. The limiting behavior is: a) $B \to \infty : \sin\alpha \to 0$; b) $B \to 0 : \cos^2\alpha \to {}^1\!/_2$. The upper two levels intersect at $B \approx 16.7$ T.

Because $l = 0$ there is no orbital magnetic moment and $g_J = g_e \simeq 2$. For $F = I + s = 1$ we observe the splitting into a triplet and using Eq. (5.120) we calculate for the effective g factor $g_F \simeq 1$. For the stretched spin state $m_F = 1$ the Zeeman shift is given by $\Delta E^{\text{Z}}_{F,m_F} \simeq \mu_B B$, which is exactly what we expect because neglecting the nuclear spin we are left with the contribution of the electron spin, $\Delta E^{\text{Z}}_{s,m_s} = g_e \mu_B m_s B \simeq \mu_B B$.

Accurate expressions for the field dependence of the hyperfine levels are obtained with Eqs. (5.92) and (5.105). For nuclear spin $I = \frac{1}{2}$, electronic spin $s = \frac{1}{2}$ and $\gamma_I > 0$ we obtain (see Fig. 5.8)

$$\varepsilon_\pm(m_F) = \begin{cases} +\frac{1}{4}a_{\text{hfs}} \pm \frac{1}{2}\mu_- B & m_F = \pm 1 \\ -\frac{1}{4}a_{\text{hfs}} \pm \frac{1}{2}a_{\text{hfs}}\sqrt{1 + (B/B_{\text{hfs}})^2} & m_F = 0\,, \end{cases} \tag{5.132}$$

The hyperfine field is $B_{\text{hfs}} = a_{\text{hfs}}/\mu_+ \simeq 507$ Gauss. For $m_F = 0$ the coupled eigenstates are[1]

$$|F_{\max}, 0\rangle = +\cos\alpha|\uparrow ⇟\rangle + \sin\alpha\,|\downarrow ⇞\rangle, \tag{5.133a}$$

$$|F_{\min}\,, 0\rangle = -\sin\alpha\,|\uparrow ⇟\rangle + \cos\alpha|\downarrow ⇞\rangle. \tag{5.133b}$$

The phase factors determining the sign of the coefficients are chosen in accordance with the phase convention for the $(I = \frac{1}{2} \times S = \frac{1}{2})$ Clebsch-Gordan decomposition of Appendix J.5.1. The field dependence of the coefficients is given by Eq. (5.99), which reduces for small magnetic fields to

$$A^2_\pm \simeq \tfrac{1}{2}\left[1 \pm (B/B_{\text{hfs}})^2 + \cdots\right] \qquad (B \ll |B_{\text{hfs}}|). \tag{5.134}$$

Using this procedure we find for the coupled states $|a\rangle$, $|b\rangle$, $|c\rangle$ and $|d\rangle$ (in order of increasing energy)

$$\begin{aligned} |d\rangle &= |\uparrow ⇞\rangle \\ |c\rangle &= +\cos\alpha|\uparrow ⇟\rangle + \sin\alpha|\downarrow ⇞\rangle \\ |b\rangle &= |\downarrow ⇟\rangle \\ |a\rangle &= -\sin\alpha|\uparrow ⇟\rangle + \cos\alpha|\downarrow ⇞\rangle. \end{aligned} \tag{5.135}$$

---

[1] In this notation $|\uparrow\rangle = |s, +\frac{1}{2}\rangle$ and $|\downarrow\rangle = |s, -\frac{1}{2}\rangle$ refer to the electron spin and $|⇞\rangle = |I, +\frac{1}{2}\rangle$ and $|⇟\rangle = |I, -\frac{1}{2}\rangle$, to the proton spin.

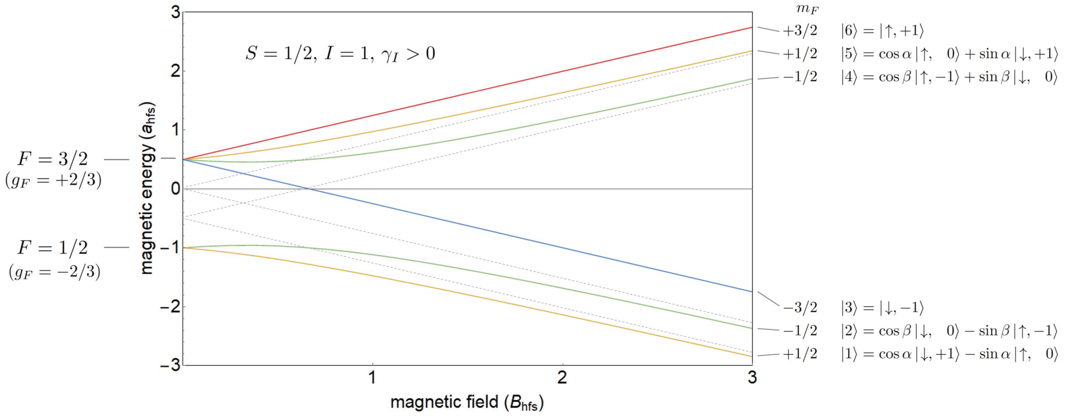


Figure 5.9: Hyperfine structure of the ${}^2S_{1/2}$ electronic ground states of deuterium (D) and ^{6}Li. The magnetic field is given in units of the hyperfine crossover field, 175 Gauss for D and 81.4 Gauss for ^{6}Li. The dotted lines correspond to the high-field asymptotes - see Eq. (5.111). Note that the lower manifold ($F = \frac{1}{2}$) is inverted because its g-factor is negative ($g_F = -2/3$). The limiting behavior is: a) $B \to \infty : \sin\alpha, \sin\beta \to 0$; b) $B \to 0 : \cos^2\alpha, \cos^2\beta \to 2/3,\ 1/3$.

For $B \to 0 : \alpha = \pi/4$, *i.e.* $\sin^2\alpha = \cos^2\alpha = {}^1\!/\!{}_2$ and $|a\rangle = |0,0\rangle$, $|b\rangle = |1,-1\rangle$, $|c\rangle = |1,0\rangle$ and $|d\rangle = |1,+1\rangle$. For $B \to \infty : \alpha \to 0$, *i.e.* $\sin\alpha \simeq \alpha \simeq 0$ and $\cos\alpha \simeq 1$.

5.5.2 Deuterium (^{2}H) and ^{6}Li in the electronic ground state ${}^2S_{1/2}$ $(I = 1)$

The heavy hydrogen isotope deuterium (^{2}H) has one electron. Like hydrogen the ground state configuration is $1s$ and the spin-orbit type is $1\,{}^2S_{1/2}$. The ^{6}Li atom in its electronic ground state has configuration $1s^22s$ and is of the spin-orbit type $2\,{}^2S_{1/2}$. With orbital angular momentum $L = 0$, electronic spin $s = \frac{1}{2}$ and nuclear spin $I = 1$ the quantum number for the total angular momentum can have the values $F = I \pm \frac{1}{2} = \frac{3}{2}, \frac{1}{2}$. Both elements have a positive gyromagnetic ratio. Hence, the Fermi contact interaction is positive ($a_{\rm hfs} > 0$) and the hyperfine structure regular; i.e., the manifold with the highest multiplicity $\left(F = \frac{3}{2}\right)$ has the highest energy. Since $\gamma_I > 0$ the energies of the states are given by

$$\varepsilon_\pm(m_F) = \begin{cases} -g_I\mu_N m_F B - \frac{1}{4}a_{\rm hfs} + \frac{3}{4}a_{\rm hfs}(1 \pm B/B_{\rm hfs}) & m_F = \pm\frac{3}{2} \\ -g_I\mu_N m_F B - \frac{1}{4}a_{\rm hfs} \pm \frac{3}{4}a_{\rm hfs}\sqrt{1 + \frac{4}{3}m_F B/B_{\rm hfs} + (B/B_{\rm hfs})^2} & m_F = \pm\frac{1}{2}\,. \end{cases} \tag{5.136}$$

The hyperfine crossover field $B_{\rm hfs} = (I + \frac{1}{2})a_{\rm hfs}/\mu_+ \simeq 175$ Gauss in the case of deuterium and 81.4 Gauss in the case of ^{6}Li (see Table 5.1). The Breit-Rabi diagram for this case is shown in Fig. 5.9. The upper manifold is regular $\left(g_F = \frac{2}{3}\right)$. The lower manifold $(F = \frac{1}{2})$ is inverted because its g factor is negative $\left(g_F = -\frac{2}{3}\right)$. The phase factors determining the sign of the coefficients are in accordance with the $(1 \times 1/2)$ Clebsch-Gordan decomposition of Appendix J.5.1. Because $m_F = m_I + m_s$ is a good quantum number the eigenstates can be expressed for any field in the form

$$|F_{\rm max}, m_F\rangle = +\cos\alpha|\uparrow; 1, \left(m_F - \tfrac{1}{2}\right)\rangle + \sin\alpha\,|\downarrow; 1, \left(m_F + \tfrac{1}{2}\right)\rangle \tag{5.137a}$$

$$|F_{\rm min}\,, m_F\rangle = -\sin\alpha\,|\uparrow; 1, \left(m_F - \tfrac{1}{2}\right)\rangle + \cos\alpha|\downarrow; 1, \left(m_F + \tfrac{1}{2}\right)\rangle. \tag{5.137b}$$

Here we use the notation of Eq. (5.93a). The field dependence of the coefficients is given by Eq. (5.97), which reduces for small magnetic fields to

$$A_\pm^2 \simeq \frac{1}{2}\left(1 \pm \tfrac{2}{3}m_F\right)\left[1 \pm B/B_{\rm hfs}\left(1 \mp \tfrac{2}{3}m_F\right) + \cdots\right] \quad (B \ll |B_{\rm hfs}|)\,. \tag{5.138}$$

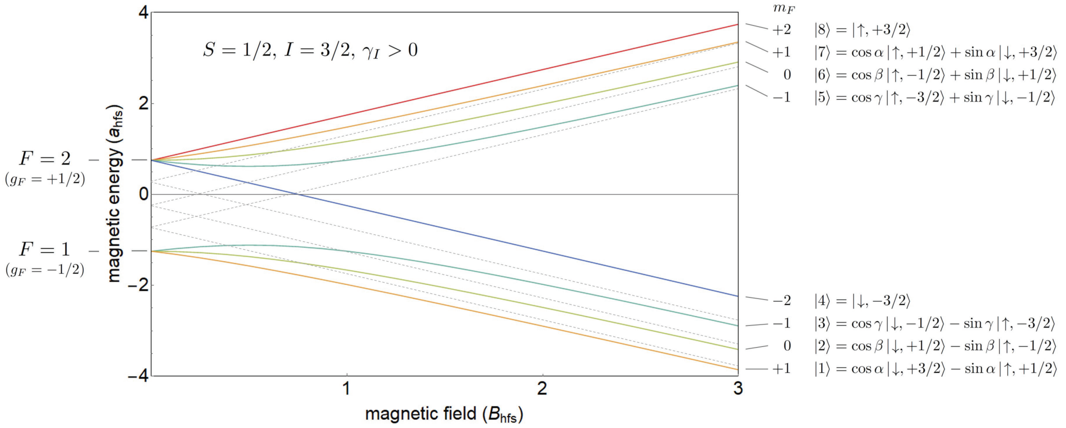


Figure 5.10: Hyperfine structure of $^2S_{1/2}$ electronic ground states with $I=\frac{3}{2}$. The magnetic field is given in units of the hyperfine crossover field, 287 G for ^{7}Li, 633 G for ^{23}Na, 165 G for ^{39}K, 90 G for ^{41}K and 2.44 kG for ^{87}Rb. The dotted lines correspond to the high-field asymptotes - see Eq. (5.111). Note that the lower manifold ($F=1$) is inverted because its g-factor is negative, $g_F=-\frac{1}{2}$. The limiting behavior is: a) $B\to\infty:\sin\alpha,\sin\beta,\sin\gamma\to 0$; b) $B\to 0:\cos^2\alpha,\cos^2\beta,\cos^2\gamma\to 3/4,2/4,1/4$.

5.5.3 The alkali atoms ^{7}Li, ^{23}Na, ^{39}K, ^{41}K and ^{87}Rb in the electronic ground state $^2S_{1/2}$ $(I=3/2)$

The electronic ground state configuration of these alkali atoms is [He]2s for ^{7}Li, [Ne]3s for ^{23}Na, [Ar]4s for ^{39}K and ^{41}K, and [Kr]5s for ^{87}Rb. In all cases the term type is $^2S_{1/2}$. With orbital angular momentum $L=0$, electronic spin $s=\frac{1}{2}$ and nuclear spin $I=\frac{3}{2}$ the quantum number for the total angular momentum can have the values $F=I\pm\frac{1}{2}=2,1$. All elements mentioned have a positive gyromagnetic ratio. Hence, the Fermi contact interaction is positive ($a_{\rm hfs}>0$) and the hyperfine structure regular; i.e., the manifold with the highest multiplicity ($F=2$) has the highest energy. For nuclear spin $I=\frac{3}{2}$, electronic spin $s=\frac{1}{2}$ and $\gamma_I>0$ the energies of the coupled states are given by

$$\varepsilon_\pm(m_F)=\begin{cases}-g_I\mu_N m_F B-\frac{1}{4}a_{\rm hfs}+a_{\rm hfs}(1\pm B/B_{\rm hfs}), & m_F=\pm2\\ -g_I\mu_N m_F B-\frac{1}{4}a_{\rm hfs}\pm a_{\rm hfs}\sqrt{1+m_F B/B_{\rm hfs}+(B/B_{\rm hfs})^2}, & m_F=\pm1\\ -g_I\mu_N m_F B-\frac{1}{4}a_{\rm hfs}\pm a_{\rm hfs}\sqrt{1+(B/B_{\rm hfs})^2}, & m_F=0\,.\end{cases}\tag{5.139}$$

The hyperfine crossover field $B_{\rm hfs}=(I+\frac{1}{2})a_{\rm hfs}/\mu_+\simeq 287$ G in the case of ^{7}Li, 633 G in the case of ^{23}Na, 165 G in the case of ^{39}K, 90 G in the case of ^{41}K and 2.44 kG in the case of ^{87}Rb (see Table 5.1). The Breit-Rabi diagram for this case is shown in Fig. 5.10. The upper manifold $(F=2)$ is regular $\left(g_F=\frac{1}{2}\right)$. The lower manifold $(F=1)$ is inverted because its g factor is negative $\left(g_F=-\frac{1}{2}\right)$. The phase factors determining the sign of the coefficients are in accordance with the $(3/2\times 1/2)$ Clebsch-Gordan decomposition of Appendix J.5.1. Because $m_F=m_I+m_s$ is a good quantum number the eigenstates can be expressed for any field in the form

$$|F_{\rm max},m_F\rangle=+\cos\alpha_+|\uparrow;\tfrac{3}{2},\left(m_F-\tfrac{1}{2}\right)\rangle+\sin\alpha_+|\downarrow;\tfrac{3}{2},\left(m_F+\tfrac{1}{2}\right)\rangle\tag{5.140a}$$

$$|F_{\rm min}\,,m_F\rangle=-\sin\alpha_+|\uparrow;\tfrac{3}{2},\left(m_F-\tfrac{1}{2}\right)\rangle+\cos\alpha_+|\downarrow;\tfrac{3}{2},\left(m_F+\tfrac{1}{2}\right)\rangle.\tag{5.140b}$$

Here we use the notation of Eq. (5.93a). For $0<|m_F|<2$ the field dependence of the coefficients is given by Eq. (5.97), which reduces for small magnetic fields to

$$A_\pm^2\simeq\tfrac{1}{2}\left(1\pm\tfrac{1}{2}m_F\right)\left[1\pm B/B_{\rm hfs}\left(1\mp\tfrac{1}{2}m_F\right)+\cdots\right]\qquad(B\ll|B_{\rm hfs}|)\,.\tag{5.141}$$

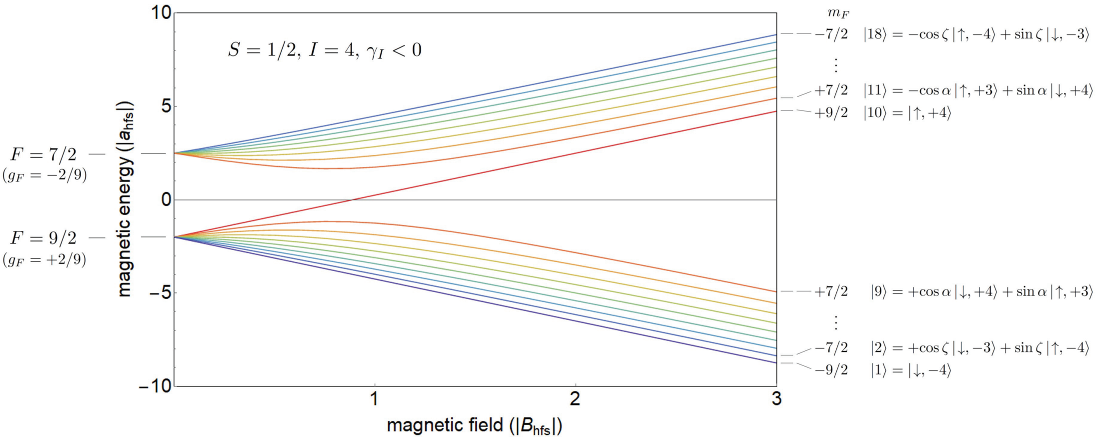


Figure 5.11: Example of an inverted hyperfine structure: the electronic ground state ${}^2S_{1/2}$ of ^{40}K $(I = 4, \gamma_N < 0)$. The magnetic field is given in units of the hyperfine crossover field, $|B_{\rm hfs}| = 459$ G for ^{40}K. Note that the manifold with the smallest multiplicity $(F = \frac{7}{2})$ is also inverted because its g-factor is negative, $g_F = -2/9$. The limiting behavior is: a) $B \to \infty : \sin\alpha, \cdots, \sin\zeta \to 0$; b) $B \to 0 : \cos\alpha, \cdots, \cos\zeta \to 8/9, \cdots, 1/9$.

For $m_F = 0$ the coefficients are given by Eq. (5.99) and become in low field

$$A_\pm^2 \simeq \tfrac{1}{2}\left[1 \pm (B/B_{\rm hfs})^2 + \cdots\right] \qquad (B \ll |B_{\rm hfs}|)\,. \tag{5.142}$$

5.5.4 Potassium-40 in the electronic ground state ${}^2S_{1/2}\,(I = 4)$ - negative hyperfine shift

The final example to be discussed in this section on the hyperfine interaction in hydrogen-like systems is the electronic ground state of ^{40}K. This case is special because the ^{40}K nucleus has a *negative* gyromagnetic ratio, which implies that also the Fermi contact interaction is negative ($a_{\rm hfs} < 0$). Like the other atomic isotopes of potassium the electronic ground-state configuration is [Ar]4s with term type ${}^2S_{1/2}$. With orbital angular momentum $L = 0$, electronic spin $s = \frac{1}{2}$ and nuclear spin $I = 4$ the quantum number for the total angular momentum can have the values $F = I \pm \frac{1}{2} = \frac{9}{2}, \frac{7}{2}$. The manifold with the lowest multiplicity $\left(F = \frac{7}{2}\right)$ has the highest energy because the hyperfine shift is *negative* $(a_{\rm hfs} < 0)$. This is called an inverted hyperfine structure. For nuclear spin $I = 4$, electronic spin $s = \frac{1}{2}$ and $\gamma_I < 0$ the energies are given by

$$\varepsilon_\pm(m_F) = \begin{cases} -g_I\mu_N m_F B - \frac{1}{4}a_{\rm hfs} + \frac{9}{4}a_{\rm hfs}(1 \pm B/\,|B_{\rm hfs}|), & m_F = \pm\frac{9}{2} \\ -g_I\mu_N m_F B - \frac{1}{4}a_{\rm hfs} \pm \frac{9}{4}\,|a_{\rm hfs}|\,\sqrt{1 + \frac{4}{9}m_F B/B_{\rm hfs} + (B/B_{\rm hfs})^2} & |m_F| = \frac{1}{2}, \cdots \frac{7}{2}\,. \end{cases} \tag{5.143}$$

The hyperfine crossover field $B_{\rm hfs} = (I + \frac{1}{2})a_{\rm hfs}/\mu_+ \simeq -459$ G (see Table 5.1). The Breit-Rabi diagram for this case is shown in Fig. 5.11. An example of a corresponding Stern-Gerlach image is shown in Fig. 5.12.

The upper manifold $(F = \frac{7}{2})$ is inverted because its g factor is negative $\left(g_F = -\frac{2}{9}\right)$. The lower manifold $(F = \frac{9}{2})$ is regular $\left(g_F = \frac{2}{9}\right)$. The phase factors determining the sign of the coefficients are in accordance with the $(4 \times 1/2)$ Clebsch-Gordan decomposition of Appendix J.5.1. Because $m_F = m_I + m_s$ is a good quantum number the eigenstates can be expressed for any field in the form

$$|F_{\rm min}\,, m_F\rangle = -\cos\alpha|\uparrow; 4, \left(m_F - \tfrac{1}{2}\right)\rangle + \sin\alpha\,|\downarrow; 4, \left(m_F + \tfrac{1}{2}\right)\rangle \tag{5.144a}$$

$$|F_{\rm max}, m_F\rangle = +\sin\alpha\,|\uparrow; 4, \left(m_F - \tfrac{1}{2}\right)\rangle + \cos\alpha|\downarrow; 4, \left(m_F + \tfrac{1}{2}\right)\rangle. \tag{5.144b}$$

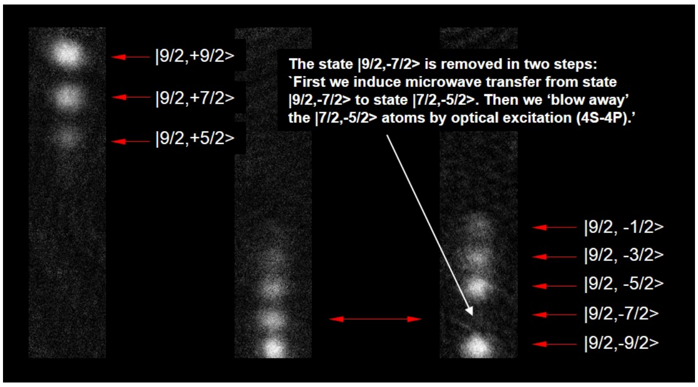


Figure 5.12: Stern-Gerlach imaging of the hyperfine states of ^{40}K. From left to right the atoms are first transferred (using radio waves) from the upper half to the lower half of the lower hyperfine manifold ($F = \frac{9}{2}$ - see Fig. 5.11). Subsequently, the atoms are removed from the $|\frac{9}{2}, -\frac{7}{2}\rangle$ state in two steps: first they are transferred from the state $|\frac{9}{2}, -\frac{7}{2}\rangle$ to the state $|\frac{7}{2}, -\frac{5}{2}\rangle$ of the upper hyperfine manifold from which they are "blown away" with resonant light on the $4S \to 4P$ transition.

Here we use the notation of Eq. (5.93b). The field dependence of the coefficients is given by Eq. (5.97), which reduces for small magnetic fields to

$$A_\pm^2 \simeq \tfrac{1}{2}\left(1 \pm \tfrac{2}{9} m_F\right)\left[1 \pm B/B_{\mathrm{hfs}}\left(1 \mp \tfrac{2}{9} m_F\right) + \cdots\right] \qquad (B \ll |B_{\mathrm{hfs}}|)\,. \tag{5.145}$$

6

Helium-like atoms

Once we have more than one electron orbiting the nucleus the Hamiltonian for the electronic motion rapidly becomes more complicated. Aside from the attraction by the nuclear charge we have to account for the electrostatic repulsion among the electrons. At a more refined level we not only have to deal with the *spin-orbit interaction* between the spin and the orbit of the individual electrons but also with the interaction of the spin of one electron with the orbit of the other electrons; i.e., the *spin-other-orbit interaction.* Likewise we have *magnetic-dipole interaction* between the spins of the various electrons and *magnetic-orbit-orbit interaction* between the orbital moments of the electrons. Needless to say that even a glance at the Hamiltonian shows that we are facing a formidable task.

Fortunately a lot can be understood about many-electron atoms without pursuing the accuracy possible for hydrogen. Actually, bearing in mind that the electrostatic interaction is by far the strongest interaction in the atom, it is pointless to consider refinements before we have a theory at hand in which the electronic charge is distributed over the atom in the best possible way. If the electronic wavefunction is known this requirement is of course implicitly satisfied but, unfortunately, for atoms with more than one electron exact analytical solutions cannot be obtained. We find that at some point we have to rely on some semiclassical approximation to obtain a result for the electrostatic energy. To our advantage the electronic repulsion favors the spherical form, which is important from the theoretical point of view as it acts to preserve the *central symmetry.* Importantly, the distribution of the charge is strongly affected by a quantum phenomenon completely absent in hydrogenic atoms: *exchange.* The exchange phenomenon is not captured in the Hamiltonian but has to be added to the theory as a completely new element. It strongly affects all many-electron systems. In atoms it tends to align the electron spins in a parallel fashion, in molecules anti-parallel, in many-body systems like gases and solids it gives rise to various forms of magnetism.

Exchange emerges as an effective electrostatic interaction resulting from kinematic correlations between the electrons. These correlations have a purely quantum mechanical origin and only occur between *identical particles.* As these are *indistinguishable* the Hamiltonian is manifestly invariant under exchange of two of these particles [33]. So the exchange operator commutes with the Hamiltonian which implies that the symmetry under exchange is a conserved quantity (see Appendix F.2.4). The eigenstates of the exchange operator are either *symmetric* or *antisymmetric* [33]. Importantly, linear combinations of these (so called mixed statistics) are *not* observed. To distinguish between the two cases the particles are referred to as *bosons* (symmetric) or *fermions* (antisymmetric) [33]. As the quantum correlations affect the probability of occupation of the single particle states, their presence demands new statistics, *quantum statistics*, *Bose-Einstein statistics* (short: Bose statistics) for bosons and *Fermi-Dirac statistics* (short: Fermi statistics) for fermions [33]. Electrons behave like fermions. This was discovered by Wolfgang Pauli, when he realized that the magnetic fine structure of the optical spectra of the elements could be explained by excluding atomic states in which more than one electron in a given spin state would occupy a given atomic orbital. He formulated this observation in his famous *exclusion principle*: double occupation of a given single-electron state has to be excluded [81]. Bose statistics plays no role in theoretical descriptions of atoms simply because the electrons are fermions and the atom only has a single nucleus. By studying the helium atom we meet the phenomenology essential for the conception of the Pauli principle.

6.1 Heliogenic atoms

To introduce the physics of atoms with more than one electron we turn to the element helium, the archetype of the *heliogenic* atoms. The notion *heliogenic* refers to atomic systems with *exactly two* elementary charges orbiting about the nucleus. Aside from the helium isotopes ^{3}He and ^{4}He, also the negative ion of hydrogen, H$^-$, and the positive ions Li$^+$, Be^{++}, etc., as well as exotic atoms like muonic helium belong to this class.

We start by calculating the *principal structure* of a heliogenic atom/ion with two electrons. In constructing the Hamiltonian, the essential difference with the hydrogenic case is that we include the electrostatic repulsion between the electrons,

$$\mathcal{H}=\sum_{i=1,2}\left(-\frac{\hbar^2}{2m_e}\boldsymbol{\nabla}_i^2-\frac{Ze^2}{4\pi\varepsilon_0 r_i}\right)+\frac{e^2}{4\pi\varepsilon_0 r_{12}}. \tag{6.1}$$

The first term consists of the Schrödinger Hamiltonians of two electrons orbiting a nucleus of charge Ze pinned in space (infinite mass limit), with m_e the electron rest mass and $\mathbf{r}_i$ the position of electron i relative to the atomic center; ε_0 is the *electric permittivity of vacuum*. The second term represents the electrostatic repulsion, with $r_{12}=|\mathbf{r}_1-\mathbf{r}_2|$ the inter-electronic distance. Small corrections related to the fine- and hyperfine structure are not considered in this Hamiltonian.

Before starting the actual work we have a look at the symmetry properties of $\mathcal{H}$. Like the Schrödinger Hamiltonian also $\mathcal{H}$ is invariant under space inversion, which means that the parity of the electronic states is conserved in time (see Appendix F.2.4). As $\mathcal{H}$ is also invariant under rotation of the spatial coordinates about the origin, also the total orbital angular momentum $\mathbf{L}=\boldsymbol{l}_1+\boldsymbol{l}_2$ and its projection L_z are conserved quantities,

$$[\mathbf{L}^2,\mathcal{H}]=0 \quad \text{and} \quad [L_z,\mathcal{H}]=0. \tag{6.2}$$

In other words, L and M_L are good quantum numbers, also in the presence of the repulsion term in the Hamiltonian. The commutation of $\mathbf{L}^2$ and L_z with the repulsion term is demonstrated explicitly in Problem 6.1. The scattering of the electrons reflects itself in the absence of commutation of $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ with $\mathcal{H}$ (see Problem 6.3).

To calculate the energy levels of the atom we start by crudely neglecting the electrostatic repulsion term of Eq. (6.1). In this way we regain a structure similar to that of the hydrogen atom. We obtain the hydrogenic eigenstates $|nlm_l\rangle$ and the eigenvalues are given by $\mathcal{H}$

$$E_n=-\alpha^2 m_e c^2\frac{Z^2}{2n^2}=-hcR_\infty\frac{Z^2}{n^2}. \tag{6.3}$$

Here n,l,m_l are the quantum numbers of the principal atomic structure and R_∞ is the *Rydberg constant* for an electron bound to a nucleus of infinite mass. As, in the given approximation, the Hamiltonian for the electron pair does not contain the Coulomb coupling between the orbital motions of the electrons the pair wavefunction $\psi_a(\mathbf{r}_1,\mathbf{r}_2)=(\mathbf{r}_1,\mathbf{r}_2|u,v)$ can be represented by products of two hydrogenic states $|u\rangle=|nlm_l\rangle$ and $|v\rangle=|n'l'm_{l'}\rangle$,

$$\psi_a(\mathbf{r}_1,\mathbf{r}_2)=\varphi_u(\mathbf{r}_1)\,\varphi_v(\mathbf{r}_2). \tag{6.4}$$

The corresponding energy of the *non-interacting* electronic pair follows from solving the Schrödinger equation for the noninteracting hamiltonian

$$\mathcal{H}_0=\sum_{i=1,2}\left(-\frac{\hbar^2}{2m_e}\boldsymbol{\nabla}_i^2-\frac{Ze^2}{4\pi\varepsilon_0 r_i}\right). \tag{6.5}$$

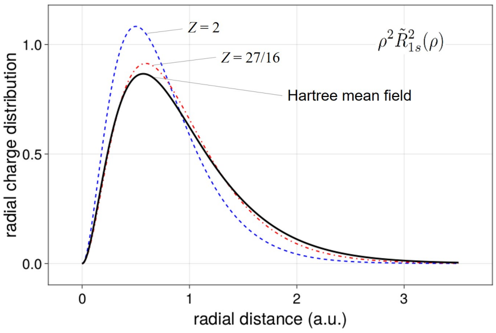


Figure 6.1: Radial distribution function for the electrons in the ground state of the helium atom. The blue dashed line corresponds to a $1s^2$ configuration without interaction. The red dash-dotted line shows the effect of inflation of the charge distribution caused by the Coulomb repulsion. This curve corresponds to an effective atomic number $Z_{1s} = 27/16$. The black curve is the Hartree result.

The solutions are given by

$$E_a^{(0)} = E_{n_u} + E_{n_v} = -hcR_\infty \left(\frac{Z^2}{n_u^2} + \frac{Z^2}{n_v^2} \right), \tag{6.6}$$

which is simply the sum of the energies of the individual electrons in the Coulomb field of the bare nucleus. For example, the helium atom in its electronic ground state consists of a doubly charged nucleus ($Z = 2$) and two equivalent electrons in the *configuration* $1s^2$, both occupying the $|u\rangle = |nlm\rangle = |100\rangle$ orbital of *hydrogenic helium* shown as the solid line in Fig. 6.1. With Eq. (6.6) we calculate for the total energy of the electronic ground state

$$E_{1s^2}^{(0)} = -8hcR_\infty \qquad (Z = 2,\ n = 1), \tag{6.7}$$

which is twice the energy of *hydrogenic helium*.

Problem 6.1. Show that for the orbital angular momentum $\mathbf{L} = \boldsymbol{l}_1 + \boldsymbol{l}_2$, where $\boldsymbol{l}_\nu = -i\hbar\,(\mathbf{r}_\nu \times \boldsymbol{\nabla}_\nu)$ with $\nu \in \{1, 2\}$, the following commutation relations hold:

$$[L_z, 1/r_{12}] = 0 \quad \text{and} \quad [\mathbf{L}^2, 1/r_{12}] = 0,$$

where $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ is the inter-electronic distance.

Solution. The first commutation relation will be derived for the more general case of an arbitrary cartesian component, $[L_i, 1/r_{12}] = 0$ with $i \in \{x, y, z\}$; i.e., $[\mathbf{L}, 1/r_{12}] = 0$. Since $\mathbf{L}$ is a linear differential operator acting on the variables $\mathbf{r}$ and $\mathbf{r}'$ we find

$$[\mathbf{L}, 1/r_{12}]\psi = \mathbf{L}(\psi/r_{12}) - (1/r_{12})\mathbf{L}\psi = \psi\mathbf{L}(1/r_{12}) = \psi\,(\boldsymbol{l}_1 + \boldsymbol{l}_2)\,(1/r_{12}) = 0,$$

where we used $\boldsymbol{l}_1(1/r_{12}) = -\boldsymbol{l}_2(1/r_{12})$ as follows with Eq. (N.21). The second commutation relation follows with the aid of the first relation. Using the Einstein summation convention we find

$$[\mathbf{L}^2, 1/r_{12}] = [L_iL_i, 1/r_{12}] = [L_i, 1/r_{12}]L_i + L_i[L_i, 1/r_{12}] = 0. \qquad \square$$

6.1.1 Electrostatic repulsion versus screening

Not surprisingly the approximation of non-interacting electrons is rather poor. Electrostatic repulsion tends to inflate the $1s$ orbitals as illustrated by the dashed line in Fig. 6.1. This reduces the *binding energy* of the electrons because the average distance to the nucleus increases. Another way to look at this is that one electron will partly *screen* the nuclear charge for the other electron. This effectively reduces the nuclear charge Ze, which also results in a reduction of the binding energy. Using Eq. (6.6), i.e., without the interaction term, we calculate for the *first ionization energy* of the atom 54.4 eV (4 Rydberg), whereas the measured value is only 24.6 eV. Clearly, the screening is important.

To calculate the effect of screening we turn to perturbation theory; it may be that for this purpose the repulsive term is not sufficiently small but at least it will reveal the trend. We first make the Hamiltonian dimensionless by multiplying Eq. (6.1) with $m_e a_0^2/\hbar^2 = 1/\left(2hcR_\infty\right)$, thus switching to Hartree atomic units,

$$\mathcal{H} = \mathcal{H}_0 + \frac{1}{\rho_{12}}, \tag{6.8}$$

where

$$\mathcal{H}_0 = \sum_{i=1,2} h_0^{(i)}, \quad \text{with} \quad h_0^{(i)} = -\tfrac{1}{2}\boldsymbol{\nabla}_i^2 - \frac{Z}{\rho_i}. \tag{6.9}$$

Note that $h_0^{(i)}$ is the Schrödinger Hamiltonian of electron $i \in \{1,2\}$ in the Coulomb field of the bare nucleus, $\rho_i = r_i/a_0$ being its distance to the nucleus (in Bohr radii) and $\boldsymbol{\nabla}_i$ the gradient operator with respect to $\boldsymbol{\rho}_i$. We will treat $\mathcal{H}_0$ as the unperturbed Hamiltonian. The second term of Hamiltonian (6.8),

$$\mathcal{H}' = \frac{1}{\rho_{12}}, \tag{6.10}$$

acts as the perturbation by the electrostatic repulsion.

To first-order in perturbation theory, $\Delta\varepsilon_{1s} = \Delta\varepsilon_{1s}^{(1)}$, the energy shift is given by

$$\Delta\varepsilon_{1s} = \langle 1s, 1s|\mathcal{H}'|1s, 1s\rangle. \tag{6.11}$$

Let us postpone the actual evaluation of this expression until Section 6.4.3 and simply use the result:

$$\Delta\varepsilon_{1s} = \int_0^\infty U_{1s}(\rho)\tilde{R}_{1s}^2(\rho)\rho^2 d\rho, \tag{6.12}$$

where $\tilde{R}_{1s}(\rho) = Z^{3/2}2e^{-Z\rho}$ is the $1s$ radial wavefunction and

$$U_{1s}(\rho) = \frac{1}{\rho}\left[1 - (1 + Z\rho)e^{-2Z\rho}\right], \tag{6.13}$$

is the expectation value of the Coulomb repulsion energy of the atom with one electron fixed at position ρ. Note that the form (6.12) suggests to interpret $U_{1s}(\rho)$ semiclassically as some kind of effective potential field for *one* of the $1s$ electrons (the *spectator electron*) in the electrostatic field of the other $1s$ electron (the *screening electron*),

$$\Delta\varepsilon_{1s} = \langle 1s|U_{1s}(\rho)|1s\rangle. \tag{6.14}$$

The effective potential $U_{1s}(\rho)$ is called the *potential energy of screening* (short: *screening potential*), a notion from classical electrostatics used to account for the field associated with a charge-density distribution. In the present context the charge density arises from a semiclassical approximation in which the charge of the $1s$ electron, $-e$, is distributed in space proportional to the probability density of the wavefunction. In the upper panel of Fig. 6.2*a* the screening potential is plotted for the helium ground state. It is instructive to write the screening potential in the form

$$U_{1s}(\rho) = \sigma_{1s}(\rho)/\rho, \tag{6.15}$$

where $\sigma_{1s}(\rho)$ is the *screening charge* at distance ρ from the nucleus. This screening charge can be regarded as a generalization of the *screening constant* introduced in Section 4.6.2.1. As shown in

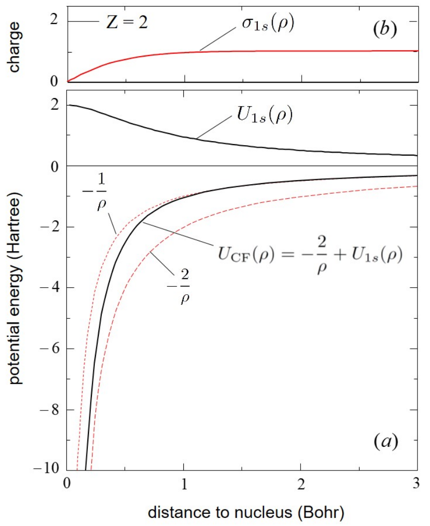


Figure 6.2: (a) Effective central field potential $U_{\rm CF}(\rho) = Z_{1s}(\rho)/\rho$ and screening potential $U_{1s}(\rho)$ for the helium atom ($Z = 2$) in its ground state. For comparison also the $-1/\rho$ hydrogenic Coulomb potential is shown (short dash) as well as the $-2/\rho$ potential of the unscreened helium nucleus (long dash), demonstrating that for an electron at distances $\rho \gg 1$ a full elementary charge of the nucleus is screened by "the other" electron; (b) Screening charge $\sigma_{1s}(\rho)$ corresponding to the screening potential $U_{1s}(\rho)$.

Fig. 6.2b $\sigma_{1s}(\rho)$ grows from *zero* at $\rho = 0$ to *unity* for $\rho \gg 1$. Accordingly, far from the nucleus $U_{1s}(\rho)$ falls off like $1/\rho$. Apparently, for the spectator electron at large distance a full elementary charge of the nucleus is screened by the other electron, which is intuitively correct. Close to the nucleus the screening potential reaches the constant value Z,

$$U_{1s}(\rho) = Z - \tfrac{2}{3}Z^3\rho^2 + \cdots \quad \text{for } \rho \ll 1. \tag{6.16}$$

The constant Z represents the *mean electrostatic repulsion energy* (in atomic units) between the spectator electron close to the nucleus and the surrounding cloud of electronic charge. The precise value depends on the charge distribution but it being constant implies that the screening can be neglected close to the nucleus (where the nuclear scalar potential diverges as $-Z/\rho$). This is in accordance with the absence of screening charge, $\sigma_{1s}(0) = 0$, and certainly also intuitively correct.

Combining $U_{1s}(\rho)$ with the potential energy of the spectator electron in the Coulomb field of the nucleus we obtain the *effective central potential field* shown in the lower panel of Fig. 6.2a,

$$U_{\rm CF}(\rho) = -\frac{Z}{\rho} + U_{1s}(\rho) = -\frac{1}{\rho}\left[Z - \sigma_{1s}(\rho)\right] = -\frac{Z_{1s}(\rho)}{\rho}. \tag{6.17}$$

The function $Z_{1s}(\rho)$ is the *effective nuclear charge* for a $1s$ electron at distance r from the nucleus as introduced in Section 4.4.2,

$$Z_{1s}(\rho) = (Z - \sigma_{1s}) = \begin{cases} Z - Z\,\rho + \tfrac{2}{3}Z^3\rho^3 + \cdots & \text{for } \rho \ll 1 \\ 1 + \cdots & \text{for } \rho \gg 1. \end{cases} \tag{6.18}$$

Sometimes preference is given to a multiplicative screening correction to the Coulomb field,

$$U_{\rm CF}(\rho) = -\frac{Z}{\rho}\tilde{F}(\rho), \tag{6.19}$$

The correction factor,

$$\tilde{F}(\rho) \equiv Z(\rho)/Z, \tag{6.20}$$

is called the screening function for a given charge-density distribution $Z(\rho)$. A famous example is the Thomas-Fermi function [109, 41].

So, screening by an s electron conserves the *central symmetry* (hence also the rotational structure derived in Chapter 1 for central potentials). The potential $U_{\mathrm{CF}}(\rho)$ is the first example of a class of electrostatic potentials which are *central* but do not have the familiar $1/\rho$ dependence of the Coulomb potential. Close to the nucleus the effective nuclear charge, $Z(\rho)$, approaches the bare nuclear charge, Z. Far from the nucleus, $Z(\rho)$ approaches *unity* because the nucleus can be screened by all electrons but not by the spectator electron itself. Although the concept of an effective potential is attractive from the computational point of view it should be remembered that it represents a semiclassical approximation in which *correlations* in the relative motion of the electrons are neglected. Such correlations have to be present in view of the strong Coulomb repulsion between two $1s$ electrons at short distance.[1] The concept of the central electrostatic potential is the stepping stone toward the description of many-electron atoms and we return to this concept in Chapter 8.

To conclude this section we evaluate the screening integral (6.12),

$$\Delta\varepsilon_{1s} = \int_0^\infty \left[1 - (1 + Z\rho)e^{-2Z\rho}\right]\left[Z^{3/2}2e^{-Z\rho}\right]^2 \rho d\rho = \tfrac{5}{8}Z. \tag{6.21}$$

Hence, the total energy of a helium-like atom is found to be

$$E^{(1)}_{1s^2} = E^{(0)}_{1s^2} + \Delta E_{1s} = hcR_\infty\left[-2Z^2 + \tfrac{5}{4}Z\right]. \tag{6.22}$$

Here we restored the dimension by multiplying with the Hartree, $E_H = 2hcR_\infty$. For the helium ground state ($Z = 2$) this implies $E^{(1)}_{1s^2} = -(11/2)hcR_\infty = -74.8$ eV. The energy shift ΔE_{1s} of the ground state level is illustrated in the Term diagrams of Fig. 6.3. Adding the perturbation reduces the binding by 34.4 eV, which is the screening effect we are looking for. Comparing the energy obtained for the ground state of He with the energy of the ground state of the He$^+$ ion, $E_{1s} = -4hcR_\infty = -54.4$ eV, we see that to first-order in perturbation theory the ionization energy of the second electron is $-E^{(0)}_{1s} - \Delta E_{1s} = hcR_\infty\left[Z^2 - \tfrac{5}{4}Z\right] = 20.4$ eV, less than the measured value of 24.6 eV but much better than the 54.4 eV obtained in zero order (without screening). Apparently the screening correction of 34.4 eV overestimates the experimental value (29.8 eV) by 4.2 eV (14%), actually not bad for a first try but of course poor as compared to the results obtained for hydrogen.

6.1.2 Variational calculation

At this point we have established that first order perturbation theory is inadequate to provide a satisfactory description of even the simplest many-electron atom. We could proceed by calculating the second order terms but this turns out to make little sense because the convergence of the perturbation series is too slow. Apparently, the effective potential $U_{1s}(\rho)$ in combination with the $1s$ wavefunction of the unperturbed atom does not properly represent the actual charge distribution in the atom. This observation suggests to try a different approach and use Eq. (6.11) as an Ansatz to minimize the energy using the variational procedure of Appendix I.2.1. We vary the $1s$ wavefunction,

$$\tilde{R}_{1s}(\rho) = Z^{3/2}2e^{-Z\rho},$$

treating Z as the variational parameter. Interestingly, note that in this variational procedure both the $1s$ shape and the normalization remain conserved. To avoid confusion with the variational parameter (Z) the nuclear charge number is denoted by $\mathcal{Z}$. Using Eqs. (2.52) and (2.49) we obtain

$$E^{(0)}_{1s^2} = 2\langle 1s| - \frac{1}{2\rho}\frac{\partial^2}{\partial\rho^2}\rho + \frac{l(l+1)}{2\rho^2} - \frac{\mathcal{Z}}{\rho}|1s\rangle = 2\left(Z^2 - 2\mathcal{Z}Z\right)hcR_\infty. \tag{6.23}$$

---

[1] Note that we tacitly assume a *paired* spin state (Pauli principle satisfied).

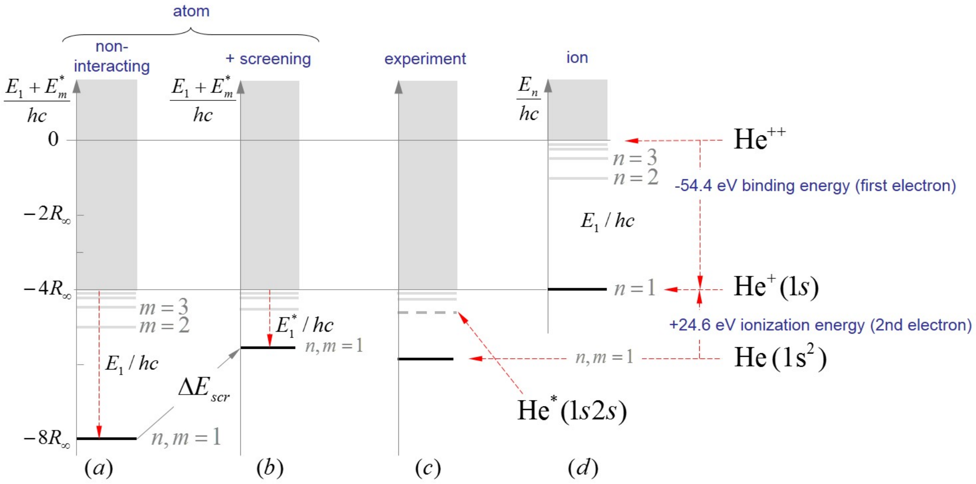


Figure 6.3: Energy level diagram for the ground state of *helium* showing the effect of electrostatic screening: (a) $1s^2$ configuration without the electrostatic interaction between the electrons; (b) $1s^2$ configuration with electrostatic screening $\boldsymbol{l}_1(1/r_{12}) = -\boldsymbol{l}_2(1/r_{12})$ as follows with Eq. (N.21) calculated by first-order perturbation theory (E_m^* is the binding energy of the second electron); (c) diagram experimental values; (d) $1s$ configuration of the He$^+$ ion. Comparison with experiment for the helium ground state shows that even first-order perturbation theory gives a reasonable estimate for the electrostatic screening. Also the energy of *metastable triplet helium* (He*) is indicated.

Adding to this the screening energy as given by Eq. (6.21) we find for the total energy

$$E_{1s^2}^{(1)} = E_{1s^2}^{(0)} + \Delta E_{1s} = 2\left[Z^2 + (5/8 - 2\mathcal{Z})Z\right] hcR_\infty. \tag{6.24}$$

Taking the derivative with respect to Z we find that the energy $E_{1s^2}^{(1)}$ reaches its minimum value for $Z = Z_{\text{eff}} \equiv \mathcal{Z} - 5/16$. Substituting this result into Eq. (6.24) we obtain

$$E_{1s^2}^{(1)} = -2\left(\mathcal{Z} - 5/16\right)^2 hcR_\infty. \tag{6.25}$$

For the helium ground state ($\mathcal{Z} = 2$) we find $E_{1s^2}^{(1)} = -2(27/16)^2 hcR_\infty = -77.4$ eV, which corresponds to a binding energy of 23.0 eV and $Z_{1s} = 27/16 = 1.69$. The corresponding wavefunction is shown as the dashed line in Fig. 6.1. As expected the $1s$ orbitals are inflated in comparison to the unscreened states. This variational result still deviates from the experimental value by 1.6 eV but is clearly better than the 4.2 eV discrepancy obtained by first-order perturbation theory. The screening correction is reduced to 31.4 eV, still overestimating the experimental value by 5.4%.

6.1.3 The hydrogen negative ion H$^-$

The substantial improvement in binding energy obtained with the variational method is encouraging, in particular since there is no reason why the $1s$ shape of the wavefunction would be the best. Hence, further improvement is to be expected with more sophisticated variational wavefunctions. We discuss this for the hydrogen negative ion, H^-, which is a negatively charged helium-like atom consisting of two electrons bound to a proton. In chemistry the H^- ion is sometimes called the *hydride ion*. Experimentally, H^- is found to be weakly bound with a binding energy

$$E_{1s^2}^{(\text{exp})} \approx -1.055\, hcR_\infty, \tag{6.26}$$

i.e., the second electron is bound by 0.75 eV to the neutral atom. As hydrogen is the most abundant atom in the universe the existence of a stable negative ion of hydrogen is of obvious astrophysical significance [25]. In theoretical physics the ion is not of less significance because its existence cannot be demonstrated without taking into account the electronic correlations.

Neglecting the correlations, we regain (by first-order perturbation theory) the result (6.22) for the binding energy of the ground state of helium-like atoms,

$$E_{1s^2}^{(1)} = hcR_\infty \left[-2Z^2 + \tfrac{5}{4}Z\right]. \tag{6.27}$$

For the H^- ion ($Z = 1$) this implies $E_{1s^2}^{(1)} = -\frac{3}{4}\,hcR_\infty$, which is larger than $E_{1s} = -hcR_\infty$ of the neutral hydrogen atom. Hence, according to first-order perturbation theory the proton cannot bind two electrons. To have binding of two electrons one requires $E_{1s^2}^{(1)} < E_{1s} = -hcR_\infty$ (which is the case for $Z > 1.08$). With the variational principle we can do better than first-order perturbation theory; with Eq. (6.25) we find for $\mathcal{Z} = 1$

$$E_{1s^2}^{(1)} = -2\,(11/16)^2\,hcR_\infty = -0.945\,31\,hcR_\infty. \tag{6.28}$$

Although this variational result is better it is still not good enough to bind the ion.

To achieve binding the calculation has to account for the presence of correlations in the electron motion because these tend to increase the average distance between the electrons. This obviously lowers the electrostatic repulsion and, therefore increases the binding energy. The key idea of accounting for such correlations (even without sacrificing the central symmetry) was published in a famous paper by Chandrasekhar [25]. The average distance between the electrons can be increased by writing the two-electron wavefunction as the product state of two s electrons with *different* Bohr radius,

$$\psi(\rho_1, \rho_2) = e^{-a\rho_1 - b\rho_2} + e^{-b\rho_1 - a\rho_2}. \tag{6.29}$$

Using the variational principle (see Appendix I.2.1) the lowest energy for the ground state is obtained for $a = 1.03925$ and $b = 0.28309$ and the binding energy is

$$E_{1s^2}^{(\mathrm{var})} \approx -1.026\,6\,hcR_\infty. \tag{6.30}$$

The actual calculation represents a substantial two-dimensional variational problem, which is not reproduced here. The result is quite intuitive. One electron remains almost hydrogenic with its Bohr radius only slightly compressed. The other electron is largely inflated (the Bohr radius increases by a factor 3.5). This is exactly what we expect. As the nuclear charge is strongly shielded by the first electron, the second electron has its largest probability far from the nucleus and has to be weakly bound. It may speak for itself that by choosing a still more sophisticated variational wavefunction the ground state energy can be further lowered. In this way high precision can be obtained for helium-like atoms [37].

6.1.4 Effective potential and self-consistent mean field

The success of the variational method calls for a systematic approach to optimize the shape of the ground state wavefunction in the presence of screening. What is the best shape? To answer this question, let us suppose that the ground state is a product state of the form $|u_1, u_2) = |u_1\rangle_1 \otimes |u_2\rangle_2$, with electron-1 in s-orbital u_1 and electron-2 in s-orbital u_2; i.e., for the time being the electrons are treated as distinguishable. In this notation the energy of the state $|u_1, u_2)$ is given by the expectation value of the Hamiltonian (6.8),

$$\varepsilon = (u_1, u_2|\mathcal{H}|u_1, u_2) = \sum_{i=1}^{2} \langle u_i|h_0|u_i\rangle + (u_1, u_2|\frac{1}{\rho_{12}}|u_1, u_2). \tag{6.31}$$

We dropped the particle index on the Schrödinger operators, $h_0^{(i)} \to h_0$, with $i \in \{1, 2\}$, because (for identical particles) there can be no confusion about the integral to be evaluated. Since the orbitals

u_1 and u_2 are still to be determined the expectation value ε can be regarded as a functional of u_1 and u_2. To optimize the orbitals u_i we minimize the energy functional (6.31)

$$H_i(u_1,u_2) = \langle u_i|h_0|u_i\rangle + (u_1,u_2|\frac{1}{\rho_{12}}|u_1,u_2) \tag{6.32}$$

by variation of $\langle u_i|$ using the variational procedure of Appendix I.1.3 starting from hydrogenic s orbitals. Introducing the effective potential of screening of electron i by electron j, with $i \neq j \in \{1,2\}$,

$$U_{\rm scr}(\rho_i) = \langle u_j|1/\rho_{ij}|u_j\rangle \tag{6.33}$$

the functional $H_i(u_1,u_2)$ can be written in the form

$$H_i(u_1,u_2) = \langle u_i|h(\rho_i)|u_i\rangle, \tag{6.34}$$

where

$$h(\rho_i) = h_0^{(i)} + U_{\rm scr}(\rho_i). \tag{6.35}$$

For s orbitals $U_{\rm scr}(\rho_i)$ is a central potential as will be shown in Section 6.4.3. To conserve the normalization of the states, H_i has to be minimized under the constraint $\langle u_i|u_i\rangle = 1$, with $i \in \{1,2\}$. Hence, the lagrangian for the variation of $\langle u_i|$ is

$$\mathcal{L}_i(u_1,u_2,\lambda_i) = \langle u_i|h(\rho_i)|u_i\rangle + \lambda_i\left[\langle u_i|u_i\rangle - 1\right], \tag{6.36}$$

where λ_i is the Lagrange multiplier for the constraint $\langle u_i|u_i\rangle = 1$, with $i \in \{1,2\}$.[1] As discussed in Appendix I.1.3, after minimization of the lagrangian (8.71) under variation of $\langle u_i|$ the state $|u_i\rangle$ is found as the solution with eigenvalue $\varepsilon_i = -\lambda_i$ of an effective single-electron Schrödinger equation,

$$h\,u_i(\boldsymbol{\rho}) = \varepsilon_i u_i(\boldsymbol{\rho}). \tag{6.37}$$

This expression corresponds to a set of 2 integro-differential equations, one for each electron and coupled by the screening integrals $U_{\rm scr}(\rho_1)$ and $U_{\rm scr}(\rho_2)$.

Note that the ground state energy is not simply the sum of the optimized single electron energies $\varepsilon \neq \varepsilon_1 + \varepsilon_2$ because the functionals H_1 and H_2 both contain the same repulsion term. Comparing with Eq. (6.31) we find that the energy of the ground state follows from

$$\varepsilon = \varepsilon_1 + \varepsilon_2 - (u_1,u_2|1/\rho_{12}|u_1,u_2). \tag{6.38}$$

To solve Eqs. (6.37) we start with hydrogenic $1s$ wavefunctions, $u_i(\boldsymbol{\rho}) = \langle\boldsymbol{\rho}|u_i\rangle = R_{1s}^{(i)}(\rho)Y_0^0(\hat{\mathbf{r}})$, and calculate the screening potentials. With these potentials we can solve Eqs. (6.37) numerically and obtain improved expressions for the wavefunctions. This, of course, gives rise to an inconsistency because the new wavefunctions imply different screening, modifying the effective single-electron Schrödinger equation to be solved. Thus the procedure has to be repeated until a *self-consistent solution* is reached for the *mean field.* The principal drawback of this approach is that by constructing the semiclassical effective field, correlations in the relative motion of the electrons are neglected. The mean field approximation is best suited to describe large atoms because the concept of a local charge density becomes exact in the classical limit. Mean field methods based on variation of the wavefunction are known as Hartree and Hartree-Fock theories (see Chapter 9).

Actually, the emergence of the electrostatic interaction as the dominant interaction in the atom suggests another semiclassical approach in which the wavefunction does not appear at all but a quantum statistical model is used to calculate the distribution of the electronic charge (or, equivalently, the electrostatic potential). Statistical methods for these semiclassical quantities are called *density functional* or *potential functional* methods [39]. They work best for atoms with a large number of electrons because in this case the concept of a local charge density is best justified. This being said statistical methods fall outside the scope of this course. Whatever the quantity being varied (wavefunction, density distribution or effective potential), the general idea is that the calculation of the atomic energy levels can be strongly improved by the proper optimization procedure.

---

[1] Note that for this lagrangian the normalization of the orbitals is enforced but their orthogonality is *not.*

6.2 The helium ground state in a magnetic field

A lot can be learned by analyzing the behavior of the ^{4}He atom in an externally applied magnetic field $\mathbf{B}$. Remarkably, this is because little happens! A small diamagnetic shift may be observed but no Zeeman splitting, in spite of the presence of two electrons each carrying a Bohr magneton. *A priori* this is surprising. As we are dealing with s orbitals the orbital angular momentum of the atom is zero. Because the nuclear spin of ^{4}He is also zero all angular momentum has to come from the electronic spin. The conserved quantity is the total angular momentum $\mathbf{S} = \mathbf{s}_1 + \mathbf{s}_2$. Because $s_1 = s_2 = 1/2$ we have two possibilities for coupling of these spins, *parallel* ($S = 1$) or *anti-parallel* ($S = 0$). In a sufficiently weak probe field we expect Zeeman coupling of the total spin to the field as described by a Hamiltonian of the spin-Zeeman type,

$$\mathcal{H}_Z = g_S \mu_B B \left(S_z / \hbar \right), \tag{6.39}$$

where g_S is the g factor for the total spin $\mathbf{S}$. The Zeeman shifts are obtained with first-order perturbation theory, just as in the case of the low-field limit of the fine- or hyperfine structure,

$$E_z = g_S \mu_B M_S B. \tag{6.40}$$

Hence for $S = 1$ we expect the ground state to split into a triplet. For $S = 0$ we expect a singlet, no splitting and no Zeeman shift. As mentioned above *only the singlet is observed.* Apparently the triplet option, in which both electrons are in the same (orbital + spin) state, has to be *excluded.* Also in strong magnetic fields the degeneracy is not lifted, which points to $\mathbf{s}_1 \cdot \mathbf{s}_2$ coupling because for this type of coupling $M_S = m_{s_1} + m_{s_2}$ is conserved in accordance with the commutation relation $[S_z, \mathbf{s}_1 \cdot \mathbf{s}_2] = 0$. Pauli generalized similar observations into his *exclusion principle.*

6.3 Exchange degeneracy and Pauli principle

If two electrons are in *different* orbitals, $|u\rangle = |nlm_l\rangle$ and $|v\rangle = |n'l'm_{l'}\rangle$, the same pair energy is obtained irrespective of which of the two electrons is in state $|u\rangle$ and which in state $|v\rangle$,

$$(u, v|\mathcal{H}_0|u, v) = E_{n_u} + E_{n_v} \tag{6.41a}$$

$$(v, u|\mathcal{H}_0|v, u) = E_{n_v} + E_{n_u}. \tag{6.41b}$$

This phenomenon is called *exchange degeneracy.* It implies that any linear combination of the type

$$\psi_{uv}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{|c_1|^2 + |c_2|^2}} \left\{ c_1 \varphi_u(\mathbf{r}_1) \varphi_v(\mathbf{r}_2) + c_2 \varphi_v(\mathbf{r}_1) \varphi_u(\mathbf{r}_2) \right\} \tag{6.42}$$

represents a properly normalized energy eigenstate of the pair; hence, for any choice of c_1 and c_2 we can construct an orthogonal pair state with the same energy.

Note that if the electrons would have a slightly different mass their hydrogenic energies would be different, $E_{n_u 1} \neq E_{n_u 2}$, and the exchange degeneracy would be absent, $E_{n_u 1} + E_{n_v 2} \neq E_{n_v 1} + E_{n_u 2}$. Exchange degeneracy arises if the particles to be exchanged cannot be distinguished. As a consequence the Hamiltonian (6.1) is invariant under interchange of the electrons.

Let us analyze how the exchange degeneracy (6.41a) is *lifted* by the Coulomb term (6.10) using perturbation theory for a doubly degenerate level (see Appendix H.3). Like in Section 6.1 we split the Hamiltonian (6.8) in the form

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}', \tag{6.43}$$

where $\mathcal{H}_0$ is again as *unperturbed* Hamiltonian and the electrostatic repulsion, $\mathcal{H}' = \rho_{12}^{-1}$, is the perturbation. For each pair of hydrogenic states $|u\rangle$ and $|v\rangle$ an exchange degenerate pair can be defined, $|a) = |u, v)$ and $|b) = |v, u)$. It is straightforward to show that $\mathcal{H}'_{aa} = (a|\mathcal{H}'|a) = (b|\mathcal{H}'|b) =$

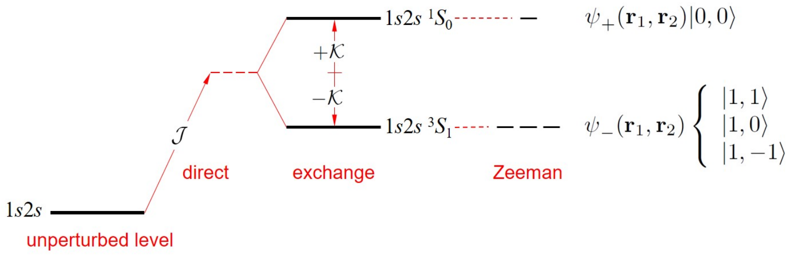


Figure 6.4: Exchange tends to "align" the electron spins in atoms. The energy level diagram shows the direct and exchange shifts for helium atoms with two nonequivalent electrons. Because $\mathcal{K}$ is always positive the antisymmetric orbital has always the lowest energy. Experiments show that the lowest pair state is always a spin triplet and the highest pair state a spin singlet.

$\mathcal{H}'_{bb} \equiv \mathcal{J}$. Furthermore we have $\mathcal{H}'_{ab} = (a|\mathcal{H}'|b) = (b|\mathcal{H}'|a) = \mathcal{H}'_{ba} \equiv \mathcal{K}$, as follows from the hermiticity of $\mathcal{H}'$. The overlap integrals S_{ab} introduced in Appendix H.3 are zero because the hydrogenic eigenstates are orthogonal, $S_{ab} = (a|b) = (b|a) = S_{ba} = 0$. Solving the secular equation (H.70) we find the result of first order perturbation theory of a two-fold degenerate level (*symmetric case* of *strong coupling*),

$$E_{\pm}^{(1)} = E_{uv}^{(0)} + \Delta E_{\pm}, \tag{6.44}$$

where $\Delta E_{\pm}$ can be written in the form of Eqs. (G.38) and (G.55),

$$\Delta E_{\pm} = \mathcal{J} \pm \mathcal{K}. \tag{6.45}$$

The $\Delta E_{\pm}$ represent the energy level shifts caused by the electrostatic repulsion. The integrals

$$\mathcal{J}(u,v) = (u,v|\rho_{12}^{-1}|u,v) \tag{6.46a}$$

$$\mathcal{K}(u,v) = (u,v|\rho_{12}^{-1}|v,u) \tag{6.46b}$$

are called the *direct* ($\mathcal{J}$) and the *exchange* ($\mathcal{K}$) contribution to the energy shift. Note the properties

$$\mathcal{J}(u,v) = \mathcal{J}(v,u) \quad \text{and} \quad \mathcal{K}(u,v) = \mathcal{K}(v,u). \tag{6.47}$$

The eigenstates corresponding to the shifts (6.45) are

$$\psi_{u,v}^{\pm}(\mathbf{r}_1,\mathbf{r}_2) = \sqrt{\tfrac{1}{2}}\left\{\varphi_u(\mathbf{r}_1)\,\varphi_v(\mathbf{r}_2) \pm \varphi_v(\mathbf{r}_1)\,\varphi_u(\mathbf{r}_2)\right\}, \tag{6.48}$$

which correspond to the symmetric (+) and antisymmetric (−) linear combinations of the pair states $|u,v)$ and $|v,u)$. The situation is sketched in Fig. 6.4. The matrix elements $\mathcal{J}$ and $\mathcal{K}$ are called the *Coulomb integrals of the electrostatic repulsion.*

Two important features should be pointed out for the antisymmetric pair state $|\psi_-\rangle$. First we note that $|\psi_-\rangle$ has the lowest energy. This is intuitively clear because the electrostatic repulsion is largest when the electrons come close to each other and precisely this probability is excluded by the antisymmetric form. The probability of finding two electrons at the same position is identically zero. This follows directly by substituting $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$ into the antisymmetric form of the pair wavefunctions (6.48),

$$\psi_{u,v}^{-}(\mathbf{r},\mathbf{r}) = \sqrt{\tfrac{1}{2}}\left\{\varphi_u(\mathbf{r})\,\varphi_v(\mathbf{r}) - \varphi_v(\mathbf{r})\,\varphi_u(\mathbf{r})\right\} \equiv 0. \tag{6.49}$$

Hence the electrostatic repulsion is smallest in the antisymmetric case. This increases the binding energy of the electron pair and lowers the total energy of atoms in antisymmetric states with respect to atoms in the corresponding symmetric states.

A second point to note holds for two electrons in the same orbital, $|u\rangle = |v\rangle$, for instance the $1s^2$ configuration of the helium ground state discussed in Section 6.1. In this case the antisymmetric wavefunction is also identically zero, whatever the positions of the two electrons,

$$\psi^-_{u,u}(\mathbf{r}_1, \mathbf{r}_2) = \sqrt{\tfrac{1}{2}}\,\{\varphi_u(\mathbf{r}_1)\,\varphi_u(\mathbf{r}_2) - \varphi_u(\mathbf{r}_1)\,\varphi_u(\mathbf{r}_2)\} \equiv 0. \tag{6.50}$$

For two electrons in the same orbital only the symmetric state is an option,

$$\psi_{u,u}(\mathbf{r}_1, \mathbf{r}_2) = \varphi_u(\mathbf{r}_1)\,\varphi_u(\mathbf{r}_2)\,. \tag{6.51}$$

Note that this wavefunction is both symmetrized and normalized to start with. Actually, the explicit symmetrization (6.48) yields the *wrong* normalization as $\psi^\pm$ were derived for the case $u \neq v$; i.e., for the presence of exchange degeneracy. For electrons in the same orbital exchange degeneracy is absent and the energy shift caused by electrostatic repulsion between the electrons is given by the direct shift only

$$\Delta E_{\mathrm{scr}} = \mathcal{J} = (u, u|\mathcal{H}'|u, u). \tag{6.52}$$

Thus we arrived at two important conclusions: (a) two electrons in an antisymmetric orbital pair state have *zero probability* to be found at the same position; (b) two electrons in the same orbital state are represented by the simple product state (6.51) and not by Eq. (6.42). Pauli noticed that in *two-electron atoms* (Pauli studied the alkaline-earth elements) the antisymmetric orbital is always a spin triplet ($S = 1$) whereas the symmetric orbital always a spin singlet ($S = 0$).[1] This is illustrated in Fig. 6.4. As the spin triplet is symmetric and the spin singlet antisymmetric under exchange of the two electrons (see Problem 6.2), this shows that the overall electron pair state (orbital + spin) must be antisymmetric under exchange of the two electrons; i.e., electrons are fermionic particles. Pauli summarized these and other observations in his famous exclusion principle: no two electrons can occupy the same (orbital + spin) state [82].

The symmetric spin state is energetically favored over the antisymmetric spin state. Phenomenologically, this emerges as the result of some *effective magnetic interaction* which tends to align the spins in atoms. This effective interaction is called the *exchange interaction.* It is much stronger than a truly magnetic interaction because its origin is *electrostatic* and not magnetic at all.

Problem 6.2. Show that the spin triplet is symmetric and the spin singlet antisymmetric under exchange of two $s = \frac{1}{2}$ particles.

Solution. The symmetry of the total spin state under exchange of the two $s = \frac{1}{2}$ particles is found by Clebsch-Gordan decomposition,

$$|S, M\rangle = \sum_{m_1, m_2} |s, m_1; s, m_2\rangle\langle s, m_1; s, m_2|S, M\rangle. \tag{6.53}$$

Using Table J.5.1 for the Clebsch-Gordan coefficients (in the Condon and Shortley phase convention) we obtain

$$\left.\begin{array}{ll} |1,+1\rangle = |\uparrow\uparrow\rangle & \\ |1,0\rangle \;\;= \sqrt{1/2}\,[|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle] & \\ |1,-1\rangle = |\downarrow\downarrow\rangle & \end{array}\right\} \quad (S=1) \\ |0,0\rangle \;\;= \sqrt{1/2}\,[|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle] \qquad (S=0)\,, \tag{6.54}$$

which demonstrates the requested symmetry. □

---

[1] The other combination (symmetric orbital state with spin singlet and antisymmetric orbital state with spin triplet) is not observed although this would be expected given the angular momentum addition rules.

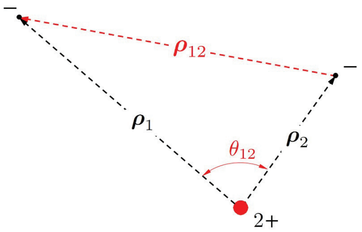


Figure 6.5: Doubly-charged nucleus with electrons at positions indicated by the vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ (enclosing the angles θ_{12}) as well as the vector for the relative position $\mathbf{r}_{12}$.

6.4 Expressions for the Coulomb integrals

Rather than evaluating a special case we derive in this section general expressions fothe eigenstatesr the diagonal and off-diagonal matrix elements of the electrostatic repulsion between atomic orbitals. In Hartree atomic units the electrostatic repulsion between two electrons is given by

$$\mathcal{H}'(\rho_{12}) = \frac{1}{\rho_{12}} = \frac{1}{(\rho_1^2 + \rho_2^2 - 2\rho_1\rho_2\cos\theta_{12})^{1/2}}. \tag{6.55}$$

Here the cosine rule is used to express the relative distance between the electrons, $\rho_{12} = |\boldsymbol{\rho}_1 - \boldsymbol{\rho}_2|$, in terms of the absolute positions $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$ relative to the nucleus and the angle θ_{12} enclosed by the directions $\hat{\mathbf{r}}_1$ and $\hat{\mathbf{r}}_2$ (see Fig. 6.5). This is important for the calculation of matrix elements because the electronic wavefunctions are defined in terms of the variables $\boldsymbol{\rho}_1$ and $\boldsymbol{\rho}_2$. Recalling the radial averages (2.49) we expand Eq. (6.55) in powers of the ratio $(\rho_</\rho_>)$, where $\rho_< = \min\{\rho_1, \rho_2\}$ is the lesser and $\rho_> = \max\{\rho_1, \rho_2\}$ the greater of ρ_1 and ρ_2,

$$\frac{1}{\rho_{12}} = \frac{1}{\rho_>}\sum_{k=0}^{\infty}\left(\frac{\rho_<}{\rho_>}\right)^k P_k(\cos\theta_{12}) \to \frac{1}{\rho_>}\ \text{for}\,\rho_< \ll \rho_>. \tag{6.56}$$

The functions $P_k(u)$ are Legendre polynomials of order k (see Appendix L.9). What remains to be done is to express the relative angle θ_{12} into absolute position angles $\hat{\mathbf{r}}_1$ and $\hat{\mathbf{r}}_2$ of the individual electrons. This is done with the aid of the spherical harmonic *addition theorem* (L.56),

$$P_k(\cos\theta_{12}) = \frac{4\pi}{2k+1}\sum_{m=-k}^{k} Y_k^{m*}(\hat{\mathbf{r}}_1)Y_k^m(\hat{\mathbf{r}}_2). \tag{6.57}$$

Substituting this expression into the expansion (6.56) the perturbation takes the form of a *multipole expansion*, which is suited for evaluation with hydrogenic wavefunctions,

$$\frac{1}{\rho_{12}} = \frac{1}{\rho_>}\sum_{k=0}^{\infty}\left(\frac{\rho_<}{\rho_>}\right)^k \frac{4\pi}{2k+1}\sum_{m=-k}^{k} Y_k^{m*}(\hat{\mathbf{r}}_1)Y_k^m(\hat{\mathbf{r}}_2). \tag{6.58}$$

This form is optimal for the evaluation of the Coulomb integrals (6.46). Using the multipole expansion the integrals $\mathcal{J}$ and $\mathcal{K}$ can be expressed in the form

$$\mathcal{J} = (nlm_l; n'l'm_{l'}|\frac{1}{\rho_{12}}|nlm_l; n'l'm_{l'}) = \sum_{k=0}^{\infty} a^k(lm_l; l'm_{l'})F^k(nl; n'l') \tag{6.59a}$$

$$\mathcal{K} = (nlm_l; n'l'm_{l'}|\frac{1}{\rho_{12}}|n'l'm_{l'}; nlm_l) = \sum_{k=0}^{\infty} b^k(lm_l; l'm_{l'})G^k(nl; n'l'), \tag{6.59b}$$

The coefficients $a^k(lm_l; l'm_{l'})$ and $b^k(lm_l; l'm_{l'})$ represent the angular parts and $F^k(nl; n'l')$ and $G^k(nl; n'l')$ the radial parts of the *direct* ($\mathcal{J}$) and *exchange* ($\mathcal{K}$) contribution to the energy shift.

Table 6.1: List of a^k and b^k coefficients for the ss, sp, pp, sd, pd, dd and ff configurations. For reasons of transparency the common denominators of the elements of the table are only printed at the top of the columns (of more than 3 rows); $\{m_l, m_{l'}\} = \{\pm a, \mp b\}$ stands for $\{m_l, m_{l'}\} = \{\{+a, -b\}\ or\ \{-a, +b\}\}$.

			$a^k(lm_l; l'm_{l'}) = a^k(l'm_{l'}; lm_l)$				$b^k(lm_l; l'm_{l'}) = b^k(l'm_{l'}; lm_l)$			
	m_l	m_l'	$k=0$	$k=2$	$k=4$	$k=6$	$k=0$	$k=2$	$k=4$	$k=6$
ss	0	0	1				1			
pp	±1	±1	1	1/25			1	1/25		
	±1	0	1	−2			0	3		
	±1	∓1	1	1			0	6		
	0	0	1	4			1	4		
sd	0	±2	1				0	1/5		
	0	±1	1				0	1/5		
	0	0	1				0	1/5		
dd	±2	±2	1	4/49	1/441		1	4/49	1/441	
	±2	±1	1	−2	−4		0	6	5	
	±2	0	1	−4	6		0	4	15	
	±2	∓1	1	−2	−4		0	0	35	
	±2	∓2	1	4	1		0	0	70	
	±1	±1	1	1	16		1	1	16	
	±1	0	1	2	−24		0	1	30	
	±1	∓1	1	1	16		0	6	40	
	0	0	1	4	36		1	4	36	
ff	±3	±3	1	25/225	9/1089	25/184041	1	25/225	9/1089	25/184041
	±3	±2	1	0	−21	−150	0	25	30	175
	±3	±1	1	−15	3	375	0	10	54	700
	±3	0	1	−20	18	−500	0	0	63	2100
	±3	∓1	1	−15	3	375	0	0	42	5250
	±3	∓2	1	0	−21	−150	0	0	0	11550
	±3	∓3	1	25	9	25	0	0	0	23100
	±2	±2	1	0	49	900	1	0	49	900
	±2	±1	1	0	−7	−2250	0	15	32	2625
	±2	0	1	0	−42	3000	0	20	3	5600
	±2	∓1	1	0	−7	−2250	0	0	14	9450
	±2	∓2	1	0	49	900	0	0	70	12600
	±1	±1	1	9	1	5625	1	9	1	5625
	±1	0	1	12	6	−7500	0	2	15	8750
	±1	∓1	1	9	1	5625	0	24	40	10500
	0	0	1	16	36	10000	1	16	36	10000
	m_l	m_l'	$k=0$	$k=2$	$k=4$		$k=1$	$k=3$	$k=5$	
sp	0	±1	1				1/3			
	0	0	1				1/3			
pd	±1	±2	1	2/35			6/15	3/245		
	±1	±1	1	−1			3	9		
	±1	0	1	−2			1	18		
	±1	∓1	1	−1			0	30		
	±1	∓2	1	2			0	45		
	0	±2	1	−4			0	15		
	0	±1	1	2			3	24		
	0	0	1	4			4	27		

It is hard to overemphasize the importance of Eqs. (6.59) for the rest of the course. The Coulomb integrals show that, in principle, *all* hydrogenic states are coupled by the Coulomb interaction. However, as the electrons can only scatter under conservation of energy and angular momentum, only a hand full of Coulomb integrals actually play a role. For ground state atoms, s, p, d and f electrons are most relevant and the corresponding angular integrals can be collected on a single page (see Table 6.1). As the electrons only scatter effectively at close encounter the Coulomb integrals also fall off with increasing value of k,

$$\frac{1}{\rho_{12}} = \frac{1}{\rho_>}\sum_{k=0}^{\infty}\left(\frac{\rho_<}{\rho_>}\right)^k P_k(\cos\theta_{12}) \;\Rightarrow\; \frac{1}{\rho_{12}} \simeq \frac{1}{\rho_>} \;\text{ for } \rho_< \ll \rho_> . \tag{6.60}$$

This is readily verified in Tables 6.1 and 6.2.

Problem 6.3. Show that the orbital angular momenta $\boldsymbol{l}_i$ of the electrons $i = 1, 2$ in a helium atom are not conserved under the Hamiltonian (8.1),

$$[\boldsymbol{l}_i^2, \mathcal{H}]\psi \neq 0 \quad \text{and} \quad [l_{iz}, \mathcal{H}]\psi \neq 0,$$

Solution. We demonstrate this for $\boldsymbol{l}_1^2$ and l_{1z}. Since these operators are conserved under the Schrödinger Hamiltonian, it is sufficient to show that (for arbitrary position $\mathbf{r}_2$ of the other electron) they do not commute with the electrostatic repulsion term,

$$[\boldsymbol{l}_1^2, 1/r_{12}]\psi = \boldsymbol{l}_1^2(1/r_{12})\psi - (1/r_{12})\boldsymbol{l}_1^2\psi = \psi\boldsymbol{l}_1^2(1/r_{12}) \neq 0.$$

Using the multipole expansion (6.58) and the eigenvalue relation (1.52) we find for arbitrary $\hat{\mathbf{r}}_2$,

$$\begin{aligned}\boldsymbol{l}_1^2\frac{1}{r_{12}} &= \frac{1}{\rho_>}\sum_{k=0}^{\infty}\left(\frac{\rho_<}{\rho_>}\right)^k \frac{4\pi}{2k+1}\sum_{m=-k}^{k}\boldsymbol{l}_1^2\, Y_k^{m*}(\hat{\mathbf{r}}_1)Y_k^m(\hat{\mathbf{r}}_2)\\ &= \frac{1}{\rho_>}\sum_{k=0}^{\infty}\left(\frac{\rho_<}{\rho_>}\right)^k (2k+1)\hbar^2 P_k(\cos\theta_{12}) \neq 0.\end{aligned}$$

Likewise, the relation $[l_{iz}, \mathcal{H}]\psi \neq 0$ follows from $[l_{iz}(1/r_{12}) \neq 0$. □

6.4.1 Angular integrals

The angular integrals can be expressed in $3j$ symbols (see Problem 6.4),

$$\begin{aligned}a^k(lm_l; l'm_{l'}) &= \frac{4\pi}{2k+1}\sum_{m=-k}^{k}\langle lm_l|Y_k^{m*}(\hat{\mathbf{r}}_1)|lm_l\rangle\langle l'm_{l'}|Y_k^m(\hat{\mathbf{r}}_2)|l'm_{l'}\rangle\\ &= (-)^{m_l+m_{l'}}(2l+1)(2l'+1)\begin{pmatrix} l & k & l\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & k & l\\ -m_l & 0 & m_l\end{pmatrix}\begin{pmatrix} l' & k & l'\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l' & k & l'\\ -m_{l'} & 0 & m_{l'}\end{pmatrix} \quad (6.61\text{a})\\ b^k(lm_l; l'm_{l'}) &= \frac{4\pi}{2k+1}\sum_{m=-k}^{k}\langle l'm_{l'}|Y_k^{m*}(\hat{\mathbf{r}}_1)|lm_l\rangle\langle lm_l|Y_k^m(\hat{\mathbf{r}}_2)|l'm_{l'}\rangle\\ &= (2l+1)(2l'+1)\begin{pmatrix} l & k & l'\\ 0 & 0 & 0\end{pmatrix}^2\begin{pmatrix} l & k & l'\\ -m_l & (m_l-m_{l'}) & m_{l'}\end{pmatrix}^2 \geq 0 \quad (6.61\text{b})\end{aligned}$$

where, in view of the properties of the $3j$ symbols,

$$a^k(lm_l; l'm_{l'}) = 0 \;\text{ unless } k = \text{even and } 0 \leq k \leq 2l_< \tag{6.62}$$

$$b^k(lm_l; l'm_{l'}) = 0 \;\text{ unless } k + l + l' = \text{even and } |l - l'| \leq k \leq l + l', \tag{6.63}$$

with $l_< = \min\{l, l'\}$ (the lesser of l and l'). The a^k coefficients can be positive, negative or zero. The b^k coefficients are positive or zero. Note that a^k and b^k coincide for electrons in the same state of angular momentum: $a^k(lm_l;lm_l) = b^k(lm_l;lm_l)$. Note further the symmetry under state reversal: $a^k(lm_l;l'm_{l'}) = a^k(l'm_{l'};lm_l)$ and $b^k(lm_l;l'm_{l'}) = b^k(l'm_{l'};lm_l)$. The a^k and b^k coefficients are readily calculated with a symbolic manipulation program like *Mathematica*. For the configurations s^2, sp, p^2, sd, pd, d^2 and f^2 the results are given in Table 6.1. The special cases $a^k(00;lm_l)$ and $b^k(00;lm_l)$ are discussed in Problems 6.11 and 6.12.

Problem 6.4. Derive equation 6.61a:

$$a^k(lm_l;l'm_{l'}) = (-)^{m_l+m_{l'}}(2l+1)(2l'+1)\begin{pmatrix} l & k & l \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l & k & l \\ -m_l & 0 & m_l \end{pmatrix}\begin{pmatrix} l' & k & l' \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l' & k & l' \\ -m_{l'} & 0 & m_{l'} \end{pmatrix}.$$

Solution. Using Eq. (L.53) and Eq. (L.59) for the integral over three spherical harmonics we obtain

$$\langle lm_l|Y_k^{m*}(\hat{\mathbf{r}}_1)|lm_l\rangle = (-1)^{m_l+m}(2l+1)\sqrt{\frac{(2k+1)}{4\pi}}\begin{pmatrix} l & k & l \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l & k & l \\ -m_l & -m & m_l \end{pmatrix}$$
$$\langle l'm_{l'}|Y_k^{m}(\hat{\mathbf{r}}_2)|l'm_{l'}\rangle = (-1)^{m_{l'}}(2l'+1)\sqrt{\frac{(2k+1)}{4\pi}}\begin{pmatrix} l' & k & l' \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l' & k & l' \\ -m_{l'} & m & m_{l'} \end{pmatrix}.$$

Multiplying these two expressions we find

$$\frac{4\pi}{2k+1}\langle lm_l|Y_k^{m*}(\hat{\mathbf{r}}_1)|lm_l\rangle\langle l'm_{l'}|Y_k^{m}(\hat{\mathbf{r}}_2)|l'm_{l'}\rangle =$$
$$= (-1)^{m_l+m_{l'}+m}(2l+1)(2l'+1)\begin{pmatrix} l & k & l \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l & k & l \\ -m_l & -m & m_l \end{pmatrix}\begin{pmatrix} l' & k & l' \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l' & k & l' \\ -m_{l'} & m & m_{l'} \end{pmatrix}.$$

which is only nonzero for $0 \le k \le 2l_<$ and $m = 0$ because $\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = 0$ unless $|j_1 - j_3| \le j_2 \le j_1 + j_3$ and $m_1 + m_2 + m_3 = 0$ (see Appendix J). □

Problem 6.5. Show that

$$a^0(lm_l;l'm_{l'}) = 1.$$

Solution. Using Eq. (J.16) we have

$$(-)^{m_l}(2l+1)\begin{pmatrix} l & 0 & l \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l & 0 & l \\ -m_l & 0 & m_l \end{pmatrix} = 1.$$

Using this relation twice in Eq. (6.61a) we obtain $a^0(lm_l;l'm_{l'}) = 1$. □

Problem 6.6. Show that

$$\sum_{m=-l}^{l} a^k(lm;l'm') = 0 \text{ for } k \neq 0.$$

Solution. This result follows immediately by applying the shell-summation formula (J.20) to Eq. (J.16)

$$\begin{pmatrix} l & k & l \\ 0 & 0 & 0 \end{pmatrix}\sum_{m=-l}^{l}(-)^m\begin{pmatrix} l & k & l \\ -m & 0 & m \end{pmatrix} = 0.$$ □

Problem 6.7. Show that

$$\sum_{m=-l}^{l} b^k(lm;l'm') = (2l+1)\begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2.$$

6.4.2 Radial integrals

The direct $\left(F^k\right)$ and exchange $\left(G^k\right)$ radial integrals are given by

$$F^k(nl;n'l') = \int_0^\infty \left\{ \frac{1}{\rho_1^{k+1}} \int_0^{\rho_1} \rho_2^k \left[\tilde{R}_{nl}(\rho_2)\right]^2 \rho_2^2 d\rho_2 + \right.$$
$$\left. + \rho_1^k \int_{\rho_1}^\infty \frac{1}{\rho_2^{k+1}} [\tilde{R}_{nl}(\rho_2)]^2 \rho_2^2 d\rho_2 \right\} \left[\tilde{R}_{n'l'}(\rho_1)\right]^2 \rho_1^2 d\rho_1 \tag{6.64a}$$

$$G^k(nl;n'l') = \int_0^\infty \left\{ \frac{1}{\rho_1^{k+1}} \int_0^{\rho_1} \rho_2^k \tilde{R}_{nl}(\rho_2)\tilde{R}_{n'l'}(\rho_2)\rho_2^2 d\rho_2 + \right.$$
$$\left. + \rho_1^k \int_{\rho_1}^\infty \frac{1}{\rho_2^{k+1}} \tilde{R}_{nl}(\rho_2)\tilde{R}_{n'l'}(\rho_2)\rho_2^2 d\rho_2 \right\} \tilde{R}_{nl}(\rho_1)\tilde{R}_{n'l'}(\rho_1)\rho_1^2 d\rho_1. \tag{6.64b}$$

Note that the radial integrals coincide for equivalent electrons, $F^k(nl;nl) = G^k(nl;nl)$. Furthermore, it is important to note that (like the a^k and b^k coefficients) the radial integrals are symmetric under state reversal, $F^k(nl;n'l') = F^k(n'l';nl)$ and $G^k(nl;n'l') = G^k(n'l';nl)$. The integrals can be reformulated in a more convenient and insightful form by introducing two *screening potentials*,

$$U_F^k(\rho) = \frac{1}{\rho^{k+1}} \int_0^\rho \varrho^k \left[\tilde{R}_{nl}(\varrho)\right]^2 \varrho^2 d\varrho + \rho^k \int_\rho^\infty \frac{1}{\varrho^{k+1}} \left[\tilde{R}_{nl}(\varrho)\right]^2 \varrho^2 d\varrho \tag{6.65a}$$

$$U_G^k(\rho) = \frac{1}{\rho^{k+1}} \int_0^\rho \varrho^k \tilde{R}_{nl}(\varrho)\tilde{R}_{n'l'}(\varrho)\, \varrho^2 d\varrho + \rho^k \int_\rho^\infty \frac{1}{\varrho^{k+1}} \tilde{R}_{nl}(\varrho)\tilde{R}_{n'l'}(\varrho)\, \varrho^2 d\varrho, \tag{6.65b}$$

where $U_F^k(\rho)$ is called the potential for *direct screening* and $U_G^k(\rho)$ the potential for *exchange screening* (note the change of notation: $\rho_1 \to \rho$; $\rho_2 \to \varrho$). With the aid of these potentials the radial integrals take the convenient form

$$F^k(nl;n'l') = \int_0^\infty U_F^k(\rho) \left[\tilde{R}_{n'l'}(\rho)\right]^2 \rho^2 d\rho, \tag{6.66a}$$

$$G^k(nl;n'l') = \int_0^\infty U_G^k(\rho)\tilde{R}_{nl}(\rho)\tilde{R}_{n'l'}(\rho)\rho^2 d\rho. \tag{6.66b}$$

For hydrogenic wavefunctions these integrals are conveniently calculated with a symbolic manipulation program like *Mathematica*.

Equivalent electrons

Some valuable exact solutions can be obtained for equivalent electrons. These are interesting in their own right but also prove valuable as benchmarks for numerical calculations. For equivalent electrons the integrals (6.66) coincide and simplify to the form (see Problem 6.8)

$$F^k(nl;nl) = 2\int_0^\infty d\rho \int_0^\rho d\varrho \frac{\varrho^{k+2}}{\rho^{k-1}} [\tilde{R}_{nl}(\varrho)\tilde{R}_{nl}(\rho)]^2. \tag{6.67}$$

Note that this expression is positive definite, $F^k(nl;nl) > 0$. Changing to the variables $x = 2Z\rho/n$ and $y = 2Z\rho\varrho/n$ this can be rewritten as

$$F^k(nl;nl) = (2Z/n)\frac{2}{\mathcal{N}_{nl}^2} \int_0^\infty dx\, x^{2l+1-k} e^{-x} [L_{n-l-1}^{2l+1}(x)]^2 \int_0^x dy\, y^{2l+2+k} e^{-y} [L_{n-l-1}^{2l+1}(y)]^2. \tag{6.68}$$

Since $L_0^{2l+1}(x) \equiv 1$ further simplification occurs for the case $l = n-1$; i.e., for the configurations $1s^2$, $2p^2$, $3d^2$ and $4f^2$. In these cases the Coulomb integrals further simplify to

$$F^k(nl;nl) = (2Z/n)\frac{2}{\mathcal{N}_{nl}^2} \int_0^\infty dx\, x^{2l+1-k} e^{-x} \int_0^x dy\, y^{2l+2+k} e^{-y}. \tag{6.69}$$

Table 6.2: Coulomb integrals for orbitals corresponding to a zero-order Laguerre polynomial.

$F^k(nl^2)/Z$	$k=0$	$k=2$	$k=4$	$k=6$
$1s^2$	$\frac{5}{8}$			
$2p^2$	$\frac{93}{512}$	$\frac{45}{512}$		
$3d^2$	$\frac{3965}{46\,080}$	$\frac{2093}{46\,080}$	$\frac{1365}{46\,080}$	
$4f^2$	$\frac{184331}{3670016}$	$\frac{103275}{3670016}$	$\frac{69003}{3670016}$	$\frac{51051}{3670016}$

These integrals are easily evaluated in *Mathematica*. For the configurations $1s^2$, $2p^2$, $3d^2$ and $4f^2$ the results are summarized in Table 6.2. Note that the following inequality is satisfied:

$$F^k(nl^2) < F^{k-2}(nl^2). \tag{6.70}$$

Importantly, this inequality holds for any radial wavefunction; i.e., also in the non-hydrogenic case. Note that Problem 6.9 is helpful in this context.

Problem 6.8. Derive Eq. (6.67).

Solution. Starting from Eq. (6.66a) and (6.65a) we find for equivalent electrons

$$F^k(nl;nl) = \int_0^\infty d\rho \int_0^\rho d\varrho \frac{\varrho^k}{\rho^{k+1}} [\tilde{\chi}_{nl}(\varrho)\tilde{\chi}_{nl}(\rho)]^2 + \int_0^\infty d\rho \int_\rho^\infty d\varrho \frac{\rho^k}{\varrho^{k+1}} [\tilde{\chi}_{nl}(\varrho)\tilde{\chi}_{nl}(\rho)]^2,$$

where $\tilde{\chi}_{nl}(\rho) = \rho\tilde{R}_{nl}(\rho)$. Next we change the order of integration of the second term

$$\int_0^\infty d\rho \int_\rho^\infty d\varrho \frac{\rho^k}{\varrho^{k+1}} [\tilde{\chi}_{nl}(\varrho)\tilde{\chi}_{nl}(\rho)]^2 = \int_0^\infty d\varrho \int_0^\varrho d\rho \frac{\rho^k}{\varrho^{k+1}} [\tilde{\chi}_{nl}(\varrho)\tilde{\chi}_{nl}(\rho)]^2.$$

This leads to the desired result by a change of dummy indices, $\rho \to \varrho$ and vice versa. □

Problem 6.9. Show that $U_F^k(\rho) > U_F^{k+1}(\rho)$.

Solution. We rewrite $U_F^k(\rho) - U_F^{k+1}(\rho)$ using the substitution $z = x/\rho$:

$$U_F^k(\rho) - U_F^{k+1}(\rho) = \rho^2 \int_0^1 \left[\tilde{R}_{nl}(\rho z)\right]^2 z^{k+2}(1-z)\,dz + \rho^2 \int_1^\infty \left[\tilde{R}_{nl}(\rho z)\right]^2 \left(\frac{1}{z}\right)^k (z-1)\,dz > 0$$

This is positive definite because

$$\begin{aligned} \left[\tilde{R}_{nl}(\rho z)\right]^2 z^{k+2}(1-z) &\geq 0 \quad \text{for } 0 \leq z \leq 1 \\ \left[\tilde{R}_{nl}(\rho z)\right]^2 (1/z)^k (z-1) &\geq 0 \quad \text{for } z \geq 1. \end{aligned}$$

□

Problem 6.10. Show that for equivalent electrons the exchange integral is positive definite, $\mathcal{K} > 0$.

Solution. For equivalent electrons we have $G^k(nl;nl) = F^k(nl;nl)$ and Eq. (6.59b) can be rewritten in the form

$$\mathcal{K}(nlm_l;nlm_l') = \sum_{k=0}^\infty b^k(lm_l;lm_l')F^k(nl;nl).$$

Since $b^k(lm_l;l'm_{l'}) \geq 0$ and $F^k(nl;nl) > 0$ for all values of k, we have $\mathcal{K}(nlm_l;nlm_l') > 0$. □

6.4.3 The ground state of helium $^1S_0$

As a first example we analyze the ground state of ^{4}He, which consists of two electrons in the $1s^2$ configuration and $Z = 2$. The energy shift caused by the repulsion between the electrons in the same state is to first order in perturbation theory given by

$$\Delta\varepsilon_{1s} \simeq (1s,1s|\mathcal{H}'|1s,1s) = \mathcal{J}. \tag{6.71}$$

Using Eq. (6.59a) this can be written in the form

$$\Delta\varepsilon_{1s} = a^0(00;00)F^0(1s;1s). \tag{6.72}$$

Since $a^0(00;00) = 1$, see Problem 6.11, we find for the energy shift of one $1s$ electron by the other $1s$ electron

$$\Delta\varepsilon_{1s} = F^0(1s;1s) = \langle 1s|U_F^0(\rho)|1s\rangle. \tag{6.73}$$

Here $U_F^k(\rho)$ is the potential of direct screening of one of the $1s$ electrons by the other; i.e., $U_F^0(\rho) = U_{1s}(\rho)$ in the notation of Section 6.1.1. Substituting $\tilde{R}_{1s}(\rho) = Z^{3/2}2e^{-Z\rho}$ in Eq. (6.65a) for the wavefunction of the screening electron in a helium-like atom with nuclear charge Ze, we find with $k = 0$,

$$U_F^0(\rho) = \frac{1}{\rho}\int_0^\rho \left[\tilde{R}_{1s}(\varrho)\right]^2 \varrho^2 d\varrho + \int_\rho^\infty \frac{1}{\varrho}\left[\tilde{R}_{1s}(\varrho)\right]^2 \varrho^2 d\varrho = \frac{1}{\rho}\left[1 - e^{-2Z\rho}(1+Z\rho)\right] = U_{1s}(\rho). \tag{6.74}$$

Substituting this expression into Eq. (6.66a) we obtain with $Z = 2$ for the direct integral

$$\Delta\varepsilon_{1s} = F^0(1s;1s) = 32\int_0^\infty \left[1 - e^{-4\rho}(1+2\rho)\right] e^{-4\rho}\rho d\rho = \tfrac{5}{4}. \tag{6.75}$$

Note that this result also can be obtained directly from Table 6.2 because the $1s^2$ configuration involves only zero-order Laguerre polynomials. Restoring the dimension the energy shift is found to be $\Delta E_{1s} \simeq \frac{5}{2}hcR_\infty = 34.0$ eV, to be compared with the value 29.8 eV based on experiment. This shows that the perturbation theory overestimates the screening by 14% as was discussed in Section 6.1.1. It is straightforward to show with the aid of the closure approximation (see Section H.2) that the second order correction overshoots the proper value, which means that the multipole expansion (6.58) converges too slowly for an accurate determination of the screening. The best approximation for uncorrelated electrons is obtained with the Hartree method, yielding a total energy of -2.8617 Hartree (-77.87 eV) to be compared with -2.7450 Hartree (-74.83 eV) obtained by first-order perturbation theory and the experimental value of -2.903 Hartree (-78.99 eV).

6.4.4 The ground state of metastable triplet helium $^3S_1$

The level diagram of the $1snl$ excited state configurations of helium is sketched in Fig. 6.6. Our special interest concerns the $1s2s$ configuration, a typical example of a lifted exchange degeneracy in a system with two non-equivalent electrons. The $^3S_1$ excited state of this configuration with total spin $S = 1$ is famous because it is metastable against de-excitation to the $^1S_0$ ground state with a radiative lifetime of 8000 s. The atom is called metastable triplet helium with the notation He*. Neglecting the electrostatic repulsion between the electrons the energy of the system is given by

$$E^{(0)}_{1s\,2s} = -hcR_\infty\left(\frac{Z^2}{n^2} + \frac{Z^2}{n'^2}\right) = -5hcR_\infty \qquad (Z=2,\ n=1,\ n'=2)\,. \tag{6.76}$$

The radial distribution functions of the $\tilde{R}_{1s}(\rho)$ and $\tilde{R}_{2s}(\rho)$ radial wavefunctions are shown as the drawn lines in Fig. 6.7. The $2s$ orbital appears as a charged halo enclosing the $1s$ orbital with little

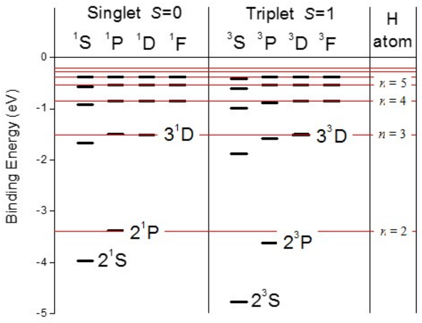


Figure 6.6: Energy-level diagram of the $1snl$ excited state configurations of helium. *Left*: singlet levels ($S = 0$) also known as *parahelium*; *Right*: triplet levels ($S = 1$) also known as *orthohelium.* Note that the highly-excited states give rise to hydrogenic energy levels, which means the screened nuclear charge is unity.

overlap in the charge distribution. Therefore, in view of the discussion of electrostatic screening in Section 6.1.1 we expect this screening to be more or less maximal for the $2s$ outer electron ($Z_{2s} \simeq Z-1$) and minimal for the $1s$ inner electron ($Z_{1s} \simeq Z$). The $2s$ charge distribution for the case $Z_{2s} = 1$ is shown as the dashed line in Fig. 6.7.

The energy shift caused by the repulsion between the electrons is to first order in perturbation theory given by Eq. (6.45), $\Delta E_{-} = \mathcal{J} - \mathcal{K}$, where in view of Eq. (6.61) only the $k = 0$ terms contribute to the direct and exchange integrals,

$$\mathcal{J} = (2hcR_\infty)\, a^0(00;00) F^0(1s;2s) \tag{6.77a}$$

$$\mathcal{K} = (2hcR_\infty)\, b^0(00;00) G^0(1s;2s). \tag{6.77b}$$

Note that $a^0(00;00) = b^0(00;00) = 1$ - see Problems 6.11 and 6.12. Substituting the expressions for the $\tilde{R}_{1s}(\rho)$ and $\tilde{R}_{2s}(\rho)$ hydrogenic wavefunctions we find, using Eqs. (6.65) with $k = 0$, for the $1s$ direct screening and the $1s2s$ exchange screening potentials

$$U_F^0(\rho) = (1/\rho)\left[1 - e^{-2Z\rho}(1 + Z\rho)\right] = U_{1s}(\rho) \tag{6.78a}$$

$$U_G^0(\rho) = (2/27)\,\sqrt{2}Ze^{-3Z\rho/2}\,[2 + 3Z\rho] = X_{1s2s}(\rho). \tag{6.78b}$$

Substituting these expressions into Eqs. (6.66) we obtain with $Z = 2$ for the direct and exchange radial integrals of the He* atom

$$F^0(1s;2s) = \int_0^\infty U_F^0(\rho)\left[\tilde{R}_{2s}(\rho)\right]^2 \rho^2 d\rho = 34/81 \tag{6.79a}$$

$$G^0(1s;2s) = \int_0^\infty U_G^0(\rho)\tilde{R}_{1s}(\rho)\tilde{R}_{2s}(\rho)\rho^2 d\rho = 32/729. \tag{6.79b}$$

Hence, the energy shift is found to be $\Delta\varepsilon_{2s} = F^0(1s;2s) - G^0(1s;2s)$. Restoring the units this becomes $\Delta E_{2s} = (548/729)\,hcR_\infty = 10.2$ eV and the estimate for the energy of the state to first order in perturbation theory is $E^{(1)}_{1s\,2s} = E^{(0)}_{1s\,2s} + \Delta E_{\rm scr} = -57.8$ eV. The experimental value is $E_{1s\,2s} = -59.2$ eV. Comparing to the energy -54.4 eV of the He$^+$ ion the state is calculated to be bound with an ionization energy of 3.4 eV. The experimental value for the ionization energy of ^{4}He* is 4.8 eV (see Fig. 6.6). Thus, like for the $1s^2$ configuration, also in the $1s2s$ case the perturbation theory underestimates the screening, 15.9% of 8.8 eV for ^{4}He*. For the singlet-triplet splitting we calculate $2\mathcal{K} = (64/729)\,2hcR_\infty = 2.39$ eV, to be compared with the experimental value

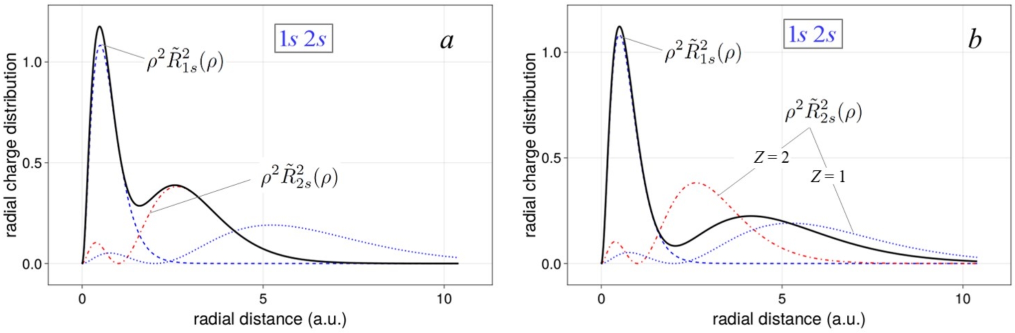


Figure 6.7: Radial distribution function for the $1s2s$ excited state configuration of the helium atom: (a) without screening; (b) with screening. Note that the $1s$ orbital is hardly affected but the $2s$ orbital is inflated as a result of screening of the nuclear charge by the $1s$ orbital. For comparison also the case of complete screened $2s$ orbital ($Z = 1$) is included.

$2\mathcal{K} = 0.80$ eV. The modest agreement is not surprising since the $\tilde{R}_{2s}(\rho)$ wavefunction does not have the long-range behavior corresponding to an effective nuclear charge $Z_{2s} = 1$. Obviously, this leaves room for a lot of improvement but this falls outside the scope of the course.

Problem 6.11. Show that for the configuration $ns\,n'l$ the a^k coefficient is given by

$$a^k(00;lm_l) = \begin{cases} 1 \text{ for } k = 0 \\ 0 \text{ for } k > 0 \end{cases}.$$

Solution. In view of Eqs. (6.61) $a^k(00, lm_l) = 0$ for $k > 0$. For $k = 0$ we have

$$a^k(00;lm_l) = (2l+1)\begin{pmatrix} 0\,0\,0 \\ 0\,0\,0 \end{pmatrix}^2 \begin{pmatrix} l\;0\;l \\ 0\,0\,0 \end{pmatrix} \begin{pmatrix} l & 0 & l \\ -m_l & 0 & m_l \end{pmatrix} = 1$$

because (see Appendix J)

$$\begin{pmatrix} 0\,0\,0 \\ 0\,0\,0 \end{pmatrix} = 1 \quad \text{and} \quad \sqrt{2l+1}\begin{pmatrix} l & 0 & l \\ -m_l & 0 & m_l \end{pmatrix} = (-1)^{-1}.$$ ☐

Problem 6.12. Show that for the configuration $ns\,n'l$ the b^k coefficient is given by

$$(2l+1)b^k(00;lm_l) = \begin{cases} 1 \text{ for } k = l \\ 0 \text{ for } k \neq l \end{cases}.$$

Solution. In view of Eqs. (6.61) $b^k(00, lm_l) = 0$ for $k \neq l$. For $k = l$ we have

$$b^k(00;lm_l) = (2l+1)\begin{pmatrix} 0\;l\;l \\ 0\,0\,0 \end{pmatrix}^2 \begin{pmatrix} 0 & l & l \\ 0 & -m_l & m_l \end{pmatrix}^2 = \frac{1}{(2l+1)},$$

because (see Appendix J)

$$\sqrt{2l+1}\begin{pmatrix} 0 & l & l \\ 0 & -m_l & m_l \end{pmatrix} = (-1)^{l-m_l}.$$ ☐

6.4.5 Helium-like atoms - energy levels

It is instructive to compare the energy-level diagram of helium with those of helium-like atoms. In Fig. 6.8 this is done for the alkaline-earth elements. The level energies are determined by optical spectroscopy. Note that the helium-like behavior is evident; in all cases we have a $n\,^1S_0$ electronic ground state and a similar structure of excited states featuring only singlet and triplet terms. The

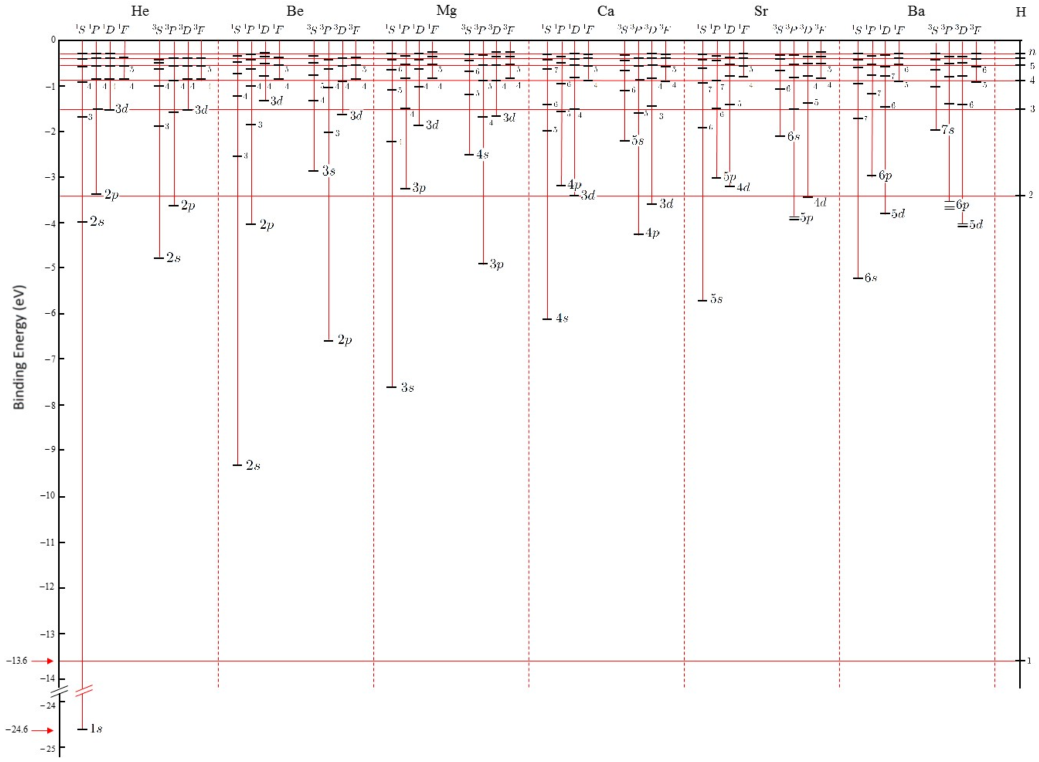


Figure 6.8: Energy levels of helium and helium-like atoms (the alkaline earth elements) for *nl* excitation of *one* of the *s* electrons from the $^1S_0$ ground state. The data are taken from the NIST Atomic Spectra Database Levels Data [66]. In the Sr data the fine-structure splitting of the $5\,^3P$ term is just visible in the plot; for Ba this is the case for both the $6\,^3P$ and $5\,^3D$ terms. Levels associated with two-electron excitations are not included in the diagram.

latter points to effective two-electron behavior. Apparently, the principal effect of the core electrons is to screen the nucleus. This shows up as a reduction of the ground-state binding energy with increasing Z. Aside from the strong similarities there are also important differences. For instance, only in the helium case ($n = 1$) we find a metastable $(n+1)\,^3S_1$ level (He*). In all other cases the lowest triplet term is $n\,^3P$ or even $n\,^3D$. The exceptional case of helium can be traced back to the non-existence of $1p$ orbitals. Optical transitions between singlet and triplet terms are referred to as *intercombination lines*. These lines are weak because the transitions are spin forbidden. The helium $2\,^3S_1$ state is special because the $1\,^1S_0 \leftrightarrow 2\,^3S_1$ transition is doubly forbidden (spin as well as dipole forbidden). With increasing Z some of the intercombination lines become weakly allowed. This has to do with the increasing importance of spin-orbit coupling as will be explained in Section 9.5 (crossover from LS coupling to jj coupling). The growing importance of fine-structure effects are even visible on the rough energy scale of Fig. 6.8, where the fine-structure splitting becomes just visible in the *strontium* ($Z = 38$) and *barium* ($Z = 56$) data.

7

Many-electron wavefunctions

To introduce the subject of many-electron wavefunctions, we first recall the consequences of the indistinguishability of identical particles in quantum mechanics. Two particles are called identical if there is no physical way to establish whether or not the particles have been exchanged. This condition is satisfied for particles with identical internal structure because, in quantum mechanics, it is fundamentally impossible to label a particle by tracking its motion along an orbit. The latter is only possible in classical mechanics as it requires the simultaneous determination of position and momentum. The state of the pair is represented by the pair wavefunction

$$\psi\left(\mathbf{r}_1, \sigma_1; \mathbf{r}_2, \sigma_2\right), \tag{7.1}$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the *position coordinates* and σ_1 and σ_2 the *spin coordinates*, respectively (see Appendix F.1.4). The squared modulus of the wavefunction,

$$|\psi\left(\mathbf{r}_1, \sigma_1; \mathbf{r}_2, \sigma_2\right)|^2, \tag{7.2}$$

corresponds to the probability of observing one particle (say particle 1) at position $\mathbf{r}_1$ in spin state σ_1 with the other particle (say particle 2) at position $\mathbf{r}_2$ in spin state σ_2. With this procedure we labeled the particles by the position of observation. It does *not* mean that we observe a labeled particle. The latter is only possible for particles with a difference in internal structure (the "label"). Next we allow for interchange of particles before doing the measurement. Formally this is realized by introducing the *exchange operator* P_{12},

$$P_{12}\psi\left(\mathbf{r}_1, \sigma_1; \mathbf{r}_2, \sigma_2\right) \equiv \psi\left(\mathbf{r}_2, \sigma_2; \mathbf{r}_1, \sigma_1\right). \tag{7.3}$$

As "tracking" is fundamentally impossible in quantum mechanics, we have no physical means to establish whether or not two identical particles have been exchanged. This means that the probability to observe particle 2 at position $\mathbf{r}_1$ in spin state σ_1 with particle 1 at position $\mathbf{r}_2$ in spin state σ_2 is equal to that of observing particle 1 at position $\mathbf{r}_1$ in spin state σ_1 with particle 2 at position $\mathbf{r}_2$ in spin state σ_2 :

$$|\psi\left(\mathbf{r}_1, \sigma_1; \mathbf{r}_2, \sigma_2\right)|^2 = |\psi\left(\mathbf{r}_2, \sigma_2; \mathbf{r}_1, \sigma_1\right)|^2. \tag{7.4}$$

This shows that P_{12} has to be *norm-conserving*,

$$\langle\psi|P_{12}^{\dagger}P_{12}|\psi\rangle = \langle\psi|\psi\rangle = 1, \tag{7.5}$$

but also that the particle indices, 1 and 2, have no physical significance.

Let us have a closer look at the exchange operator. Because P_{12} is norm-conserving, we have

$$P_{12}^{\dagger}P_{12} = \mathbb{1} \tag{7.6}$$

and

$$P_{12}|\psi\rangle = e^{-i\theta}|\psi\rangle. \tag{7.7}$$

Furthermore, exchanging the particles twice must leave the pair state unchanged. Rewriting the phase angle in the form $\theta = n\,\pi + \varphi$ (which can be done without loss of generality) we find from Eq. (7.7)

$$P_{12}^2|\psi\rangle = e^{-i2\varphi}|\psi\rangle. \tag{7.8}$$

So the question arises for what value of φ the pair state is invariant under all possible ways to exchange the particles twice. For 3D systems (and in the absence of topological excitations)[1] this is the case only for $\varphi = 0$; i.e., for

$$P_{12}^2 = \mathbb{1}. \tag{7.9}$$

Writing $P_{12}^{\dagger} = P_{12}^{\dagger}P_{12}^2 = P_{12}$ we see that P_{12} is *hermitian*; i.e., the eigenvalues of P_{12} are real an take the values ±1. In view of the definition (7.3) these eigenvalues correspond to pair wavefunctions which are either *symmetric* (+1) or *antisymmetric* (−1) under exchange of the particles.

As, by definition, the pair Hamiltonian $\mathcal{H}$ is invariant under exchange of identical particles, P_{12} commutes with $\mathcal{H}$, which implies that P_{12} and $\mathcal{H}$ share a complete set of eigenstates. Therefore, the eigenstates of P_{12} span the full Hilbert space of the pair and P_{12} is not only hermitian but also an *observable*. In nature, particles of a given species are always found in states showing the same symmetry under exchange, corresponding to *only one* of the eigenvalues of P_{12}. Apparently, for *identical* particles the pair wavefunction has to be an *eigenfunction* of the exchange operator; i.e., the exchange symmetry is conserved in time. If the wavefunction is symmetric under exchange of two particles the particles are called *bosons*, if antisymmetric they are called *fermions.* The division of particles into bosons and fermions points to a rigorously enforced underlying principle, which is found in relativistic quantum field theory in the form of the *spin-statistics theorem.* Discussion of this fundamental theorem falls outside the scope of this course. We accept it as a postulate in which bosons always carry *integral* total *spin* and fermions *half-integral* total spin. Taking into account the spin-statistics theorem, the action of the exchange operator is summarized by the expression

$$P_{12}\psi\left(\mathbf{r}_1,\sigma_1;\mathbf{r}_2,\sigma_2\right) = e^{-i2\pi s}\psi\left(\mathbf{r}_1,\sigma_1;\mathbf{r}_2,\sigma_2\right), \tag{7.10}$$

where s is the (integral or half-integral) spin of the particle.

Problem 7.1. Show that any pair wavefunction can be written as the sum of a part symmetric and a part antisymmetric under exchange of the particles of the pair.

Solution. For any pair state we have $|\psi\rangle = \frac{1}{2}\left(\mathbb{1} + P_{12}\right)|\psi\rangle + \frac{1}{2}(\mathbb{1} - P_{12})\,|\psi\rangle$, where P_{12} is the exchange operator, $P_{12}^2 = \mathbb{1}$. The first term is symmetric, $P_{12}\left(\mathbb{1} + P_{12}\right)|\psi\rangle = \left(P_{12} + P_{12}^2\right)|\psi\rangle = \left(\mathbb{1} + P_{12}\right)|\psi\rangle$, and the second term is antisymmetric, $P_{12}\left(\mathbb{1} - P_{12}\right)|\psi\rangle = \left(P_{12} - P_{12}^2\right)|\psi\rangle = -\left(\mathbb{1} - P_{12}\right)|\psi\rangle$. □

7.1 Quantum correlations, Pauli principle and Slater determinants

Let us specialize to the case of two $s = \frac{1}{2}$ fermions in the pair state $\psi\left(\mathbf{r}_1,\sigma_1;\mathbf{r}_2,\sigma_2\right)$. Eq. (7.10) shows that the pair state must be antisymmetric under exchange of the particles,

$$\psi\left(\mathbf{r}_2,\sigma_2;\mathbf{r}_1,\sigma_1\right) = -\psi\left(\mathbf{r}_1,\sigma_1;\mathbf{r}_2,\sigma_2\right). \tag{7.11}$$

For two fermions in the same spin state ($\sigma_2 = \sigma_1 = \sigma$) and at the same position ($\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$) this condition becomes $\psi\left(\mathbf{r},\sigma;\mathbf{r},\sigma\right) = -\psi\left(\mathbf{r},\sigma;\mathbf{r},\sigma\right)$. So, it follows that these fermions have zero

---

[1] The general case was analyzed by Jon Leinaas and Jan Myrheim in 1977 [71]. In 2D (or 3D in the presence of vorticity), the phase angle θ corresponds to a rotation angle, with $\theta_1 = n\,\pi + \varphi$ being equivalent to $\theta_2 = -(n\,\pi - \varphi)$. This implies that exchange with $\theta_1 + \theta_2 = 2\varphi \neq 0$ is a physical option. It gives rise to quantum statistics with its own type of identical particle called anyon (with fractional spin defined by $\theta = 2\pi s$) [121].

probability to be at the same position; i.e., *they avoid each other.* Apparently, for two fermions in the same spin state the motion is correlated as a result of the symmetry conditions. These kinematic correlations are purely quantum mechanical in origin and clearly *not* the result of potential forces (e.g., electrostatic) between the particles. The correlations affect the distribution over position and momentum with major consequences for statistical treatments of quantum many-body systems (in our case Fermi-Dirac statistics). In particular, the condition that two fermions in the same spin state cannot be at the same position excludes the possibility that they are in the same state. This is known as the *Pauli exclusion principle.*

7.1.1 Antisymmetric pair states for electrons

As an important example of $s=\frac{1}{2}$ fermions we consider the case of two electrons. With one electron in state $|\alpha_1\rangle$ and the other in state $|\alpha_2\rangle$, the standard method to form a pair state is by constructing the product state

$$|\alpha_1,\alpha_2) \equiv |\alpha_1\rangle_1 \otimes |\alpha_2\rangle_2, \tag{7.12}$$

where the ket $|\alpha_\kappa\rangle_i$ represents electron i in state $|\alpha_\kappa\rangle$. However, this state state does not qualify as a proper two-electron state because it is not antisymmetric under exchange as is required for fermionic particles. The simplest pair state with the proper symmetry has the form of a 2×2 determinant,[1]

$$|\alpha_1,\alpha_2\rangle = \sqrt{\tfrac{1}{2}}\,[|\alpha_1,\alpha_2) - |\alpha_2,\alpha_1)] = \sqrt{\tfrac{1}{2}}\begin{vmatrix} |\alpha_1\rangle_1 & |\alpha_2\rangle_1 \\ |\alpha_1\rangle_2 & |\alpha_2\rangle_2 \end{vmatrix}. \tag{7.13}$$

By construction determinants are antisymmetric under exchange of the rows (particle index) or, equivalently, the columns (state index),

$$P_{12}|\alpha_1,\alpha_2\rangle = -|\alpha_2,\alpha_1\rangle. \tag{7.14}$$

For equal states, $|\alpha_1\rangle = |\alpha_2\rangle$, the determinant vanishes in accordance with the Pauli principle.

Now we apply this symmetrization procedure to *spinorbitals*, the single electron eigenstates. These are defined by specifying both the orbital u and the spin state ($|\uparrow\rangle$ or $|\downarrow\rangle$ of the electron. In search for a compact notation we decorate the orbital with a dot for *spin-up* and with a bar for *spin-down*,

$$|\dot{u}\rangle \equiv |u\uparrow\rangle \equiv |u\rangle \otimes |\uparrow\rangle \qquad \text{and} \qquad |\overline{u}\rangle \equiv |u\downarrow\rangle \equiv |u\rangle \otimes |\downarrow\rangle\,. \tag{7.15}$$

In this notation any properly symmetrized state of two electrons in the orbitals u and v can be written as a linear combination of an orthonormal set of four 2×2 determinants,

$$|\dot{u},\dot{v}\rangle\,, |\dot{u},\overline{v}\rangle\,, |\overline{u},\dot{v}\rangle\,, |\overline{u},\overline{v}\rangle\,. \tag{7.16}$$

Because the sign of the determinants depends on the ordering of the rows and columns we fix the phase by convention. We shall use alphabetic ordering of the orbitals (u precedes v) followed by *up* precedes *down.*

7.1.1.1 Coulomb repulsion integrals for electron pairs

The determinants of the basis (7.16) are properly symmetrized but they are not the singlet and triplet eigenstates obtained for the Coulomb repulsion energy in the helium atom (see Section 6.3). The reason is that the 2×2 determinants are eigenstates of the exchange operator P_{12} but *not* of the atomic hamiltonian of two electrons moving in the central potential $-Z(\rho)/\rho$,

$$\mathcal{H} = \sum_{i=1}^{2}\left[-\tfrac{1}{2}\boldsymbol{\nabla}_i^2 - Z(\rho_i)/\rho_i\right] + \rho_{12}^{-1}. \tag{7.17}$$

---

[1] Note the use of curved brackets for product states and angled brackets for symmetrized states.

As all electrons are identical and the Coulomb repulsion only depends of the interparticle distance, $\mathcal{H}$ is manifestly invariant under exchange of the electrons. Hence, P_{12} commutes with $\mathcal{H}$ and the operators P_{12} and $\mathcal{H}$ must share a complete set of eigenstates. This means that the energy eigenstates can be written in a form antisymmetric under exchange of the electrons. The difficulty is that, in general, this cannot be done using a single 2×2 determinant. In other words, $\mathcal{H}$ is not diagonal in the basis (7.16). Not surprisingly, this is caused by the perturbation $\mathcal{H}' = \rho_{12}^{-1}$, which in matrix notation takes the form

$$\begin{array}{cccc} \langle \dot{u}, \dot{v}| & \langle \dot{u}, \overline{v}| & \langle \overline{u}, \dot{v}| & \langle \overline{u}, \overline{v}| \end{array} \tag{7.18}$$

$$\mathcal{H}' = \begin{array}{c} \begin{pmatrix} \mathcal{J}-\mathcal{K} & & & \\ & \mathcal{J} & -\mathcal{K} & \\ & -\mathcal{K} & \mathcal{J} & \\ & & & \mathcal{J}-\mathcal{K} \end{pmatrix} \end{array} \begin{array}{l} |\dot{u}, \dot{v}\rangle \\ |\dot{u}, \overline{v}\rangle \\ |\overline{u}, \dot{v}\rangle \\ |\overline{u}, \overline{v}\rangle \end{array} \tag{7.19}$$

Let us verify this for the matrix elements of the *spin-up* stretched state $|\dot{u}, \dot{v}\rangle = |u, v\rangle\, |\upharpoonleft\upharpoonright)$. Expanding the determinant, $|\dot{u}, \dot{v}\rangle = \frac{1}{\sqrt{2}}\left[|uv) - |vu)\right] |\upharpoonleft\upharpoonright)$, we obtain for the diagonal contribution

$$\langle \dot{u}, \dot{v}|\, \rho_{12}^{-1}\, |\dot{u}, \dot{v}\rangle = \tfrac{1}{2}\left[(uv|\rho_{12}^{-1}|uv) - (uv|\rho_{12}^{-1}|vu) - (vu|\rho_{12}^{-1}|uv) + (vu|\rho_{12}^{-1}|vu)\right] = \mathcal{J} - \mathcal{K}, \tag{7.20}$$

where $\mathcal{J}$ and $\mathcal{K}$ are the direct and exchange integrals defined in Eqs. (6.46a) and (6.46b), respectively. The off-diagonal contributions of $|\dot{u}, \dot{v}\rangle$ are all zero because because at least one of the electrons of the hermitian conjugates is in a down-spin state. The matrix elements of the non-stretched states are calculated in Problems 7.2 and 7.3. Diagonalizing the Hamiltonian we obtain the singlet and triplet eigenvalues

$$E_{0,0}(u, v) = \mathcal{J}(u, v) + \mathcal{K}(u, v) = \Delta E_{+} \tag{7.21a}$$

$$\left.\begin{array}{r} E_{1,1}(u, v) \\ E_{1,0}(u, v) \\ E_{1,-1}(u, v) \end{array}\right\} = \mathcal{J}(u, v) - \mathcal{K}(u, v) = \Delta E_{-}. \tag{7.21b}$$

These energies correspond to the Coulomb repulsion shift with respect to the energy of two noninteracting electrons moving in the central potential $-Z(\rho)/\rho$. These shifts were also obtained (for $Z(\rho) \equiv 2$) in Section 6.3. The corresponding eigenstates are (in standard ordering)

$$\begin{aligned} |\psi_{u,v}^{-}; 1, 1\rangle &= |\dot{u}, \dot{v}\rangle \\ |\psi_{u,v}^{-}; 1, 0\rangle &= \tfrac{1}{\sqrt{2}}\left[|\dot{u}, \overline{v}\rangle + |\overline{u}, \dot{v}\rangle\right] \qquad\qquad |\psi_{u,v}^{+}; 0, 0\rangle = \tfrac{1}{\sqrt{2}}\left[|\dot{u}, \overline{v}\rangle - |\overline{u}, \dot{v}\rangle\right]. \\ |\psi_{u,v}^{-}; 1, -1\rangle &= |\overline{u}, \overline{v}\rangle \end{aligned} \tag{7.22}$$

Note that these eigenstates are also eigenstates of the total spin, $\mathbf{S} = \mathbf{s}_1 + \mathbf{s}_2$, which is not surprising because spin is conserved by the hamiltonian (7.17). This reveals the possibility of circumventing the diagonalization by generating the determinants as a product of properly symmetrized spin states and orbital states; e.g., the basis $\{|LM_L SM_S\rangle\}$ or the basis $\{|LSJM_J\}$ as these are also eigenstates of the exchange operator. We return to this approach in Chapter 8.

Problem 7.2. Show that for the matrix elements of the type $\langle \dot{u}\overline{v}|\rho_{12}^{-1}|\dot{u}\overline{v}\rangle$ the *exchange* terms vanish,

$$\langle \dot{u}\overline{v}|\rho_{12}^{-1}|\dot{u}\overline{v}\rangle = (uv|\rho_{12}^{-1}|uv) = \mathcal{J}.$$

Solution. Using the determinant $|\dot{u}, \overline{v}\rangle = \frac{1}{\sqrt{2}}\left[|\dot{u}\overline{v}) - |\overline{v}\dot{u})\right]$ the matrix element becomes

$$\begin{aligned} \langle \dot{u}\overline{v}|\rho_{12}^{-1}|\dot{u}\overline{v}\rangle &= \tfrac{1}{2}\left[(\dot{u}\overline{v}|\rho_{12}^{-1}|\dot{u}\overline{v}) - (\dot{u}\overline{v}|\rho_{12}^{-1}|\overline{v}\dot{u}) - (\overline{v}\dot{u}|\rho_{12}^{-1}|\dot{u}\overline{v}) + (\overline{v}\dot{u}|\rho_{12}^{-1}|\overline{v}\dot{u})\right] \\ &= \tfrac{1}{2}\left[(uv|\rho_{12}^{-1}|uv) + (vu|\rho_{12}^{-1}|vu)\right] = \mathcal{J}, \end{aligned}$$

where $\mathcal{J}$ is the direct integral (6.46a). Note that $(\dot{u}\overline{v}|\rho_{12}^{-1}|\overline{v}\dot{u}) = (\upharpoonleft\downharpoonright|\downharpoonleft\upharpoonright)\,\mathcal{K} = 0$ and $(\overline{v}\dot{u}|\rho_{12}^{-1}|\dot{u}\overline{v}) = (\downharpoonleft\upharpoonright|\upharpoonleft\downharpoonright)\,\mathcal{K} = 0$, where $\mathcal{K}$ is the exchange integral (6.46b). □

Problem 7.3. Show that for the matrix elements of the type $\langle\dot{u},\overline{v}|\rho_{12}^{-1}|\dot{v},\overline{u}\rangle$ the *direct* terms vanish,

$$\langle\dot{u}\overline{v}|\rho_{12}^{-1}|\dot{v}\overline{u}\rangle=(uv|\rho_{12}^{-1}|vu)=\mathcal{K}.$$

Solution. Using the determinants $|\dot{u}\overline{v}\rangle=\frac{1}{\sqrt{2}}\left[|\dot{u}\overline{v})-|\overline{v}\dot{u})\right]$ and $|\dot{v}\overline{u}\rangle=\frac{1}{\sqrt{2}}\left[|\dot{v}\overline{u})-|\overline{u}\dot{v})\right]$ the matrix element becomes

$$\begin{aligned}\langle\dot{u}\overline{v}|\rho_{12}^{-1}|\dot{u}\overline{v}\rangle&=\tfrac{1}{2}\left[(\dot{u}\overline{v}|\rho_{12}^{-1}|\dot{v}\overline{u})-(\dot{u}\overline{v}|\rho_{12}^{-1}|\overline{u}\dot{v})-(\overline{v}\dot{u}|\rho_{12}^{-1}|\dot{v},\overline{u})+(\overline{v}\dot{u}|\rho_{12}^{-1}|\overline{u}\dot{v})\right]\\&=\tfrac{1}{2}\left[(uv|\rho_{12}^{-1}|vu)+(vu|\rho_{12}^{-1}|uv)\right]=\mathcal{K},\end{aligned}$$

where $\mathcal{K}$ is the exchange integral (6.46b). Note further that $(\dot{u}\overline{v}|\rho_{12}^{-1}|\overline{u},\dot{v})=(\uparrow\downarrow|\downarrow\uparrow)\,\mathcal{J}=0$ and $(\overline{v}\dot{u}|\rho_{12}^{-1}|\dot{v},\overline{u})=(\downarrow\uparrow|\uparrow\downarrow)\,\mathcal{J}=0$ is the direct integral (6.46a). □

7.1.2 Spinorbitals and Slater determinants

An important tool to deal with many-electron systems is the *Slater determinant.* Its discovery by John Slater marks a true milestone in many-body physics [103]. The Slater determinant is the simplest generalization of the product wavefunction with the proper symmetry under interchange of any two electrons and consistent the Pauli principle. For N fermions the Slater determinant is written as

$$\psi_\alpha(\mathbf{r}_1,\sigma_1;\cdots;\mathbf{r}_N,\sigma_N)=\sqrt{\frac{1}{N!}}\begin{vmatrix}\varphi_{\alpha_1}(\mathbf{r}_1,\sigma_1)&\cdots&\varphi_{\alpha_N}(\mathbf{r}_1,\sigma_1)\\\vdots&\cdots&\vdots\\\varphi_{\alpha_1}(\mathbf{r}_N,\sigma_N)&\cdots&\varphi_{\alpha_N}(\mathbf{r}_N,\sigma_N)\end{vmatrix}.\tag{7.23}$$

Indeed, the property of determinants to vanish when two columns or two rows are equal assures that the wavefunction vanishes when two electrons are in the same state α or share the same (position and spin) coordinates $(\mathbf{r}_i,\sigma_i)$, while exchanging two rows or two columns yields the minus sign required for anti-symmetric wavefunctions. In Dirac notation the antisymmetrized form of N fermions in states $\alpha_1,\cdots,\alpha_N$ is given by

$$|\psi_\alpha\rangle\equiv|\alpha_1,\cdots,\alpha_N\rangle\equiv\sqrt{\frac{1}{N!}}\sum_P(-1)^pP|\alpha_1,\cdots\alpha_N),\tag{7.24}$$

where

$$|\alpha_1,\cdots\alpha_N)\equiv|\alpha_1\rangle_1\otimes|\alpha_2\rangle_2\otimes\cdots|\alpha_N\rangle_N\equiv|\psi_\alpha)\tag{7.25}$$

is the N-body ordered product of the single-particle states $|\alpha_\kappa\rangle_i$, where $\kappa\in\{1,\cdots,N\}$ is the *state index* and $i\in\{1,\cdots,N\}$ the *particle index.* The sum runs over all permutations P while accounting for their *parity* by the phase factor $(-1)^p$. Here p stands for the number of *transpositions* (binary interchanges) required to realize the permutation P starting from an initial ordering fixed by convention. The *parity* is called *even/odd* depending on p being even/odd. As the sum runs over all permutations, it makes no difference whether we permute all particles or permute all states of the particles. We choose the permutation operator P *to act on the state index* (κ) and *not on the particle index* (i). With this choice, the interchange of the states of particles 1 and 2 is written as

$$P|\alpha_1,\alpha_2,\cdots\alpha_N)=|\alpha_2,\alpha_1,\cdots\alpha_N)=|\alpha_2\rangle_1\otimes|\alpha_1\rangle_2\otimes\cdots|\alpha_N\rangle_N\,.\tag{7.26}$$

To assure a uniquely defined sign of the Slater determinants we shall adopt the *standard ordering convention* of atomic configurations (see later). The state labeling $\alpha_1,\cdots,\alpha_N$ refers to the *spinorbitals* introduced in Dirac notation in Eqs. (7.15). The spinorbital $\varphi_{u\uparrow}(\mathbf{r}_i,\sigma_i)=\varphi_u(\mathbf{r})\,\chi_\uparrow(\sigma)$ corresponds to spin-up electron in the motional state $\varphi_u(\mathbf{r})$.

7.1.3 Slater determinants - notations and ordering convention

The rows of the Slater determinant (7.23) differ only in the particle index and the columns only in the state index. Note that is a matter of convention because determinants are invariant under transposition - see property M.21. Whatever our preference and without loss of information, Eq. (7.23) can be written in the short-hand form for the singlets.

$$\psi_\alpha\left(\mathbf{r}_1,\sigma_1;\cdots;\mathbf{r}_N,\sigma_N\right)=\det(\alpha_1,\cdots,\alpha_N). \tag{7.27}$$

We will use the Dirac notation, in which the N-electron state $|\psi_\alpha\rangle=|\alpha_1,\cdots,\alpha_N\rangle$ is constructed from one-electron spinorbitals

$$|\alpha_\kappa\rangle=|n^\kappa l^\kappa m_l^\kappa s^\kappa m_s^\kappa\rangle=|u_\kappa\rangle\otimes|s^\kappa m_s^\kappa\rangle, \tag{7.28}$$

where $\kappa\in\{1,\cdots N\}$ is the *state index*. Using the arrow notation for spin, the N-electron state for the spinorbitals $|u{\uparrow}\rangle$, $|v{\uparrow}\rangle,\cdots|w{\downarrow}\rangle$ can be written as (again without loss of information)

$$|\psi_\alpha\rangle=|u{\uparrow},v{\uparrow},,\cdots,w{\downarrow}\rangle\equiv|\dot{u},\dot{v},\cdots,\bar{w}\rangle, \tag{7.29}$$

where the spin "up" and "down" are indicated by putting a dot or a bar, respectively.

For *equivalent* electrons the states differ in the magnetic quantum numbers,

$$|\alpha_\kappa\rangle=|m_l^\kappa m_s^\kappa\rangle_{nl}. \tag{7.30}$$

For two equivalent s electrons the only possible determinant consistent with the Pauli principle is $|\dot{0},\bar{0}\rangle_{ns^2}\equiv|\upharpoonleft\!\downharpoonright\rangle_{ns^2}$, *e.g.* the ground state of the helium atom is [1]

$$|\psi_{\mathrm{He}}\rangle=|1\dot{s},1\bar{s}\rangle\equiv|\dot{0},\bar{0}\rangle_{1s^2}\equiv|\upharpoonleft\!\downharpoonright\rangle_{1s^2}. \tag{7.31}$$

If we have more than one occupied shell consisting the full state is represented by a joined determinant of order equal to the total number of electrons. Thus we obtain for *beryllium*

$$|\psi_{\mathrm{Be}}\rangle=|1\dot{s},1\bar{s},2\dot{s},2\bar{s}\rangle\equiv|(\dot{0},\bar{0})_{1s^2}(\dot{0},\bar{0})_{2s^2}\rangle=|(\upharpoonleft\!\downharpoonright)_{1s^2}(\upharpoonleft\!\downharpoonright)_{2s^2}\rangle \tag{7.32}$$

and for neon

$$|\psi_{\mathrm{Ne}}\rangle=|1\dot{s},1\bar{s};2\dot{s},2\bar{s};2\dot{p}_1,2\bar{p}_1,2\dot{p}_0,2\bar{p}_0,2\dot{p}_{-1},2\bar{p}_{-1}\rangle \tag{7.33a}$$

$$=|(\dot{0},\bar{0})_{1s^2}(\dot{0},\bar{0})_{2s^2}(\dot{1},\bar{1},\dot{0},\bar{0},-\dot{1},-\bar{1})_{2p^6}\rangle\equiv|(\upharpoonleft\!\downharpoonright)_{1s^2}(\upharpoonleft\!\downharpoonright)_{2s^2}(\upharpoonleft\!\downharpoonright\ \upharpoonleft\!\downharpoonright\ \upharpoonleft\!\downharpoonright)_{2p^6}\rangle. \tag{7.33b}$$

In this way we can write the state of any closed shell configuration in the form of a single Slater determinant: the determinant for which the states satisfy the exclusion principle (i.e., the determinant is non-zero). The example of Strontium is given in Fig. 7.1. Note that the notation with the arrows is the most compact and used in the well-known configuration plots for the various elements. The state notation

$$|2\dot{p}_1,2\bar{p}_1,2\dot{p}_0,2\bar{p}_0,2\dot{p}_{-1},2\bar{p}_{-1}\rangle=|\dot{1},\bar{1},\dot{0},\bar{0},-\dot{1},-\bar{1}\rangle_{2p^6} \tag{7.34}$$

is written conform the *standard ordering convention* for the sign of the determinant: high M_L precedes low M_L and spin up precedes spin down in the *non-permuted* product state $|\dot{1},\bar{1},\dot{0},\bar{0},-\dot{1},-\bar{1})$.

With partially filled shells the situation is more complicated because - in general - they cannot be represented by a single Slater determinant. In the case of two-electron atoms we are dealing with two valence electrons in ns^2, np^2, nd^2 or nf^2 configurations orbiting about an atomic core presenting at large distance from the nucleus a net positive charge $2e$. For the ns^2 configurations

---

[1] Sliced arrows are used for convenience of notation but have no physical significance.

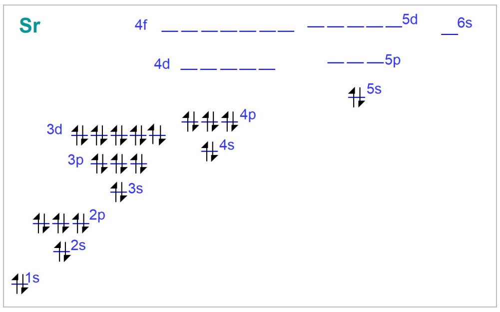


Figure 7.1: Configuration plot for the ground state of Strontium. Note that the arrow notation is a shorthand notation for Slater determinants as explained in the text. The slicing of the arrows has no significance other than offering a compact notation.

one distinguishes *helium-like* atoms and *alkaline-earth-like* atoms.[1] The helium-like atoms were discussed in Chapter 6. Alkaline-earth-like atoms are many-electron atoms consisting of two outer electrons in a ns^2 configuration and a doubly charged central core, thus resembling a two-electron atom like helium. Aside from the alkaline-earth elements (group II) atoms Be, Mg, Ca, Sr and Ba this class also includes the ions of the "three electron atoms" (group III atoms), B^+, Al^+, Ga^+, In^+ and Ti^+. The group IV atoms have configuration np^2 (C, Si, Ge, Sn, Pb). The nd^2 atoms (Ti, Zr, Hf, Rf) are transition metals. For each electron in a nl^2 configuration there are $\nu = 2(2l+1)$ possible states $|nlm_l sm_s\rangle$. For non-equivalent electrons this implies ν^2 possible pair states; i.e., each electron can be paired with an electron in any of the ν possible states. For equivalent electrons we have to exclude the occurrence of two electrons in the same state. This leaves

$$\frac{1}{2}\nu(\nu-1) = \frac{\nu!}{2!(\nu-2)!} = \binom{\nu}{2} \tag{7.35}$$

combinations consistent with the Pauli principle. This procedure is readily generalized to x equivalent electrons in a nl^x configuration. To avoid identical pairs we have to choose x out of ν, which can be done in

$$\binom{\nu}{x} = \frac{\nu!}{x!(\nu-x)!} \tag{7.36}$$

different ways as is well known from combinatorics. Thus, Eq. (7.36) provides the total number of different x-body states consistent with the Pauli principle in nl^x configurations. We return to these cases in Chapter 9.

Problem 7.4. Show that Slater determinants are invariant under unitary transformations.

Solution. To demonstrate this property we write the Slater determinant $|\psi_{\alpha'}\rangle$ in the form

$$|\psi_{\alpha'}\rangle = \sqrt{\frac{1}{N!}} \begin{vmatrix} |\alpha'_1\rangle_1 & \cdots & |\alpha'_1\rangle_N \\ \vdots & \dots & \vdots \\ |\alpha'_N\rangle_1 & \cdots & |\alpha'_N\rangle_N \end{vmatrix} = \sqrt{\frac{1}{N!}}\,\det(\alpha').$$

Next we consider the subspace defined by the spinorbitals $|\alpha'_1\rangle, \cdots |\alpha'_N\rangle$. Let $|\alpha_1\rangle, \cdots |\alpha_N\rangle$ be an alternative basis for the same subspace. Using the closure relation we define the basis transformation

$$|\alpha'_j\rangle = \sum_{i=1}^{N} |\alpha_i\rangle\langle\alpha_i|\alpha'_j\rangle = \sum_{i=1}^{N} U_{ij}|\alpha_i\rangle.$$

---

[1] The *alkaline-earth elements* are to be distinguished from the *alkaline earths*, which are the oxides of the elements. Historically, the notion earths was used for water-insoluble substances that do not burn in fire.

Since $\langle\alpha'_j|\alpha_i\rangle = \langle\alpha_i|\alpha'_j\rangle^*$ this is a unitary transformation. Applied to the matrix α' we have

$$\alpha' = U\alpha = \begin{pmatrix} \langle\alpha_1|\alpha'_1\rangle & \cdots & \langle\alpha_N|\alpha'_1\rangle \\ \vdots & \ldots & \vdots \\ \langle\alpha_1|\alpha'_N\rangle & \cdots & \langle\alpha_N|\alpha'_N\rangle \end{pmatrix} \begin{pmatrix} |\alpha_1\rangle_1 & \cdots & |\alpha_1\rangle_N \\ \vdots & \ldots & \vdots \\ |\alpha_N\rangle_1 & \cdots & |\alpha_N\rangle_N \end{pmatrix}.$$

Using the properties of determinants (see Appendix M.2) we find

$$\det(\alpha') = \det(U\alpha) = \det(U)\det(\alpha) = e^{i\xi}\det(\alpha),$$

where ξ is a real number. This expression shows that (up to a phase factor) the Slater determinant is invariant under unitary transformation, $|\psi_{\alpha'}\rangle = e^{i\xi}|\psi_\alpha\rangle$; the phase factor is of no consequence because it does not affect the expectation values. □

7.1.4 Slater determinants - parity and electric-dipole matrix elements

In Section 2.5.1 we found that electric-dipole transitions require a change parity of the atomic state. This selection rule was derived for one-electron atoms. For many-electron atoms this can be different. Therefore, we reconsider the dipole operator as well as the parity of many-electron states.

The *electric-dipole operator* of an N electron atom is given by

$$\mathbf{D} = -\,e\sum_{i=1}^{N}\mathbf{r}_i = -ea\rho\sum_{i=1}^{N}\hat{\mathbf{r}}_i. \tag{7.37}$$

Changing the radius vectors of all electrons from $\mathbf{r}_i$ to $-\mathbf{r}_i$, with $i \in \{1, N\}$, we find that the total dipole operator has *odd* parity; i.e., also in the many-electron electron case electric-dipole transitions require a change parity of the atomic state.

Turning to the states, we note that Slater determinants are linear combinations of products of single particle spinorbitals. As the parity of spinorbital $|\alpha_\kappa\rangle$ is $(-1)^{l_\kappa}$, the parity of the full Slater determinant is given by

$$(-1)^{\sum_\kappa l_\kappa} = \prod_\kappa (-1)^{l_\kappa}. \tag{7.38}$$

This expression shows that the parity of Slater determinant is even unless $\sum_\kappa l_\kappa = \text{odd}$. Phrased differently, the parity is even except for atoms with an odd number of electrons in odd-parity states. Only in the case of one-electron atoms the total angular momentum is decisive for the parity of the many-electron state. For this reason, in the case of many-electron atoms, the parity of *odd-parity* total angular momentum states is explicitly indicated in the spectroscopic notation by an upper index o. In the case of even parity the notation remains unchanged; e.g.,

$$1s^2 2p\,{}^2P \to 1s^2 2p\,{}^2P^o \tag{7.39a}$$

$$npnd\,{}^3D \to npnd\,{}^3D^o \quad \text{and} \quad np^2\,{}^3D \to np^2\,{}^3D. \tag{7.39b}$$

7.2 Matrix elements of operators between Slater determinants

Dealing with systems of more than one electron means dealing with many-body operators and antisymmetrized N-body states. At first sight it may seem a complicating factor to deal with wavefunctions in the form of Slater determinants and in a way it is, but fortunately a number of simplifying expressions may be derived relating the properties of N-body states to those of single electrons, pairs, triples, *etc.*. It is the purpose of the present section to derive a number of such expressions for future use. We consider the matrix elements for common operator types in combination with antisymmetric N-electron basis states $|\psi_\alpha\rangle = |\alpha_1, \cdots, \alpha_N\rangle$, where the α_κ, with $\kappa \in \{1, \cdots, N\}$, correspond to one-electron spinorbitals.

7.2.1 One-body operators

First we consider the matrix elements of a sum of *one-body operators*,

$$Q_1 = \sum_{i=1}^{N} q_i, \tag{7.40}$$

where the operator q_i acts only on the electron with index $i \in \{1, \cdots, N\}$. A good example of an operator of this type is the Schrödinger Hamiltonian for electrons in the unscreened Coulomb field of the nucleus. Using Eq. (7.24) the most general matrix element of Q_1 with determinantal wavefunctions is of the form

$$\langle \psi_{\alpha'} | Q_1 | \psi_\alpha \rangle = \frac{1}{N!} \sum_{i=1}^{N} \sum_{P,P'} (-1)^{p+p'} (\psi_{\alpha'} | P' q_i P | \psi_\alpha), \tag{7.41}$$

where $|\psi_\alpha) \equiv |\alpha_1, \cdots, \alpha_N)$ is the non-symmetrized product state in standard order; p and p' are the parities of the permutations P and P', respectively. In view of the orthogonality of the spinorbitals $|\alpha_\kappa\rangle$ the matrix elements $(\psi_{\alpha'} | P' q_i P | \psi_\alpha) \equiv (\alpha_1, \cdots, \alpha_{\kappa'}, \cdots \alpha_N | P' q_i P | \alpha_1, \cdots, \alpha_\kappa, \cdots \alpha_N)$ are zero for all possible permutations P and P' whenever $|\psi_\alpha\rangle$ and $|\psi_{\alpha'}\rangle$ differ in *more than one* pair of spinorbitals.

- *Off-diagonal matrix elements differing in one pair.* Let $|\psi_\alpha\rangle$ and $|\psi_{\alpha'}\rangle$ differ in one pair of spin orbitals, α_κ and $\alpha_{\kappa'}$. Then, for every value of $i \in \{1, \cdots, N\}$ there are $(N-1)!$ permutations $P'|\alpha_1, \cdots, \alpha_{\kappa'}, \cdots \alpha_N)$ for which $\alpha_{\kappa'}$ appears on position i. For each of these permutations there is exactly one permutation $P|\alpha_1, \cdots, \alpha_\kappa, \cdots \alpha_N)$ for which also α_κ appears on position i while, in addition, $(\psi_{\alpha'} | P' q_i P | \psi_\alpha) \neq 0$. For this permutation the N-body matrix elements $(\psi_{\alpha'} | P' q_i P | \psi_\alpha)$ reduce to a one-body matrix element involving only electron i,

  $$(\psi_{\alpha'} | P' q_i P | \psi_\alpha) = \langle \alpha_{\kappa'} | q_i | \alpha_\kappa \rangle_i. \tag{7.42}$$

  Importantly, the sign $(-1)^{p+p'}$ corresponding to P and P' has the same value for all values of i, $+1$ or -1 depending on the ordering convention. Further, also the matrix elements $\langle \alpha_{\kappa'} | q_i | \alpha_\kappa \rangle_i$ are the same for all electrons; hence, without loss of generality we can drop the particle index and write[1]

  $$\langle \alpha_{\kappa'} | q_i | \alpha_\kappa \rangle_i = \langle \alpha_{\kappa'} | q | \alpha_\kappa \rangle. \tag{7.43}$$

  Summing $(\psi_{\alpha'} | P' q_i P | \psi_\alpha)$ over $i \in \{1, \cdots, N\}$ and over all permutations P, P' we arrive at $N(N-1)! = N!$ nonzero contributions of equal value $(-1)^{p+p'} \langle \alpha_{\kappa'} | q | \alpha_\kappa \rangle$ and Eq. (7.41) reduces to a single one-body matrix element,

  $$\langle \psi_{\alpha'} | Q_1 | \psi_\alpha \rangle = \langle \alpha_{\kappa'} | q | \alpha_\kappa \rangle \frac{1}{N!} \sum_{i=1}^{N} \sum_{P}{}' (-1)^{p+p'} = \pm \langle \alpha_{\kappa'} | q | \alpha_\kappa \rangle, \tag{7.44}$$

  where the sign depends on the ordering convention. The prime indicates that the permutation sum excludes, for a given value of i, all permutations P in which α_κ appears on position i but $(\psi_{\alpha'} | P' q_i P | \psi_\alpha) = 0$. We thus established that an operator of the type Q_1 can couple two one-electron states. Note that the corresponding matrix element does not depend on the particle index. This reflects the indistinguishability of identical particles.

- *Diagonal matrix elements.* In this case $|\psi_\alpha\rangle = |\psi_{\alpha'}\rangle$. Adopting the same approach as for the off-diagonal elements we note that in the diagonal case the N-body matrix elements

---

[1] For example for the hydrogenic ns orbitals we calculate $\langle ns | Z/\rho_i | ns \rangle = \int |\phi_{ns}(\rho_i)|^2 \rho_i d\rho_i = Z^2/n^2$.

are only nonzero if $P' = P$. Again we find that for every value of $i \in \{1, \cdots, N\}$ there are $(N-1)!$ permutations $P|\alpha_1, \cdots, \alpha_\kappa, \cdots \alpha_N)$ for which α_κ appears on position i while $(\psi_\alpha|Pq_iP|\psi_\alpha) \neq 0$ but in the present case this is possible *for each value of* $\kappa \in \{1, \cdots, N\}$. For each set of $(N-1)!$ permutations the N-body matrix elements reduce to the same one-body matrix element, $(\psi_\alpha|Pq_iP|\psi_\alpha) = (-1)^{2p}\langle\alpha_\kappa|q|\alpha_\kappa\rangle = \langle\alpha_\kappa|q|\alpha_\kappa\rangle$. Summing $(\psi_\alpha|Pq_iP|\psi_\alpha)$ over $i \in \{1, \cdots, N\}$ and over all permutations P we have *this time for each* $\kappa \in \{1, \cdots, N\}$ again $N(N-1)! = N!$ nonzero terms of equal value $\langle\alpha_\kappa|q|\alpha_\kappa\rangle$ and Eq. (7.44) is replaced by

$$\langle\psi_\alpha|Q_1|\psi_\alpha\rangle = \sum_{\kappa=1}^{N}\langle\alpha_\kappa|q|\alpha_\kappa\rangle. \tag{7.45}$$

This result is intuitively clear: the expectation value of the operator Q_1 represents the sum of the contributions of all spin-orbitals contained in the many-body state $|\alpha_1, \cdots, \alpha_N\rangle$. Again the result is independent of the particle indices as expected for identical particles.

7.2.2 Two-body operators

Next we consider the matrix elements of a sum of *two-body operators*,

$$Q_2 = \tfrac{1}{2}\sum_{i,j=1}^{N}{}' q_{ij} = \sum_{i>j}^{N} q_{ij}, \tag{7.46}$$

where the operator q_{ij} acts only on the electrons with indices $i, j \in \{1, \cdots, N\}$ and the sum runs over all *different* pairs. The prime excludes the case $i = j$ and the factor $\frac{1}{2}$ corrects for double counting of pairs. Alternatively one can sum over all pairs with $i > j$. A good example of such an operator is the electrostatic repulsion (6.10) between electrons in an atom. Using Eq. (7.24) the most general matrix element of Q_2 with determinantal wavefunctions is of the form

$$\langle\psi_{\alpha'}|Q_2|\psi_\alpha\rangle = \frac{1}{N!}\frac{1}{2}\sum_{i,j=1}^{N}{}'\sum_{P,P'}(-1)^{p+p'}(\psi_{\alpha'}|P'q_{ij}P|\psi_\alpha), \tag{7.47}$$

where $|\psi_\alpha) \equiv |\alpha_1, \cdots, \alpha_N)$ is the N-body product state in standard order. In view of the orthogonality of the spin-orbitals we have $(\psi_{\alpha'}|P'q_{ij}P|\psi_\alpha) = 0$ for all possible permutations P and P' whenever $|\psi_\alpha\rangle$ and $|\psi_{\alpha'}\rangle$ differ in *more than two* pairs of spin orbitals.

- *Off-diagonal matrix elements differing in two pairs.* Let $|\psi_\alpha\rangle$ and $|\psi_{\alpha'}\rangle$ differ by two pairs of spinorbitals, $\alpha_\kappa, \alpha_\nu$ and $\alpha_{\kappa'}, \alpha_{\nu'}$, respectively. For each pair of indices $i, j \in \{1, \cdots, N\}$ with $i \neq j$, there are $2\,(N-2)!$ permutations $|\alpha_1, \cdots, \alpha_\kappa, \alpha_\nu, \cdots \alpha_N) = P|\alpha_1, \cdots \alpha_N)$ for which α_κ appears on the position i and α_ν on j (those we call P_1) or *vice versa* (those we call P_2). For each of these permutations there is exactly *one* permutation, P_1' or P_2', for which $\alpha_{\kappa'}$ and $\alpha_{\nu'}$ appear, respectively, on the position i and j in $(\psi_{\alpha'}|P'$, while in addition, $(\psi_{\alpha'}|P'q_{ij}P|\psi_\alpha) \neq 0$. With these *two* permutations the N-body matrix element reduces to *two* two-body matrix element for the electrons i and j. The first is

$$\begin{aligned}(\psi_{\alpha'}|P_1'q_{ij}P_1|\psi_\alpha) &= (-1)^{p+p_1}\left(\alpha_{\kappa'}, \alpha_{\nu'}|q_{ij}|\alpha_\kappa, \alpha_\nu\right)_{i,j}\\ &= (-1)^{p+p_1}(\alpha_{\kappa'}, \alpha_{\nu'}|q_{12}|\alpha_\kappa, a_\nu).\end{aligned} \tag{7.48}$$

The permutation P_2 gives $(-1)^{p+p_2}(\alpha_{\kappa'}, \alpha_{\nu'}|q_{12}|\alpha_\nu, a_\kappa)$. The sign $(-1)^{p+p_2} = -(-1)^{p+p_1}$ has the same value for all values of i. Summing $(\psi_{\alpha'}|P'q_{ij}P|\psi_\alpha)$ over the $\frac{1}{2}N(N-1)$ different pairs $i, j \in \{1, \cdots, N\}$ and over all permutations P, P' we obtain $\frac{1}{2}N\,(N-1)\,2(N-2)! = N!$

equal nonzero terms and Eq. (7.47) becomes

$$\langle\psi_{\alpha'}|Q_2|\psi_\alpha\rangle = [(\alpha_{\kappa'},\alpha_{\nu'}|q_{12}|\alpha_\kappa,\alpha_\nu) - (\alpha_{\kappa'},\alpha_{\nu'}|q_{12}|\alpha_\nu,\alpha_\kappa)]\,\frac{1}{N!}\frac{1}{2}\sum_{i,j=1}^{N}{}'\sum_{P}{}'(-1)^{p+p_1'} \tag{7.49}$$

$$= \pm\,[(\alpha_{\kappa'},\alpha_{\nu'}|q_{12}|\alpha_\kappa,\alpha_\nu) - (\alpha_{\kappa'},\alpha_{\nu'}|q_{12}|\alpha_\nu,\alpha_\kappa)]\,, \tag{7.50}$$

with the overall sign depending on the ordering convention. The prime on the permutation sum indicates that the sum excludes, for given values of i and j, all permutations P in which α_κ and α_ν appear on the positions i or j but $(\psi_{\alpha'}|P'q_iP|\psi_\alpha) = 0$. We thus found that an operator of the type Q_2 can couple two pairs of spinorbitals and that the corresponding matrix elements do not depend on the particle indices of the electrons involved.

- *Off-diagonal matrix elements differing in one pair.* Let $|\psi_\alpha\rangle$ and $|\psi_{\alpha'}\rangle$ differ by one pair of spin orbitals, α_ν and $\alpha_{\nu'}$. For each pair $i,j\in\{1,\cdots,N\}$ with $i\neq j$ and each $\kappa\in\{1,\cdots,N\}$ we have, like in the previous case, $2\,(N-2)!$ permutations $|\alpha_1,\cdots,\alpha_\kappa,\alpha_\nu,\cdots\alpha_N) = P|\psi_\alpha)$ for which α_κ appears on the position i and α_ν on j (those we call again P_1) or *vice versa* (those we call again P_2). For each of these permutations there is exactly *one* permutation, P_1' or P_2', for which α_κ and $\alpha_{\nu'}$ appear, respectively, on the position i and j in $(\psi_{\alpha'}|P'$, while in addition, $(\psi_{\alpha'}|P'q_{ij}P|\psi_\alpha)\neq 0$. The corresponding *two* two-body matrix elements are $(\psi_{\alpha'}|P'q_{ij}P|\psi_\alpha) = (-1)^{p+p_1}(\alpha_\kappa,\alpha_{\nu'}|q_{12}|\alpha_\kappa,\alpha_\nu)$ and $-(-1)^{p+p_1}(\alpha_\kappa,\alpha_{\nu'}|q_{12}|\alpha_\nu,\alpha_\kappa)$. Summing $(\psi_{\alpha'}|P'q_{ij}P|\psi_\alpha)$ over the $\frac{1}{2}N(N-1)$ different electron pairs (i,j) and over all permutations P,P' we have *this time for each* $\kappa\in\{1,\cdots,N\}$ again $\frac{1}{2}N\,(N-1)\,2(N-2)! = N!$ equal terms and Eq. (7.50) is replaced by

$$\langle\psi_{\alpha'}|Q_2|\psi_\alpha\rangle = \pm\sum_{\kappa=1}^{N}\,[(\alpha_\kappa,\alpha_{\nu'}|q_{12}|\alpha_\kappa,\alpha_\nu) - (\alpha_\kappa,\alpha_{\nu'}|q_{12}|\alpha_\nu,\alpha_\kappa)]\,, \tag{7.51}$$

with the overall sign depending again on the ordering convention. We found that an operator of the type Q_2 can also give rise to the coupling of one pair of spinorbitals.

- *Diagonal matrix elements.* In this case $|\psi_\alpha\rangle = |\psi_{\alpha'}\rangle$. Continuing with the approach adopted above we note that for each pair of electron indices $i,j\in\{1,\cdots,N\}$ with $i\neq j$ and each pair of state indices $\kappa,\nu\in\{1,\cdots,N\}$ there are $2\,(N-2)!$ permutations $|\alpha_1,\cdots,\alpha_\kappa,\alpha_\nu,\cdots\alpha_N) = P|\psi_\alpha)$ for which α_κ appears on the position i and α_ν on j (those we call again P_1) or *vice versa* (those we call again P_2). For each of these permutations there is exactly *one* permutation, P_1' or P_2', for which α_κ and α_ν appear, respectively, on the position i and j in $(\psi_{\alpha'}|P'$, while in addition, $(\psi_{\alpha'}|P'q_{ij}P|\psi_\alpha)\neq 0$. The corresponding *two* two-body matrix elements give rise to a *direct* and an *exchange* contribution, given by $(\psi_\alpha|P'q_{ij}P|\psi_\alpha) = (\alpha_\kappa,\alpha_\nu|q_{12}|\alpha_\kappa,\alpha_\nu)$ and $-(\alpha_\kappa,\alpha_\nu|q_{12}|\alpha_\kappa,\alpha_\nu)$, respectively. Summing $(\psi_\alpha|P'q_{ij}P|\psi_\alpha)$ over the $\frac{1}{2}N(N-1)$ different electron pairs (i,j) and over all permutations P,P' we have *this time for each* $\kappa,\nu\in\{1,\cdots,N\}$, with $\kappa\neq\nu$, again $\frac{1}{2}N\,(N-1)\,2(N-2)! = N!$ equal contributions and obtain

$$\langle\psi_\alpha|Q_2|\psi_\alpha\rangle = \tfrac{1}{2}\sum_{\kappa,\nu=1}^{N}\,[(\alpha_\kappa,\alpha_\nu|q_{12}|\alpha_\kappa,\alpha_\nu) - (\alpha_\kappa,\alpha_\nu|q_{12}|\alpha_\nu,\alpha_\kappa)]\,. \tag{7.52}$$

Interestingly, for $\nu=\kappa$ the direct and exchange terms are equal. Therefore, the terms with $\nu=\kappa$ can be excluded from the summation,

$$\begin{aligned}\langle\psi_\alpha|Q_2|\psi_\alpha\rangle &= \tfrac{1}{2}\sum_{\kappa,\nu=1}^{N}{}'\,[(\alpha_\kappa,\alpha_\nu|q_{12}|\alpha_\kappa,\alpha_\nu) - (\alpha_\kappa,\alpha_\nu|q_{12}|\alpha_\nu,\alpha_\kappa)]\,,\\ &\equiv \sum_{\kappa>\nu}^{N}\,[(\alpha_\kappa,\alpha_\nu|q_{12}|\alpha_\kappa,\alpha_\nu) - (\alpha_\kappa,\alpha_\nu|q_{12}|\alpha_\nu,\alpha_\kappa)]\,.\end{aligned} \tag{7.53}$$

which shows that the expectation value of the operator Q_2 represents the sum of the contributions of the $\frac{1}{2}N(N-1)$ possible pairs of states.

Problem 7.5. Two electrons are said to be in a *paired spin state* if one of the electrons has spin up, $|\uparrow\rangle$, and the other spin down, $|\downarrow\rangle$. Show that for *paired* spins the exchange terms cancel for a spin-independent two-body interaction,

$$(\alpha_\kappa, \alpha_\nu | q_{12} | \alpha_\nu, \alpha_\kappa) = 0 \text{ for } m_s^\kappa \neq m_s^\nu.$$

Solution. For a spin-independent two-body interaction the spin dependence factors out,

$$(\alpha_\kappa, \alpha_\nu | q_{12} | \alpha_\nu, \alpha_\kappa) = (u_\kappa, u_\nu | q_{12} | u_\nu, u_\kappa)\delta_{m_s^\kappa, m_s^\nu}.$$

Hence, for paired spins ($m_s^\kappa \neq m_s^\nu$) the matrix element vanishes. □

7.3 Occupation number representation

7.3.1 Introduction

The notation of the previous section calls for simplification. This is realized by introducing *construction operators* which satisfy an algebra that enforces the quantum statistics. The first construction operators were introduced by Paul Dirac in 1927 [34]. Starting from Maxwell's equations, Dirac quantized the electromagnetic field by treating the eigenmodes of the field as independent harmonic oscillators. The excitation level of the oscillator represents the mode occupation of the field. The raising (lowering) operators of the oscillator serve to construct the field by creation (annihilation) of photons, the quanta of the radiation field, which occupy the modes. The commutation relations between the operators define the algebra that enforces the Bose statistics of the field. This marks the start of quantum field theory. In the same year Pascual Jordan and Oskar Klein showed that the method could be extended to describe quantum many-body systems of bosons satisfying the Schrödinger equation [62]. Adapting the algebra, Jordan and Wigner further extended the method to describe quantum many-body systems of interacting fermionic particles [63]. The above sequence of seminal papers is not complete without the name of Vladimir Fock, who emphasized in 1932 the use of field operators (construction operators for configuration space) [44]. This approach leads to an operator identity resembling the Schrödinger equation, which explains the unfortunate name *second quantization* for the *construction operator formalism*. In following sections we give a concise introduction into the construction operator formalism for quantum many-body systems. As electrons are fermions, we focus on the fermionic case. For a systematic introduction the lecture notes of Jan de Boer are recommended [30].

7.3.2 Number states in the N-body Hilbert space

We start by introducing a new representation, the *occupation number representation*. In this representation we keep track of the occupation of the spinorbitals. To introduce this notation we start with a seemingly innocent change in notation of the Slater determinants,

$$|\psi_\alpha\rangle \equiv |\alpha_r, \cdots, \alpha_s, \cdots, \alpha_t\rangle = |1_r, \cdots, 1_s, \cdots, 1_t\rangle, \tag{7.54}$$

where by $1_r, \cdots$ we indicate that we have 1 electron in the spinorbitals $|\alpha_r\rangle, \cdots$, given in standard order. If we have N electrons we have

$$1_r + \cdots + 1_s + \cdots + 1_t = N. \tag{7.55}$$

If the state $|\alpha_\nu\rangle$ is *not* occupied we are of course free to write 0_ν to indicate an empty orbital. At first sight this may not appear as particularly useful but, as we shall see, an enormous simplification of notation can be realized by introducing a representation in which we specify the occupation numbers of all possible spinorbitals (given in standard order). In this representation an *arbitrary* N-electron Slater determinant $|\psi_\gamma\rangle$ is written as

$$|\psi_\gamma\rangle \equiv |n_1, n_2, \cdots\rangle, \tag{7.56}$$

where $n_s \in \{0, 1\}$ is the occupation number of the state $|\alpha_s\rangle$, with $n_r + n_s + \cdots + n_t = N$. In this notation the states are called *number states*, which are the basis states of the occupation number representation (see next section). The N-electron number states have normalization

$$\langle n_1', n_2', \cdots |n_1, n_1, \cdots\rangle = \delta_{n_1,n_1'}\delta_{n_2,n_2'}\cdots \tag{7.57}$$

and closure

$$\sum_{n_1,n_2\cdots}{}' |n_1, n_2, \cdots\rangle \langle n_1, n_2, \cdots| = \mathbb{1}, \tag{7.58}$$

where $n_1, n_2, \cdots$ is either 0 or 1 and the prime indicates that the sum over all occupations equals the total number of particles, $n_1 + n_2 + \cdots = N$. This is called closure within $\mathbf{H}^N$.

7.3.3 Number states in Grand Hilbert space - construction operators

An important generalization of number states is obtained by interpreting the occupation numbers $n_1, n_2, \cdots$ as the eigenvalues of the *number operators* $\hat{n}_1, \hat{n}_2, \cdots$,

$$\hat{n}_s|n_1, n_2, \cdots, n_s, \cdots\rangle = n_s|n_1, n_2, \cdots, n_s, \cdots\rangle, \tag{7.59}$$

where $n_1, n_2, \cdots$ is either 0 or 1. The eigenstates $|n_1, n_2, \cdots, n_s, \cdots\rangle$ form the basis of the *occupation number representation.* With this definition the expectation value of $\hat{n}_s$ is *exclusively* determined by the occupation number n_s of the state $|s\rangle$; it is independent of the occupation of all other states; i.e., independent of N. Therefore, the number operators may be interpreted as acting in a *Grand Hilbert space*, also known as *Fock space*, which is the direct sum of the Hilbert spaces of all possible atom number states of a gas cloud, including the vacuum,

$$\mathbf{H}^{Gr} = \mathbf{H}^0 \oplus \mathbf{H}^1 \oplus \cdots \oplus \mathbf{H}^N \oplus \cdots .$$

By adding an atom we shift from $\mathbf{H}^N$ to $\mathbf{H}^{N+1}$, analogously we shift from $\mathbf{H}^N$ to $\mathbf{H}^{N-1}$ by removing an atom. As long as this does not affect the occupation of the single-particle state $|s\rangle$ the operator $\hat{n}_s$ yields the same result. Hence, the *number states* from $\mathbf{H}^N$ may be reinterpreted as number states $|n_1, n_2, \cdots\rangle$ within $\mathbf{H}^{Gr}$ by specifying the occupations of *all* single-particle states. For example, for the N-electron state $|\psi_\alpha\rangle \equiv |\alpha_r, \cdots, \alpha_s, \cdots, \alpha_t\rangle$ we have $n_r = 1$, $n_s = 1$, $\cdots$, $n_t = 1$ and all other occupation numbers are zero.

The basic operators in Grand Hilbert space are the *construction operators.* In particular the operator $\hat{a}_s^\dagger$ is called a *creation operator* and defined by

$$\hat{a}_s^\dagger|1_r, \cdots, 0_s, \cdots, 1_t\rangle = |1_s, 1_r, \cdots, 0_s, \cdots, 1_t\rangle = (-1)^p\, |1_r, \cdots, 1_s, \cdots, 1_t\rangle \tag{7.60a}$$

$$\hat{a}_s^\dagger|1_r, \cdots, 1_s, \cdots, 1_t\rangle = 0, \tag{7.60b}$$

where $(-)^p$ yields $+1$ or -1 depending on whether it takes an even or an odd permutation P between *occupied* states to bring the occupation number 1_s to the empty position in the normal order of states; with the second line we assure that the state $|\alpha_s\rangle$ cannot be occupied more than once. Likewise the operator $\hat{a}_s$ is called an *annihilation operator* and defined by

$$\hat{a}_s\,|1_r, \cdots, 1_s, \cdots, 1_t\rangle = (-1)^p\, \hat{a}_s\,|1_s, 1_r, \cdots, 0_s, \cdots, 1_t\rangle = (-1)^p\, |1_r, \cdots, 0_s, \cdots, 1_t\rangle \tag{7.61a}$$

$$\hat{a}_s\,|1_r, \cdots, 0_s, \cdots, 1_t\rangle = 0, \tag{7.61b}$$

where $(-)^p$ yields $+1$ or -1 depending on whether it takes an even or an odd permutation P between *occupied* states to bring the occupation number 1_s the first position in the Slater determinant (without changing the order of the other states); the second line assures that an electron cannot be cannot be removed from the state $|\alpha_s\rangle$ more than once. Let us have a look at a few examples: with the creation operators we obtain $\hat{a}_s^\dagger\,|0_q, 0_s, \cdots\rangle = +\,|0_q, 1_s, \cdots\rangle$, $\hat{a}_s^\dagger\,|1_q, 0_s, \cdots\rangle = -\,|1_q, 1_s, \cdots\rangle$, where we used the alphabetic order to represent the normal order. Analogously, we obtain with the annihilation operators $\hat{a}_s\,|0_q, 1_s, \cdots\rangle = |0_q, 0_s, \cdots\rangle$ and $\hat{a}_s\,|1_q, 1_s, \cdots\rangle = -\,|1_q, 0_s, \cdots\rangle$.

The basic actions of the construction operators can be summarized as

$$\hat{a}_s^\dagger|n_s, \cdots\rangle \equiv \sqrt{n_s+1}|n_s+1, \cdots\rangle \tag{7.62a}$$

$$\hat{a}_s|n_s, \cdots\rangle \equiv \sqrt{n_s}|n_s-1, \cdots\rangle, \tag{7.62b}$$

with $n_s \in \{0, 1\}$. The creation operators transform an antisymmetrized N-body eigenstate in $\mathbf{H}^N$ into an antisymmetrized $N+1$ body eigenstate in $\mathbf{H}^{N+1}$. Analogously, the annihilation operators transform an antisymmetrized N-body eigenstate in $\mathbf{H}^N$ into an antisymmetrized $N-1$ body eigenstate in $\mathbf{H}^{N-1}$. Further, we introduce the commutation relations [1]

$$\{\hat{a}_r, \hat{a}_s^\dagger\} = \delta_{rs}\ ; \qquad \{\hat{a}_r, \hat{a}_s\} = \{\hat{a}_r^\dagger, \hat{a}_s^\dagger\} = 0, \tag{7.63}$$

which serve to enforce the Pauli principle as embodied by Eq. (7.60b) and reflect the antisymmetry of the Slater determinants. From Eqs. (7.62) and (7.63) follow the following properties of the construction operators and the number states:

- the construction operators $\hat{a}_s^\dagger$ and $\hat{a}_s$ are hermitian conjugates,

  $$\langle n_s+1|\hat{a}_s^\dagger|n_s\rangle = \langle n_s|\hat{a}_s|n_s+1\rangle^* = \sqrt{n_s+1}. \tag{7.64}$$

  Hence, *when acting on the bra side* $\hat{a}_s^\dagger$ and $\hat{a}_s$ change their role, $\hat{a}_s^\dagger$ becomes the annihilation operator and $\hat{a}_s$ the creation operator. Note that Eq. (7.64) only has physical significance for $n_s = 0$ because double occupation of electronic states violates the Pauli principle.

- the number operator is given by $\hat{n}_s = \hat{a}_s^\dagger\hat{a}_s$,

  $$\hat{n}_s|n_s, \cdots\rangle = \hat{a}_s^\dagger\hat{a}_s|n_s, \cdots\rangle \equiv \sqrt{n_s}\hat{a}_s^\dagger|n_s-1, \cdots\rangle = n_s|n_s, \cdots\rangle. \tag{7.65}$$

- an absent particle cannot be annihilated. This is embodied by Eq. (7.61b),

  $$\hat{a}_s|n_s, \cdots\rangle \equiv \sqrt{n_s}|n_s-1, \cdots\rangle = 0 \quad \text{for} \quad n_s = 0. \tag{7.66}$$

- the Pauli principle is satisfied. This is embodied by Eq. (7.60b),

  $$\hat{a}_s\hat{a}_s^\dagger|n_s, \cdots\rangle = (1-\hat{a}_s^\dagger\hat{a}_s)|n_s, \cdots\rangle = (1-n_s)|n_s, \cdots\rangle = 0 \quad \text{for} \quad n_s = 1. \tag{7.67}$$

- *any* N-body state $|\psi_\gamma\rangle = |1_r, \cdots, 1_s, \cdots, 1_t\rangle$ can be constructed by repetitive use of a set of creation operators

  $$|\psi_\gamma\rangle = \prod_{s\in\gamma} \hat{a}_s^\dagger\,|0\rangle\,, \tag{7.68}$$

  where the index $s \in \gamma = \{q, \cdots, r, \cdots, t\}$ points to the set of one-body states to be populated in normal order and $|0\rangle \equiv |0_1, 0_2, \cdots\rangle$ is called the *vacuum state*.

- also the vacuum state is unit normalized,

  $$\langle 0|0\rangle = \left\langle 1_s|\hat{a}_s^\dagger\hat{a}_s|1_s\right\rangle = \langle 1_s|\hat{n}_s|1_s\rangle = \langle s|s\rangle = 1, \tag{7.69}$$

  where α_s is an arbitrary orbital.

---

[1] Note that we use the curly bracket convention for the anti-commutator: $\{a, b\} = ab + ba$.

- the number states created from the vacuum are *antisymmetric*. This follows from the commutation relation $\{\hat{a}_r^\dagger, \hat{a}_s^\dagger\} = 0$,

$$|1_r, 1_s\rangle = \hat{a}_r^\dagger \hat{a}_s^\dagger \,|0\rangle = -\hat{a}_s^\dagger \hat{a}_r^\dagger \,|0\rangle = -|1_s, 1_r\rangle. \tag{7.70}$$

Thus we have obtained the *occupation number representation.* By extending $\mathbf{H}^N$ to $\mathbf{H}^{Gr}$ the definition of the number states and their normalization has remained unchanged,

$$\langle n_1', n_2', \cdots | n_1, n_2, \cdots\rangle = \delta_{n_1,n_1'}\delta_{n_2,n_2'} \cdots . \tag{7.71}$$

Importantly, by turning to $\mathbf{H}^{Gr}$ the condition on particle conservation is lost. This has the very convenient consequence that in the closure relation (7.58) the restricted sum may be replaced by an *unrestricted* sum, thus allowing for all possible values of N,

$$\sum_{n_1,n_2\cdots} |n_1, n_2, \cdots\rangle \, \langle n_1, n_2, \cdots| = \mathbb{1}, \tag{7.72}$$

where $n_1, n_2, \cdots$ can take the values 0 or 1. This is called closure within $\mathbf{H}^{Gr}$.

Problem 7.6. Show that the following commutation relations hold

$$\left[\hat{n}_q, \hat{a}_s^\dagger\right] = +\hat{a}_s^\dagger \delta_{qs} \quad \text{and} \quad [\hat{n}_q, \hat{a}_s] = -\hat{a}_s \delta_{qs} \,. \tag{7.73}$$

7.3.4 Operators in the occupation number representation

The real added value of the occupation number representation is in the calculation of the interaction terms of the atomic Hamiltonian. In the present section we show that the extension of one-body operators into Grand Hilbert space is given by

$$Q_1 = \sum_i q_i \quad \rightarrow \quad \hat{Q}_1 = \sum_{s,s'} \langle s'|q|s\rangle \hat{a}_{s'}^\dagger \hat{a}_s, \tag{7.74}$$

where the sum in the expression for Q_1 runs over the *particle index* and the sums in the expression for $\hat{Q}_1$ run over the *state index* of all one-body states (occupied or not) and $\hat{a}_s^\dagger$ and $\hat{a}_s$ are the construction operators that change the occupation of the state $|\alpha_s\rangle$. Likewise the extension of two-body operators is given by

$$Q_2 = \tfrac{1}{2} \sum_{i,j}{}' q_{ij} \quad \rightarrow \quad \hat{Q}_2 = \tfrac{1}{2} \sum_{t,t'} \sum_{s,s'} (s', t'|q_{12}|s, t) \hat{a}_{s'}^\dagger \hat{a}_{t'}^\dagger \hat{a}_t \hat{a}_s. \tag{7.75}$$

To demonstrate the validity of these simple expressions we rederive the expression of Sections 7.2.1 and 7.2.2. As we shall see, in the occupation representation these expressions are obtained with a minimal effort by straightforward application of the operator algebra. The key idea is based on the orthonormality relation

$$\langle \psi_{\alpha'} | \psi_\alpha \rangle = \langle n_1', n_2', \cdots | n_1, n_2, \cdots\rangle = \delta_{n_1,n_1'}\delta_{n_2,n_2'} \cdots . \tag{7.76}$$

In the language of the occupation number representation for N-electron atoms this means that the inner product is zero unless every state that is created on the ket side is annihilated on the bra side,

$$\langle \alpha_1, \cdots \alpha_N | \alpha_1, \cdots \alpha_N \rangle = \langle 0|\hat{a}_N \cdots \hat{a}_1 \hat{a}_1^\dagger \cdots \hat{a}_N^\dagger|0\rangle = 1. \tag{7.77}$$

7.3.4.1 One-body operators:

- *Diagonal matrix elements.*

$$\langle\psi_\alpha|\hat{Q}_1|\psi_\alpha\rangle = \sum_{s,s'}\langle s'|q|s\rangle\langle 0|\hat{a}_N\cdots\hat{a}_1\hat{a}_{s'}^\dagger\hat{a}_s\hat{a}_1^\dagger\cdots\hat{a}_N^\dagger|0\rangle.$$

  In view of the diagonality any one-body state that is destroyed has to be recreated, otherwise the matrix element will be zero by orthonormality of the one-body states. This means that only the following terms contribute:

$$\hat{Q}_1 = \sum_s\langle s|q|s\rangle\hat{a}_s^\dagger\hat{a}_s = \sum_s\langle s|q|s\rangle\hat{n}_s, \tag{7.78}$$

  which is readily evaluated by summing the contributions of the occupied states $\{|\kappa\rangle\}$,

$$\langle\psi_\alpha|\hat{Q}_1|\psi_\alpha\rangle = \sum_\kappa\langle\alpha_\kappa|q|\alpha_\kappa\rangle. \tag{7.79}$$

  Here we regained Eq. (7.45).

- *Off-diagonal matrix elements differing in one pair of states* (the pair $\langle\kappa'|$ and $|\kappa\rangle$). The only way in which such matrix elements can be nonzero is if we destroy the state $|\kappa\rangle$ and create the state $|\kappa'\rangle$. This means that only one term contributes,

$$\hat{Q}_1 = \langle\kappa'|q|\kappa\rangle\hat{a}_{\kappa'}^\dagger\hat{a}_\kappa, \tag{7.80}$$

  which is readily evaluated to yield,

$$\langle\psi_{\alpha'}|\hat{Q}_1|\psi_\alpha\rangle = (-)^{p_\kappa+p_{\kappa'}}\langle\alpha_{\kappa'}|q|\alpha_\kappa\rangle, \tag{7.81}$$

  which coincides with the result (7.44) with p_κ and $p_{\kappa'}$ representing the orders of the permutations that bring the states $|\alpha_\kappa\rangle$ and $|\alpha_{\kappa'}\rangle$ to the first position in the Slater determinant.

- *Off-diagonal matrix elements differing in more than one pair of states.* In this case we always have

$$\langle\psi_{\alpha'}|\hat{Q}_1|\psi_\alpha\rangle = 0 \tag{7.82}$$

  because the operator $\hat{Q}_1$ can at most replace one state by an other.

7.3.4.2 Two-body operators:

- *Diagonal matrix elements* $\langle\psi_\alpha|\hat{Q}_2|\psi_\alpha\rangle$. In view of the diagonality, any one-body state that is destroyed has to be recreated. If we annihilate two one-body states, we have two options, first annihilate $|s\rangle$ and then $|t\rangle$ or *vice versa* (note that $a_s a_t|\psi_\alpha\rangle \neq 0$ implies $a_t a_s|\psi_\alpha\rangle \neq 0$). Exploiting this symmetry we obtain the sum of two terms, which leaves us with a single option to recreate the states. We implement this option by first recreating $\langle t|$ and subsequently $\langle s|$ (the inverse order yields the same result by renaming $s \leftrightarrow t$). Summing over all different pairs s and t we obtain

$$\hat{Q}_2 = \tfrac{1}{2}\sum_{t,s}{}'[(s,t|q_{12}|s,t)\hat{a}_s^\dagger\hat{a}_t^\dagger\hat{a}_t\hat{a}_s + (s,t|q_{12}|t,s)\hat{a}_s^\dagger\hat{a}_t^\dagger\hat{a}_s\hat{a}_t], \tag{7.83}$$

  where the prime on the summation stands for $t \neq s$. Switching to the number operators, $\hat{n}_t = \hat{a}_t^\dagger\hat{a}_t$, and using the fermionic commutation relations $\{\hat{a}_s, \hat{a}_t^\dagger\} = 0$ and $\{\hat{a}_t, \hat{a}_s\} = 0$ we have $\hat{a}_s^\dagger \hat{a}_t^\dagger \hat{a}_t \hat{a}_s = \hat{a}_s^\dagger \hat{a}_s \hat{a}_t^\dagger \hat{a}_t = \hat{n}_s \hat{n}_t$ and $\hat{a}_s^\dagger \hat{a}_t^\dagger \hat{a}_s \hat{a}_t = -\hat{a}_s^\dagger \hat{a}_s \hat{a}_t^\dagger \hat{a}_t = \hat{n}_s \hat{n}_t$ the expression for $\hat{Q}_2$ becomes

$$\hat{Q}_2 = \tfrac{1}{2} \sum_{t,s}{}' \left[(s,t|q_{12}|s,t) - (s,t|q_{12}|t,s)\right] \hat{n}_s \hat{n}_t. \tag{7.84}$$

This operator is readily evaluated by summing the contributions of the occupied pairs of states $\{|\alpha_\kappa\rangle, |\alpha_\nu\rangle\}$,

$$\langle \psi_\alpha | \hat{Q}_2 | \psi_\alpha \rangle = \tfrac{1}{2} \sum_{\kappa,\nu}{}' \left[(\alpha_\kappa, \alpha_\nu | q_{12} | \alpha_\kappa, \alpha_\nu) - (\alpha_\kappa, \alpha_\nu | q_{12} | \alpha_\nu, \alpha_\kappa)\right]. \tag{7.85}$$

Here we regained Eq. (7.53).

- *Off-diagonal matrix elements differing in one pair of states* (the pair $\langle \nu'|$ and $|\nu\rangle$). If we annihilate two occupied one-body states, we have two options, first $|\kappa\rangle$ and then $|\nu\rangle$ or *vice versa*, before creating the state $\langle \nu'|$ and recreating $\langle \kappa|$. This means that for each value of κ only two terms contribute:

$$\hat{Q}_2 = \sum_{\kappa}{}' [(\kappa, \nu'|q_{12}|\kappa, \nu) \hat{a}_\kappa^\dagger \hat{a}_{\nu'}^\dagger \hat{a}_\nu \hat{a}_\kappa + (\kappa, \nu'|q_{12}|\nu, \kappa) \hat{a}_\kappa^\dagger \hat{a}_{\nu'}^\dagger \hat{a}_\kappa \hat{a}_\nu]. \tag{7.86}$$

Switching to the number operators, $\hat{n}_s = \hat{a}_s^\dagger \hat{a}_s$, and using the commutation relations $\{\hat{a}_\kappa^\dagger, \hat{a}_{\nu'}^\dagger\} = \{\hat{a}_\kappa^\dagger, \hat{a}_\nu\} = 0$ (note that $\nu' \neq \kappa$ and $\nu \neq \kappa$) we have $\hat{a}_\kappa^\dagger \hat{a}_{\nu'}^\dagger \hat{a}_\nu \hat{a}_\kappa = \hat{a}_{\nu'}^\dagger \hat{a}_\nu \hat{n}_\kappa$ and $\hat{a}_\kappa^\dagger \hat{a}_{\nu'}^\dagger \hat{a}_\kappa \hat{a}_\nu = -\hat{a}_{\nu'}^\dagger \hat{a}_\nu \hat{n}_\kappa$ this becomes

$$\langle \psi_{\alpha'} | Q_2 | \psi_\alpha \rangle = (-)^{p_\kappa + p_\nu} \sum_{\kappa} \left[(\alpha_\kappa, \alpha_\nu | q_{12} | \alpha_\kappa, \alpha_\nu) - (\alpha_\kappa, \alpha_\nu | q_{12} | \alpha_\nu, \alpha_\kappa)\right], \tag{7.87}$$

which coincides with the result (7.51) with p_ν and $p_{\nu'}$ representing the orders of the permutations that bring the states $|\alpha_\nu\rangle$ and $|\alpha_\kappa\rangle$, respectively, to the first column position in the Slater determinant.

- *Off-diagonal matrix elements differing in two pairs of states* (the pairs $\langle \kappa'|, |\kappa\rangle$ and $\langle \nu'|, |\nu\rangle$). If we annihilate two occupied one-body states, we have two option, first $|\kappa\rangle$ and then $|\nu\rangle$ or *vice versa*, before creating the states $\langle \nu'|$ and $\langle \kappa'|$. This means that only two terms contribute:

$$\hat{Q}_2 = (\kappa', \nu'|q_{12}|\kappa, \nu) \hat{a}_{\kappa'}^\dagger \hat{a}_{\nu'}^\dagger \hat{a}_\nu \hat{a}_\kappa + (\kappa', \nu'|q_{12}|\nu, \kappa) \hat{a}_{\kappa'}^\dagger \hat{a}_{\nu'}^\dagger \hat{a}_\kappa \hat{a}_\nu. \tag{7.88}$$

Switching to the number operators, $\hat{n}_s = \hat{a}_s^\dagger \hat{a}_s$, and using the commutation relations $\{\hat{a}_{\kappa'}^\dagger, \hat{a}_{\nu'}^\dagger\} = \{\hat{a}_{\kappa'}^\dagger, \hat{a}_\nu\} = 0$ (note that $\nu' \neq \kappa'$ and $\nu \neq \kappa'$) we have $\hat{a}_{\kappa'}^\dagger \hat{a}_{\nu'}^\dagger \hat{a}_\nu \hat{a}_\kappa = \hat{a}_{\nu'}^\dagger \hat{a}_\nu \hat{a}_{\kappa'}^\dagger \hat{a}_\kappa$ and $\hat{a}_{\kappa'}^\dagger \hat{a}_{\nu'}^\dagger \hat{a}_\kappa \hat{a}_\nu = -\hat{a}_{\nu'}^\dagger \hat{a}_\nu \hat{a}_{\kappa'}^\dagger \hat{a}_\kappa$ this becomes

$$\langle \psi_{\alpha'} | Q_2 | \psi_\alpha \rangle = (-)^{p_\kappa + p_\nu} \left[(\alpha_\kappa, \alpha_\nu | q_{12} | \alpha_\kappa, \alpha_\nu) - (\alpha_\kappa, \alpha_\nu | q_{12} | \alpha_\nu, \alpha_\kappa)\right], \tag{7.89}$$

which coincides with the result (7.50) with p_ν and $p_{\nu'}$ representing the orders of the permutations that bring the states $|\alpha_\nu\rangle$ and $|\alpha_{\kappa'}\rangle$, respectively, to the first column position in the Slater determinant.

- *Off-diagonal matrix elements differing in more than two pairs of states.* In this case we always have

$$\langle \psi_{\alpha'} | \hat{Q}_2 | \psi_\alpha \rangle = 0 \tag{7.90}$$

because the operator $\hat{Q}_2$ can annihilate and create at most replace two states.

7.4 Angular momentum of N-electron Slater determinants

7.4.1 Total orbital angular momentum L

Having established a procedure to calculate the matrix elements of one-body and two-body operators for properly symmetrized N-electron basis states $|\psi_\alpha\rangle = |\alpha_1, \cdots, \alpha_N\rangle$, where $|\alpha_\kappa\rangle = |n^\kappa l^\kappa m_l^\kappa s m_s^\kappa\rangle$ with $\kappa \in \{1, \cdots N\}$ and $s = \frac{1}{2}$ being one-electron eigenstates of the spin-orbital type, we are in a position to determine the properties of the total orbital angular momentum operator

$$\mathbf{L} = \sum_{i=1}^{N} \boldsymbol{l}_i \tag{7.91}$$

on the basis $\{|\psi_\alpha\rangle\}$. Obvious questions involve the action of the operators L_z, $L_\pm$ and $\mathbf{L}^2$ on a given many-body basis state $|\psi_\alpha\rangle$. Using the closure relation the action of any operator G can be written in the form of an expansion in the basis functions

$$G|\psi_\alpha\rangle = \sum_{\alpha'} |\psi_{\alpha'}\rangle\langle\psi_{\alpha'}|G|\psi_\alpha\rangle \tag{7.92}$$

and we shall use this relation to investigate L_z, $L_\pm$ and $\mathbf{L}^2$. Importantly, we shall find that Slater determinants are not necessarily eigenstates of the operators $L_\pm$ and $\mathbf{L}^2$.

- *The operator* L_z. The operator $L_z = \sum_i l_z^{(i)}$ is of the diagonal one-body type. It conserves the angular momentum and its projection along the z axis is diagonal in the $\{|\psi_\alpha\rangle\}$ representation. Thus, the expansion (7.92) reduces to a single term

$$L_z|\psi_\alpha\rangle = |\psi_\alpha\rangle\langle\psi_\alpha|L_z|\psi_\alpha\rangle = |\psi_\alpha\rangle \sum_{\kappa=1}^{N} \langle l^\kappa m_l^\kappa|l_z|l^\kappa m_l^\kappa\rangle\langle s m_s^\kappa|s m_s^\kappa\rangle = M_L^\alpha \hbar\, |\psi_\alpha\rangle, \tag{7.93}$$

where we factored out the spin part, $\langle s m_s^\kappa|s m_s^\kappa\rangle = 1$, and where

$$M_L^\alpha = \sum_{\kappa=1}^{N} m_l^\kappa \tag{7.94}$$

is the magnetic quantum number of the total orbital angular momentum of the state $|\psi_\alpha\rangle$.

- *The operator* $L_\pm$. The operator $L_\pm = \sum_i l_\pm^{(i)}$ is of the off-diagonal one-body type. It raises or lowers the m_l^κ of *one* of the electrons of the state $|\psi_\alpha\rangle$ by one unit of angular momentum. Thus, if $\{|\psi_\gamma\rangle\}$ represents the subset (of N elements) of $\{|\psi_{\alpha'}\rangle\}$ for which m_l^κ (and M_L^α) change by one, the expansion (7.92) reduces to

$$\begin{aligned} L_\pm|\psi_\alpha\rangle &= \sum_\gamma |\psi_\gamma\rangle\langle\psi_\gamma|L_\pm|\psi_\alpha\rangle = \sum_{\kappa=1}^{N} |\psi_\alpha(m_l^\kappa \pm 1)\rangle\langle l^\kappa(m_l^\kappa \pm 1)|l_\pm|l^\kappa m_l^\kappa\rangle \\ &= \sum_{\kappa=1}^{N} \sqrt{l^\kappa(l^\kappa+1) - m_l^\kappa(m_l^\kappa \pm 1)}\,\hbar\,|\psi_\alpha(m_l^\kappa \pm 1)\rangle. \end{aligned} \tag{7.95}$$

Again the spin part has been factored out, $\langle s m_s^\kappa|s m_s^\kappa\rangle = 1$. With the notation $|\psi_\gamma\rangle = |\psi_\alpha(m_l^\kappa \pm 1)\rangle$, with $\kappa \in \{1, \cdots N\}$, we mean that the state $|\alpha_\kappa\rangle = |n^\kappa l^\kappa m_l^\kappa s m_s^\kappa\rangle$ has been replaced by $|\alpha'_\kappa\rangle = |n^\kappa l^\kappa (m_l^\kappa \pm 1) s m_s^\kappa\rangle$. With this choice of ordering of the many-body state the phase factor is always unity, $(-1)^{2p} = 1$. The meaning of $|\psi_\alpha(m_l^\kappa \pm 1)\rangle = |\alpha'_\kappa\rangle$ is best expressed in the occupation number representation

$$|\alpha'_\kappa\rangle = \hat{a}^\dagger_{\alpha'_\kappa}\hat{a}_{\alpha_\kappa}|\alpha_\kappa\rangle. \tag{7.96}$$

- *The operator* $\mathbf{L}^2$. This operator $\mathbf{L}^2 = (\sum_i \boldsymbol{l}_i)^2$ can be written in the form

$$\mathbf{L}^2 = \sum_i \boldsymbol{l}_i^2 + {\sum_{i,j}}' \boldsymbol{l}_i \cdot \boldsymbol{l}_j = \sum_i \boldsymbol{l}_i^2 + \frac{1}{2} {\sum_{i,j}}' [2 l_z^{(i)} l_z^{(j)} + (l_+^{(i)} l_-^{(j)} + l_-^{(i)} l_+^{(j)})]. \tag{7.97}$$

Since $l_-^{(i)}$ and $l_+^{(j)}$ commute for $i \neq j$ and the indices run over the same set of values, $i, j \in \{1, \cdots N\}$, the expression for $\mathbf{L}^2$ reduces to

$$\mathbf{L}^2 = \sum_i \boldsymbol{l}_i^2 + {\sum_{i,j}}' [l_z^{(i)} l_z^{(j)} + l_+^{(i)} l_-^{(j)}]. \tag{7.98}$$

The first term is of the diagonal one-body type and the second of the diagonal two-body type. The remaining term is of the off-diagonal two-body type with differences in two pairs. In view of the off-diagonality of the $l_+^{(i)} l_-^{(j)}$ operator, the direct terms vanish. Along the same lines as above the action of the operator $\mathbf{L}^2$ is given by

$$\begin{aligned} \mathbf{L}^2 |\psi_\alpha\rangle = &\left\{ \sum_{\kappa=1}^{N} l^\kappa (l^\kappa + 1) + {\sum_{\kappa,\kappa'=1}^{N}}' m_l^\kappa m_l^{\kappa'} \right\} \hbar^2 \, |\psi_\alpha\rangle + \\ &+ {\sum_{\kappa,\kappa'=1}^{N}}' \sqrt{l^\kappa (l^\kappa + 1) - m_l^\kappa (m_l^\kappa + 1)} \sqrt{l^{\kappa'} (l^{\kappa'} + 1) - m_l^{\kappa'} (m_l^{\kappa'} - 1)} \, \hbar^2 \, |\psi_\alpha (m_l^\kappa + 1)(m_l^{\kappa'} - 1)\rangle, \end{aligned} \tag{7.99}$$

where the notation

$$|\psi_\gamma\rangle = |\psi_\alpha (m_l^\kappa + 1)(m_l^{\kappa'} - 1)\rangle \equiv |\psi_\alpha (m_l^{\kappa'} - 1)(m_l^\kappa + 1)\rangle, \tag{7.100}$$

with $\kappa, \kappa' \in \{1, \cdots N\}$, indicates that the state $|\alpha_\kappa\rangle = |n^\kappa l^\kappa m_l^\kappa s m_s^\kappa\rangle$ has been replaced by $|\alpha'_\kappa\rangle = |n^\kappa l^\kappa (m_l^\kappa + 1) s m_s^\kappa\rangle$ and the state $|\alpha_{\kappa'}\rangle = |n^{\kappa'} l^{\kappa'} m_l^{\kappa'} s m_s^{\kappa'}\rangle$ by $|\alpha'_{\kappa'}\rangle = |n^{\kappa'} l^{\kappa'} (m_l^{\kappa'} - 1) s m_s^{\kappa'}\rangle$. Note that

$$|\psi_\alpha (m_l^\kappa + 1)(m_l^{\kappa'} - 1)\rangle = -|\psi_\alpha\rangle \quad \text{for } m_l^{\kappa'} - m_l^\kappa = 1. \tag{7.101}$$

The meaning of $|\psi_\alpha (m_l^\kappa + 1)(m_l^{\kappa'} - 1)\rangle$ is best expressed in the occupation number representation

$$|\psi_\alpha (m_l^\kappa + 1)(m_l^{\kappa'} - 1)\rangle = |\alpha'_\kappa, \alpha'_{\kappa'}\rangle = \hat{a}^\dagger_{\alpha'_\kappa} \hat{a}^\dagger_{\alpha'_{\kappa'}} \hat{a}_{\alpha_{\kappa'}} \hat{a}_{\alpha_\kappa} |\alpha_\kappa, \alpha_{\kappa'}\rangle. \tag{7.102}$$

There are $\binom{N}{2}$ different states of this type and these correspond to all different pair combinations of $\kappa, \kappa' \in \{1, \cdots N\}$ with $\kappa \neq \kappa'$. Note that for these states $\langle\psi_\gamma|L_z|\psi_\gamma\rangle = \langle\psi_\alpha|L_z|\psi_\alpha\rangle$. As demonstrated in Problem 7.8 Eq. (7.99) can be strongly simplified for *stretched* states of N equal angular momenta l, where both the total angular momentum and its projection are maximized ($L = M_L = Nl$).

7.4.1.1 Example:

Let us further clarify the notation $|\psi_\alpha (m_l^\kappa + 1)(m_l^{\kappa'} - 1)\rangle$. For the state

$$|\psi_\alpha\rangle = |\alpha_\kappa, \alpha_{\kappa'}\rangle = |m_l^\kappa, m_l^{\kappa'}\rangle_{nd^2} = |\dot{1}, \dot{0}\rangle_{nd^2} \tag{7.103}$$

we have $|\alpha_\kappa\rangle = |n^\kappa l^\kappa m_l^\kappa s m_s^\kappa\rangle = |n, 2, 1, \frac{1}{2}, \frac{1}{2}\rangle$ and $|\alpha_{\kappa'}\rangle = |n^{\kappa'} l^{\kappa'} m_l^{\kappa'} s m_s^{\kappa'}\rangle = |n, 2, 0, \frac{1}{2}, \frac{1}{2}\rangle$. The states

$$|\psi_\gamma\rangle = |\psi_\alpha (m_l^\kappa + 1)(m_l^{\kappa'} - 1)\rangle = |\alpha'_\kappa, \alpha'_{\kappa'}\rangle = |\dot{2}, -\dot{1}\rangle_{nd^2} \tag{7.104a}$$

$$|\psi'_\gamma\rangle = |\psi_\alpha (m_l^\kappa - 1)(m_l^{\kappa'} + 1)\rangle = |\alpha'_\kappa, \alpha'_{\kappa'}\rangle = |\dot{0}, \dot{1}\rangle_{nd^2} = -|\dot{1}, \dot{0}\rangle_{nd^2} \tag{7.104b}$$

correspond to, respectively

$$\left.\begin{aligned} |\alpha'_\kappa\rangle &= |n^\kappa l^\kappa (m_l^\kappa + 1) s m_s^\kappa\rangle = |n, 2, 2, \tfrac{1}{2}, \tfrac{1}{2}\rangle \\ |\alpha'_{\kappa'}\rangle &= |n^{\kappa'} l^{\kappa'} (m_l^{\kappa'} - 1) s m_s^{\kappa\prime}\rangle = |n, 2, -1, \tfrac{1}{2}, \tfrac{1}{2}\rangle \end{aligned}\right\} \tag{7.105a}$$

$$\left.\begin{aligned} |\alpha'_\kappa\rangle &= |n^\kappa l^\kappa (m_l^\kappa - 1) s m_s^\kappa\rangle = |n, 2, 0, \tfrac{1}{2}, \tfrac{1}{2}\rangle \\ |\alpha'_{\kappa'}\rangle &= |n^{\kappa'} l^{\kappa'} (m_l^{\kappa'} + 1) s m_s^{\kappa\prime}\rangle = |n, 2, +1, \tfrac{1}{2}, \tfrac{1}{2}\rangle \end{aligned}\right\}. \tag{7.105b}$$

Problem 7.7. Calculate the total orbital angular momentum of the Slater determinant $|\dot{1}, \dot{0}\rangle_{np^2}$.

Solution. For two equivalent electrons we have $N = 2$ and the expression for $\mathbf{L}^2$ becomes

$$\begin{aligned} \mathbf{L}^2|\dot{1}, \dot{0}\rangle_{np^2} = &\left\{2l(l+1) + 2m_l^{(1)} m_l^{(2)}\right\} \hbar^2\, |\dot{1}, \dot{0}\rangle_{np^2} \\ &+ \sqrt{l(l+1) - m_l^{(1)}(m_l^{(1)} + 1)}\sqrt{l(l+1) - m_l^{(2)}(m_l^{(2)} - 1)}\, \hbar^2\, |\psi_\alpha(m_l^{(1)} + 1)(m_l^{(2)} - 1)\rangle \\ &+ \sqrt{l(l+1) - m_l^{(2)}(m_l^{(2)} + 1)}\sqrt{l(l+1) - m_l^{(1)}(m_l^{(1)} - 1)}\, \hbar^2\, |\psi_\alpha(m_l^{(2)} + 1)(m_l^{(1)} - 1)\rangle \end{aligned}$$

Substituting $l^{(1)} = l^{(2)} = l = 1$, $m_l^{(1)} = 1$ and $m_l^{(2)} = 0$ we calculate $m_l^{(1)} m_l^{(2)} = 0$ and

$$\begin{array}{lll} m_l^{(1)}(m_l^{(1)} + 1) = 2 & m_l^{(2)}(m_l^{(2)} - 1) = 0 & |\psi_\alpha(m_l^{(1)} + 1)(m_l^{(2)} - 1)\rangle = |\dot{2}, -\dot{1}\rangle_{np^2} \\ m_l^{(2)}(m_l^{(2)} + 1) = 0 & m_l^{(1)}(m_l^{(1)} - 1) = 0 & |\psi_\alpha(m_l^{(2)} + 1)(m_l^{(1)} - 1)\rangle = -|\dot{1}, \dot{0}\rangle_{np^2} \end{array}$$

Thus we obtain

$$\mathbf{L}^2|\dot{1}, \dot{0}\rangle_{np^2} = \{2l(l+1) - 2\}\hbar^2\, |\dot{1}, \dot{0}\rangle_{np^2} = 2\hbar^2\, |\dot{1}, \dot{0}\rangle_{np^2}$$

and with $L(L+1) = 2$ we calculate $L = 1$. □

Problem 7.8. Show that eigenstates with $L = M_L = Nl$, i.e. with stretched total orbital angular momentum, can be represented by a single N-body Slater determinant $|\psi_\alpha^{ll}\rangle = |\alpha_1, \cdots, \alpha_N\rangle$ of spin-orbitals $|\alpha_\kappa\rangle = |n^\kappa l l s m_s^\kappa\rangle$ with $\kappa \in \{1, \cdots N\}$. For $N > 2$ the spin-orbitals must differ in the quantum numbers n^κ and/or m_s^κ.

Solution. Setting $l^\kappa = m_l^\kappa = l$ in Eq. (7.93) and (7.99) these expression evaluate to

$$\begin{aligned} &L_z|\psi_\alpha^{ll}\rangle = Nl\hbar\, |\psi_\alpha^{ll}\rangle \\ &\mathbf{L}^2|\psi_\alpha^{ll}\rangle = \left[N\, l(l+1) + N(N-1)l^2\right] \hbar^2 |\psi_\alpha^{ll}\rangle = Nl(Nl+1)\hbar^2 |\psi_\alpha^{ll}\rangle, \end{aligned}$$

where $Nl = L = M_L$ and $|\psi_\alpha^{ll}\rangle$ is the Slater determinant we are looking for. Analogous expressions can be derived for stretched spin states. □

7.4.1.2 Commutation relations

Starting from $\mathbf{L} = \sum_{i=1}^N \boldsymbol{l}_i$ and $L_z = \sum_{i=1}^N l_z^{(i)}$ it is straightforward to derive the following commutation relations (see Problems 7.9 and 7.10)

$$[\mathbf{L}^2, \boldsymbol{l}_i^2] = 0 \text{ and } [L_z, \boldsymbol{l}_i^2] = 0 \tag{7.107a}$$

$$[\mathbf{L}^2, l_z^{(i)}] \neq 0 \text{ but } [\mathbf{L}^2, L_z] = 0. \tag{7.107b}$$

Problem 7.9. Prove the commutation relations (7.107b)

$$[\mathbf{L}^2, l_\nu^{(i)}] \neq 0 \text{ but } [\mathbf{L}^2, L_\nu] = 0 \;\text{ for } i \in \{1, \cdots N\} \text{ and } \nu \in \{x, y, z\}.$$

Note that this is a generalization of the commutation relation (4.9).

Solution. Let us choose $\nu = z$ (proof analogous for $\nu \in \{x, y\}$). Using Eq. (7.98) we find

$$\begin{aligned}[\mathbf{L}^2, l_z^{(i)}] &= \sum_j [\boldsymbol{l}_j^2, l_z^{(i)}] + \sum_{j,k}{}' \left([l_x^{(j)} l_x^{(k)}, l_z^{(i)}] + [l_y^{(j)} l_y^{(k)}, l_z^{(i)}] + [l_z^{(j)} l_z^{(k)}, l_z^{(i)}] \right) \\ &= [\boldsymbol{l}_i^2, l_z^{(i)}] + \sum_j{}' \left(l_x^{(j)} [l_x^{(i)}, l_z^{(i)}] + l_y^{(j)} [l_y^{(i)}, l_z^{(i)}] \right) + \sum_k{}' \left([l_x^{(i)}, l_z^{(i)}] l_x^{(k)} + [l_y^{(i)}, l_z^{(i)}] l_y^{(k)} \right),\end{aligned}$$

where the prime indicates that $j \neq i \neq k$. Changing dummy index $k \to j$ and using the commutation relation (1.36) this results in [

$$[\mathbf{L}^2, l_z^{(i)}] = 2 \sum_j{}' \left(l_x^{(j)} [l_x^{(i)}, l_z^{(i)}] + l_y^{(j)} [l_y^{(i)}, l_z^{(i)}] + l_z^{(j)} [l_z^{(i)}, l_z^{(i)}] \right) = 2\hbar \sum_j{}' \left(-l_x^{(j)} l_y^{(i)} + l_y^{(j)} l_x^{(i)} \right) \neq 0.$$

This shows that $\mathbf{L}^2$ and $l_z^{(i)}$do *not* commute. Summing the $l_z^{(i)}$ to L_z we obtain

$$[\mathbf{L}^2, L_z] = \sum_i [\mathbf{L}^2, l_z^{(i)}] = 2\hbar \sum_{i,j}{}' \left(l_y^{(j)} l_x^{(i)} - l_x^{(j)} l_y^{(i)} \right) = 0$$

because i and j run over the same range of values $i, j \in \{1, \cdots N\}$. □

Problem 7.10. Prove the commutation relations (7.107a)

$$[\mathbf{L}^2, \boldsymbol{l}_i^2] = 0 \text{ and } [L_\nu, \boldsymbol{l}_i^2] = 0 \quad \text{for } i \in \{1, \cdots N\} \text{ and } \nu \in \{x, y, z\}.$$

Note that this is a generalization of the commutation relation (4.7).

Solution. Using Eq. (7.97) we have

$$[\mathbf{L}^2, \boldsymbol{l}_i^2] = \sum_j [\boldsymbol{l}_j^2, \boldsymbol{l}_i^2] + \sum_{j,k}{}' [\boldsymbol{l}_j \cdot \boldsymbol{l}_k, \boldsymbol{l}_i^2]$$

The first term vanishes because $\sum_j [\boldsymbol{l}_j^2, \boldsymbol{l}_i^2] = [\boldsymbol{l}_i^2, \boldsymbol{l}_i^2] = 0$. Also the second term is zero because

$$\begin{aligned}\sum_{j,k}{}' [\boldsymbol{l}_j \cdot \boldsymbol{l}_k, \boldsymbol{l}_i^2] &= \sum_{\nu=x,y,z} \sum_{j,k}{}' \left([l_\nu^{(j)}, \boldsymbol{l}_i^2] l_\nu^{(k)} + l_\nu^{(j)} [l_\nu^{(k)}, \boldsymbol{l}_i^2] \right), \\ &= \sum_{\nu=x,y,z} \sum_k{}' [l_\nu^{(i)}, \boldsymbol{l}_i^2] l_\nu^{(k)} + \sum_{\nu=x,y,z} \sum_j{}' l_\nu^{(j)} [l_\nu^{(i)}, \boldsymbol{l}_i^2] = 0\end{aligned}$$

because $[l_\nu^{(i)}, \boldsymbol{l}_i^2] = 0$ for $\nu \in \{x, y, z\}$; the prime indicates that $j \neq i \neq k$. Hence, $\mathbf{L}^2$ and $\boldsymbol{l}_i^2$ indeed commute. In view of the definition of L_ν we have

$$[L_\nu, \boldsymbol{l}_i^2] = \sum_j [l_\nu^{(j)}, \boldsymbol{l}_i^2] = [l_\nu^{(i)}, \boldsymbol{l}_i^2] = 0.$$ □

7.4.2 Total electronic spin S

The properties of the total spin angular momentum operator

$$\mathbf{S} = \textstyle\sum_i \boldsymbol{s}_i \tag{7.108}$$

acting on the basis $\{|\psi_\alpha\rangle\}$ is rapidly derived from the expressions for the orbital angular momentum derived above. The spin case is somewhat simpler because the spin of the individual electrons is always $s = \frac{1}{2}$ and the magnetic quantum number can only take the values $m_s = \pm 1/2$. Thus we obtain

- *The operator* $S_z = \sum_i s_z^{(i)}$. In analogy with Eq. (7.93) we find

  $$S_z |\psi_\alpha\rangle = M_S^\alpha \hbar \, |\psi_\alpha\rangle, \tag{7.109}$$

  where $M_S^\alpha = \sum_\kappa m_s^\kappa$ is the total magnetic quantum number of the state $|\psi_\alpha\rangle$.

- *The operator* $S_\pm = \sum_i s_\pm^{(i)}$. In analogy with Eq. (7.95) we find

$$S_\pm|\psi_\alpha\rangle = \sum_\kappa \sqrt{s(s+1) - m_s^\kappa(m_s^\kappa \pm 1)}\,\hbar\,|\psi_\alpha(m_s^\kappa \pm 1)\rangle. \tag{7.110}$$

- *The operator* $\mathbf{S}^2 = (\sum_i \boldsymbol{s}_i)^2$. In analogy with Eq. (7.99) we find

$$\begin{aligned}\mathbf{S}^2|\psi_\alpha\rangle = &\left\{\sum_\kappa s(s+1) + {\sum_{\kappa,\kappa'}}' m_s^\kappa m_s^{\kappa'}\right\}\hbar^2\,|\psi_\alpha\rangle + \\ &+ {\sum_{\kappa,\kappa'}}' \sqrt{s(s+1) - m_s^\kappa(m_s^\kappa + 1)}\sqrt{s(s+1) - m_s^{\kappa'}(m_s^{\kappa'} - 1)}\,\hbar^2\,|\psi_\alpha(m_s^\kappa + 1)(m_s^{\kappa'} - 1)\rangle.\end{aligned} \tag{7.111}$$

7.4.2.1 Commutation relations

Analogous to the commutation relations for the orbital angular momentum it is straightforward to derive the following commutation relations for the total electronic spin

$$[\mathbf{S}^2, \boldsymbol{s}_i^2] = 0 \quad \text{and} \quad [S_z, \boldsymbol{s}_i^2] = 0; \qquad [\mathbf{S}^2, s_z^{(i)}] \neq 0 \quad \text{but} \quad [\mathbf{S}^2, S_z] = 0. \tag{7.112}$$

7.4.3 Total electronic angular momentum J

To conclude this chapter we also consider the total electronic angular momentum,

$$\mathbf{J} = \textstyle\sum_i(\boldsymbol{l}_i + \boldsymbol{s}_i). \tag{7.113}$$

Since $\mathbf{J} = \mathbf{L} + \mathbf{S}$ we can derive the following commutation relations,

$$[\mathbf{J}^2, \boldsymbol{l}_i^2] = 0 \text{ and } [J_z, \boldsymbol{l}_i^2] = 0. \tag{7.114}$$

Problem 7.11. Prove the commutation relations (7.114)

$$[\mathbf{J}^2, \boldsymbol{l}_i^2] = 0 \text{ and } [J_z, \boldsymbol{l}_i^2] = 0 \;\text{ for } i \in \{1, \cdots N\}.$$

Solution. Using the first commutation relation (7.107a) we obtain

$$[\mathbf{J}^2, \boldsymbol{l}_i^2] = [\mathbf{L}^2 + \mathbf{S}^2 + \mathbf{L}\cdot\mathbf{S}, \boldsymbol{l}_i^2] = [\mathbf{L}\cdot\mathbf{S}, \boldsymbol{l}_i^2]$$

and using the second commutation relation (7.107a) we find

$$\begin{aligned}[\mathbf{L}\cdot\mathbf{S}, \boldsymbol{l}_i^2] &= [L_xS_x + L_yS_y + L_zS_z), \boldsymbol{l}_i^2] \\ &= [L_x, \boldsymbol{l}_i^2]S_x + [L_y, \boldsymbol{l}_i^2]S_y + [L_z, \boldsymbol{l}_i^2]S_z = 0.\end{aligned}$$

□

8

Many-electron atoms

The description of atoms with more than two electrons builds on the concepts introduced for helium: *central symmetry*, *screening*, *Pauli principle* and *exchange*. Just as for helium also for many-electron atoms we start from a Hamiltonian which is invariant under space inversion and rotation about the origin; i.e., the parity of the electronic wavefunction as well as the total orbital angular momentum are conserved in time. Recalling from the helium chapter the concept of electrostatic screening and the convenience of the central field, it is an excellent Ansatz that also in the many-electron case central forces dominate the motion of the individual electrons. This *central field approximation* allows the separation of variables, which makes it possible to continue the practice of describing the atoms in terms of single-electron orbitals specified by the quantum numbers n and l. This means that the atomic state can be characterized by specifying the *electron configuration.* The configuration with the lowest energy (strongest binding) corresponds to the ground state. For hydrogenic atoms the possible configurations are: $1s$, $2s$, $2p$, $3s$, $3p$, $3d, \cdots$. So, in hindsight - the central field approximation provides the justification for treating the alkali atoms as *hydrogen-like* atoms as was done in previous chapters (in Section 4.6.2 for the fine structure of *alkali-like* atoms; in Chapter 5 for their hyperfine structure). For atoms with more than one electron the Pauli principle puts a constraint on the occupation of the orbitals. In this way the helium configurations become $1s^2$, $1s2s$, $1s2p$, $2s^2, \cdots$, for lithium we have $1s^22s$, $1s^22p, \cdots$, and so forth. So, in the configuration we group the electrons into *shells* of *equivalent* electrons, where electrons are called equivalent if they carry the same quantum numbers n and l (irrespective of the magnetic quantum numbers m_l and m_s). A shell containing the maximum number of electrons consistent with the Pauli principle is called *closed* or *fully filled.* The electrons in closed shells are called *core electrons.* A vacant electron in a shell is called a *hole.* The electrons outside closed shells are called *valence electrons.*

As was established before, the Coulomb interaction dominates by orders of magnitude ($\alpha^2 \approx 5 \times 10^{-5}$) over all other terms in the atomic Hamiltonian. Therefore, to calculate the ground state energy the electronic charge has to be distributed across the atom as accurately as possible. In atomic theory this is usually done with a self-consistent mean field method like the Hartree method introduced in the helium chapter. Other approaches are *density-functional theory* and *potential-functional theory* [39]. These are variational methods in which the variation is not done at the level of the wavefunction but by varying the probability charge density or its potential field, respectively. The archetype of such statistical theories is the Thomas-Fermi method [109, 41]. Density functional theory plays an important role in chemistry and condensed matter physics, more in general in the physics of many-body systems which are too big for an *ab initio* quantum mechanical approach like the self-consistent mean field method [64].

In this chapter we explore the *central field* concept and introduce the *central field approximation.* In Section 4.7 we already showed how this concept is used to describe *alkali-like* atoms. In the present chapter we deepen the discussion by introducing the *Hartree-* and *Hartree-Fock methods* for calculating the central field.

8.1 Hamiltonian for N-electron atoms

The Hamiltonian for the principal structure of a N-electron atom is given by a generalization of the Hamiltonian (6.1),

$$\mathcal{H}=\sum_{i=1}^{N}\left(-\frac{\hbar^2}{2m_e}\nabla_i^2-\frac{Ze^2}{4\pi\varepsilon_0 r_i}\right)+\tfrac{1}{2}\sum_{i,j}^{N}{}'\frac{e^2}{4\pi\varepsilon_0 r_{ij}} \tag{8.1}$$

The first sum is the Hamiltonian for N noninteracting electrons orbiting about a nucleus of charge Ze and infinite mass; m_e is the electron rest mass and $\mathbf{r}_i$ the position of electron i relative to the atomic center. The second term accounts pairwise for the electrostatic repulsion between the electrons, with $r_{ij}=|\mathbf{r}_i-\mathbf{r}_j|$ being the distance between the electrons i and j. The factor $\frac{1}{2}$ in front of the double summation is to correct for double counting and the prime indicates that the case $i=j$ is excluded. For neutral atoms we have $N=Z$.

Like the Schrödinger Hamiltonian $\mathcal{H}$ is invariant under space inversion, which means that the parity of the electronic states is conserved in time (see Appendix F.2.4). Likewise, $\mathcal{H}$ is invariant under rotation of the spatial coordinates about the origin, so the total orbital angular momentum $\mathbf{L}=\boldsymbol{l}_1+\cdots+\boldsymbol{l}_N$ and its projection L_z are conserved quantities (cf. Problem 8.1),

$$[\mathbf{L}^2,\mathcal{H}]=0\quad\text{and}\quad[L_z,\mathcal{H}]=0. \tag{8.2}$$

Although spin is absent in the Hamiltonian (8.1), the electrons carry of course spin. So, also the total spin $\mathbf{S}=\boldsymbol{s}_1+\cdots+\boldsymbol{s}_N$ and its projection S_z are conserved quantities under the Hamiltonian (8.1). Hence, also the sum,

$$\mathbf{J}=\mathbf{L}+\mathbf{S}=\sum_i(\boldsymbol{l}_i+\boldsymbol{s}_i)=\sum_i\boldsymbol{j}_i, \tag{8.3}$$

and its projection J_z are conserved,

$$[\mathbf{J}^2,\mathcal{H}]=0\quad\text{and}\quad[J_z,\mathcal{H}]=0. \tag{8.4}$$

The spin plays a major role because the Hamiltonian is invariant under exchange of any two electrons; i.e., it commutes with the permutation operator,

$$[P,\mathcal{H}]=0. \tag{8.5}$$

As the electrons are fermions, quantum statistics imposes the condition that the N-electron wavefunction be antisymmetric under exchange of any two electrons. As was demonstrated for helium atoms this is possible only for certain combinations of L and S. As the corresponding wavefunctions differ in electrostatic repulsion their energy levels are typically separated by electron volts. This firmly locks the N-electron atom into (magnetic) states of given L and S, in which L,M_L,S,M_S,J,M_J are good quantum numbers and the parity under exchange of electrons is conserved. In this way the electrostatic interaction gives rise to strong coupling (*exchange coupling*) of both the orbital angular momenta and the spins of the individual electrons into the total angular momenta $\mathbf{L}=\sum_i\boldsymbol{l}_i$ and $\mathbf{S}=\sum_i\boldsymbol{s}_i$, respectively. This electrostatic origin makes the exchange interaction into the strongest magnetic coupling mechanism in atomic physics.

Problem 8.1. Show that the total orbital angular momentum $\mathbf{L}=\boldsymbol{l}_1+\cdots+\boldsymbol{l}_N$, where $\boldsymbol{l}_\nu=-i\hbar(\mathbf{r}_\nu\times\boldsymbol{\nabla}_\nu)$ with $\nu\in\{1,\ldots,N\}$, is a conserved quantity under Hamiltonian (8.1).

Solution. The electrostatic repulsion is pairwise additive. Thus we require

$$\frac{1}{2}\sum_{i,j}^{N}{}'[L_z,1/r_{ij}]=0\quad\text{and}\quad\frac{1}{2}\sum_{i,j}^{N}{}'[\mathbf{L}^2,1/r_{ij}]=0,$$

where $r_{ij}=|\mathbf{r}_i-\mathbf{r}_j|$ is the distance between the electron i and j. This condition is satisfied as follows by applying Problem 6.1 to all individual pairs. ☐

8.1.1 Central field approximation

Recalling the concept of screening of the nuclear charge as discussed for the helium atom we rewrite the Hamiltonian (8.1) in the form

$$\mathcal{H}=\sum_{i=1}^{N}\left(-\frac{\hbar^2}{2m_e}\nabla_i^2-\frac{Ze^2}{4\pi\varepsilon_0 r_i}+\mathcal{V}_{\rm scr}(r_i)\right)+\tfrac{1}{2}\sum_{i,j}^{N}{}'\frac{e^2}{4\pi\varepsilon_0 r_{ij}}-\sum_{i=1}^{N}\mathcal{V}_{\rm scr}(r_i), \tag{8.6}$$

where for each electron i we both added and subtracted a *centrally symmetric* potential energy term $\mathcal{V}_{\rm scr}(r_i)$, which can be regarded as a generalization of the screening potential (6.13) introduced for helium. By this procedure the Hamiltonian separates into two parts, $\mathcal{H}=\mathcal{H}_{\rm CF}+\mathcal{H}'$.

The first term,

$$\mathcal{H}_{\rm CF}\equiv\sum_{i=1}^{N}\left(-\frac{\hbar^2}{2m_e}\nabla_i^2+\mathcal{V}_{\rm CF}(r_i)\right), \tag{8.7}$$

is called the *central field Hamiltonian.* It is the sum of N one-electron Hamiltonians, a generalization of the Hamiltonian (6.5) but it accounts for screening because

$$\mathcal{V}_{\rm CF}(r_i)=-\frac{Ze^2}{4\pi\varepsilon_0 r_i}+\mathcal{V}_{\rm scr}(r_i)=-\frac{e^2}{4\pi\varepsilon_0}\frac{Z_{\rm eff}(r_i)}{r_i} \tag{8.8}$$

is the *potential energy* of electron i in the (centrally symmetric) *electrostatic potential* of the nucleus screened by the other $(N-1)$ electrons. It is the *effective central field potential* for the N-electron problem, a generalization of the Eq. (6.17) used for helium. The quantity $Z_{\rm eff}(r_i)$ is the *effective nuclear charge* for electron i at distance r_i from the nucleus introduced in Eq. (4.101). Close to the nucleus the screening is negligible and $\mathcal{V}_{\rm CF}(r)$ must approach the potential energy of an electron in the unscreened nuclear charge Ze. Far from the nucleus all but the Rydberg charge $Z_c e$ is shielded. This imposes onto the $Z_{\rm eff}(r)$ the boundary conditions $\lim_{r\to 0}Z_{\rm eff}(r)=Z$ and $\lim_{r\to\infty}Z_{\rm eff}(r)=Z_c$, with $Z_c=1$ for neutral atoms.

The second term of Eq. (8.6),

$$\mathcal{H}'\equiv\tfrac{1}{2}\sum_{i,j=1}^{N}{}'\frac{e^2}{4\pi\varepsilon_0 r_{ij}}-\sum_{i=1}^{N}\mathcal{V}_{\rm scr}(r_i), \tag{8.9}$$

represents the residual Coulomb interactions that cannot be included in the *centrally symmetric* screening field. This term can be regarded as a (non-central) perturbation on $\mathcal{H}_{\rm CF}$ and accounts for *Coulomb correlations* in the relative motion of the electrons. For optimal screening potentials this perturbation is minimal. Neglecting $\mathcal{H}'$ we obtain the *central field approximation*,

$$\mathcal{H}\simeq\mathcal{H}_{\rm CF}. \tag{8.10}$$

A first estimate for the importance of the Coulomb correlations can be obtained by perturbation theory with respect to the central field solutions.

The *central-field approximation* preserves the separation of radial from angular variables while accounting for most of the screening. This makes it possible to replace the hydrogenic basis by a new basis in which the hydrogenic state labeling $|n,l,m_l\rangle$ is preserved but the radial wavefunctions are modified to account for the screening. In this way the atomic state can be approximated by a single *electron configuration.* For instance, for *one-electron atoms* the possible configurations remain $1s$, $2s$, $2p$, $3s$, $3p$, $3d,\cdots$ but their energy differs dramatically from the hydrogenic values. In principle, the configuration with the lowest energy (strongest binding) corresponds to the ground state, much like in non-interacting electron atoms (see Section 8.3).

8.2 Systems with more than two identical particles

In the previous chapters we used Slater determinants as the properly symmetrized basis states to calculate the repulsion energy in atoms. With this approach the Coulomb interaction between any two electrons is expressed in terms of the Coulomb integrals $\mathcal{J}$ and $\mathcal{K}$ for which analytic expressions are given in Section 6.4. As $\mathcal{H}$ is not diagonal in this representation (cf. Section 7.1.1.1), the energy eigenstates are (in general) linear combinations of Slater determinants.

We now ask for a representation in which the Hamiltonian (8.1) is diagonal for properly symmetrized eigenstates of the total angular momentum. Our reference point is the symmetrization procedure of Chapter 7. For a system of two spins we recall that the total spin $\mathbf{S} = \mathbf{s}_1 + \mathbf{s}_2$ can be written as a linear combination of the eigenstates of the uncoupled representation. These are the product states $\{|sm_1\rangle\otimes, sm_2\rangle\}$, which span a $(2s+1)\times(2s+1) = 4$ dimensional subspace of Hilbert space. Alternatively, we can turn to the coupled representation in which the eigenstates of $\mathbf{S}$ are the singlet state (antisymmetric) and 3 triplet states (symmetric) as shown in Problem 6.2. In this way the 4-dimensional subspace is reduced into the sum of two irreducible subspaces of lower dimension (a 3-dimensional subspace for the triplet and a 1-dimensional subspace for the singlet). As is demonstrated in Problem 7.1, this reduction is achieved by projecting out the symmetric and antisymmetric parts of an arbitrary spin state.

8.2.1 Permutation symmetry

We now ask for a generalized reduction procedure suitable for systems of more than two particles. Since $\mathcal{H}$ commutes with P we look for eigenstates that conserve the permutation symmetry. For the simplest form of permutation, exchange of two particles, we found two symmetry classes (symmetric and antisymmetric). For permutations involving more than two particles we shall find that other symmetry classes emerge (partly symmetric and partly antisymmetric). In the language of group theory each permutation belongs to one of the symmetry classes of the *permutation*, or *symmetric*, *group*, S_n.[1]

To properly introduce the subject we include a mathematical intermezzo on permutations [95]. The discussion starts by considering a collection of n *objects* as a set of n elements $X = \{x_1, \ldots, x_n\}$. For our purpose X can be a set of particles (e.g., electrons) or a set of quantum mechanical states (e.g., electronic orbitals). A *permutation* of the set X is a *one-to-one* map of X *onto* itself, $P : X \to X$.[2] The permutation defined by $x_i \mapsto P(x_i)$, with $i \in \{1, \ldots, n\}$, is realized by mapping of the *indices* $i \mapsto p_i$, where $p_i \in \{1, \ldots, n\}$.

The permutation P and its inverse P^{-1} are written in the form

$$P = \begin{pmatrix} 1 & 2 & \cdots & n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} \qquad P^{-1} = \begin{pmatrix} p_1 & p_2 & \cdots & p_n \\ 1 & 2 & \cdots & n \end{pmatrix}. \tag{8.11}$$

Note that the *ordering of the columns is irrelevant* in this notation. This makes the permutation P into a *unordered* set of n *independent* maps between the indices. By ordering of this set we can simplify the notation. We demonstrate this for a *cyclic* permutation (short: *cycle*). A permutation is cyclic if it can be written in the following *ordered* form:

$$\begin{pmatrix} 1 & q_2 & q_3 & \cdots & q_{\ell-1} & q_\ell \\ q_2 & q_3 & q_4 & \cdots & q_\ell & 1 \end{pmatrix} = (1\, q_2 \cdots q_{\ell-1} q_\ell), \tag{8.12}$$

In this ordering/notation we keep track of how one index is mapped onto the next until we return at the first index (here $q_1 = 1$). Note that on the r.h.s. the bottom row has been left out without loss

---

[1] The set of all permutations of n objects constitutes a group of $n!$ elements. The unit element is *no* permutation, for each permutation there exists an *inverse*, the set is *closed* under permutation and successive permutations are *associative*. The number of elements is called the *order* of the group.

[2] In mathematics such a map is called an *automorphism*.

of information. This is known as the *cycle notation.* The number of elements of a cycle is called the *length* of the cycle, ℓ in the case of Eq. (8.12). Cycles are at least of *length* 1 and at most of *length* n. A cycle of length ℓ is called an *ℓ-cycle.* A 2-cycle is called a *transposition*, i.e., a permutation which exchanges two elements of X while keeping all others fixed. Such a permutation is evidently equal to its inverse. A cycle of length 1 is called a *unit cycle.*

Let us demonstrate how we find the cycle notation for a cyclic permutation from S_6,

$$P = \begin{pmatrix} 1\,2\,3\,4\,5\,6 \\ 3\,5\,2\,1\,6\,4 \end{pmatrix} = \begin{pmatrix} 1\,3\,2\,5\,6\,4 \\ 3\,2\,5\,6\,4\,1 \end{pmatrix} = (1\,3\,2\,5\,6\,4) \tag{8.13a}$$

$$P^{-1} = \begin{pmatrix} 3\,5\,2\,1\,6\,4 \\ 1\,2\,3\,4\,5\,6 \end{pmatrix} = \begin{pmatrix} 3\,1\,4\,6\,5\,2 \\ 1\,4\,6\,5\,2\,3 \end{pmatrix} = (3\,1\,4\,6\,5\,2) = (4\,6\,5\,2\,3\,1). \tag{8.13b}$$

Note how we first rearrange the columns to bring the permutation in the standard ordering (8.12) and subsequently drop the second row to arrive at the cycle notation. In the following example from S_6 the permutation can be written as the product of three *disjoint* cycles,[1]

$$P = \begin{pmatrix} 1\,2\,3\,4\,5\,6 \\ 3\,1\,2\,5\,4\,6 \end{pmatrix} = \left(\begin{array}{c|c|c} 1\,3\,2 & 4\,5 & 6 \\ 3\,2\,1 & 5\,4 & 6 \end{array}\right) = (1\,3\,2)(4\,5)(6) = (1\,3\,2)(4\,5), \tag{8.14}$$

where (132) is a 3-cycle, (4 5) a 2-cycle and (6) a unit cycle. Unit cycles may be omitted because they do not affect the permutation. Note that *disjoint cycles commute* (permutations are invariant under rearrangement of the columns).

Successive permutations are written as an *operator product*, $P = P_\beta P_\alpha$ (P_α applied first, P_β second). In general, permutations do *not* commute, $P_\alpha P_\beta \neq P_\beta P_\alpha$. We demonstrate this for two successive permutations from S_4,

$$P_\alpha = \begin{pmatrix} 1\,2\,3\,4 \\ 3\,4\,1\,2 \end{pmatrix} = (1\,3)(2\,4) \qquad P_\beta = \begin{pmatrix} 1\,2\,3\,4 \\ 2\,4\,3\,1 \end{pmatrix} = (1\,2\,4)(3). \tag{8.15}$$

Evaluating the products $P_\beta P_\alpha$ and $P_\alpha P_\beta$ we find

$$P_\beta P_\alpha = \begin{pmatrix} 1\,2\,3\,4 \\ 2\,4\,3\,1 \end{pmatrix}\begin{pmatrix} 1\,2\,3\,4 \\ 3\,4\,1\,2 \end{pmatrix} = \begin{pmatrix} 1\,2\,3\,4 \\ 3\,1\,2\,4 \end{pmatrix} = \begin{pmatrix} 1\,3\,2\,4 \\ 3\,2\,1\,4 \end{pmatrix} = (1\,3\,2)(4) \tag{8.16a}$$

$$P_\alpha P_\beta = \begin{pmatrix} 1\,2\,3\,4 \\ 3\,4\,1\,2 \end{pmatrix}\begin{pmatrix} 1\,2\,3\,4 \\ 2\,4\,3\,1 \end{pmatrix} = \begin{pmatrix} 1\,2\,3\,4 \\ 4\,2\,1\,3 \end{pmatrix} = \begin{pmatrix} 1\,4\,3\,2 \\ 4\,3\,1\,2 \end{pmatrix} = (1\,4\,3)(2) \neq P_\beta P_\alpha. \tag{8.16b}$$

In cycle notation this becomes compactly

$$P_\beta P_\alpha = (1\,2\,4)(1\,3)(2\,4) = (1\,3\,2). \tag{8.17a}$$

$$P_\alpha P_\beta = (1\,3)(2\,4)(1\,2\,4) = (1\,4\,3) \neq P_\beta P_\alpha. \tag{8.17b}$$

We now arrived at an important point. In the example of Eq, (8.14) we found that the permutation P could be written as a product of *disjoint* cycles. Without entering into formal group theory we point out that this is a general property of permutations: *every permutation can be written as a product of disjoint cycles.* This decomposition is *unique* and defines the symmetry classes of S_n. It follows that for a decomposition into m cycles of length $\lambda_1, \lambda_2, \ldots, \lambda_m$, the lengths add up to n; i.e., $\lambda_1 + \lambda_2 \cdots + \lambda_m = n$.

Let us return to the physics and demonstrate how to apply two subsequent transpositions, $P_{13} = (13)$ followed by $P_{12} = (12)$, on the product state of 3 particles $|abc)$. Like in Chapter 7 we use the convention of permuting states rather than the particle indices,

$$P_{12}P_{13}|abc) = P_{12}|cba) = |bca). \tag{8.18}$$

---

[1] Two sets are called *disjoint* if they have no elements in common.

Note that in cycle notation we have $(12)(13) = (1\,3\,2)$. Therefore, we can equally well apply $P_{132} = (1\,3\,2)$ on the product state $|abc)$. However, since permutations of more than two particles are not equal to their inverse, the correct use of the permutation convention deserves attention. Permuting again states rather than particle indices, the permutation P_{132} shifts the state at position 1 (short: state 1) to position 3 (puts particle 3 into the state of particle 1), state 3 to position 2 and state 2 to position 1,

$$P_{132}|abc) = |bca) \qquad P_{231}|bca) = |abc), \tag{8.19}$$

where $P_{231} = P_{132}^{-1}$ is the inverse permutation. Note that we correctly reproduce the result of Eq. (8.18).

Problem 8.2. Show for an example from S_4 that $P^4 = \mathbb{1}$ and $P^3 = P^{-1}$.

Solution. We show this for the cycle $P = (1\,4\,3\,2)$

$$P = (1\,4\,3\,2) = \begin{pmatrix} 1\,2\,3\,4 \\ 4\,1\,2\,3 \end{pmatrix}$$

$$P^2 = \begin{pmatrix} 1\,2\,3\,4 \\ 4\,1\,2\,3 \end{pmatrix} \begin{pmatrix} 1\,2\,3\,4 \\ 4\,1\,2\,3 \end{pmatrix} = \begin{pmatrix} 1\,2\,3\,4 \\ 3\,4\,1\,2 \end{pmatrix}$$

$$P^3 = \begin{pmatrix} 1\,2\,3\,4 \\ 4\,1\,2\,3 \end{pmatrix} \begin{pmatrix} 1\,2\,3\,4 \\ 3\,4\,1\,2 \end{pmatrix} = \begin{pmatrix} 1\,2\,3\,4 \\ 2\,3\,4\,1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 4\,1\,2\,3 \\ 1\,2\,3\,4 \end{pmatrix} = P^{-1}$$

$$P^4 = \begin{pmatrix} 1\,2\,3\,4 \\ 4\,1\,2\,3 \end{pmatrix} \begin{pmatrix} 1\,2\,3\,4 \\ 2\,3\,4\,1 \end{pmatrix} = \begin{pmatrix} 1\,2\,3\,4 \\ 1\,2\,3\,4 \end{pmatrix} = \mathbb{1}\,.$$

8.2.2 Young diagrams and Young tableaux

The search for the symmetry classes of S_n leads us to the *Young diagrams* of combinatorics, which offer a convenient (graphical) approach of partitioning n objects into m parts [95]. The simplest form of partitioning is *integer partitioning*: any positive integer n can be partitioned into a sum of integers of increasing value,

$$n = \lambda_1 + \lambda_2 \cdots + \lambda_m, \tag{8.20}$$

with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$. The ordered set of m integers $\lambda \equiv [\lambda_1, \lambda_2, \ldots, \lambda_m]$ is called the *partition* λ of the integer n into m integer parts.

Importantly, notice that *each partition λ corresponds to one of the symmetry classes of S_n* because decomposition of n objects into m disjoint cycles is equivalent to integer partitioning of n into m parts. Since there is (in general) more than a single way of partitioning n into m parts, it is useful to denote the set of all partitions of n into m parts by $\mathbb{P}(n,m)$. By summing over all values of m we obtain the set of all partitions of n,

$$\mathbb{P}(n) = \sum_{m=1}^{n} \mathbb{P}(n,m). \tag{8.21}$$

For up to $n = m = 7$ the integer partitions are summarized in Table 8.1.

The integer partition $\lambda \in \mathbb{P}(n,m)$ can be represented graphically in the form of the *Young diagram* Y_λ consisting of m rows of λ_1, $\lambda_2, \ldots, \lambda_m$ *left-justified* square *cells*. For instance, we find two Young diagrams for the partitions of 4 cells over 2 rows

$$\mathbb{P}[4,2] \to \lambda = \begin{cases} [3,1] \to \ydiagram{3,1} \equiv Y_{[3,1]} \\ [2,2] \equiv [2^2] \to \ydiagram{2,2} \equiv Y_{[2^2]}. \end{cases} \tag{8.22}$$

Table 8.1: Partitions of up to 7 objects over up to 7 parts. The columns contain *all* partitions of n.

$\mathbb{P}(n,m)$	objects (n)						
parts (m)	1	2	3	4	5	6	7
1	[1]	[2]	[3]	[4]	[5]	[6]	[7]
2		[1,1]	[2,1]	[3,1], [2,2]	[4,1], [3,2]	[5,1], [4,2], [3,3]	[6,1], [5,2], [4,3]
3			[1,1,1]	[2,1,1]	[3,1,1], [2,2,1]	[4,1,1], [3,2,1], [2,2,2]	[5,1,1], [4,2,1], [3,3,1]
4				[1,1,1,1]	[2,1,1,1]	[3,1,1,1], [2,2,1,1]	[4,1,1,1], [3,2,1,1], [2,2,2,1]
5					[1,1,1,1,1]	[2,1,1,1,1]	[3,1,1,1,1], [2,2,1,1,1]
6						[1,1,1,1,1,1]	[2,1,1,1,1,1]
7							[1,1,1,1,1,1,1]

It is a simple procedure to draw all diagrams of $\mathbb{P}(n)$. For $\mathbb{P}(4)$ we have $\mathbb{P}(4,1)$, $\mathbb{P}(4,2)\,(2\times)$, $\mathbb{P}(4,3)$ and $\mathbb{P}(4,4)$ and find 5 shapes,

$$\begin{array}{|c|c|c|c|}\hline \ & \ & \ & \ \\ \hline \end{array} \qquad \begin{array}{|c|c|c|}\hline \ & \ & \ \\ \hline \ \\ \cline{1-1} \end{array} \qquad \begin{array}{|c|c|}\hline \ & \ \\ \hline \ & \ \\ \hline \end{array} \qquad \begin{array}{|c|c|}\hline \ & \ \\ \hline \ \\ \cline{1-1} \ \\ \cline{1-1} \end{array} \qquad \begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \ \\ \hline \ \\ \hline \end{array} \; . \tag{8.23}$$

Since Young diagrams are invariant under *horizontal* permutation of the cells in the rows, we identified all symmetry classes of S_4. Furthermore, we found one *purely horizontal*, one *purely vertical* and several *mixed* partitions. This is typical for all S_n (mixed partitions only for $n \geq 3$).

To deal with the invariance of many-body states under exchange of two identical particles we label cells of the Young diagrams with particle/state indices $i = 1, \cdots, n$. From the $n!$ different ways of doing this we select *standard ordering.* This is a labeling convention in which the index increases horizontally from left to right (for symmetrization) and vertically from top to bottom (for antisymmetrization). After labeling, the *Young diagram* Y_λ is referred to as the *Young tableaux* Θ_λ. For example, for the partition $\lambda = [2,1^2]$ we find

$$Y_{[2,1^2]} \;\rightarrow\; \Theta_{[2,1^2]} \equiv \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} 4 \\ \cline{1-1} \end{array} \;, \quad P_{34}\Theta_{[2,1^2]} = \Theta_{[2,1^2]}^{(3\,4)} \equiv \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 4 \\ \cline{1-1} 3 \\ \cline{1-1} \end{array} \;, \quad P_{234}\Theta_{[2,1^2]} = \Theta_{[2,1^2]}^{(2\,3\,4)} \equiv \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 4 \\ \cline{1-1} 2 \\ \cline{1-1} \end{array} \;.$$

Hence, for mixed partitions the standard ordering is not unique. In the first tableau, $\Theta_{[2,1^2]}$, the index increases in *unit steps.* This is called *normal* ordering. The other tableaux are generated from $\Theta_{[2,1^2]}$ by relabeling the cells using the cyclic permutations $(3\,4)$ and $(2\,3\,4)$, respectively. The labeling convention serves to avoid duplication of equivalent tableaux.

To illustrate the procedure we derive all Young tableaux of S_1, S_2, S_3 and S_4:

- S_1: one particle ($n = 1$): one diagram; the horizontal and vertical partitioning coincide,

$$[1] \equiv \begin{array}{|c|}\hline \ \\ \hline \end{array} \;\rightarrow\; \begin{array}{|c|}\hline 1 \\ \hline \end{array} \tag{8.24}$$

- S_2: two particles ($n = 2$): two diagrams; only purely horizontal and purely vertical partitioning (this is the case of exchange in a two particle system),

$$[2] \equiv \begin{array}{|c|c|}\hline \ & \ \\ \hline \end{array} \;\rightarrow\; \begin{array}{|c|c|}\hline 1 & 2 \\ \hline \end{array} \tag{8.25a}$$

$$[1,1] \equiv [1^2] \equiv \begin{array}{|c|}\hline \ \\ \hline \ \\ \hline \end{array} \;\rightarrow\; \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline \end{array} \tag{8.25b}$$

- S_3: three particles ($n=3$): three diagrams; aside from purely horizontal/vertical partitioning we also have one mixed partition with two possible tableaux,

$$[3] \equiv \begin{array}{|c|c|c|}\hline & & \\\hline\end{array} \rightarrow \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\\hline\end{array} \tag{8.26a}$$

$$[2,1] \equiv \begin{array}{|c|c|}\hline & \\\hline \multicolumn{1}{|c|}{} & \multicolumn{1}{c}{} \\\cline{1-1}\end{array} \rightarrow \begin{array}{|c|c|}\hline 1 & 2 \\\hline \multicolumn{1}{|c|}{3} & \multicolumn{1}{c}{} \\\cline{1-1}\end{array}, \quad \begin{array}{|c|c|}\hline 1 & 3 \\\hline \multicolumn{1}{|c|}{2} & \multicolumn{1}{c}{} \\\cline{1-1}\end{array} \tag{8.26b}$$

$$[1,1,1] \equiv [1^3] \equiv \begin{array}{|c|}\hline \\\hline \\\hline \\\hline\end{array} \rightarrow \begin{array}{|c|}\hline 1 \\\hline 2 \\\hline 3 \\\hline\end{array} \tag{8.26c}$$

- S_4: four particles ($n=4$): five diagrams; aside from purely horizontal/vertical partitioning we have 3 mixed partitions with 3, 2 and 3 tableaux, respectively,

$$[4] \equiv \begin{array}{|c|c|c|c|}\hline & & & \\\hline\end{array} \rightarrow \begin{array}{|c|c|c|c|}\hline 1 & 2 & 3 & 4 \\\hline\end{array} \tag{8.27a}$$

$$[3,1] \equiv \begin{array}{|c|c|c|}\hline & & \\\hline \multicolumn{1}{|c|}{} & \multicolumn{2}{c}{} \\\cline{1-1}\end{array} \rightarrow \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\\hline \multicolumn{1}{|c|}{4} & \multicolumn{2}{c}{} \\\cline{1-1}\end{array}, \quad \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\\hline \multicolumn{1}{|c|}{3} & \multicolumn{2}{c}{} \\\cline{1-1}\end{array}, \quad \begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\\hline \multicolumn{1}{|c|}{2} & \multicolumn{2}{c}{} \\\cline{1-1}\end{array} \tag{8.27b}$$

$$[2,2] \equiv [2^2] \equiv \begin{array}{|c|c|}\hline & \\\hline & \\\hline\end{array} \rightarrow \begin{array}{|c|c|}\hline 1 & 2 \\\hline 3 & 4 \\\hline\end{array}, \quad \begin{array}{|c|c|}\hline 1 & 3 \\\hline 2 & 4 \\\hline\end{array} \tag{8.27c}$$

$$[2,1,1] \equiv [2,1^2] \equiv \begin{array}{|c|c|}\hline & \\\hline \multicolumn{1}{|c|}{} & \multicolumn{1}{c}{} \\\cline{1-1} \multicolumn{1}{|c|}{} & \multicolumn{1}{c}{} \\\cline{1-1}\end{array} \rightarrow \begin{array}{|c|c|}\hline 1 & 2 \\\hline \multicolumn{1}{|c|}{3} & \multicolumn{1}{c}{} \\\cline{1-1} \multicolumn{1}{|c|}{4} & \multicolumn{1}{c}{} \\\cline{1-1}\end{array}, \quad \begin{array}{|c|c|}\hline 1 & 3 \\\hline \multicolumn{1}{|c|}{2} & \multicolumn{1}{c}{} \\\cline{1-1} \multicolumn{1}{|c|}{4} & \multicolumn{1}{c}{} \\\cline{1-1}\end{array}, \quad \begin{array}{|c|c|}\hline 1 & 4 \\\hline \multicolumn{1}{|c|}{2} & \multicolumn{1}{c}{} \\\cline{1-1} \multicolumn{1}{|c|}{3} & \multicolumn{1}{c}{} \\\cline{1-1}\end{array} \tag{8.27d}$$

$$[1,1,1,1] \equiv [1^4] \equiv \begin{array}{|c|}\hline \\\hline \\\hline \\\hline \\\hline\end{array} \rightarrow \begin{array}{|c|}\hline 1 \\\hline 2 \\\hline 3 \\\hline 4 \\\hline\end{array} \tag{8.27e}$$

8.2.3 Symmetrization using Young tableaux

At this point we have identified the symmetry classes of the permutation group and continue our search for the energy eigenstates. Since P not only commutes with $\mathcal{H}$ but also with $\mathbf{S}$, $\mathbf{L}$ and $\mathbf{J}$ we can find energy eigenstates which are also eigenstates of the standard representation of the angular momenta. We first ask for the joint eigenstates of P and $\mathbf{S}=\mathbf{s}_1+\cdots+\mathbf{s}_n$. Unfortunately, a complication arises immediately. Although the operator $\mathbf{S}$ is invariant under both permutation P and exchange P_{ij} of the spins, for more than two particles ($n>2$) the operators P and P_{ij} do not commute mutually for arbitrary values of $i,j \in \{1,\cdots,n\}$. This means that symmetric and antisymmetric are not the only options for the permutation symmetry of the joint eigenstates. Hence, before we can proceed we have to generalize the symmetrization procedure.

The generalization starts with the introduction of two *symmetrizers* (symmetrization operators), the *row symmetrizer* $\mathcal{S}_i$ for the *symmetrization* of row i of a Young tableau and the *column symmetrizer* $\mathcal{A}_i$ for the *antisymmetrization* of column j of a Young tableau.[1]

$$\mathcal{S}_i \equiv \sum_{P\in \mathrm{row}(i)} P \quad \text{and} \quad \mathcal{A}_j \equiv \sum_{Q\in \mathrm{col}(j)} (-1)^q Q. \tag{8.28}$$

Here $P \in \mathrm{row}(i)$ stands for a permutation among the elements of *row* i and $Q \in \mathrm{col}(j)$ for a permutation among the elements of *column* j. With $\mathcal{S}_i$ we *symmetrize row* i by summing over all $P \in \mathrm{row}(i)$. With $\mathcal{A}_i$ we *antisymmetrize column* i by summing over all $Q \in \mathrm{col}(j)$ while accounting for the *parity* of P by the phase factor $(-1)^q$. Here q stands for the number of *transpositions* (binary interchanges) required to realize the permutation P (starting from a tableaux in standard ordering). The *parity* is called *even/odd* depending on q being even/odd.

---

[1] We focus on the symmetrization aspects without entering into normalization and phase issues.

The symmetrizers are *projection operators* because symmetrizing an already symmetrized state leaves the state unchanged,

$$\mathcal{S}_i^2 = \lambda_i! \mathcal{S}_i \quad \text{and} \quad \mathcal{A}_j^2 = \mu_j! \mathcal{A}_j, \tag{8.29}$$

where λ_i and μ_j are the lengths of row i and column j, respectively. Symmetrizing and antisymmetrizing of the same row/column yields orthogonal results; i.e., the *product* of a symmetrizer and an antisymmetizer acting on the *same* row/column *vanishes.* For this reason we will do without introducing notation for antisymmetrized rows and symmetrized columns as is part of the general algebra of the symmetric group [83, 95]. Furthermore, it is straightforward to derive the following commutation rules:

$$P\mathcal{S}_i = \mathcal{S}_i P = \begin{cases} \mathcal{S}_i \text{ for } P \in \text{row}(i) \\ \mathcal{S}_i \text{ for } P \notin \text{row}(i) \end{cases} \tag{8.30}$$

$$Q\mathcal{A}_i = \mathcal{A}_i Q = \begin{cases} -\mathcal{A}_i \text{ for } Q \in \text{col}(i) \\ +\mathcal{A}_i \text{ for } Q \notin \text{col}(i) \end{cases}. \tag{8.31}$$

We next introduce a new operator, the Young operator Y, symmetrizing all n_r rows and antisymmetrizing all n_c columns,

$$Y = \prod_{i=1}^{n_r} \mathcal{S}_i \prod_{j=1}^{n_c} \mathcal{A}_j.$$

Note that repeated application of Y conserves the symmetrization,

$$Y^2 = Y. \tag{8.32}$$

This *idempotence* makes the Young operator into a *projection operator.* It is the projection operator we were looking for because by applying Y to all Young tableaux of S_n we generate all possible forms of symmetrization, from fully symmetric to fully antisymmetric.

First, we consider two particles, with particle indices $1, 2$, and occupying states a, b. Depending on the tableau the Young operator reduces to either $Y = \mathcal{S}_1 = S_{12}$ or $Y = \mathcal{A}_1 = A_{12}$, where S_{12} and A_{12} are the symmetrization and antisymmetrization operators for states a and b, respectively.[1] Applying the Young operator we regain the well known symmetric and antisymmetric pair states,

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline\end{array}\, |ab) = S_{12}\, |ab) = (\mathbb{1} + P_{12})|ab) \tag{8.33a}$$

$$= |ab) + |ba) \tag{8.33b}$$

$$\begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline\end{array}\, |ab) = A_{12}\, |ab) = (\mathbb{1} - P_{12})|ab) \tag{8.33c}$$

$$= |ab) - |ba) \tag{8.33d}$$

$$= \begin{vmatrix} |a\rangle_1 |b\rangle_1 \\ |a\rangle_2 |b\rangle_2 \end{vmatrix}. \tag{8.33e}$$

So, it is the Young diagram determines the outcome of the symmetrization procedure. For two particles the result is familiar and we can do equally well without Young operators.

For three particles the added value of the Young operator already emerges as in this case the

---

[1] Here we adopt the convention of *permuting state indices* rather than particle indices.

fully-symmetric and fully-antisymmetric states are not the only options. ,

$$\begin{aligned}\begin{array}{|c|c|c|}\hline 1&2&3\\\hline\end{array}\;|abc) &= S_{123}\;|abc) = (\mathbb{1} + P_{12} + P_{13} + P_{23} + P_{(123)} + P_{(132)})\,|abc)\\ &= |abc) + |bac) + |cba) + |acb) + |cab) + |bca) = |abc\rangle_S \end{aligned}\tag{8.34a}$$

$$\begin{array}{|c|c}\hline 1&\multicolumn{1}{c|}{2}\\\hline 3\\\cline{1-1}\end{array}\;|abc) = A_{13}S_{12}\;|abc) = (\mathbb{1} - P_{13})(\mathbb{1} + P_{12})\,|abc) = (\mathbb{1} - P_{13} + P_{12} - P_{123})\,|abc)\tag{8.34b}$$

$$= |abc) - |cba) + |bac) - |cab) = |\psi_1\rangle\tag{8.34c}$$

$$\begin{array}{|c|c}\hline 1&\multicolumn{1}{c|}{3}\\\hline 2\\\cline{1-1}\end{array}\;|abc) = A_{12}S_{13}\;|abc) = (\mathbb{1} - P_{12})(\mathbb{1} + P_{13})\,|abc) = (\mathbb{1} - P_{12} + P_{13} - P_{132})\,|abc)\tag{8.34d}$$

$$= |abc) - |bac) + |cba) - |bca) = |\psi_2\rangle\tag{8.34e}$$

$$\begin{array}{|c|c}\hline 1&\multicolumn{1}{c|}{3}\\\hline 2\\\cline{1-1}\end{array}\;|acb) = A_{12}S_{13}\;|acb) = (\mathbb{1} - P_{12})(\mathbb{1} + P_{13})\,|acb) = (\mathbb{1} - P_{12} + P_{13} - P_{132})\,|acb)\tag{8.34f}$$

$$= |acb) - |cab) + |bca) - |cba) = |\psi_2\rangle\tag{8.34g}$$

$$\begin{aligned}\begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\hline\end{array}\;|abc) &= A_{123}\;|abc) = (\mathbb{1} - P_{12} - P_{13} - P_{23} + P_{(123)} + P_{(132)})\,|abc)\\ &= |abc) - |bac) - |cba) - |acb) + |cab) + |bca)\\ &= \begin{vmatrix} |a\rangle_1 & |b\rangle_1 & |c\rangle_1\\ |a\rangle_2 & |b\rangle_2 & |c\rangle_2\\ |a\rangle_3 & |b\rangle_3 & |c\rangle_3\end{vmatrix} = |abc\rangle_A\end{aligned}\tag{8.34h}$$

Again the Young diagram determines the outcome of the symmetrization procedure. For a product state of 3 particles, $|abc)$, can be symmetrized according to 3 different symmetry classes by application of the Young operator. Two of these classes are pure (fully symmetric and fully antisymmetric) and one is mixed. For the mixed class we identified two linearly independent symmetrizations. These are not orthogonal.

For more than three particles the number of symmetry classes and independent symmetrizations grows rapidly. For S_4 - see Eqs. (8.27) - we identify 5 symmetry classes, two of which are pure (fully symmetric and fully antisymmetric) and three are mixed (each with up to three linearly independent symmetrizations). As writing down all these options has little added value, they are left out in this text. They are best computer generated anyhow.

Before we address specific cases it is good to realize that often we have more particles than states. This implies symmetrization of identical states, which limits the number of symmetry options because (for identical states) the symmetrization procedure often leads to a vanishing result. We demonstrate these features for three particles in Section 8.2.4. In view of the importance of spin-$\frac{1}{2}$ particles we list in Table 8.2 the partitions of up to 10 particles over 2 states.

8.2.4 Application to three-electron states

In this section we apply the Young partitioning theory to atoms with three-electrons. We already found in that $\mathcal{H}$ commutes with $\mathbf{L}^2$, L_z, $\mathbf{S}^2$, S_z and P. As these operators also commute mutually they share a joint basis. It is goal to find this basis in the form of the spinorbitals satisfying the Pauli principle.

8.2.4.1 Spin sector

Let us start with the spin part. For three electrons we have to couple three spins. The coupling rules tell us that by coupling two spins $s = \frac{1}{2}$ we have the options $S = 0, 1$. By coupling the third

Table 8.2: The partitions of up to 10 particles over 2 states.

n	$P_2(n)$: partitions of n particles over 2 states
2	[2], [1,1]
3	[2,1], [1,1,1]
4	[2,2], [2,1,1], [1,1,1,1]
5	[2,2,1], [2,1,1,1], [1,1,1,1,1]
6	[2,2,2], [2,2,1,1], [2,1,1,1,1], [1,1,1,1,1,1]
7	[2,2,2,1], [2,2,1,1,1], [2,1,1,1,1,1], [1,1,1,1,1,1,1]
8	[2,2,2,2], [2,2,2,1,1], [2,2,1,1,1,1], [2,1,1,1,1,1,1,1], [1,1,1,1,1,1,1,1]
9	[2,2,2,2,1], [2,2,2,2,1], [2,2,2,1,1,1], [2,2,1,1,1,1,1,1], [2,1,1,1,1,1,1,1], [1,1,1,1,1,1,1,1,1]
10	[2,2,2,2,2], [2,2,2,2,1,1], [2,2,2,1,1,1,1], [2,2,1,1,1,1,1,1], [2,1,1,1,1,1,1,1,1], [1,1,1,1,1,1,1,1,1,1]

the total spin becomes $S = \frac{1}{2}(2\times)$, $\frac{3}{2}$. Note that $S = \frac{1}{2}$ appears twice, once from coupling $S = 0$ with $s = \frac{1}{2}$, and once from coupling $S = 1$ with $s = \frac{1}{2}$.

Next we derive expressions for these spin states in the uncoupled representation starting from the three symmetry classes given in Eqs. (8.26) and (8.34). Let us start with Young tableau 8.34a (one row),

$$\begin{array}{|c|c|c|}\hline 1&2&3\\\hline\end{array} \rightarrow \begin{cases} S_{123}\,|{\uparrow\uparrow\uparrow}\rangle \sim |{\uparrow\uparrow\uparrow}\rangle & \Leftrightarrow |\frac{3}{2},\frac{3}{2}\rangle \\ S_{123}\,|{\uparrow\uparrow\downarrow}\rangle \sim |{\uparrow\uparrow\downarrow}\rangle + |{\downarrow\uparrow\uparrow}\rangle + |{\uparrow\downarrow\uparrow}\rangle & \Leftrightarrow |\frac{3}{2},\frac{1}{2}\rangle \\ S_{123}\,|{\uparrow\downarrow\downarrow}\rangle \sim |{\uparrow\downarrow\downarrow}\rangle + |{\downarrow\uparrow\downarrow}\rangle + |{\downarrow\downarrow\uparrow}\rangle & \Leftrightarrow |\frac{3}{2},-\frac{1}{2}\rangle \\ S_{123}\,|{\downarrow\downarrow\downarrow}\rangle \sim |{\downarrow\downarrow\downarrow}\rangle & \Leftrightarrow |\frac{3}{2},-\frac{3}{2}\rangle. \end{cases} \tag{8.35}$$

Here we recognize the 4 eigenstates $|S, M_S\rangle$ of total spin $S = \frac{3}{2}$ (the stretched state of 3 spins). This need not be a surprise because the symmetrization operator (being a linear combination of permutations) leaves the total spin invariant (P commutes with S^2 and S_z). Therefore, S is an irreducible tensor operator (according to the definition in Appendix K); i.e., it acts in one of the irreducible subspaces of the total spin. We continue with the Young tableau (8.34b)

$$\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array} \rightarrow \begin{cases} A_{13}S_{12}\,|{\uparrow\uparrow\uparrow}\rangle \rightarrow 0 \\ A_{13}S_{12}\,|{\uparrow\uparrow\downarrow}\rangle \sim [|{\uparrow\downarrow}\rangle - |{\downarrow\uparrow}\rangle]_{13} \otimes |{\uparrow}\rangle_2 & \Leftrightarrow |\frac{1}{2},\frac{1}{2}\rangle \\ A_{13}S_{12}\,|{\uparrow\downarrow\downarrow}\rangle \sim [|{\uparrow\downarrow}\rangle - |{\downarrow\uparrow}\rangle]_{13} \otimes |{\downarrow}\rangle_2 & \Leftrightarrow |\frac{1}{2},-\frac{1}{2}\rangle \\ A_{13}S_{12}\,|{\downarrow\downarrow\downarrow}\rangle \rightarrow 0. \end{cases} \tag{8.36}$$

Note that for three identical spin states the Young tableau vanish, as they should because this case is incompatible with $S = \frac{1}{2}$. We continue with the Young tableau (8.34d), which belongs to the same symmetry class as the tableau (8.34b). Applying the Young operator we find

$$\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array} \rightarrow \begin{cases} A_{12}S_{13}\,|{\uparrow\uparrow\uparrow}\rangle \rightarrow 0 \\ A_{12}S_{13}\,|{\uparrow\uparrow\downarrow}\rangle \sim |{\downarrow\uparrow\uparrow}\rangle - |{\uparrow\downarrow\uparrow}\rangle = [|{\downarrow\uparrow}\rangle - |{\uparrow\downarrow}\rangle]_{12} \otimes |{\uparrow}\rangle_3 & \Leftrightarrow |\frac{1}{2},\frac{1}{2}\rangle \\ A_{12}S_{13}\,|{\uparrow\downarrow\downarrow}\rangle \sim |{\uparrow\downarrow\downarrow}\rangle - |{\downarrow\uparrow\downarrow}\rangle = [|{\uparrow\downarrow}\rangle - |{\downarrow\uparrow}\rangle]_{12} \otimes |{\downarrow}\rangle_3 & \Leftrightarrow |\frac{1}{2},-\frac{1}{2}\rangle \\ A_{12}S_{13}\,|{\downarrow\downarrow\downarrow}\rangle \rightarrow 0. \end{cases} \tag{8.37}$$

The Young tableau(8.34h) corresponds to the fully-antisymmetric case (determinantal state). This state has to be identically zero because with 3 spins, 2 are necessarily identical and the determinant vanishes; e.g., we find

$$\begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\hline\end{array} \rightarrow \begin{cases} A_{123}\,|{\uparrow\uparrow\uparrow}\rangle \rightarrow 0 \\ A_{123}\,|{\uparrow\uparrow\downarrow}\rangle \rightarrow 0 \\ A_{123}\,|{\uparrow\downarrow\downarrow}\rangle \rightarrow 0 \\ A_{1123}\,|{\downarrow\downarrow\downarrow}\rangle \rightarrow 0. \end{cases} \tag{8.38}$$

8.2.4.2 Orbital sector and product states

Turning to the orbital part we ask for three-orbital states to combine with the three-spin states in such a way that the total state is antisymmetric under interchange of any two electrons. In some cases this is simple. The stretched spin states (8.35) are fully symmetric and have to be combined with fully antisymmetric orbital states. In terms of young diagrams this corresponds to

$$\begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\hline\end{array}\otimes\begin{array}{|c|c|c|}\hline 1&2&3\\\hline\end{array}\,, \tag{8.39}$$

where the horizontal bar stands for the fully-symmetric spin state and the vertical bar for the fully-antisymmetric orbital state.

Partially symmetrized spin states of the types (8.36) and (8.37) have to be combined with partially symmetrized orbital states. More precisely, labeling three particles as $1,2,3$, the *horizontally* symmetrized part of the spin state of particles $1,2,3$ has to be combined with a *horizontally* antisymmetrized orbital states the particles $1,2,3$. Likewise for the vertical part. However, since horizontal symmetrization commutes with vertical antisymmetrization,

$$S_{13}A_{12}\;|abc)=A_{13}S_{12}\;|abc)\,,$$

rather than introducing a new Young operator for horizontal antisymmetrization and vertical symmetrization we stick to the old Young operator but transpose the Young tableau about its matrix diagonal. We then combine the *horizontally* symmetrized part of the spin tableau $\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}$ with the *vertically* antisymmetrized part of the orbital tableau $\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}$. For the partially symmetrized spin states (8.36) and (8.37) this leads to the following two types of product states

$$\overset{orbit}{\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}}\otimes\overset{spin}{\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}} \tag{8.40}$$

$$\overset{orbit}{\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}\end{array}}\otimes\overset{spin}{\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}\end{array}}\,. \tag{8.41}$$

Fully symmetrized orbital states do not have to be considered because the fully antisymmetrized spin states (8.38) vanish identically.

Let us apply the above to three equivalent p electrons in the orbitals $p_1, p_{0,}p_{-1}$. The only (nonvanishing) fully antisymmetric orbital state is the determinantal state

$$\begin{array}{|c|}\hline 1\\\hline 2\\\hline 3\\\hline\end{array}\rightarrow\left\{\begin{array}{l} A_{123}\;|1,0,-1)_{np^3}\sim|1,0,-1)-|0,-1,1)-|-1,1,0)-|1,-1,0)+|-1,1,0)+|0,-1,1)\\ \qquad\qquad\sim\begin{vmatrix}|p_1\rangle_1 & |p_0\rangle_1 & |p_{-1}\rangle_1\\ |p_1\rangle_2 & |p_0\rangle_2 & |p_{-1}\rangle_2\\ |p_1\rangle_3 & |p_0\rangle_3 & |p_{-1}\rangle_3\end{vmatrix}=|1,0,-1\rangle_A\Leftrightarrow|0,0\rangle_{np^3}.\end{array}\right. \tag{8.42}$$

Note that all other states of three p orbitals are of the type $|p_a p_a p_b)$ and vanish since $\mathcal{A}_1\;|p_a p_a p_b)_{np^3}=0$ for $a,b\in\{1,0,-1\}$. The partially symmetric orbital states are

$$\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \to \begin{cases} A_{12}S_{13}\ |1,1,0)_{132} & \sim [|1,0) - |0,1)]_{13} \otimes |1)_2 \\ A_{12}S_{13}\ |1,1,-1) & \sim [|1,-1) - |-1,1)]_{13} \otimes |1)_2 \\ A_{12}S_{13}\ |1,0,0) & \sim [|1,0) - |0,1)]_{13} \otimes |0)_2 \\ A_{12}S_{13}\ |1,0,-1) & \sim [|1,0) - |0,1)]_{13} \otimes |-1)_2 - [|0,-1) - |-1,0)]_{13} \otimes |1)_2 \\ A_{12}S_{13}\ |1,-1,-1) & \sim [|1,-1) - |-1,1)]_{13} \otimes |-1)_2 \\ A_{12}S_{13}\ |0,0,-1) & \sim [|0,-1) - |-1,0)]_{13} \otimes |0)_2 \\ A_{12}S_{13}\ |0,-1,-1) & \sim [|0,-1) - |-1,0)]_{13} \otimes |-1)_2 \,. \end{cases} \tag{8.43}$$

Note that the other three p orbitals are of the type $|p_a p_a p_a)$ and vanish since $A_{12}S_{13}\ |p_a p_a p_b)_{np^3} = 0$ for $a \in \{1, 0, -1\}$.

8.2.4.3 Product states and spin orbitals

The final orbital state of three equivalent p electrons three-body product states are also antisymmetric, as required for electrons.

$$\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \to \begin{cases} A_{12}S_{13}\ |1,1,0)_{132} & \sim |1,0\rangle_{13} \otimes |1\rangle_2 \\ A_{12}S_{13}\ |1,1,-1) & \sim |1,-1\rangle_{13} \otimes |1\rangle_2 \\ A_{12}S_{13}\ |1,0,0) & \sim |1,0\rangle_{13} \otimes |0\rangle_2 \\ A_{12}S_{13}\ |1,0,-1) & \sim |1,0\rangle_{13} \otimes |-1\rangle_3 - |0,-1\rangle_{12} \otimes |1\rangle_3 \\ A_{12}S_{13}\ |1,-1,-1) & \sim |1,-1\rangle_{13} \otimes |-1\rangle_2 \\ A_{12}S_{13}\ |0,0,-1) & \sim |0,-1\rangle_{13} \otimes |0\rangle_2 \\ A_{12}S_{13}\ |0,-1,-1) & \sim |0,-1\rangle_{13} \otimes |-1\rangle_2 \end{cases} \tag{8.44}$$

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \to \begin{cases} A_{13}S_{12}\ |\uparrow\uparrow\uparrow) \to 0 & \\ A_{13}S_{12}\ |\uparrow\uparrow\downarrow) \sim |\uparrow,\downarrow\rangle_{13} \otimes |\uparrow\rangle_2 & \Leftrightarrow |\tfrac{1}{2}, \tfrac{1}{2}\rangle \\ A_{13}S_{12}\ |\uparrow\downarrow\downarrow) \sim |\uparrow,\downarrow\rangle_{13} \otimes |\downarrow\rangle_2 & \Leftrightarrow |\tfrac{1}{2}, -\tfrac{1}{2}\rangle \\ A_{13}S_{12}\ |\downarrow\downarrow\downarrow) \to 0. & \end{cases} \tag{8.45}$$

$$\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \otimes \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \to \begin{cases} |1,0\rangle_{12} \otimes |1\rangle_3 \otimes [|\uparrow\uparrow\downarrow) - |\downarrow\uparrow\uparrow)] = \left|\dot{1},\dot{0}\right\rangle_{12} \otimes |\bar{1}\rangle_3 & \sim \\ |1,-1\rangle_{12} \otimes |1\rangle_3 \otimes [|\uparrow\uparrow\downarrow) - |\downarrow\uparrow\uparrow)] & \sim \\ |1,0\rangle_{12} \otimes |0\rangle_3 \otimes [|\uparrow\uparrow\downarrow) - |\downarrow\uparrow\uparrow)] & \sim \\ |1,0\rangle_{12} \otimes |-1\rangle_3 \otimes [|\uparrow\uparrow\downarrow) - |\downarrow\uparrow\uparrow)] - |0,-1\rangle_{12} \otimes |1\rangle_3 \otimes [|\uparrow\uparrow\downarrow) - |\downarrow\uparrow\uparrow)] & \sim \\ |1,-1\rangle_{12} \otimes |-1\rangle_3 \otimes [|\uparrow\uparrow\downarrow) - |\downarrow\uparrow\uparrow)] & \sim \\ |0,-1\rangle_{12} \otimes |0\rangle_3 \otimes [|\uparrow\uparrow\downarrow) - |\downarrow\uparrow\uparrow)] & \sim \\ |0,-1\rangle_{12} \otimes |-1\rangle_3 \otimes [|\uparrow\uparrow\downarrow) - |\downarrow\uparrow\uparrow)] & \sim \end{cases}$$

$$\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \otimes \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \rightarrow \left\{\begin{array}{ll}
[|1,0,1) - |0,1,1)] \otimes [|{\uparrow\uparrow\downarrow}) - |{\downarrow\uparrow\uparrow})] & \sim \big|\dot{\imath},\dot{0},\bar{1}\big) - \big|\dot{0},\dot{\imath},\bar{1}\big) - \big|\bar{1},\dot{0},\dot{\imath}\big) + \big|\bar{0},\dot{\imath},\dot{\imath}\big) \\
 & \\
 & \begin{vmatrix} \big|\dot{\imath}\big\rangle_1 & \big|\dot{0}\big\rangle_1 & |\bar{1}\rangle_1 \\ \big|\dot{\imath}\big\rangle_2 & \big|\dot{0}\big\rangle_2 & |p_{-1}\rangle_2 \\ \big|\dot{\imath}\big\rangle_3 & |p_0\rangle_3 & |\bar{1}\rangle_3 \end{vmatrix} \\
[|1,1,0) - |0,1,1)] \otimes [|{\uparrow\uparrow\downarrow}) - |{\downarrow\uparrow\uparrow})] & \sim \big|\dot{\imath},\dot{\imath},\bar{0}\big) - \big|\bar{1},\dot{\imath},\dot{0}\big) - \big|\dot{0},\dot{\imath},\bar{1}\big) + \big|\bar{0},\dot{\imath},\dot{\imath}\big) \\
 & \\
|1,-1,1) - |-1,1,1) \otimes [|{\uparrow\uparrow\downarrow}) - |{\downarrow\uparrow\uparrow})] & \sim \big|\dot{\imath},-1,1\big) - |-1,1,1) \\
|1,0,0) - |0,1,0) \otimes [|{\uparrow\uparrow\downarrow}) - |{\downarrow\uparrow\uparrow})] & \sim \big|\dot{\imath},\dot{0},0\big) - |0,1,0) \\
|1,0,-1) - |0,1,-1) - |0,-1,1) - |-1,0,1) & \sim \big|\dot{\imath},\dot{0},-1\big) - \big|\dot{\imath},1,-1\big) - \big|\dot{\imath},-1,1\big) - \big|-\dot{\imath},0,1\big) \\
|1,-1,-1) - |-1,1,-1) \otimes [|{\uparrow\uparrow\downarrow}) - |{\downarrow\uparrow\uparrow})] & \sim \big|\dot{\imath},-1,-1\big) - \big|-\dot{\imath},1,-1\big) \\
|0,-1,0) - |-1,0,0) \otimes [|{\uparrow\uparrow\downarrow}) - |{\downarrow\uparrow\uparrow})] & \sim \big|\dot{\imath},-1,0\big) - \big|-\dot{\imath},0,0\big) \\
|0,-1,-1) - |-1,0,-1) \otimes [|{\uparrow\uparrow\downarrow}) - |{\downarrow\uparrow\uparrow})] & \sim \big|\dot{\imath},-1,-1\big) - \big|-\dot{\imath},0,-1\big)
\end{array}\right. \tag{8.46}$$

$$\begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline\end{array} \otimes \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline\end{array} \rightarrow \left\{\begin{array}{ll}
|1,0,-1\rangle_A \otimes |{\uparrow\uparrow\uparrow}) = \big|\dot{\imath},\dot{0},-\dot{\imath}\big\rangle_A & \Leftrightarrow \big|{}^4S;0,\tfrac{3}{2}\big\rangle \\
|1,0,-1\rangle_A \otimes [|{\uparrow\uparrow\downarrow}) - |{\downarrow\uparrow\uparrow})] = \big|\dot{\imath},\dot{0},-\bar{1}\big\rangle_A - \big|\bar{1},\dot{0},-\dot{\imath}\big\rangle_A & \Leftrightarrow \big|{}^4S;0,\tfrac{1}{2}\big\rangle \\
|1,0,-1\rangle_A \otimes [|{\uparrow\downarrow\downarrow}) - |{\downarrow\downarrow\uparrow})] = \big|\dot{\imath},\bar{0},-\bar{1}\big\rangle_A - \big|\bar{1},\bar{0},-\dot{\imath}\big\rangle_A & \Leftrightarrow \big|{}^4S;0,-\tfrac{1}{2}\big\rangle \\
|1,0,-1\rangle_A \otimes |{\downarrow\downarrow\downarrow}) = |\bar{1},\bar{0},-\bar{1}\rangle_A & \Leftrightarrow \big|{}^4S;0,-\tfrac{3}{2}\big\rangle
\end{array}\right. \tag{8.47}$$

Note that the antisymmetry is obtained in terms of spinorbitals (but not in the form of a *single* determinant).

$$\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \otimes \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \rightarrow \left\{\begin{array}{l}
[|1,0,-1) - |0,1,-1) + |-1,0,1) - |0,-1,1)] \otimes [|{\uparrow\uparrow\downarrow}) - |{\downarrow\uparrow\uparrow})] = \big|\dot{\imath},\dot{0},-\bar{1}\big\rangle_A - \big|\bar{1},\dot{0},-\dot{\imath}\big\rangle_A \\
\big|\dot{\imath},\dot{0},-\bar{1}\big) - \big|\dot{0},\dot{\imath},-\bar{1}\big) + \big|-\dot{\imath},\dot{0},\bar{1}\big) - \big|\dot{0},-\dot{\imath},\bar{1}\big) + \big|\bar{1},\dot{0},-\dot{\imath}\big) - \big|\bar{0},\dot{\imath},-\dot{\imath}\big) + \big|-\bar{1},\dot{0},\dot{\imath}\big) - \big|\bar{0},-\dot{\imath},\dot{\imath}\big) = \big|\dot{\imath},\dot{0}, \\
\big|\dot{\imath},\dot{0},-\bar{1}\big) - \big|\dot{0},\dot{\imath},-\bar{1}\big) - \big|-\bar{1},\dot{0},\dot{\imath}\big) - \big|\dot{\imath},-\bar{1},\dot{0}\big) + \big|-\bar{1},\dot{\imath},\dot{0}\big) + \big|\dot{\imath},\dot{0},-\bar{1}\big) + 2\big|-\bar{1},\dot{0},\dot{\imath}\big) + \big|\dot{\imath},-\bar{1},\dot{0}\big) - \big|-\bar{1},
\end{array}\right.$$

8.2.5 Configuration interaction

However, although in many cases a single configuration provides a good Ansatz for the atomic structure, collisional correlations (for so far not accounted for by the mean field) give rise to mixing of the configurations. This is known as *configuration interaction.* Not rarely, this results in significant shifts of the single-configuration energy levels.

To elucidate the mixing of configurations we return to the N-electron Hamiltonian (8.1). Note that, in view of the mutual Coulomb repulsion, the orbital angular momenta of the individual electrons, $\boldsymbol{l}_1 \cdots \boldsymbol{l}_N$, do not commute with $\mathcal{H}$ (see Problem 6.3). This points to mixing of the state $|n,l,m_l\rangle$ with, in principle, all other states $|n',l,m_l\rangle$. The contain the mixing a bit we recall that the Coulomb repulsion can only mix states of given $L, M_L, S, M_S J, M_J$ and of the same parity because these are good quantum numbers of $\mathcal{H}$. Although the total orbital angular momentum is a conserved quantity, see Problem 8.1, this is not the case for t

Problem 8.3. Show that the orbital angular momentum $\boldsymbol{l}_i$ of an arbitrary electron from a multi-electron atom is not conserved under the Hamiltonian (8.1),

$$[\boldsymbol{l}_i^2, \mathcal{H}]\psi \neq 0 \quad \text{and} \qquad [l_{iz}, \mathcal{H}]\psi \neq 0.$$

8.3 Non-interacting electron atoms

A lot can be said about many-electron atoms even by ignoring the interaction between the electrons. For *non-interacting-electron* atoms only the Coulomb interaction with the nucleus remains and the atoms can be constructed using configurations of hydrogenic wavefunctions. The configuration energies are determined by the energies of the single-electron hydrogenic levels. The Pauli principle is included to assure that no orbital level is occupied more than twice. Many of the configurations will be degenerate but adding electrons in order of increasing orbital angular momentum the Z dependence of the configuration energy has the following order:

$$\begin{array}{ll}
[\mathrm{H}] = 1s & [\mathrm{He}] = 1s^2 \\
[\mathrm{Li}] = [\mathrm{He}]2s,\ [\mathrm{Be}] = [\mathrm{He}]2s^2,\ [\mathrm{Bo}] = [\mathrm{He}]2s^22p^1, \cdots, & [\mathrm{Ne}] = [\mathrm{He}]2s^22p^6 \\
[\mathrm{Na}] = [\mathrm{Ne}]3s,\ [\mathrm{Na}] = [\mathrm{Ne}]3s^2,\ [\mathrm{Al}] = [\mathrm{Ne}]3s^23p^1,\ \cdots, & [\mathrm{Ar}] = [\mathrm{Ne}]3s^23p^6\,.
\end{array}$$

This is in agreement with the periodic system of Mendeleev (see Table A.1) but at $Z = 18$ this agreement ends. At $Z = 19$, rather than the expected configuration

$$[\mathrm{Ne}]3s^23p^63d$$

we find in the periodic system [K] = [Ar]$4s$. In this respect only the first three rows of the periodic table show the character of non-interacting-electron atoms. Of course we could try to add fine structure corrections but such refinements makes no sense because even the dominant corrections (electrostatic repulsion and exchange) are not included.

Despite the obvious shortcomings, non-interacting electron atoms provide valuable information about the scaling of atomic properties with Z. Since the degeneracy of states of given principal quantum number is n^2 - see Eq. (2.34) - the total number of electrons in *closed shells*, N, can be expressed in terms of $n_{\max}$, which is the principal quantum number of the highest occupied shell,

$$N = \sum_{n=1}^{n_{\max}} 2n^2 = \tfrac{1}{3} n_{\max}(n_{\max}+1)(2n_{\max}+1). \tag{8.48}$$

For neutral atoms the number of electrons is equal to the number of protons in the nucleus,

$$N = Z. \tag{8.49}$$

Along similar lines we can obtain an expression for the total binding energy of *closed shell* atoms. Using Hartree atomic units we calculate

$$-\varepsilon = \sum_{n=1}^{n_{\max}} 2n^2 \frac{Z^2}{2n^2} = Z^2 n_{\max}. \tag{8.50}$$

Eliminating $n_{\max} = \varepsilon/Z^2$ from Eqs. (8.48) and (8.50) we can express ε as a function of N and Z (see Problem 8.4)

$$-\varepsilon/Z^2 = -\tfrac{1}{2} + (3N/2)^{1/3} + \tfrac{1}{12}(3N/2)^{-1/3} + \cdots\,. \tag{8.51}$$

Comparing this expression with a numerical inversion of Eq. (8.48) we find that approximating ε/Z^2 by the first three terms of the expansion is accurate to the ppm level for $N \geq 2$; i.e., the three-term approximation holds for all many-electron atoms or ions of the periodic system. We recall that Eq. (8.51) was derived for closed shell atoms but, as $n_{\max}$ is eliminated from the expressions, it also offers an interpolation procedure to interpolate for half-filled shells. For *neutral* atoms ($N = Z$) Eq. (8.51) becomes

$$-\varepsilon = -\tfrac{1}{2}Z^2 + c_{7/3}\,Z^{7/3} + c_{5/3}\,Z^{5/3} + \cdots, \tag{8.52}$$

where $c_{7/3} = (3/2)^{1/3} \approx 1.145$ and $c_{5/3} = (3/2)^{-1/3}/12 \approx 7.280 \times 10^{-2}$. For instance, for the ground state of ^{4}He we calculate $\epsilon = -4.00012$ Hartree, which is already an excellent approximation of the exact value (-4 Hartree for non-interacting electrons). It only gets better for larger Z.

Problem 8.4. Derive the expansion (8.51).

Solution. We first rewrite the sum (8.48) in the form

$$N = \tfrac{2}{3}(n_{\max} + \tfrac{1}{2})^3 - \tfrac{1}{6}(n_{\max} + \tfrac{1}{2}),$$

which is equivalent to $n_{\max} + \frac{1}{2} = (3N/2)^{1/3}\left[1 + \frac{1}{4}(n_{\max} + \frac{1}{2})/(3N/2)\right]^{1/3}$. Expanding the cubic root we obtain

$$n_{\max} = -\tfrac{1}{2} + (3N/2)^{1/3} + \tfrac{1}{12}(n_{\max} + \tfrac{1}{2})(3N/2)^{-2/3} + \cdots .$$

Iterating $n_{\max}$ we obtain the desired expression. □

8.4 Hartree method

In Section 6.2 we introduced the concept of the self-consistent mean-field to better describe the helium atom. The Hartree mean field is a generalization of this approach. The key idea is to minimize the electrostatic energy of the atom by systematically redistributing the electronic charge. It represents an early method of dealing with screening of the nuclear charge by the core electrons and was introduced by Hartree in 1928 [56, 57]. In this section we derive the Hartree equations with the aid of the *variational principle*, following the analysis by Fock and by Slater in 1930 [104, 43]. The Hartree method is quite powerful but has the shortcoming of not accounting for exchange. By using properly symmetrized many-electron wavefunctions this shortcoming can be repaired. This results in the Hartree-Fock theory and will be discussed in Section 8.5. In the present section we focus on the *Hartree method* because it offers a concise introduction into the structure of self-consistent mean field calculations without the added complexity of exchange.

To discuss the Hartree method we turn to (Hartree) atomic units by setting $m_e = e = \hbar = 1/4\pi\varepsilon_0 = 1$ in the Hamiltonian (8.1),

$$\mathcal{H} = \sum_{i=1}^{N} h_0^{(i)} + \tfrac{1}{2}\sum_{k,l=1}^{N}{}' \frac{1}{\rho_{kl}}. \tag{8.53}$$

Here we recognize the notation introduced for helium in which

$$h_0^{(i)} = -\tfrac{1}{2}\boldsymbol{\nabla}_i^2 - \frac{Z}{\rho_i} \tag{8.54}$$

represents the Schrödinger Hamiltonian of electron i in the bare field of the nucleus - see Eq. (6.9). Recalling the use of product wavefunctions for the description of the ground state of the helium atom we again look for ground state solutions of the Schrödinger equation in the form of a product state of single-electron orbitals. This is the approach of Hartree [56, 57]. In Dirac notation the product state is written as

$$|\psi_u) = |u_1, \cdots u_N) \equiv |u_1\rangle_1 \otimes |u_2\rangle_2 \otimes \cdots |u_N\rangle_N , \tag{8.55}$$

where particle 1 is in state u_1, particle 2 in state u_2, *etc.*; by convention the state of particle 1 always appears at the first (most left) position, the state of particle 2 always at the second position, *etc.*. Sometimes, when addressing a subset of the particles, we cannot use this implicit convention and attach the particle index and ordering explicitly,

$$|u_\kappa, u_\nu)_{i,j} \equiv |u_\kappa\rangle_i \otimes |u_\nu\rangle_j . \tag{8.56}$$

The indices κ and ν are called *state* indices, and i and j *particle* indices. In the product state (8.55) each electron is assigned a particular orbital in accordance with the electron configuration. Exchange effects are left out of consideration. This is of course an approximation but irrelevant for our goal of introducing the Hartree method. To first order in perturbation theory the energy of the N-electron state $|\psi_u)$ is given by

$$\varepsilon(u_1,\cdots u_N)=(\psi_u|\mathcal{H}|\psi_u)=\sum_{i=1}^{N}\langle u_i|h_0^{(i)}|u_i\rangle_i+\tfrac{1}{2}\sum_{k,l=1}^{N}{}'(u_ku_l|\frac{1}{\rho_{kl}}|u_ku_l)_{kl}. \tag{8.57}$$

Note that in this expression the particle index and state index coincide, a direct consequence of Eq. (8.55). Since the outcome of the integrals only depends on the orbitals involved we can drop the particle indices and write the expression in the form

$$\varepsilon(u_1,\cdots u_N)=\sum_{\kappa=1}^{N}\langle u_\kappa|h_0|u_\kappa\rangle+\tfrac{1}{2}\sum_{\kappa,\nu=1}^{N}{}'(u_\kappa u_\nu|\frac{1}{\rho_{12}}|u_\kappa u_\nu). \tag{8.58}$$

Applying the variational procedure of Appendix I.1.3 we require that the energy be *stationary* under arbitrary *norm-conserving* variations of any of the orbitals u_κ while keeping the other orbitals fixed [104, 43]. For this purpose we consider the energy $\varepsilon(u_1,\cdots u_N)$ as a functional of the electronic orbitals. For the orbital u_κ this functional is

$$H_\kappa(u_1,\cdots,u_N)=\langle u_\kappa|h_0|u_\kappa\rangle+\tfrac{1}{2}\sum_{\nu=1}^{N}{}'\left[(u_\kappa u_\nu|\frac{1}{\rho_{12}}|u_\kappa u_\nu)+(u_\nu u_\kappa|\frac{1}{\rho_{12}}|u_\nu u_\kappa)\right], \tag{8.59}$$

where the prime indicates $\nu\neq\kappa$. Using the equality $(u_\kappa u_\nu|1/\rho_{12}|u_\kappa u_\nu)=(u_\nu u_\kappa|1/\rho_{12}|u_\nu u_\kappa)$ the functional becomes

$$H_\kappa(u_1,\cdots,u_N)=\langle u_\kappa|h_0|u_\kappa\rangle+\sum_{\nu=1}^{N}{}'(u_\kappa u_\nu|\frac{1}{\rho_{12}}|u_\kappa u_\nu), \tag{8.60}$$

which represents the sum of the unscreened *one-electron* contribution plus the *pair* contribution of the Coulomb interaction between the u_κ electron and all other electrons. The total energy $\varepsilon(u_1,\cdots u_N)$ is stationary if the functionals $H_\kappa(u_1,\cdots,u_N)$ are simultaneously stationary under variation of the u_κ. The *optimized* values are denoted by

$$\varepsilon=\varepsilon(u_1,\cdots,u_N)|_{\rm opt}\quad\text{and}\quad\varepsilon_\kappa=H_\kappa(u_1,\cdots,u_N)|_{\rm opt}. \tag{8.61}$$

Importantly, the total energy is *not* simply sum of the $H_\kappa(u_1,\cdots,u_N)|_{\rm opt}$,

$$\varepsilon\neq\sum_{\kappa=1}^{N}\varepsilon_\kappa=\sum_{\kappa=1}^{N}\langle u_\kappa|h_0|u_\kappa\rangle+\sum_{\kappa,\nu=1}^{N}{}'(u_\kappa u_\nu|\frac{1}{\rho_{12}}|u_\kappa u_\nu), \tag{8.62}$$

because this would amount to double counting of the interactions. The proper expression is found by subtracting the interaction energy,

$$\varepsilon=\sum_{\kappa=1}^{N}\varepsilon_\kappa-\tfrac{1}{2}\sum_{\mu,\nu=1}^{N}{}'(u_\mu u_\nu|\frac{1}{\rho_{12}}|u_\mu u_\nu). \tag{8.63}$$

This expression also follows from a comparison of Eqs. (8.58) and (8.60).

Adopting the notation of Eq. (6.46a) we can write Eq. (8.63) in the form

$$\varepsilon=\sum_{\kappa=1}^{N}\varepsilon_\kappa-\tfrac{1}{2}\sum_{\mu,\nu=1}^{N}{}'\mathcal{J}(u_\kappa,u_\nu) \tag{8.64}$$

and the optimized form of Eq. (8.60) becomes

$$\varepsilon_\kappa = \langle u_\kappa | h_0 | u_\kappa \rangle + {\sum_{\nu=1}^{N}}' \mathcal{J}(u_\kappa, u_\nu). \tag{8.65}$$

Substituting this expression into Eq. (8.64) we regain Eq. (8.58) but now in the form

$$\varepsilon = \sum_{\kappa=1}^{N} \langle u_\kappa | h_0 | u_\kappa \rangle + \tfrac{1}{2} {\sum_{\kappa,\nu=1}^{N}}' \mathcal{J}(u_\kappa, u_\nu). \tag{8.66}$$

At this point we could use the procedure of Section 6.4 to evaluate the integrals $\mathcal{J}(u_\kappa, u_\nu)$ using hydrogenic wavefunctions and even obtain analytic expressions. However, there is no reason why hydrogenic orbitals would be the optimal choice in many-electron atoms. On the contrary. Therefore, we proceed with the derivation of the Hartree equations. First we write $H_\kappa(u_1, \cdots, u_N)$ as the expectation value of the one-electron Hamiltonian h_κ for an electron in the state $|u_\kappa\rangle$,

$$H_\kappa(u_1, \cdots, u_N) = \langle u_\kappa | h_\kappa | u_\kappa \rangle, \tag{8.67}$$

where

$$h_\kappa = -\tfrac{1}{2}\boldsymbol{\nabla}^2 - \frac{Z}{\rho} + J_\kappa(\rho). \tag{8.68}$$

Here $J_\kappa(\rho)$ represents the potential energy of screening at position ρ approximated by an average of the Coulomb energy over the orbitals of all *other* electrons (i.e., skipping the one in the orbital κ),

$$J_\kappa(\rho_1) = {\sum_{\nu=1}^{N}}' \langle u_\nu | \frac{1}{\rho_{12}} | u_\nu \rangle. \tag{8.69}$$

This is a generalization of the screening function (6.14) derived for the screening of one s electron by the other s electron in helium. The average amounts to a *local density approximation* in which the effects of *correlations* in the relative motion of the electrons are neglected. In view of this average the $U_\kappa(\rho)$ are called *mean fields* or *effective fields.*

$$\langle u_\kappa | J_\kappa | u_\kappa \rangle = {\sum_{\nu=1}^{N}}' \mathcal{J}(u_\kappa, u_\nu). \tag{8.70}$$

We are now prepared to determine the mean-field energy for the ground state with the aid of the variational principle. The lagrangian for the variation of H_κ under the constraint $\langle u_\kappa | u_\kappa \rangle = 1$ is given by

$$\mathcal{L}_\kappa(u_1, \cdots, u_N, \lambda_\kappa) = \langle u_\kappa | h_\kappa | u_\kappa \rangle + \lambda_\kappa \left[\langle u_\kappa | u_\kappa \rangle - 1 \right], \tag{8.71}$$

where λ_κ is the Lagrange multiplier. Note that in this lagrangian the orthogonality of the orbitals is *not* enforced by a constraint. As discussed in Appendix I.1.3, for hermitian operators it suffices to minimize the lagrangian (8.71) through variation of $\langle u_\kappa|$,

$$\delta \mathcal{L}_\kappa = \langle \delta u_\kappa | h_\kappa | u_\kappa \rangle + \lambda_\kappa \langle \delta u_\kappa | u_\kappa \rangle = 0. \tag{8.72}$$

With this procedure the orbital u_κ is found as the stationary solution corresponding to the eigenvalue $\varepsilon_\kappa = -\lambda_\kappa$ of the effective single-electron Schrödinger equation

$$h_\kappa | u_\kappa \rangle = \varepsilon_\kappa | u_\kappa \rangle. \tag{8.73}$$

This expression represents a set of N differential equations known as the *Hartree equations*, one for each orbital u_κ, and coupled through the mean field potentials. This result is quite intuitive. The

Schrödinger equation for electronic motion in the bare nuclear field (as used for hydrogenic atoms) is replaced by a set of Hartree equation for electrons moving in the *mean field* of the nucleus screened by the other electrons,

$$\left[-\tfrac{1}{2}\boldsymbol{\nabla}^2 - \frac{Z}{\rho} + J_\kappa(\rho)\right] u_\kappa(\boldsymbol{\rho}) = \varepsilon_\kappa u_\kappa(\boldsymbol{\rho}). \tag{8.74}$$

These equations have to be solved iteratively by numerical integration until self-consistency is reached between the orbitals u_κ and the corresponding screening functions 8.69. The energy ε_κ can be interpreted as the energy (in atomic units) required for the instant removal of an electron in the orbital $|u_\kappa\rangle$. In experiments the observed ionization energy will generally be different because the removal of one electron will also affect the screening of the other electrons. We return to the interpretation of ε_κ in Section 8.5.4 when discussing *Koopmans' theorem* in the context of the Hartree-Fock equations.

Example: Because the central field Hamiltonian is invariant under rotations, n, l and m_l remain good quantum numbers and, like in the case of hydrogen, the orbital wavefunctions must be of the general form

$$\psi_{nlm}(\boldsymbol{\rho}) = P_{nl}(\rho) Y_l^m(\hat{\mathbf{r}}). \tag{8.75}$$

The Hartree wavefunctions $P_{nl}(\rho)$ satisfy the radial wave equation

$$\left[\frac{1}{2}\left(-\frac{d^2}{d\rho^2} - \frac{2}{\rho}\frac{d}{d\rho} + \frac{l(l+1)}{\rho^2}\right) - \frac{Z}{\rho} + J_{nl}(\rho)\right] P_{nl}(\rho) = \varepsilon_{nl} P_{nl}(\rho) \tag{8.76}$$

and differ of course from the hydrogenic one because the screening distorts the electrostatic potential into a non-Coulomb form. The set of N equations (8.73) are called the *Hartree equations*, one for each electron. They have to be solved iteratively using numerical integration. Starting from the hydrogenic wavefunctions $\tilde{R}_{nl}(\rho)$ of the electron configuration we obtain an educated guess for the potential energy of screening $U_{\mathrm{scr}}(\rho)$ of the electron under consideration in the screened field of the nucleus. Then we can solve the Hartree equations and obtain new radial wavefunctions $P_{nl}(\rho)$ for the electrons and recalculate the screening integral. This procedure is repeated until self consistency is reached.

8.5 Hartree-Fock method

The Hartree method has been used a lot in the past in spite of the neglect of exchange. Today, its practical importance for atomic structure calculations has essentially vanished because the relative simplicity of the Hartree wavefunctions is no longer important from the computational point of view. Without much additional effort one can replace the Hartree wavefunctions by Hartree-Fock wavefunctions, which properly take into account the required antisymmetry of the many-electron wavefunction under interchange of two electrons.

8.5.1 Hamiltonian

To introduce the Hartree-Fock method we return to the Hartree Hamiltonian (8.53) for N electrons with mutual repulsion moving in the electrostatic field of the nucleus of charge Ze,

$$\mathcal{H} = \sum_{i=1}^{N} h_0^{(i)} + \tfrac{1}{2}\sum_{i,j=1}^{N}{}' \frac{1}{\rho_{ij}}. \tag{8.77}$$

We recall from Chapters 6 and 8 that the repulsion term can be interpreted as giving rise to a mostly centrally symmetric screening potential about the nucleus. Following up on this result we start the discussion by simply presuming that also in the present case central symmetry dominates the motion of the individual electrons. This enables us to use separation of variables, which implies that the individual electrons occupy spinorbitals of well-defined angular momentum (just as in the hydrogenic wavefunctions). Importantly, $\mathcal{H}$ is invariant under rotation of the spatial coordinates about the origin. Therefore, the *total* orbital angular momentum ($\mathbf{L}$) is a conserved quantity (i.e., L and M_L are good quantum numbers - see also Problem 6.1). Moreover, as $\mathcal{H}$ is spin independent, also the total spin ($\mathbf{S}$) is conserved and S and M_S are good quantum numbers. In other words, the Hamiltonian $\mathcal{H}$ is diagonal in the $|LM_L, SM_S\rangle$ representation. Later we shall analyze how the spin-orbit interaction breaks this symmetry.

Knowing the number of electrons we use the periodic system to guess the electron configuration. To assure orthogonality of the single-electron wavefunctions the N-electron state $|\psi_\alpha\rangle$ is constructed in the form of a Slater determinant

$$|\psi_\alpha\rangle = |\alpha_1, \cdots \alpha_N\rangle \tag{8.78}$$

of one-electron spinorbitals

$$|\alpha_\kappa\rangle = |n^\kappa l^\kappa m_l^\kappa s m_s^\kappa\rangle = |u_\kappa\rangle \otimes |s m_s^\kappa\rangle, \tag{8.79}$$

with $\kappa \in \{1, \cdots N\}$ being the *state index*. As the electron configuration of partially filled shells leaves freedom in choosing the magnetic quantum numbers m_l and m_s usually several Slater determinants qualify for our purpose and the ground state will generally be a linear combination of those.

The Hartree-Fock method is a mean-field method in which *a single* Slater determinant of spin orbitals, $|\psi_\alpha\rangle$, is used as a variational trial function to search for the electronic ground-state energy of many-electron atoms [104, 43]. In Hartree atomic units the energy of the state $|\psi_\alpha\rangle$ is given by

$$\varepsilon = \varepsilon(\alpha_1, \cdots \alpha_N) = \langle\psi_\alpha|\mathcal{H}|\psi_\alpha\rangle = \sum_{i=1}^{N} \langle\psi_\alpha|h_0^{(i)}|\psi_\alpha\rangle + \tfrac{1}{2} \sum_{i,j=1}^{N}{}' \langle\psi_\alpha|\frac{1}{\rho_{ij}}|\psi_\alpha\rangle. \tag{8.80}$$

This expression is valid to first order in perturbation theory. The Schrödinger Hamiltonians $h_0^{(i)}$ in the first summation are operators of the one-body diagonal type and the Coulomb repulsions $1/\rho_{ij}$ in the second summation are of the two-body type. To first order in perturbation theory we are only interested in the diagonal matrix elements of the latter. Thus, it follows with the aid of Eqs. (7.45) and (7.53) that the energy of the state $|\psi_\alpha\rangle$ can be written as summations over the state indices,

$$\varepsilon = \sum_{\kappa=1}^{N} \langle u_\kappa|h_0|u_\kappa\rangle + \tfrac{1}{2} \sum_{\kappa,\nu=1}^{N}{}' \left[(u_\kappa, u_\nu|\frac{1}{\rho_{12}}|u_\kappa, u_\nu) - (u_\kappa, u_\nu|\frac{1}{\rho_{12}}|u_\nu, u_\kappa)\, \delta_{m_s^\kappa, m_s^\nu} \right]. \tag{8.81}$$

Note that the spinorbitals are separated in their spin and orbital parts, $|\alpha_\kappa\rangle = |u_\kappa\rangle \otimes |s m_s^\kappa\rangle$. The last term of Eq. (8.81) is the exchange term and shows that the Coulomb repulsion is reduced for electrons in the same spin state because these cannot be at the same position; they avoid close proximity as a result of the quantum correlations of Fermi-Dirac statistics.

8.5.2 Configuration mixing

Importantly, the Coulomb repulsions $1/\rho_{ij}$ can also couple the Slater determinant $|\psi_\alpha\rangle$ of the configuration α to Slater determinants $|\psi_\beta\rangle$ of the configurations β as long as these configurations differ in not more than two spinorbitals. Thus, the configurations become coupled in accordance with Eq. (H.28) of second-order time-independent perturbation theory,

$$|\psi_n\rangle \simeq |\psi_\alpha\rangle + \sum_{\beta}{}' |\phi_\beta\rangle \frac{\langle\psi_\beta|\frac{1}{2}\sum_{i,j}' \rho_{ij}^{-1}\,|\psi_\alpha\rangle}{E_\alpha^0 - E_\beta^0}. \tag{8.82}$$

Fortunately, in many cases the configuration energies E_α^0 and E_β^0 differ sufficiently to render the second-order admixture negligible. Therefore we shall restrict the discussion to first order.

8.5.3 Hartree-Fock equations

To best approximate the energy of the ground state we search for the set $\{\alpha_\kappa\}$ of N spinorbitals that minimize $\langle\psi_\alpha|\mathcal{H}|\psi_\alpha\rangle$ under the constraints $\langle\alpha_\kappa|\alpha_\kappa\rangle = 1$ and $\langle\alpha_\kappa|\alpha_\nu\rangle = 0$ for $\kappa \neq \nu$; i.e., in addition to the conservation of normalization (as required for the Hartree equations) we also enforce the orthogonality of all orbitals in order to maintain the structure of the Slater determinants [104, 43]. For this purpose we consider the energy $\varepsilon(\alpha_1, \cdots \alpha_N)$, see Eqs. (8.80) and (8.81), as a functional of the spinorbitals $\alpha_1, \cdots \alpha_N$. In particular, for the spinorbital α_κ the functional dependence on the other spinorbitals (α_ν, with $\nu \neq \kappa$) is given by

$$H_\kappa(\alpha_1, \cdots \alpha_N) \equiv \langle u_\kappa|h_0|u_\kappa\rangle + \sum_{\nu=1}^{N}{}' \left[(u_\kappa, u_\nu|\frac{1}{\rho_{12}}|u_\kappa, u_\nu) - (u_\kappa, u_\nu|\frac{1}{\rho_{12}}|u_\nu, u_\kappa)\,\delta_{m_s^\kappa, m_s^\nu} \right], \qquad (8.83)$$

where the prime indicates $\nu \neq \kappa$. Repeating the procedure of Section 8.4 the double summation in Eq. (8.81) has been reduced to a single summation Eq. (8.83) and the factor 1/2 disappears; i.e., we sum over all *pairs of orbitals* involving one electron in the spinorbital $|\alpha_\kappa\rangle$. The result is the sum of the unscreened *one-electron* contribution plus the *pair* contribution of the Coulomb interaction between the electron in state $|\alpha_\kappa\rangle$ and all other electrons while taking into account the exchange correlations. Note that the latter only play a role for $m_s^\kappa = m_s^\nu$; i.e., among electrons in the same spin state (cf. Problem 7.5).

To optimize the state $|u_\kappa\rangle$ we construct the lagrangian for the variation of $H_\kappa(\alpha_1, \cdots \alpha_N)$ under the constraints $\langle\alpha_\kappa|\alpha_\kappa\rangle = \langle u_\kappa|u_\kappa\rangle = 1$ and $\langle\alpha_\nu|\alpha_\kappa\rangle = \langle u_\nu|u_\kappa\rangle\langle m_s^\nu|m_s^\kappa\rangle = 0$, where $u_\nu, u_\kappa \in \{u_1, \cdots u_N\}$,

$$\mathcal{L}_\kappa(\alpha_1, \cdots \alpha_N, \lambda_{\kappa 1}, \cdots \lambda_{\kappa N}) = H_\kappa(\alpha_1, \cdots \alpha_N) + \sum_\nu \lambda_{\kappa\nu}\left[\langle u_\nu|u_\kappa\rangle\langle m_s^\nu|m_s^\kappa\rangle - \delta_{\nu,\kappa}\right]. \qquad (8.84)$$

Here $\lambda_{\kappa\kappa}$ is the Lagrange multiplier for the normalization constraint $\langle u_\kappa|u_\kappa\rangle = 1$ and $\lambda_{\nu\kappa}$, with $\nu \neq \kappa$, is the Lagrange multiplier for the orthogonality constraint $\langle u_\nu|u_\kappa\rangle\langle m_s^\nu|m_s^\kappa\rangle = 0$. Furthermore, comparing Eq. (8.84) with its hermitian conjugate we find that $\lambda_{\nu\kappa} = \lambda_{\kappa\nu}^*$, which means that the matrix $(\lambda_{\nu\kappa})$ is hermitian. Applying the procedure of Appendix I.1.3 it suffices to minimize the lagrangian (8.84) with respect to $\langle u_\kappa|$,

$$\begin{aligned}\delta\mathcal{L}_\kappa = \langle\delta u_\kappa|h_0|u_\kappa\rangle & \\ + \sum_\nu{}' & \left[(\delta u_\kappa, u_\nu|\frac{1}{\rho_{12}}|u_\kappa, u_\nu) - (\delta u_\kappa, u_\nu|\frac{1}{\rho_{12}}|u_\nu, u_\kappa)\,\delta_{m_s^\kappa, m_s^\nu} \right] \\ & + \sum_\nu \lambda_{\kappa\nu}[\langle\delta u_\nu|u_\kappa\rangle\langle m_s^\nu|m_s^\kappa\rangle - \delta_{\nu,\kappa}]. \end{aligned} \qquad (8.85)$$

With this procedure the orbital u_κ is found as the solution of a set of N coupled integro-differential equations, the *Hartree-Fock equations*,

$$h_0 u_\kappa(\boldsymbol{\rho}_1) + \sum_\nu{}' \left[\langle u_\nu|\frac{1}{\rho_{12}}|u_\nu\rangle u_\kappa(\boldsymbol{\rho}_1) - \langle u_\nu|\frac{1}{\rho_{12}}|u_\kappa\rangle\,\delta_{m_s^\kappa, m_s^\nu}\,u_\nu(\boldsymbol{\rho}_1)\right] = -\sum_\nu \lambda_{\nu\kappa}\langle m_s^\nu|m_s^\kappa\rangle\,u_\kappa(\boldsymbol{\rho}_1). \qquad (8.86)$$

Just like the Hartree equations, the Hartree-Fock equations have to be solved iteratively until self-consistency is reached between the orbitals u_κ and their screening potentials. The solution is not unique because the determinant $|\psi_\alpha\rangle = |\alpha_1, \cdots \alpha_N\rangle$ is invariant under a unitary transformation U of the spinorbitals (see Problem 7.4). This freedom can be exploited to diagonalize the matrix $(\lambda_{\kappa\nu})$, which is always possible because it is hermitian. If we suppose that this goal is realized by the operator $\tilde{U}$ and results in the states $|\tilde{u}_\kappa\rangle = \tilde{U}|u_\kappa\rangle$, with $\kappa \in \{1, \cdots N\}$, Eq. (8.86) is replaced by

$$h_0\tilde{u}_\kappa(\boldsymbol{\rho}_1) + \sum_\nu{}' \left[\langle\tilde{u}_\nu|\frac{1}{\rho_{12}}|\tilde{u}_\nu\rangle\tilde{u}_\kappa(\boldsymbol{\rho}_1) - \langle\tilde{u}_\nu|\frac{1}{\rho_{12}}|\tilde{u}_\kappa\rangle\,\delta_{m_s^\kappa, m_s^\nu}\,\tilde{u}_\nu(\boldsymbol{\rho}_1)\right] = \varepsilon_\kappa\tilde{u}_\kappa(\boldsymbol{\rho}_1), \qquad (8.87)$$

where we defined $\varepsilon_\kappa \equiv -\lambda_\kappa$. Since $\langle\tilde{u}_\nu|\rho_{12}^{-1}|\tilde{u}_\nu\rangle = \langle u_\nu|\rho_{12}^{-1}|u_\nu\rangle$ this set of equations reduces to the Hartree equations (8.74) by omitting the exchange terms. Since we can drop the tilde in Eqs. (8.87) we shall continue from this point on (without loss of generality) by using the symbol u_κ also for the diagonalized equations. In hindsight we come to the conclusion that the orthogonality condition is not an intrinsic requirement of the Hartree-Fock method but could have been omitted by a clever choice of basis in the first place.

The total energy $\varepsilon(\alpha_1,\cdots\alpha_N)$ is stationary if the functionals $H_\kappa(\alpha_1,\cdots\alpha_N)$ are simultaneously stationary under variation of the u_κ. The *optimized* values are denoted by

$$\varepsilon = \varepsilon(\alpha_1,\cdots\alpha_N)|_{\text{opt}} \quad \text{and} \quad \varepsilon_\kappa = H_\kappa(\alpha_1,\cdots\alpha_N)|_{\text{opt}}. \tag{8.88}$$

Importantly, the total energy is *not* simply the sum of the $H_\kappa(\alpha_1,\cdots\alpha_N)|_{\text{opt}}$,

$$\varepsilon \neq \sum_{\kappa=1}^{N}\varepsilon_\kappa = \sum_{\kappa=1}^{N}\langle u_\kappa|h_0|u_\kappa\rangle + {\sum_{\kappa,\nu=1}^{N}}' \left[(u_\kappa,u_\nu|\frac{1}{\rho_{12}}|u_\kappa,u_\nu) - (u_\kappa,u_\nu|\frac{1}{\rho_{12}}|u_\nu,u_\kappa)\,\delta_{m_s^\kappa,m_s^\nu}\right], \tag{8.89}$$

because this would amount to double counting of the interactions. Hence, the proper expression is obtained by subtracting the interaction energy,

$$\varepsilon = \sum_{\kappa=1}^{N}\varepsilon_\kappa - \tfrac{1}{2}{\sum_{\kappa,\nu=1}^{N}}' \left[\mathcal{J}(u_\kappa,u_\nu) - \mathcal{K}(u_\kappa,u_\nu)\,\delta_{m_s^\kappa,m_s^\nu}\right]. \tag{8.90}$$

In this notation, the optimized form of Eq. (8.83) becomes

$$\varepsilon_\kappa = \langle u_\kappa|h_0|u_\kappa\rangle + {\sum_{\nu=1}^{N}}' \left[\mathcal{J}(u_\kappa,u_\nu) - \mathcal{K}(u_\kappa,u_\nu)\,\delta_{m_s^\kappa,m_s^\nu}\right]. \tag{8.91}$$

Substituting this expression into (8.90) we regain the form of Eq. (8.81) for the total energy,

$$\varepsilon = \sum_{\kappa=1}^{N}\langle u_\kappa|h_0|u_\kappa\rangle + \tfrac{1}{2}{\sum_{\kappa,\nu=1}^{N}}' \left[\mathcal{J}(u_\kappa,u_\nu) - \mathcal{K}(u_\kappa,u_\nu)\,\delta_{m_s^\kappa,m_s^\nu}\right], \tag{8.92}$$

but with the hydrogenic states replaced by Hartree-Fock-optimized orbitals. Recalling Problem 6.10 we know that (for equivalent electrons) the exchange integral is positive definite, $\mathcal{K}(u_\kappa,u_\nu) > 0$. Therefore, the energy is minimal if S is maximized, which is known as Hund's first rule (see Section 9.1). Phenomenologically the exchange terms manifest themselves as an interaction that tends to align the spins. This is called *ferromagnetic exchange*. It is much stronger than typical magnetic interactions as it originates from spin-dependent differences in electrostatic screening of the nuclear charge.

8.5.4 Koopmans' theorem

The energy ε_κ of an electron in the spinorbital $|\alpha_\kappa\rangle$ of a many-electron atom represents the total energy of this electron when taking into account the Coulomb field of the nucleus and the screening by all other electrons of the atom (in the mean field approximation). Note that this is the energy for *instant* removal of an electron from this state. This is equal to the *ionization* energy of the electron under the conjecture that rearrangement of the remaining charge (during removal of the electron) may be neglected,

$$\varepsilon_\kappa = (\varepsilon_{\text{atom}} - \varepsilon_{\text{ion}}) + \cdots . \tag{8.93}$$

This relation is known as *Koopmans' theorem* [65]. The theorem is best satisfied for ionization of inner core electrons in large atoms.

Koopmans' theorem is important because it enables a comparison between Hartree-Fock theory and experiment. To obtain the theorem in a more formal way we consider the total energy ε_N of an N-electron atom. The energy of the ion created by removal of one electron is denoted by ε_{N-1}. We presume the central field approximation to be well justified in both cases. In the Hartree-Fock approximation the energies of atom and ion can be expressed, respectively, as

$$\varepsilon_N = \sum_{\mu=1}^{N} \langle u_\mu | h_0 | u_\mu \rangle + \tfrac{1}{2} \sum_{\mu,\nu=1}^{N}{}' \left[\mathcal{J}(u_\mu, u_\nu) - \mathcal{K}(u_\mu, u_\nu) \delta_{m_s^\mu, m_s^\nu} \right] \tag{8.94}$$

$$\varepsilon_{N-1} = \sum_{\mu=1}^{N-1} \langle u_\mu | h_0 | u_\mu \rangle + \tfrac{1}{2} \sum_{\mu,\nu=1}^{N-1}{}' \left[\mathcal{J}(u_\mu, u_\nu) - \mathcal{K}(u_\mu, u_\nu) \delta_{m_s^\mu, m_s^\nu} \right]. \tag{8.95}$$

In this notation it is the electron in state $|u_N\rangle$ that is removed. Subtracting the two expressions we obtain

$$\begin{aligned} \varepsilon_N - \varepsilon_{N-1} = \langle u_N | h_0 | u_N \rangle &+ \tfrac{1}{2} \sum_{\nu=1}^{N-1} \left[\mathcal{J}(u_N, u_\nu) - \mathcal{K}(u_N, u_\nu) \delta_{m_s^N, m_s^\nu} \right] \\ &+ \tfrac{1}{2} \sum_{\mu=1}^{N-1} \left[\mathcal{J}(u_\mu, u_N) - \mathcal{K}(u_\mu, u_N) \delta_{m_s^N, m_s^\nu} \right]. \end{aligned} \tag{8.96}$$

This result is obtained without loss of generality because any of the occupied orbitals $|u_\kappa\rangle$, with $\kappa \in \{1, \cdots N\}$, can be labeled $|u_N\rangle$. However, in writing these expressions we assume that the Coulomb integrals between two orbitals u_μ and u_ν are not affected by the removal of an electron from the orbital u_κ. This is known as the frozen *frozen orbital* approximation. It is only partially satisfied because the removal affects the nuclear screening for all other electrons and therefore the radial distributions of their orbitals. Using the properties of the Coulomb integrals the difference becomes - see Eq. (6.47)

$$\varepsilon_N - \varepsilon_{N-1} = \langle u_\kappa | h_0 | u_\kappa \rangle + \sum_{\nu=1}^{N-1} \left[\mathcal{J}(u_\kappa, u_\nu) - \mathcal{K}(u_\kappa, u_\nu) \delta_{m_s^N, m_s^\nu} \right] = \varepsilon_\kappa, \tag{8.97}$$

which confirms the theorem for the case of instant removal.

8.5.5 Fock operators - direct and exchange contributions

The notation of the Hartree-Fock equations can be simplified to the form

$$\left[-\tfrac{1}{2} \boldsymbol{\nabla}^2 - \frac{Z}{\rho} + J_\kappa(\rho) - K_\kappa(\rho) \right] u_\kappa(\boldsymbol{\rho}) = \varepsilon_\kappa u_\kappa(\boldsymbol{\rho}). \tag{8.98}$$

With this expression we regained the familiar form of the Hartree equations but have the advantage that the Pauli principle is implicitly satisfied. The expression has the appearance of a set of Schrödinger equations for electrons in the orbitals u_κ but formally this is not the case because the operators are *not* one-electron Hamiltonians but give rise to coupling between the equations. By comparison with Eq. (8.87) we find that the operators $J_\kappa(\rho)$ and $K_\kappa(\rho)$ are defined by the expressions

$$J_\kappa(\rho) |u_\kappa\rangle = |u_\kappa\rangle \sum_\nu{}' \langle u_\nu | \frac{1}{\rho_{12}} | u_\nu \rangle \tag{8.99a}$$

$$K_\kappa(\rho) |u_\kappa\rangle = \sum_\nu{}' |u_\nu\rangle \langle u_\nu | \frac{1}{\rho_{12}} | u_\kappa \rangle \delta_{m_s^\kappa, m_s^\nu}. \tag{8.99b}$$

The operators $J_\kappa(\rho)$ and $K_\kappa(\rho)$ represent effective operators for, respectively, the direct and exchange contributions to the Coulomb repulsion energy of an electron in the orbital u_κ. Importantly, only electrons in the same spin state as the electron under consideration contribute to the exchange. In Dirac notation the Hartree-Fock equations takes the form

$$F_\kappa|u_\kappa\rangle = \varepsilon_\kappa|u_\kappa\rangle, \tag{8.100}$$

where $F_\kappa \equiv h_0 + J_\kappa(\rho) - K_\kappa(\rho)$. The operators F_κ are called *Fock operators.* Their expectation value, $\varepsilon_\kappa = \langle u_\kappa|F_\kappa|u_\kappa\rangle$, is given by Eq. (8.91).

Hartree-Fock-Slater approximation

Slater discovered an approximation of the Hartree-Fock method that is particularly valuable to gain physical insight in the meaning of the exchange terms [105]. For two electrons in the same state ($\kappa = \nu$) the difference of the direct and the exchange terms vanishes,

$$(u_\kappa, u_\nu|\frac{1}{\rho_{12}}|u_\kappa, u_\nu) - (u_\kappa, u_\nu|\frac{1}{\rho_{12}}|u_\nu, u_\kappa)\,\delta_{m_s^\kappa, m_s^\nu} = 0 \quad \text{for } \kappa = \nu. \tag{8.101}$$

This has the interesting consequence that we can drop the prime in the derivation of the Hartree-Fock equations as the omitted term yields zero anyhow. In the position representation Eqs. (8.99) become

$$J_\kappa(\rho)u_\kappa(\boldsymbol{\rho}_1) = u_\kappa(\boldsymbol{\rho}_1)\sum_\nu \langle u_\nu|\frac{1}{\rho_{12}}|u_\nu\rangle = u_\kappa(\boldsymbol{\rho}_1)U_D(\boldsymbol{\rho}_1) \tag{8.102a}$$

$$K_\kappa(\rho)u_\kappa(\boldsymbol{\rho}_1) = \sum_\nu u_\nu(\boldsymbol{\rho}_1)\langle u_\nu|\frac{1}{\rho_{12}}|u_\kappa\rangle\delta_{m_s^\kappa, m_s^\nu} = u_\kappa(\boldsymbol{\rho}_1)U_X(\boldsymbol{\rho}_1). \tag{8.102b}$$

where $U_D(\boldsymbol{\rho}_1)$ represents an effective interaction potential, called the *direct screening potential*,

$$U_D(\boldsymbol{\rho}_1) = \sum_\nu \int \frac{1}{\rho_{12}} u_\nu^*(\boldsymbol{\rho}_2)u_\nu(\boldsymbol{\rho}_2)d\boldsymbol{\rho}_2 \tag{8.103}$$

and $U_X(\boldsymbol{\rho})$ is another effective interaction potential, known as the *exchange interaction*

$$U_X(\boldsymbol{\rho}_1) = \sum_\nu \delta_{m_s^\kappa, m_s^\nu} \frac{u_\kappa^*(\boldsymbol{\rho}_1)u_\nu(\boldsymbol{\rho}_1)}{|u_\kappa(\boldsymbol{\rho}_1)|^2} \int \frac{1}{\rho_{12}} u_\nu^*(\boldsymbol{\rho}_2)u_\kappa(\boldsymbol{\rho}_2)d\boldsymbol{\rho}_2. \tag{8.104}$$

In terms of these potentials the Hartree-Fock equations turn into the *Hartree-Fock-Slater equations*,

$$\left[-\tfrac{1}{2}\boldsymbol{\nabla}^2 - \frac{Z}{\rho} + U_D(\boldsymbol{\rho}) - U_X(\boldsymbol{\rho})\right] u_\kappa(\boldsymbol{\rho}) = \varepsilon_\kappa u_\kappa(\boldsymbol{\rho}), \tag{8.105}$$

where both $U_D(\boldsymbol{\rho})$ and $U_X(\boldsymbol{\rho})$ vanish at large distance from the nucleus. With this expression we regained the familiar form of the Hartree equations but with the advantage that the Pauli principle is implicitly satisfied.

To elucidate the significance of the effective energy potentials $U_D(\boldsymbol{\rho})$ and $U_X(\boldsymbol{\rho})$ we start by noting that the number density of electrons at position $\mathbf{r}$ is obtained by summing over the probability densities of all occupied orbitals at position $\mathbf{r}$,

$$n(\mathbf{r}) = \sum_\nu |u_\nu(\mathbf{r})|^2. \tag{8.106}$$

This is the *total electron density* at position $\mathbf{r}$. Integrating over this distribution we obtain the total number of electrons, $\int n(\mathbf{r})d\mathbf{r} = N$. Thus, the *direct* contribution to the potential energy of an electronic charge at position $\boldsymbol{\rho}_1$ in the Coulomb field of *all* electrons (including the one under consideration; i.e., the one in the orbital u_κ) can be written as

$$U_D(\boldsymbol{\rho}_1) = \sum_\nu \langle u_\nu | \frac{1}{\rho_{12}} | u_\nu \rangle = \int \frac{1}{\rho_{12}} n(\boldsymbol{\rho}_2) d\boldsymbol{\rho}_2. \tag{8.107}$$

Clearly, by summing over all orbitals we overestimate the Coulomb repulsion because we include the repulsion of the probe electron in the state $|u_\kappa\rangle$ with its own mean field. This is called the *self energy* contribution to $U_D(\boldsymbol{\rho}_1)$ and is of course nonphysical.

Next we show that the mentioned overestimate is exactly compensated by the exchange terms. For this purpose we start from the orthonormality of the electronic orbitals, $\int u_\nu^*(\mathbf{r}) u_\kappa(\mathbf{r}) d\mathbf{r} = \delta_{\nu,\kappa}$. Summing over *all* occupied orbitals (including the one under consideration) we find exactly unity,

$$\sum_\nu \int u_\nu^*(\mathbf{r}) u_\kappa(\mathbf{r}) d\mathbf{r} = 1. \tag{8.108}$$

This is not surprising because all orbitals with $\nu \neq \kappa$ are orthogonal. The expression (8.108) suggest to switch summation and integration and to introduce a quantity which is called the *electron exchange density* at position $\mathbf{r}_2$ for electrons in the spinorbital α_κ and at position $\mathbf{r}_1$,

$$n_{ex}(\mathbf{r}_1, \mathbf{r}_2) = \sum_\nu \frac{u_\nu^*(\mathbf{r}_2) u_\kappa^*(\mathbf{r}_1) u_\kappa(\mathbf{r}_2) u_\nu(\mathbf{r}_1)}{|u_\kappa(\mathbf{r}_1)|^2} \delta_{m_s^\kappa, m_s^\nu}. \tag{8.109}$$

For electrons in different spin states ($m_s^\kappa \neq m_s^\nu$) the exchange density vanishes, $n_{ex}(\mathbf{r}_1, \mathbf{r}_2) \equiv 0$. For electrons in the same spin state ($m_s^\kappa = m_s^\nu$) the exchange density is

$$n_{ex}(\mathbf{r}, \mathbf{r}) = \sum_\nu u_\nu^*(\mathbf{r}) u_\nu(\mathbf{r}) \delta_{m_s^\kappa, m_s^\nu} = n_{\uparrow\uparrow}(\mathbf{r}). \tag{8.110}$$

Note that $n_{\uparrow\uparrow}(\mathbf{r})/n(\mathbf{r})$ is the fraction of the electron density with the same spin and at the same position $\mathbf{r}$ as a probe electron in the spinorbital under consideration, $\alpha_\kappa(\mathbf{r})$. In view of Eq. (8.108) the exchange density satisfies the property

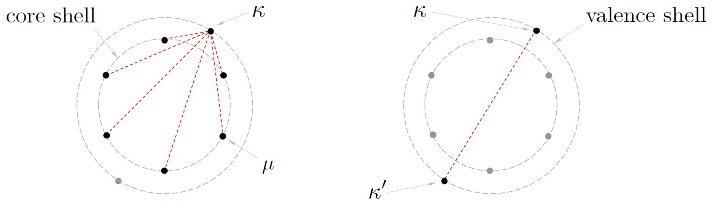
$$\int n_{ex}(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_2 = \delta_{m_s^\kappa, m_s^\nu} \frac{u_\kappa^*(\mathbf{r}_1) u_\nu(\mathbf{r}_1)}{|u_\kappa(\mathbf{r}_1)|^2} \sum_\nu \int u_\nu^*(\mathbf{r}_2) u_\kappa(\mathbf{r}_2) d\mathbf{r}_2 = \delta_{m_s^\kappa, m_s^\nu}. \tag{8.111}$$

For $\nu \neq \kappa$ the integral over $\mathbf{r}_2$ yields zero; for $\nu = \kappa$ the integral yields unity and the prefactor equals 1 provided the two electrons are in the same spin state. In other words, with this integral over the position of the second electron (which can be in *any* of the occupied states - including the one under consideration) we extract in total exactly one electron with spin parallel to that of the spinorbital under consideration). This is as it should be because two electrons in the same spin state cannot be found at the same position. Thus, in Slater's picture each electron of the atom is surrounded by an *exchange hole* (also called *Fermi hole*) in which electrons in the same spin state are excluded. The exchange contribution to the potential energy of an electronic charge at position $\mathbf{r}_1$ in the Coulomb field of *all* electrons can be written as

$$U_X(\boldsymbol{\rho}_1) = \sum_\nu \langle u_\nu | \frac{1}{\rho_{12}} | u_\kappa \rangle \frac{u_\kappa^*(\boldsymbol{\rho}_1) u_\nu(\boldsymbol{\rho}_1)}{|u_\kappa(\boldsymbol{\rho}_1)|^2} \delta_{m_s^\kappa, m_s^\nu} = \int \frac{1}{\rho_{12}} n_{ex}(\boldsymbol{\rho}_1, \boldsymbol{\rho}_2) d\boldsymbol{\rho}_2. \tag{8.112}$$

8.5.6 Energy functionals for valence electrons

In this section we have a special look at the valence electrons; i.e., the electrons in one or more *partially filled* shells $\{\alpha_\kappa\}$. Let α_κ be the spinorbital of one of these valence electrons. This electron



(a) (b)

Figure 8.1: Schematic diagram of the separation into core and valence electrons: (a) *inter-shell* interaction between the valence orbital $|\kappa\rangle$ and the core orbitals $\{\mu\}$; (b) *intra-shell* interaction between the valence orbitals $|\kappa\rangle$ and $|\kappa'\rangle$ in a partially filled shell.

will interact with the other valence electrons $\alpha_{\kappa'}$ as well as with the core electrons $\{\alpha_\mu\}$; i.e., the energy functional (8.83) can be written in the form (see Fig. 8.1)

$$H_\kappa(u_1,\cdots u_N) \equiv \langle u_\kappa|h_0|u_\kappa\rangle + \sum_{\mu}^{\text{core}} E(\alpha_\kappa,\alpha_\mu) + \sum_{\kappa'}^{\text{valence}}{}' E(\alpha_\kappa,\alpha_{\kappa'}). \tag{8.113}$$

Here $E(\alpha_\kappa,\alpha_\mu)$ is the Coulomb repulsion energy between the spinorbitals α_κ and α_μ. In terms of the direct and exchange integrals (6.46) this energy becomes

$$E(\alpha_\kappa,\alpha_\mu) = \left[\mathcal{J}(u_\kappa,u_\mu) - \mathcal{K}(u_\kappa,u_\mu)\delta_{m_s^\kappa,m_s^\mu}\right]. \tag{8.114}$$

Let us calculate the contribution to the core summation of *a single* closed shell, the shell nl (see Fig. 8.1a),

$$E_{nl^{2(2l+1)}}(\alpha_\kappa) \equiv \sum_{\mu}^{nl\,\text{shell}} E(\alpha_\kappa,\alpha_\mu). \tag{8.115}$$

The spinorbitals of this shell are $|\alpha_\mu\rangle = |nlm_lsm_s\rangle$, with $-l \leq m_l \leq l$ and $m_s = \pm\frac{1}{2}$, so the summation can be written as

$$E_{nl^{2(2l+1)}}(\alpha_\kappa) = \sum_{m_s=-1/2}^{1/2}\sum_{m_l=-l}^{l}\left[\mathcal{J}(nlm_l,n^\kappa l^\kappa m_{l^\kappa}) - \mathcal{K}(nlm_l,n^\kappa l^\kappa m_{l^\kappa})\delta_{m_s m_s^\kappa}\right] \tag{8.116}$$

$$= \sum_{m_l=-l}^{l}\left[2\mathcal{J}(nlm_l,n^\kappa l^\kappa m_{l^\kappa}) - \mathcal{K}(nlm_l,n^\kappa l^\kappa m_{l^\kappa})\right]. \tag{8.117}$$

With the aid of Eqs. (6.59) the Coulomb integrals $\mathcal{J}$ and $\mathcal{K}$ can be expressed in terms of the angular integrals a^k and b^k and the radial integrals F^k and G^k. Using the properties of a^k and b^k (see Problems 6.5, 6.6 and 6.7) the summation over m_l can be evaluated and turns out to be *independent* of m_l and m_s,

$$E_{nl^{2(2l+1)}}(\alpha_\kappa) = \sum_k\sum_{m_l=-l}^{l}\left[2a^k(lm_l,l^\kappa m_{l^\kappa})F^k(nl,n^\kappa l^\kappa) - b^k(m_l,m_{l^\kappa})G^k(nl,n^\kappa l^\kappa)\right]$$

$$= \sum_k\left[2(2l+1)F^k(nl,n^\kappa l^\kappa) - (2l+1)\begin{pmatrix} l & k & l^\kappa \\ 0 & 0 & 0\end{pmatrix}^2 G^k(nl,n^\kappa l^\kappa)\right]. \tag{8.118}$$

This expression can be written in the form of a potential energy integral using Eqs. (6.64a),

$$E_{nl^{2(2l+1)}}(\alpha_\kappa) = \int_0^\infty U_{nl}^{(\kappa)}(\rho)\left[\tilde{R}_{n^\kappa l^\kappa}(\rho)\right]^2\rho^2 d\rho, \tag{8.119}$$

where

$$U_{nl}^{(\kappa)}(\rho) \equiv \sum_{k} \left[2(2l+1)U_F^k(\rho) - (2l+1) \begin{pmatrix} l & k & l^\kappa \\ 0 & 0 & 0 \end{pmatrix}^2 \frac{\tilde{R}_{nl}(\rho)U_G^k(\rho)\tilde{R}_{n^\kappa l^\kappa}(\rho)}{\tilde{R}_{nl}(\rho)\tilde{R}_{nl}(\rho)} \right] \tag{8.120}$$

is the screening potential of the shell nl for valence electrons in the shell $n^\kappa l^\kappa$.

Repeating the above procedure for all shells of the core we write the energy functional

$$H_\kappa(u_1, \cdots u_N) \equiv \sum_{nl} \langle u_\kappa | h_0 + U_{nl}^{(\kappa)}(\rho) | u_\kappa \rangle + \sum_{\kappa'}^{\text{valence}} {}' E(\alpha_\kappa, \alpha_{\kappa'}), \tag{8.121}$$

where the sum runs over all closed shells of the atom.

An important consequence of the screening by the core being independent of the magnetic quantum numbers m_{l^κ} and m_{s^κ} of the valence electrons is that only the interactions *within* a valence shell determine which valence state has the lowest energy. The core electrons affect the binding of the valence electrons through screening of the nuclear charge and this can favor one valence shell over another but among equivalent electrons the core has no influence on the actual valence state with the lowest energy. The latter (i.e., the ground state) is determined by the Coulomb repulsion between the valence electrons (see Fig. 8.1*b*). The contribution of the valence electrons to the binding energy of the atom can be written in the form

$$\varepsilon_{val} = \sum_{\kappa} I(\alpha_\kappa) + \sum_{\kappa,\mu} E(\alpha_\kappa, \alpha_\mu) + \tfrac{1}{2} \sum_{\kappa,\kappa'}{}' E(\alpha_\kappa, \alpha_{\kappa'}), \tag{8.122}$$

where $I(\alpha_\kappa) = \langle u_\kappa | h_0 | u_\kappa \rangle$ is the hydrogenic energy of the valence electron in spinorbital α_μ and the summations represent the sum over all valence (κ) and core (μ) electrons. The core contribution to the binding energy is

$$\varepsilon_{core} = \sum_{\mu} I(\alpha_\mu) + \tfrac{1}{2} \sum_{\mu,\mu'}{}' E(\alpha_\mu, \alpha_{\mu'}). \tag{8.123}$$

By summing over the angular integrals the Coulomb repulsion energy on entire configurations can be expressed in terms of a small set of Coulomb integrals, which can be evaluated numerically starting from the Hartree-Fock solutions. Some of these compact expressions are presented in Table 8.3. Once the Coulomb integrals are known it is straightforward to calculate the binding energy of the elements. For a one-electron atom (*boron*) this is demonstrated in Problem 8.10.

8.5.7 Special cases

8.5.7.1 Closed shells

The total electrostatic repulsion energy of the electrons of a closed nl shell is obtained by adding the contributions of all electrons from the shell, $2l+1$ with spin up and $2l+1$ with spin down. While avoiding double counting, the contribution of each of the constituent electrons is found by summing over the Coulomb repulsion by “the other” $2(2l+1)-1$ electrons in the shell. In this process we have to consider the 4 possible spin combinations of spinorbitals $|n\dot{l}m; n\dot{l}m')$, $|n\dot{l}m; n\bar{l}m')$, $|n\bar{l}m; n\dot{l}m')$ and $|n\bar{l}m; n\bar{l}m')$. The corresponding spin pair repulsion energies are given by

$$\begin{aligned} E(n\dot{l}m, n\bar{l}m') = E(n\bar{l}m, n\dot{l}m') &= \mathcal{J}(nlm, nlm') \\ &= \sum_{k=0}^{\infty} a^k(lm, lm') F^k(nl^2) \end{aligned} \tag{8.124a}$$

$$\begin{aligned} E(n\dot{l}m, n\dot{l}m') = E(n\bar{l}m, n\bar{l}m') &= \mathcal{J}(nlm, nlm') - \mathcal{K}(nlm, nlm') \\ &= \sum_{k=0}^{\infty} \left[a^k(lm, lm') - b^k(lm, lm') \right] F^k(nl^2), \end{aligned} \tag{8.124b}$$

where $\mathcal{J}$ and $\mathcal{K}$ are the direct and exchange integrals given by Eqs. (6.59a) and (6.59b), respectively. Note that we used the property $F^k(nl^2) \equiv F^k(nl;nl) = G^k(nl;nl)$ as we are dealing with equivalent electrons. In principle, for like spins, we require $m \neq m'$ to satisfy the Pauli principle. However, since the direct and exchange coefficients coincide for electrons in the same state, $a^k(lm,lm) = b^k(lm,lm)$, we can drop this condition, $a^k(lm,lm') - b^k(lm,lm') = 0$ for $m = m'$. Hence, we are free to include the case $m = m'$ without affecting the core summation,

$$\begin{aligned}E_{nl^{2(2l+1)}} &= \tfrac{1}{2} \sum_{m,m'=-1}^{1} {}' E(n\dot{l}m, n\dot{l}m') + E(n\dot{l}m, n\bar{l}m') + E(n\bar{l}m, n\dot{l}m') + E(n\bar{l}m, n\bar{l}m') \\ &= \sum_{m,m'=-1}^{1} E(n\dot{l}m, n\bar{l}m') + E(n\dot{l}m, n\dot{l}m'). \end{aligned} \tag{8.125}$$

Substituting Eqs. (8.124) this becomes

$$E_{nl^{2(2l+1)}} = (2l+1)\,[2(2l+1)-1]\,F^0(nl^2) + \sum_{k=2}^{\infty} \sum_{m,m'=-1}^{1} \left[2a^k(lm,lm') - b^k(lm,lm')\right] F^k(nl^2) \tag{8.126}$$

and by applying the sum rules derived in Problems 6.6-6.7 we arive at

$$E_{nl^{2(2l+1)}} = (2l+1) \left[(4l+1)F^0(nl^2) - (2l+1) \sum_{k=2}^{\infty} \begin{pmatrix} l & k & l \\ 0 & 0 & 0 \end{pmatrix}^2 F^k(nl^2) \right]. \tag{8.127}$$

The results for closed ns^2, np^6, nd^{10}, nf^{14} shells are given in Table 8.3. The corresponding shell-averaged screening potential for the electrons in the same shell

$$\tilde{U}_F(nl,\rho) = \left[2(2l+1)^2 - l\right] \tilde{U}_F^0(nl,\rho) - (2l+1)^2 \sum_{k=2}^{\infty} \begin{pmatrix} l & k & l \\ 0 & 0 & 0 \end{pmatrix}^2 \tilde{U}_F^k(nl,\rho). \tag{8.128}$$

In particular, for $k = 0$,

$$\tilde{U}_F^0(nl,\rho) = \frac{1}{\rho} \int_0^{\rho} \left[\tilde{R}_{nl}(\varrho)\right]^2 \varrho^2 d\varrho + \int_{\rho}^{\infty} \frac{1}{\varrho} \left[\tilde{R}_{nl}(\varrho)\right]^2 \varrho^2 d\varrho = \int_0^{\infty} \frac{1}{\varrho_>} \left[\tilde{\chi}_{nl}(\varrho)\right]^2 d\varrho, \tag{8.129}$$

with $\varrho_> = \max\{\rho, \varrho\}$. The limiting behavior is given by

$$\lim_{\rho\to 0} \tilde{U}_F^0(nl,\rho) = \int_0^{\infty} \frac{1}{\varrho} \left[\tilde{\chi}_{nl}(\varrho)\right]^2 d\varrho = \tilde{U}_F^0(nl,0) \tag{8.130a}$$

$$\lim_{\rho\to\infty} \tilde{U}_F^0(nl,\rho) = \frac{1}{\rho}. \tag{8.130b}$$

Note that close to the nucleus the screening potential tends to a constant, whereas in the far field it behaves as a singly charged Coulomb potential. This was established previously for ground state helium - see Fig. 6.2.

8.5.7.2 One electron outside closed shell

This case is simpler than the previous one. The archetype of this class of electrons is a valence electron outside the electronic core. Its relative simplicity arises because the outside electron (the spectator) is not equivalent with the shell electrons. Denoting the spinorbital of the spectator by $|n'\dot{l}'m'\rangle$ the total Coulomb repulsion energy is given by

$$E_{nl^{2(2l+1)}(n'l')} = \sum_{m=-1}^{1} E(n\bar{l}m, n'\dot{l}'m') + E(n\dot{l}m, n'\dot{l}'m'), \tag{8.131}$$

Table 8.3: Coulomb repulsion energies of closed shells.

E_{ns^2}	$= F_0(ns^2)$
E_{np^6}	$= 15F_0(np^2) - 30F_2(np^2)$
$E_{nd^{10}}$	$= 45F_0(nd^2) - 70F_2(nd^2) - 630F_4(nd^2)$
$E_{nf^{14}}$	$= 91F_0(nf^2) - 420F_2(nf^2) - 1386F_4(nf^2) - 300300F_6(nf^2)$

$F_0 = F^0$; $np^6: F_2 = F^2/25$; $nd^{10}: F_2 = F^2/49,\ F_4 = F^4/441$; $nf^{14}: F_2 = F^2/225,\ F_4 = F^4/1089,\ F_6 = F^6/184041$

where we distinguished between closed shell electrons in the same spin state and those with opposite spins. As we are dealing with non-equivalent electrons, the repulsion energies are given by

$$\begin{aligned}
E(n\bar{l}m, n'\dot{l}'m') &= \mathcal{J}(nlm, n'\dot{l}'m') \\
&= \sum_{k=0}^{\infty} a^k(lm, l'm')F^k(nl; n'l') && (8.132a) \\
E(n\dot{l}m, n'\dot{l}'m') &= \mathcal{J}(nlm, n'l'm') - \mathcal{K}(nlm, n'l'm') \\
&= \sum_{k=0}^{\infty} a^k(lm, l'm')F^k(nl; n'l') - b^k(lm, l'm')G^k(nl; n'l'), && (8.132b)
\end{aligned}$$

where $\mathcal{J}$ and $\mathcal{K}$ are the direct and exchange integrals given by Eqs. (6.59a) and (6.59b), respectively. Substituting the above expressions into Eq. (8.131) we obtain

$$E_{nl^{2(2l+1)}}(n'l') = \sum_{k=0}^{\infty} \sum_{m=-1}^{1} \left[2a^k(lm, l'm')F^k(nl; n'l') - b^k(lm, l'm')G^k(nl; n'l')\right] \tag{8.133}$$

and by applying the sum rules of Problems 6.6-6.6 we arive at

$$E_{nl^{2(2l+1)}}(n'l') = 2(2l+1)\left[F^0(nl; n'l') - \tfrac{1}{2}\sum_{k=0}^{\infty} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 G^k(nl; n'l')\right]. \tag{8.134}$$

The same result is obtained starting from the spinorbital in the opposite spin state, $|n'\bar{l}'m'\rangle$. Note that, as for closed shells, the magnetic quantum number m' has been eliminated from the expression. This is an important result. Although the charge distribution of the outside electron, as expressed by $|Y_{l'}^{m'}(\theta, \phi)|^2$, depends on m', the screening by the electronic core is the same; i.e., independent of the magnetic quantum numbers m' and m_s'. This result becomes intuitively clear if we realize that the electronic core has a spherical charge distribution (and all spherical harmonics have the same normalization). This is known as *Unsöld's theorem* [115], cf. Section 9.2.

Before discussing examples we introduce a convenience notation to eliminate undesirable denominators from the expressions,

$$F_0(nl; n'l') \equiv F^0(nl; n'l') \tag{8.135}$$

$$G_k(nl; n'l') \equiv G^k(ns; n'l')/(2k+1). \tag{8.136}$$

In this notation Eq. (8.134) takes the form

$$E_{nl^{2(2l+1)}}(n'l') = 2(2l+1)F_0(nl; n'l') - \sum_{k=0}^{\infty} \frac{2l+1}{2k+1} \begin{pmatrix} l & k & l' \\ 0 & 0 & 0 \end{pmatrix}^2 G_k(nl; n'l'). \tag{8.137}$$

Actually, in view of the symmetry properties of radial integrals, we are free to change of order in the argument: $F^k(nl; n'l') = F^k(n'l'; nl)$ and $G^k(nl; n'l') = G^k(n'l'; nl)$. In this notation the results for the $ns^2n'l'$, $np^6n's$, $np^6n'p$, $nd^{10}n's$ and $nd^{10}n'p$ configutations are included in Table 8.3.

Table 8.4: Coulomb repulsion energies of closed shells plus a single electron outside the shell.

$E_{ns^2}(n'l')$	$= 2F_0(ns;n'l') - G_{l'}(ns;n'l')$
$E_{np^6}(n's)$	$= 6F_0(n's;np) - G_1(n's;np)$
$E_{np^6}(n'p)$	$= 6F_0(np;n'p) - G_0(np;n'p) - 2G_2(np;n'p)$
$E_{nd^{10}}(n's)$	$= 10F_0(n's;nd) - G_2(n's;nd)$
$E_{nd^{10}}(n'p)$	$= 10F_0(nd;n'p) - 2G_1(nd;n'p) - 3G_3(nd;n'p)$

$nl^{2(2l+1)}n'l'$: $F_0 = F^0$, $G_k = G^k/(2k+1)$

Problem 8.5. Derive an expression for the Coulomb repulsion energy of the $2s^2 2p$ configuration.

Solution. For the $2s^2 2p$ configuration only the $k = 1$ term contributes to the radial exchange summation and Eq. (8.137) reduces - with the aid of Eq. (J.23) - to

$$E_{2s^2}(2p) = 2F_0(2s;2p) - G_{l'}(2s;2p). \qquad \square$$

Problem 8.6. Derive an expression for the Coulomb repulsion energy of the $2p^6 3p$ configuration.

Solution. For the $2p^6 3s$ configuration only the $k = 1$ term contributes to the radial exchange summation and Eq. (8.137) reduces - with the aid of Eq. (J.23) - to

$$E_{2p^6}(3s) = 6F_0(2p;3s) - G_1(2p;3s). \qquad \square$$

Problem 8.7. Derive an expression for the Coulomb repulsion energy of the $2p^6 3p$ configuration.

Solution. For $2p^6 3p$ configurations only the terms with $k = 0, 2$ contribute to the radial exchange summation and Eq. (8.137) reduces - with the aid of Eqs. (J.23) and (J.25) - to

$$E_{2p^6}(3p) = 6F_0(2p;3p) - G_0(2p;3p) - 2G_2(2p;3p). \qquad \square$$

Problem 8.8. Derive an expression for the Coulomb repulsion energy of the $3d^{10} 4s$ configuration.

Solution. For the $3d^{10} 4s$ configuration only the term with $k = 2$ contributes to the radial exchange summation and Eq. (8.137) reduces - with the aid of Eq. (J.25) - to

$$E_{3d^{10}}(4s) = 10F_0(4s;3d) - G_2(4s;3d). \qquad \square$$

Problem 8.9. Derive an expression for the Coulomb repulsion energy of the $3d^{10} 4p$ configuration.

Solution. For $3d^{10} 4p$ configurations only the terms with $k = 1, 3$ contribute to the radial exchange summation and Eq. (8.137) reduces - with the aid of Eq. (J.24) - to

$$E_{3d^{10}}(n'p) = 10F_0(4p;3d) - 2G_1(4p;3d) - 3G_3(4p;3d),$$

where the prefactor 3 was determined by computer. □

8.5.7.3 Coulomb repulsion between two closed shells

The previous case is readily extended to obtain the total Coulomb repulsion energy between two closed shells. As for all electrons of a closed shell, $n'l'$, the Coulomb repulsion energy with the other shell, nl, is independent of the magnetic quantum numbers m' and m_s' and given by Eq. (8.134), the total repulsion energy between the two shells is obtained by multiplying with the number of electrons in the shell $n'l'$,

$$\begin{aligned} E_{nl^{2(2l+1)}}(n'l'^{2(2l'+1)}) &= 2(2l'+1)E_{nl^{2(2l+1)}}(n'l') && (8.138)\\ &= 4(2l+1)(2l'+1)\left[F^0(nl;n'l') - \tfrac{1}{2}\sum_{k=0}^{\infty}\begin{pmatrix} l\;k\;l' \\ 0\;0\;0 \end{pmatrix}^2 G^k(nl;n'l')\right] && (8.139)\\ &= 4(2l+1)(2l'+1)F_0(nl;n'l') - \sum_{k=0}^{\infty}\frac{2(2l'+1)(2l+1)}{2k+1}\begin{pmatrix} l\;k\;l' \\ 0\;0\;0 \end{pmatrix}^2 G_k(nl;n'l'). && (8.140) \end{aligned}$$

Note that this expression is invariant under exchange of nl and $n'l'$, as it should.

Problem 8.10. Derive an expression for the ground state energy of boron (B).

Solution. Boron is a group III atom with electron configuration $1s^22s^22p$; i.e., with the $2p$ electron as the only valence electron. The Coulomb energy of the ground state can be written as

$$E_{\mathrm{B}} = 2I(1s) + 2I(2s) + I(2p) + E_{1s^2} + E_{2s^2} + E_{1s^2}(2s^2) + E_{1s^2}(2p) + E_{2s^2}(2p),$$

where $I(nl) = \langle nl|h_0|nl\rangle$ is the hydrogenic energy of a single electron. The Coulomb repulsion energies are found with Table 6.1: $E_{1s^2} = E(1\dot{s}, 1\bar{s}) = F_0(1s^2)$, $E_{2s^2} = F_0(2s^2)$, $E_{1s^2}(2s^2) = E_{1s^2}(2\dot{s}) + E_{1s^2}(2\bar{s}) = 4F_0(2s;1s) - 2G_0(2s;1s)$, $E_{1s^2}(2p) = 2F_0(2p;1s) - G_1(2p;1s)$ and $E_{2s^2}(2p) = 2F_0(2p;2s) - G_1(2p;2s)$, with $F_0 = F^0$, $G_0 = G^0$ and $G_1 = G^1/3$. Adding the various terms the Coulomb energy becomes

$$\begin{aligned} E_{\mathrm{B}} = \varepsilon(1s^22s^22p) &= 2I(1s) + 2I(2s) + I(2p) + F_0(1s^2) + F_0(2s^2)\\ &\quad + 4F_0(2s;1s) - 2G_0(2s;1s) + 2F_0(2p;1s) - G_1(2p;1s) + 2F_0(2p;2s) - G_1(2p;2s). \quad \square \end{aligned}$$

9

Ground states of many-electron atoms

The structure of the electronic ground state of essentially all atomic elements can be predicted by a set of rules, commonly referred to under the name *Aufbau* principle and first formulated by Niels Bohr in the beginning of the 1920's [20]. It was known since the work of Dmitry Mendeleev in 1869 that the elements can be grouped systematically in order of growing nuclear charge (Z) in the form of a periodic table [72]. The Aufbau principle explains the periodic structure of the table and provides us with the electron configuration of the atomic ground state as well as with the angular momentum properties of the elements. The interpretation of the table evolved over the years but its periodic structure remained a defining feature. As shown in Table A.1 the elements are collected in seven *periods* (rows) subdivided in *groups* (columns). The periods correspond to shells of electrons with the period number defining the principal quantum number of the least bound s shell. Within a period, the group number reflects the filling of the shells. For instance, the $1s^2$ configuration of the helium ground state consists of two equivalent s electrons with $m_l = 0$ and $m_s = \pm\frac{1}{2}$ and represents the fully filled $1s$ shell. Likewise, the $1s^2 2s^2 2p^6$ configuration of the neon ground state consists of closed shells, with two equivalent s electrons in the $1s$ and the $2s$ shells with $m_l = 0$ and $m_s = \pm\frac{1}{2}$ and six equivalent p electrons in the $2p$ shell with $m_l = 0, \pm 1$ and $m_s = \pm\frac{1}{2}$. One distinguishes between short periods (group I through VIII), long periods (group 1 through 18, which include 10 groups of *transition elements*) and extra-long periods (which include an additional 14 groups of *rare-earth elements* - *lanthanides* and *actinides*).[1] The transition elements involve filling of the d shells and the rare-earth elements filling of the f shells.

Together, the shells of equivalent electrons provide the *atomic shell structure* of the atom, which is the backbone of the Aufbau principle. In X-ray spectroscopy one uses the shell notation $K, L, M, \cdots$ for the principal quantum numbers $n = 1, 2, 3, \cdots$. The rotational substructure is represented by an index; e.g., L_1, L_2, L_3. Although this notation is valuable for the innermost shells it is not practical for the outer shells of large atoms. In the latter case shells of different n can have similar binding energy. Therefore, we shall refer to shells by specifying always both n and l. When n and l no longer represent good quantum numbers the picture of a single relevant electron configuration breaks down. This happens when two configurations are approximately degenerate and a mixed configuration results in stronger binding. This being said, for most elements, a single electron configuration is a good starting point.

In ground state atoms the *closed shells* form a spherically symmetric *core* of electrons surrounded by typically one or two incomplete shells of *valence* electrons. In Section 4.6 we found that the core gives rise to substantial screening of the nuclear charge, in such a way that at large distance the valence electrons behave similar to the electrons of one-, two-, three-, $\cdots$ electron atoms with a singly-, doubly-, triply-, $\cdots$ charged nucleus, respectively. The success of describing atoms in terms of electron configurations points to the presence of a mostly *central potential field.* The related

---

[1] The *rare-earth elements* are often referred to with the plural *rare earths.* Strictly speaking this term is reserved for the oxides of these elements.

concept of a *mean field* was introduced for the exchange-free case of the helium ground state in Section 6.1.4 and formalized in the form the *Hartree method.* In Section 8.4 this method was extended to the N-electron case. The Pauli principle and exchange phenomena can be incorporated into the Hartree method by representing the electron configuration by a Slater determinant. This is the *Hartree-Fock* approach, introduced in Section 8.5.

9.1 Aufbau principle and Hund's rules

For a proper discussion of the Aufbau principle we have to address the role of the *valence electrons* because their configuration is decisive for the atomic ground state. The name *valence* points to the number of electrons participating in *chemical bonding.* This context also explains the names *donor* shell (for shells with less than half filling) and *acceptor* shell (for shells with more than half filling). As the binding energy of the valence electrons increases with shell filling, closed shells are most difficult to excite. Almost filled shells carry the angular momentum of the vacant electron(s). Such vacancies behave like electrons of positive charge and negative mass and are referred to as *holes* in the closed shell. Before addressing the possibility of spin-orbit coupling we recall that the total orbital angular momentum and the total spin,

$$\mathbf{L} = \sum_i \boldsymbol{l}_i \;\text{ and }\; \mathbf{S} = \sum_i \boldsymbol{s}_i. \tag{9.1}$$

commute with the many-electron Hamiltonian (8.1), which determines the principal structure. So, we can use the angular momentum addition rules to determine all possible values of the total orbital angular momentum, L, and the total spin, S, of a given electron configuration. Importantly, states differing in L and S differ generally strongly in binding energy because of differences in electrostatic repulsion and exchange as was demonstrated for the helium atom. With regard to the Aufbau our interest concerns the ground state. Thus we identify one or more so called LS terms, each consisting of $(2L+1)(2S+1)$ *degenerate* energy levels sharing the quantum numbers L and S. The energy of the LS terms depends on the electrostatic repulsion between the spinorbitals. For this reason we speak of *electrostatic coupling* of the single-electron orbital angular momenta $\{\boldsymbol{l}_i\}$ and of the single-electron spins $\{\boldsymbol{s}_i\}$ into states of well-defined L and S.

Closed shells can be represented by a single non-degenerate LS term ($L=0$, $S=0$). This means that only the valence electrons determine the atomic angular momentum; i.e., the valence electrons suffice to determine the LS term of the electronic ground state. The latter can be established by a sequence of three semi-empirical rules of thumb, the *Hund rules*, for which we shall establish the physical underpinning. The first two Hund rules are designed to select the LS term with smallest electrostatic repulsion between the spinorbitals:

Rule 1 Choose the maximum value of S consistent with the Pauli principle

Rule 2 Choose the maximum value of L consistent with the Pauli principle.

For a given LS term we can still have several possible values of the total electronic angular momentum J, which is the conserved quantity of the electron cloud (neglecting the hyperfine interaction with the nuclear spin) and ranges in integer steps from $|L-S|$ to $L+S$, thus adding up to the above mentioned $(2L+1)(2S+1)$ degenerate levels. The third Hund rule discriminates between these values:

$$\textit{Rule 3}\;\; \text{Choose } J = \begin{cases} J_{\min} = |L-S| \text{ for shells } \textit{less} \text{ than half-filled} \\ J_{\max} = L+S \;\text{ for shells } \textit{more} \text{ than half-filled.} \end{cases}$$

To appreciate the third Hund rule we have to understand how the degeneracy of the LS terms is lifted by the spin-orbit interaction. In one-electron atoms this originates in the velocity-induced coupling of the spin $\boldsymbol{s}$ of the electron with its own orbital angular momentum $\boldsymbol{l}$ into the total electronic angular momentum $\boldsymbol{j}$. In many-electron atoms we have to deal with competition between the electrostatic coupling and the spin-orbit coupling. This gives rise to different coupling schemes depending on Z.

9.2 Atoms with zero orbital angular momentum

At this point we can start the search for the physics underlying the Hund rules. It will be our goal to predict the angular momentum of the ground state of an arbitrary element of the periodic system. The period and group of the element provides an Ansatz for the atomic ground state configuration, adequate in most cases but with some interesting exceptions. To determine the state of angular momentum of the ground states we proceed in steps. First we consider atoms with a ground-state configuration of only closed shells. Then we look at half-filled shells. In subsequent sections we add new aspects to the analysis until we can handle the most general case. The Hund rules turn out not to be of general validity but nevertheless provide valuable guidance to sharpen our understanding of the physics that determines the ground states of atomic systems.

9.2.1 Closed shell atoms - $^1S_0$

We start with atoms composed of only closed (sub)shells. For closed shells of given values of n and l, only *a single* nonzero Slater determinant can represent the angular momentum state of the atom,

$$|\dot{l},\bar{l},\cdots,\dot{0},\bar{0},\cdots,-\dot{l},-\bar{l}\rangle_{nl^{2(2l+1)}}. \tag{9.2}$$

Moreover, the magnetic quantum numbers of the spinorbitals add up to zero,

$$M_L = 2\sum_{m_l} m_l = 0 \quad \text{and} \quad M_S = 0. \tag{9.3}$$

Hence, the angular momentum Hilbert space is one-dimensional, which means that the total orbital and the total spin angular momentum have to be zero ($L = 0$, $S = 0$); i.e., closed-shell atoms are characterized by a $^1S_0$ term. The charge distribution is *spherical* as can be made explicit with the aid of the addition theorem (L.56) for the case $\hat{\mathbf{r}} = \hat{\mathbf{r}}'$,

$$2\sum_{m=-l}^{l} |Y_l^m(\hat{\mathbf{r}})|^2 = 2\frac{(2l+1)}{4\pi}. \tag{9.4}$$

This is known as *Unsöld's theorem* [115].

There are several types of neutral atoms with only closed shells. Best known are of course the inert gases (group VIII) *helium* (He), *neon* (Ne), *argon* (Ar), *krypton* (Kr), *xenon* (Xe) and *radon* (Rn). Secondly, the *two-electron atoms*. These include the alkaline-earth elements (group II) *beryllium* (Be), *magnesium* (Mg), *calcium* (Ca), *strontium* (Sr), *barium* (Ba) and *radium* (Ra). Further, the transition elements *zinc* (Zn), *cadmium* (Cd) and *mercury* (Hg) (group 12) and the rare-earth elements *ytterbium* (Yb) and *nobelium* (No). In all these cases the *electron affinity* (i.e., the ionization energy of the negative ion with the strongest binding energy) is either zero or small as is shown in the upper panel of Fig. 9.1 (cf. Appendix B). The difference is in the *ionization energies*, which are largest for the inert gases as can be seen in the lower panel of Fig. 9.1. This difference arises from the difference in occupation of the shells. For the alkaline-earth elements only two electrons in ns shells share the Coulomb attraction of the nucleus, whereas the filled $2p$ shell (Ne) contains 6 electrons, the filled $3d$ shell (Zn) 10, *etc.*. For a growing number of electrons in a given shell, the binding increases because for neutral atoms the nuclear charge increases with the shell occupation; hence also the ionization energy. So, it takes more energy to pull an electron from

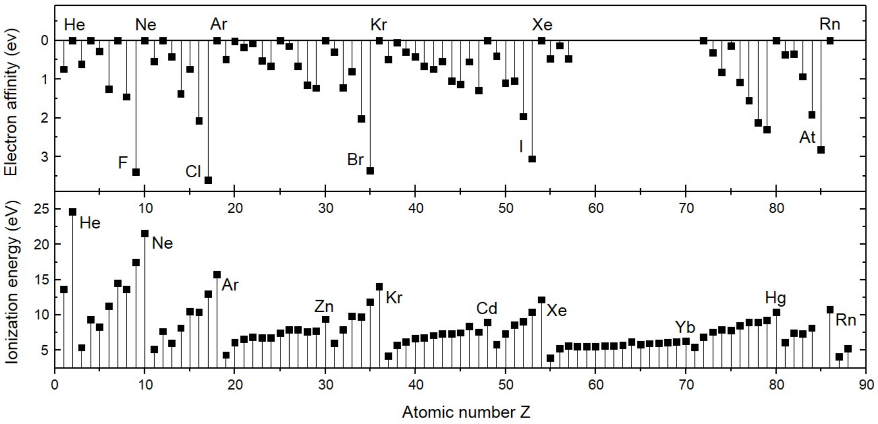


Figure 9.1: Ionization energies and electron affinities of the elements.

a shell of 6 than from a shell of 2 electrons. This effect is counteracted by the electrostatic repulsion of the electrons within a shell, which is large in small shells and small in large shells. As the relative increase in binding is largest for small Z, the differences in ionization energy are most pronounced in the first few rows of the periodic system.

For the inert gases Ne and Ar both the s shell and the p shell are closed; for Kr and Xe this holds for the s, p and d shells; for Rn for the s, p, d and f shells. As the screening efficiency by the core electrons increases with increasing Z the ionization energy of the inert gases decreases accordingly; i.e., Rn is much less *inert* than He.

9.2.2 Atoms with half-filled shells - ${}^{2J+1}S_J$

Let us next turn to atoms with a ground-state configuration of closed shells plus one *half-filled* shell of given values of n and l. In this case again only *a single* nonzero Slater determinant can represent the angular momentum state of the atom. This is the fully symmetric spin state, which can only be combined with an antisymmetric orbital state for the Pauli principle to be satisfied,

$$|\dot{l}, \cdots, \dot{0}, \cdots, -\dot{l}\rangle_{nl^{(2l+1)}}. \tag{9.5}$$

As discussed for the helium atom, antisymmetric orbital states minimize the Coulomb repulsion between the electrons, exactly as we require for atomic ground states. In other words, the total spin is maximized by quantum correlations since alignment of the spins minimizes the Coulomb repulsion within the shell. The magnetic quantum numbers of the state (9.5) add up to

$$M_L = \sum_{m_l} m_l = 0 \quad \text{and} \quad M_S = \tfrac{1}{2}(2l+1) = S. \tag{9.6}$$

Hence, also in the case of half-filled shells the orbital angular momentum subspace is one-dimensional and the total orbital angular momentum has to be zero ($L = 0$). As a consequence also the charge distribution is spherical. The ground-state terms are of the type ${}^{2J+1}S_J$, with $J = S = l + \frac{1}{2}$.

Let us look at a few examples: all ground states with three equivalent p electrons are of the type ${}^4S_{3/2}$. This holds for *nitrogen* (N), *phosphorus* (P), *arsenic* (As), *antimony* (Sb) and *bismuth* (Bi). Likewise, we have for five equivalent d electrons ${}^6S_{5/2}$. This holds for *manganese* (Mn), *technetium* (Tc) and *rhenium* (Re). For the seven equivalent f electrons of *europium* (Eu) we have ${}^8S_{7/2}$ (cf. Appendix B). *Chromium* (Cr) and *molybdenum* (Mo) are special because in these cases both an s shell and a d shell are half filled.

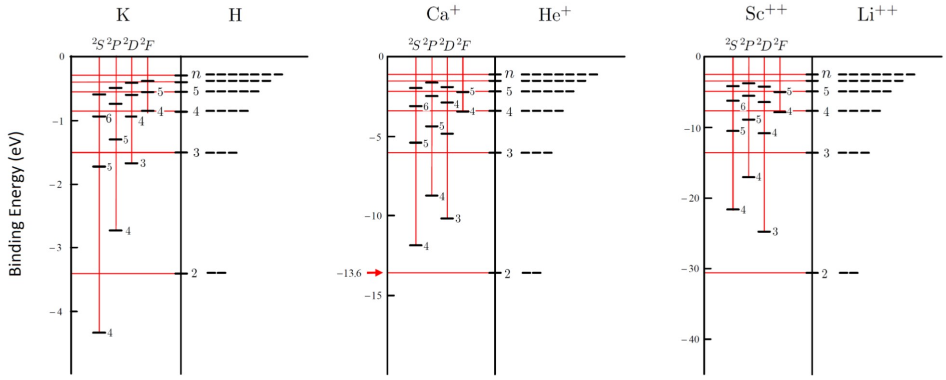


Figure 9.2: Example of three isoelectronic systems (K , Ca^+ and Sc^{++}) with one valence electron outside the core configuration of argon. With increasing nuclear charge the screening of the $4s$ electron by the $3d$ electron becomes more effective up to the point that the $3d$ electron is energetically favored; i.e., the ground state is no longer alkali like.

9.3 Atoms with one valence electron

9.3.1 Competition between electron configurations

We found that in closed shells both L and S are zero. This immediately implies that for all one-electron atoms the total angular momentum is equal to the angular momentum of the valence electron. Interestingly, in *potassium* (K) the $4s$ shell starts to be populated rather than the $3d$ shell. The same preference holds for the $5s$ shell over the both the $4d$ and $4f$ shells. This is the case of *rubidium* (Rb). In *indium* (In) the occupation of the $5p$ shell is preferred over the $4f$ shell. These preferences originate in differences of penetration of the various nl orbitals of the valences electrons into the closed shells of *core electrons.* The $R_{ns}(\rho)$ wavefunctions penetrate all the way to the nucleus. Therefore, the electron core is less effective in screening the nucleus for s electrons than for p, d, f electrons.

Anomalies occur in the periodic table when two configurations have approximately the same energy. This holds for instance for the $[\mathrm{Ar}]3d^54s^1$ and $[\mathrm{Ar}]3d^44s^2$ configurations of *chromium* (Cr) and for the $[\mathrm{Ar}]3d^{10}4s^1$ and $[\mathrm{Ar}]3d^94s^2$ configurations of *copper* (Cu). In both cases the occupation of the $3d$ shell is marginally favored over that of the $4s$ shell (cf. Appendix B). Other examples are the $[\mathrm{Kr}]4p^64d^45s$ and $[\mathrm{Kr}]4p^64d^35s^2$ configurations of *niobium* (Nb) and the $[\mathrm{Xe}]4f^{14}5d^96s$ and $[\mathrm{Xe}]4f^{14}5d^86s^2$ configurations of *platinum* (Pt). In such cases it is more difficult to estimate the ground state energy because the electronic wavefunction cannot be properly approximated by a single configuration. This phenomenon is referred to as *configuration interaction.* In theories it is accounted for by using a linear combination of configurations for the electronic state (cf. Section 8.5.2).

As an example of the arguments that come into play when discussing configuration interaction we compare the configuration of *potassium* (K) with that of *copper* (Cu). Starting from the closed shell configuration [Ar] we note that the $4s$ *electron* is favored over the $3d$ *electron.* However, starting from the closed shell configuration [Zn], we note that the $4s$ *hole* is favored over the $3d$ *hole.* How can this be if the $4s$ electron state is preferentially bound? This paradox shows that, apparently, the preferred binding of the $4s$ electron observed in K is lost in the case of Cu. This can be understood by considering the strong increase in nuclear charge along the 4^{th} period, from $Z_{\mathrm{K}} = 19$ to $Z_{\mathrm{Cu}} = 29$.

As screening is always incomplete, the increase in Z tends to increase the binding of *all* electrons but the actual increase depends on screening details of the individual orbitals:

- *Screening by the core:* As the core electrons of *potassium* $(1s, 2s, 2p, 3s)$ are least screened, the increase in binding energy with increasing Z is strongest for these electrons. This results

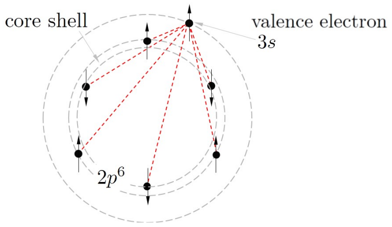


Figure 9.3: Principle of core polarization by a spin-up valence electron. The Coulomb repulsion by the core shell differs for spin-up and spin-down core electrons due to the presence (spin-up) or absence (spin-down) of the exchange contribution. In the diagram this is illustrated by splitting the $2p^6$ core shell into two shells of slightly different diameter.

in core contraction. The contraction of the $3d$ orbital (and to a lesser extent the $4p$ orbital) is more pronounced than that of the $4s$ orbital (because the $4s$ electron always probes the nucleus). This phenomenon is sometimes referred to as the collapse of the $3d$ shell. We can convince ourselves about this effect by comparing the energy levels of *potassium* (K), singly ionized *calcium* (Ca^+) and doubly ionized *scandium* (Sc^{2+}), which are *isoelectronic* (i.e., have the same number of electrons) but differ in the nuclear charge - see Fig. 9.2. Actually, the electronic ground state configurations of the transition-metal ions Sc^{2+}, Ti^{3+}, V^{4+}, *etc.*, are all $[\mathrm{Ar}]3d^1$. This shows that even a modest increase in nuclear charge from $Z_{\mathrm{K}} = 19$ to $Z_{\mathrm{Sc}} = 21$ is sufficient to remove the penetration advantage of the $4s$ electron and favor the $3d$ electron.

- *Screening by the valence electrons:* The collapsed $3d$ shell acts as an additional screening layer for the $4s$ electrons. Therefore, with growing $3d$ occupation (along the 4^{th} period) the binding of the $4s$ electron erodes by $3d$ screening up to the point that a $3d$ electron is favored. This first happens for *chromium* (Cr) and later in the period for *copper* (Cu). Beyond Cr the $4s$ electron is again favored because all $3d$ orbitals are already occupied once and screening of an additional $3d$ electron by the others is more effective at double occupation (not reduced by quantum correlations). Yet, as the increase in Z continues along the 4^{th} period the story repeats itself: with the $4s$ shell closed the $3d$ occupation resumes and the $4s$ binding erodes until the $3d$ electron is again favored at $Z = 29$. This manifests itself as the $4s$ hole in the configuration of Cu.

9.3.2 Core polarization - unrestricted Hartree-Fock method

Thus far in calculating Coulomb repulsion we presumed that all electrons of a closed shell have the same radial wavefunction. In Section 6.1.3 we briefly abandoned this restriction and found that to achieve binding of H^- the two electrons of the $1s$ shell should have different Bohr radii; i.e., half of the shell is compressed and the other half is inflated. In the case of H^- this arises from the correlation in the electron motion, which tends to keep the inter-electronic separation as large as possible. We now consider a similar phenomenon but with quantum correlations at its origin. The principle is illustrated in Fig. 9.3 where we show a schematic diagram in which the $3s$ electron of *sodium* (Na) interacts with the $2p^6$ core shell. The Coulomb repulsion energy between the valence electron and the $2p^6$ shell is given by a part which includes exchange

$$E(3\dot{s},2\dot{p}_1)+E(3\dot{s},2\dot{p}_0)+E(3\dot{s},2\dot{p}_{-1})=3F_0-3G_1, \tag{9.7}$$

where $G_1 = G^1/3$ and a part which does *not* include exchange (see Problem 7.5)

$$E(3\dot{s},2\bar{p}_1)+E(3\dot{s},2\bar{p}_0)+E(3\dot{s},2\bar{p}_{-1})=3F_0. \tag{9.8}$$

As the Coulomb repulsion is less in the former case, the up-spin part of the core will be less compressed by the valence electron than the down-spin part. To allow for this effect one has to use different orbitals for spin-up and spin-down in the Hartree-Fock optimization. This is known as the *unrestricted* Hartree-Fock method. This approach requires the use of twice as many core orbitals as in the ordinary Hartree-Fock method introduced in Section 8.5.3. The latter is called the *restricted* Hartree-Fock method because all electrons within a closed shell are forced to have the same radial wavefunction, irrespective of the spin state. An important consequence of the separation of the core shells in a spin-up and a spin-down part is the appearance of a spin density which is *globally balanced* but *locally not.* This is called local polarization of the electron core. Of particular importance is core polarization of s shells as this gives rise to a nonzero electron spin density at the nucleus, which is rigorously absent in the restricted Hartree-Fock picture. This gives rise to an *induced* Fermi contact interaction. This contact interaction can also be induced by valence electrons with $l \neq 0$. As the induced spin at the nucleus is opposite to the spin of the polarizing valence electron it tends to suppress the hyperfine splitting for parallel coupling ($j = l + s$) and to enhance the splitting for antiparallel coupling ($j = l - s$). This can give rise to a level inversion of the hyperfine levels as is illustrated for the $2\,^2P_{1/2}$ and $2\,^2P_{3/2}$ fine-structure doublet of ^{7}Li in Fig. 5.3.

9.4 Atoms with more than one valence electron - Hund's Rule 1 & 2

9.4.1 Introduction

Thus far we have been dealing with an unambiguous angular momentum state of the atom. Once we have more than one valence electron in a partially filled shell (but not half filled) matters complicate considerably because by coupling of angular momenta more than one angular momentum state can be associated with a given electron configuration. So, it takes an additional effort to determine which angular momentum state (LS term) corresponds to the electronic ground state of a given electron configuration. The first two Hund rules provide us with a rapid assessment of the LS term corresponding to the electronic ground state.

In the present section we determine for a given electron configuration which LS terms are consistent with Fermi-Dirac statistics; i.e., can be expressed as a linear combination of (nonzero) Slater determinants. Once we know the relevant LS terms we calculate their electrostatic energies and demonstrate that the first two Hund rules correctly predict (in practically all cases) the LS term of lowest electrostatic repulsion between the valence electrons, which corresponds to the electronic ground state of a given electron configuration.

9.4.2 Partially filled shells with np^2 configuration

As the first case of a partially filled shell with two electrons we consider the np^2 configuration. This is the case of the group IV atoms C, Si, Ge, Sn and Pb. In the absence of the Pauli principle the electronic pair state $|\psi_{12}\rangle$ can be expressed as a linear combination of the $(2l+1)^2(2s+1)^2 = 36$ basis states of the *uncoupled* representation $\{|nlm_{l_1}sm_{s_1}\rangle \otimes |nlm_{l_2}sm_{s_2}\rangle\}$ with $l = 1$ and $s = \frac{1}{2}$. Using a basis transformation this basis can be reexpressed into a basis of

$$\binom{6}{2} = 15 \tag{9.9}$$

fully antisymmetric and 21 partly or fully symmetric *linearly independent* states. For equivalent electrons the 21 partly or fully symmetric basis states have to be excluded on the basis of the Pauli principle. The leaves a 15-dimensional Hilbert space consistent with the Pauli principle. The 15 *linearly independent* antisymmetric states are readily identified in the Slater determinant notation. In Table 9.1 they are grouped by the values of $M_L = m_{l_1} + m_{l_2}$ and $M_S = m_{s_1} + m_{s_2}$.

Table 9.1: Example np^2 configuration: (top) uncoupled representation - the 15 linearly independent pair states consistent with the Pauli principle; (bottom) coupled representation - all LS terms (including degeneracy these yield 36 linearly independent states).

np^2		M_S			deg.
uncoupled		1	0	-1	15
M_L	2		$\lvert\dot{1},\bar{1}\rangle$		1
	1	$\lvert\dot{1},\dot{0}\rangle$	$\lvert\dot{1},\bar{0}\rangle\lvert\bar{1},\dot{0}\rangle$	$\lvert\bar{1},\bar{0}\rangle$	4
	0	$\lvert\dot{1},-\dot{1}\rangle$	$\lvert\dot{1},-\bar{1}\rangle\lvert\dot{0},\bar{0}\rangle\lvert\bar{1},-\dot{1}\rangle$	$\lvert\bar{1},-\bar{1}\rangle$	5
	-1	$\lvert\dot{0},-\dot{1}\rangle$	$\lvert\dot{0},-\bar{1}\rangle\lvert\bar{0},-\dot{1}\rangle$	$\lvert\bar{0},-\bar{1}\rangle$	4
	-2		$\lvert-\dot{1},-\bar{1}\rangle$		1

np^2		L			deg.
coupled		0	1	2	36
S	0	$^1S_0$	$^1P_1$	$^1D_2$	9
	1	$^3S_1$	$^3P_2, ^3P_1, ^3P_0$	$^3D_3, ^3D_2, ^3D_1$	27

Further insight into the role of the Pauli principle for partially filled shells is obtained by analyzing the np^2 configuration in the *coupled representation* of the total orbital angular momentum $\mathbf{L} = \boldsymbol{l}_1 + \boldsymbol{l}_2$ coupled with the total spin $\mathbf{S} = \boldsymbol{s}_1 + \boldsymbol{s}_2$ into the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$. In a potential field of central symmetry J is a good quantum number. For the np^2 configuration the total orbital angular quantum number can take the values $L = 0, 1, 2$ with possible total spin $S = 0, 1$. In the coupled representation $\{|JM_J\rangle\}$ we have 9 *singlet* states ($S = 0$) and 27 *triplet* states ($S = 1$); i.e., together again 36 atomic states can be identified. In the absence of spin-orbit coupling these are all degenerate. In the LS-term notation these 36 states are also given in Table 9.1. Three questions immediately arise:

1. which of these LS terms are consistent with the Pauli principle?
2. what is the relation between the LS terms and the Slater determinants?
3. which term has the lowest energy?

9.4.2.1 Determination of LS terms consistent with Pauli principle

To answer the first question we are guided by the knowledge that the dimension of the anti-symmetric subspace is independent of the choice of basis. Hence we search for 15 basis states in the coupled representation. This is done with the aid of a decision table. For the np^2 configurations this is Table 9.2.

We start with the angular momentum state of the largest multiplicity, i.e., the ^3D sector of the coupled representation. This sector is not contained in the anti-symmetric subspace because for $L = 2$ and $S = 1$ the ^3D sector must include the Slater determinant $|\dot{1},\dot{1}\rangle$, which corresponds to $M_L = m_{l_1} + m_{l_2} = 2$ and $M_S = m_{s_1} + m_{s_2} = 1$. But this determinant violates the Pauli principle and therefore has to be excluded.

Continuing with the triplet manifolds, we turn to the ^3P sector, which is the largest of the remaining sectors. Since $L = 1$ and $S = 1$, the ^3P sector must include the Slater determinant $|\dot{1},\dot{0}\rangle$, which corresponds to $M_L = m_{l_1} + m_{l_2} = 1$ and $M_S = m_{s_1} + m_{s_2} = 1$. Hence, we identify: $|\dot{1},\dot{0}\rangle = |^3P; M_L = 1, M_S = 1\rangle$. In Problem 7.7 we calculated L with the algebraic method of Eq. 7.99 (which requires a larger effort than the dimensional analysis of the present section). Because the determinant $|\dot{1},\dot{0}\rangle$ is consistent with the Pauli principle it must be within the anti-symmetric part of the Hilbert space. Actually, this must hold for the whole ^3P sector because all eigenstates of this sector can be generated from $|\dot{1},\dot{0}\rangle$ with the shift operators $L_\pm$ and $S_\pm$ and these preserve

Table 9.2: Decision table for np^2 configurations using the known LS-term degeneracy to determine the spin-orbit Terms consistent with the Pauli principle for two equivalent p-electrons.

M_S	M_L	Slater determinants[†]		LS terms		
		Pauli allowed:		options	accept	excl.
1	2	$\upharpoonleft\!\upharpoonleft - -$	0	^3D	-	^3D
	1	$\upharpoonleft \upharpoonleft -$	1	^3P	^3P	-
	0	$\upharpoonleft - \upharpoonleft, - \upharpoonleft\!\upharpoonleft -$	1	$^3P, ^3S$	-	^3S
0	2	$\upharpoonleft\!\downharpoonleft - -$	1	^1D	^1D	-
	1	$\upharpoonleft \downharpoonleft -$ (2×)	2	$^3P, ^1D, ^1P$	-	^1P
	0	$\upharpoonleft - \downharpoonleft$ (2×), $- \upharpoonleft\!\downharpoonleft -$	3	$^3P, ^1D, ^1S$	^1S	-

[†] The sliced-arrow notation has no significance other than being compact.

the antisymmetric symmetry - see Eq. (7.95) and question b.). Furthermore, because the ^3P sector represents a 3×3-fold degenerate manifold, also 9 linearly independent determinants are required to represent the ^3P sector for all combinations of $M_S = 1, 0, -1$ and $M_L = 1, 0, -1$. These 9 possibilities are collected in Table 9.3 along with the basis states of the ^1D and ^1S sectors, which are also Pauli-allowed (compare with Table 9.3-top).

Next we turn to the ^3S sector, for which $L = 0$ and $S = 1$. This sector has to be excluded because there is only a single Slater determinant with $M_S = 1$ and $M_L = 0$ consistent with the Pauli principle but this one was already assigned as part of the ^3P manifold. So, we proceed with the singlet sectors. First ^1D ($L = 2; S = 0$), this 1×5-fold degenerate sector includes the Slater determinant $|\dot{1}, \bar{1}\rangle$, which is allowed by the Pauli principle. Thus the ^1D sector must be part of the subspace spanned by 5 linearly independent Slater determinants corresponding to all combinations of $M_S = 0$ and $M_L = 2, 1, 0, -1, -2$ (these 5 possibilities are collected Table 9.3).

At this point we assigned $9 + 5 = 14$ of the 15 basis states, which leaves only a single non-degenerate sector; i.e., a $^1S_0$ term (also collected in Table 9.3). The same result is also obtained by continuing the procedure of Table 9.2 until we decided on all LS terms. Note that in view of the symmetry of the table only the non-positive (or the non-negative) $M_J = M_L + M_S$ values suffice to assign all LS terms. We found that only the ^3P, ^1D and ^1S sectors are consistent with the Pauli principle. In the coupled representation this implies the terms $^3P_2$, $^3P_1$, $^3P_0$, $^1D_2$ and $^1S_0$. Including their degeneracies we recover again the $5+3+1+5+1 = 15$-dimensional subspace of antisymmetric states identified at the beginning of this section (see Table 9.4).

9.4.2.2 The LS terms decomposed as linear combinations of Slater determinants

Let us turn to the second question. At this point we have identified all LS terms of the np^2 configurations consistent with the Pauli principle. It would be nice if a unique mapping would exist onto the Slater determinants. Unfortunately this is generally not the case. Exceptions are the LS-term states $|LM_L, SM_S\rangle \equiv |^{2S+1}L; M_L, M_S\rangle$ for which a single LS term combines with a single determinant in Table 9.2,

$$|^3P; 1, 1\rangle = |\dot{1}, \dot{0}\rangle_{np^2} \tag{9.10}$$

$$|^1D; 2, 0\rangle = |\dot{1}, \bar{1}\rangle_{np^2}. \tag{9.11}$$

The first example corresponds to a stretched spin state ($M_S = S = 2s = 1$), the second with a stretched orbital angular momentum state ($M_L = L = 2l = 2$). Being stretched states these examples represent eigenstates of the total spin and the total angular momentum, respectively

Table 9.3: Multiplicity analysis of the upper part of Table 9.1 to identify the 15 linearly independent LS terms of the np^2 configuration consistent with the Pauli principle. Note the 3×3 block of the ^3P sector, the 1×5 block of the ^1D sector and the 1×1 block of the ^1S sector.

$\lvert ^{2S+1}L; M_L, M_S\rangle$		M_S			deg.
(coupled)		1	0	-1	15
M_L	2		$\lvert ^1D; 2, 0\rangle$		1
	1	$\lvert ^3P; 1, 1\rangle$	$\lvert ^3P; 1, 0\rangle \lvert ^1D; 1, 0\rangle$	$\lvert ^3P; 1, -1\rangle$	4
	0	$\lvert ^3P; 0, 1\rangle$	$\lvert ^3P; 0, 0\rangle \lvert ^1S; 0, 0\rangle \lvert ^1D; 0, 0\rangle$	$\lvert ^3P; 0, -1\rangle$	5
	-1	$\lvert ^3P; -1, 1\rangle$	$\lvert ^3P; -1, 0\rangle \lvert ^1D; -1, 0\rangle$	$\lvert ^3P; -1, -1\rangle$	4
	-2		$\lvert ^1D; -2, 0\rangle$		1

(see Problem 7.8). To find the other basis states of the ^3P sector in the Hilbert space of the np^2 configuration we compare the action of the spin lowering operator S_- in both representations,

$$S_\pm|\psi_\alpha\rangle = \begin{cases} \sqrt{S(S+1) - M_S(M_S \pm 1)}\,\hbar\,|\psi_\alpha(M_S \pm 1)\rangle \\ \sum_\kappa \sqrt{s(s+1) - m_s^\kappa(m_s^\kappa \pm 1)}\,\hbar\,|\psi_\alpha(m_s^\kappa \pm 1)\rangle. \end{cases} \tag{9.12}$$

Starting from the stretched spin state (9.10) we calculate

$$\left.\begin{array}{l} S_-|^3P; 1, 1\rangle = \sqrt{2}\,\hbar\,|^3P; 1, 0\rangle \\ S_-|\dot{1}, \dot{0}\rangle_{np^2} = \hbar\,|\bar{1}, \dot{0}\rangle_{np^2} + \hbar\,|\dot{1}, \bar{0}\rangle_{np^2} \end{array}\right\} \to |^3P; 1, 0\rangle = \sqrt{\tfrac{1}{2}}\left[|\dot{1}, \bar{0}\rangle_{np^2} + |\bar{1}, \dot{0}\rangle_{np^2}\right]. \tag{9.13}$$

Although it is a bit laborious we may convince ourselves explicitly with Eqs. (7.99), (7.93), (7.111) and (7.109) that the r.h.s. of Eq. (9.13) is indeed an eigenstate of $\mathbf{L}^2, L_z, \mathbf{S}^2$ and S_z. To find the basis of the ^1D sector in the Hilbert space of the np^2 configuration we compare the action of the orbital lowering operator L_- on both representations

$$L_\pm|\psi_\alpha\rangle = \begin{cases} \sqrt{L(L+1) - M_L(M_L \pm 1)}\,\hbar\,|\psi_\alpha(M_L \pm 1)\rangle \\ \sum_\kappa \sqrt{l^\kappa(l^\kappa+1) - m_l^\kappa(m_l^\kappa \pm 1)}\,\hbar\,|\psi_\alpha(m_l^\kappa \pm 1)\rangle. \end{cases} \tag{9.14}$$

Starting from the stretched orbital angular momentum state (9.11) we obtain

$$\left.\begin{array}{l} L_-|^1D; 2, 0\rangle = 2\hbar\,|^1D; 1, 0\rangle \\ L_-|\dot{1}, \bar{1}\rangle_{np^2} = \sqrt{2}\,\hbar\,|\dot{0}, \bar{1}\rangle_{np^2} + \sqrt{2}\,\hbar\,|\dot{1}, \bar{0}\rangle_{np^2} \end{array}\right\} \to |^1D; 1, 0\rangle = \sqrt{\tfrac{1}{2}}\left[|\dot{1}, \bar{0}\rangle_{np^2} - |\bar{1}, \dot{0}\rangle_{np^2}\right]. \tag{9.15}$$

The minus sign appeared when the determinantal state $|\dot{0}, \bar{1}\rangle_{np^2}$ was brought to standard order. Note that the states $|^3P; 1, 0\rangle$ and $|^1D; 1, 0\rangle$ are orthogonal eigenstates of $\mathbf{L}^2, L_z, \mathbf{S}^2$ and S_z, both composed of the same Slater determinants.

To express the basis state $|^1S; 0, 0\rangle$ of the ^1S sector in Slater determinants we have to take a different approach because this sector does not contain a stretched state. However, we can use the L_- operator to construct the states $|^3P; 0, 0\rangle$ and $|^1D; 0, 0\rangle$, which both must be orthogonal to $|^1S; 0, 0\rangle$. For the state $|^3P; 0, 0\rangle$ we find with the aid of Eqs. (9.13) and (7.95)

$$\left.\begin{array}{r} L_-|^3P; 1, 0\rangle = \sqrt{2}\,\hbar\,|^3P; 0, 0\rangle \\ L_-\sqrt{\tfrac{1}{2}}\left[|\dot{1}, \bar{0}\rangle_{np^2} + |\bar{1}, \dot{0}\rangle_{np^2}\right] = \hbar\,|\dot{1}, -\bar{1}\rangle_{np^2} + \hbar\,|\bar{1}, -\dot{1}\rangle_{np^2} \end{array}\right\} \to$$

$$\to |^3P; 0, 0\rangle = \sqrt{\tfrac{1}{2}}\left[|\dot{1}, -\bar{1}\rangle_{np^2} + |\bar{1}, -\dot{1}\rangle_{np^2}\right]. \tag{9.16}$$

Analogously we find with the aid of Eq. (9.15) for the state $|^1D;0,0\rangle$

$$\left.\begin{aligned} L_-|^1D;1,0\rangle &= \sqrt{6}\,\hbar\,|^1D;0,0\rangle \\ L_-\sqrt{\tfrac{1}{2}}\left[|\dot{1},\bar{0}\rangle_{np^2} - |\bar{1},\dot{0}\rangle_{np^2}\right] &= 2\hbar\,|\dot{0},\bar{0}\rangle_{np^2} + \hbar\,|\dot{1},-\bar{1}\rangle_{np^2} - \hbar\,|\bar{1},-\dot{1}\rangle_{np^2} \end{aligned}\right\}\rightarrow$$

$$\rightarrow |^1D;0,0\rangle = \sqrt{\tfrac{1}{6}}\left[2|\dot{0},\bar{0}\rangle_{np^2} + |\dot{1},-\bar{1}\rangle_{np^2} - |\bar{1},-\dot{1}\rangle_{np^2}\right]. \tag{9.17}$$

What remains to be done is to find the coefficients that orthonormalize the linear combination $|^1S;0,0\rangle = a|\dot{0},\bar{0}\rangle_{np^2} + b|\dot{1},-\bar{1}\rangle_{np^2} + c|\bar{1},-\dot{1}\rangle_{np^2}$ with respect to both $|^3P;0,0\rangle$ and $|^1D;0,0\rangle$. To be orthogonal to $|^3P;0,0\rangle$ we require $\sqrt{1/2}\,b + \sqrt{1/2}\,c = 0 \Rightarrow b = -c$. To be also orthogonal to $|^1D;0,0\rangle$ the coefficients should satisfy the relation $\sqrt{2/3}\,a - \sqrt{1/6}\,c - \sqrt{1/6}\,c = 0 \Rightarrow a = c$. Thus, the normalized linear combination is

$$|^1S;0,0\rangle = \sqrt{\tfrac{1}{3}}\left[|\dot{0},\bar{0}\rangle_{np^2} - |\dot{1},-\bar{1}\rangle_{np^2} + |\bar{1},-\dot{1}\rangle_{np^2}\right]. \tag{9.18}$$

9.4.2.3 The first Hund rule

The remaining question concerns the energies of the LS terms. First of all we note that all the terms identified in Table 9.2 have the same energy in the screened Coulomb field of the nucleus,

$$E_c(np^2) = -2 \times hcR_\infty Z_{np}^2/n^2, \tag{9.19}$$

where Z_{np} has to be determined experimentally, or theoretically by numerical integration of the radial Schrödinger equation for a p-electron ($l = 1$) in the screened Coulomb field of the nucleus. Therefore, the electrostatic repulsion between the electrons determines which LS term has the lowest energy and this repulsion is smallest for the largest average separation between the electrons. In view of Pauli exclusion of close approach of identical fermions, we may argue that the symmetric spin states (the states with the largest total spin S) are energetically most favorable - see Fig. 6.4. This insight provides the physics behind the *first Hund rule*:

> The *Coulomb repulsion* between the valence electrons of an atomic ground state configuration is *minimal for the maximum (Pauli-allowed) value of the total spin*, S.

For the np^2 configuration the terms with the largest spin ($S = 1$) are the triplet terms $^3P_{2,1,0}$. The ground state of carbon is indeed a ^3P term: $^3P_0$ (cf. Appendix B). To decide among the terms $^3P_2$, $^3P_1$ and $^3P_0$ we need the *third Hund rule*. This rule originates in the spin-orbit interaction and will be discussed in Section 9.5.

Table 9.4: Coupled and uncoupled basis states for np^2 configurations. In both cases we can identify 15 linearly independent pair states spanning the antisymmetric subspace in which the Pauli principle is satisfied.

uncoupled					coupled						
M_L					$\|J,M_J\rangle$						
2	1	0	-1	-2	$^1D_2$	$^3P_2$	$^3P_1$	$^3P_0$	$^1S_0$		
$\|\dot{1},\bar{1}\rangle$	$\|\dot{1},\dot{0}\rangle$				$\|2,2\rangle$	$\|2,2\rangle$				2	
	$\|\dot{1},\bar{0}\rangle\|\bar{1},\dot{0}\rangle$	$\|\dot{1},-\dot{1}\rangle$			$\|2,1\rangle$	$\|2,1\rangle$	$\|1,1\rangle$			1	
	$\|\bar{1},\bar{0}\rangle$	$\|\dot{1},-\bar{1}\rangle\|\dot{0},\bar{0}\rangle\|\bar{1},-\dot{1}\rangle$	$\|\dot{0},-\dot{1}\rangle$		$\|2,0\rangle$	$\|2,0\rangle$	$\|1,0\rangle$	$\|0,0\rangle$	$\|0,0\rangle$	0	M_J
		$\|\bar{1},-\bar{1}\rangle$	$\|\dot{0},-\bar{1}\rangle\|\bar{0},-\dot{1}\rangle$		$\|2,-1\rangle$	$\|2,-1\rangle$	$\|1,-1\rangle$			-1	
			$\|\bar{0},-\bar{1}\rangle$	$\|-\dot{1},-\bar{1}\rangle$	$\|2,-2\rangle$	$\|2,-2\rangle$				-2	

Calculation of Term energies: Let us explicitly verify the validity of the first Hund rule by calculating the energy of the three LS terms identified above. Because in zero field the energy is independent of the magnetic quantum numbers it can be determined for any value of M_L and M_S. First we calculate the energy for the stretched states. In Hartree atomic units the electrostatic repulsion energy of the ${}^3P$ term is given by

$$E({}^3P) = \langle \dot{1}, \dot{0}| \frac{1}{\rho_{12}} |\dot{1}, \dot{0}\rangle_{np^2}. \tag{9.20}$$

This is a diagonal matrix element of a two-body operator and using Eq. (7.53) we find

$$E({}^3P) = (p_1, p_0| \frac{1}{\rho_{12}} |p_1, p_0) - (p_0, p_1| \frac{1}{\rho_{12}} |p_1, p_0) = \mathcal{J} - \mathcal{K}, \tag{9.21}$$

where $\mathcal{J} = (p_1, p_0|\rho_{12}^{-1}|p_1, p_0)$ and $\mathcal{K} = (p_0, p_1|\rho_{12}^{-1}|p_1, p_0)$ are the Coulomb integrals defined by Eqs. (6.46). For two p electrons Eq. (9.21) reduces to

$$E({}^3P) = \sum_{k=0,2} \left[a^k(p_1, p_0) F^k(np^2) - b^k(p_1, p_0) G^k(np^2) \right]. \tag{9.22}$$

Substituting the values for $a^k(p_1, p_0)$ and $b^k(p_1, p_0)$ for $k = 0, 2$ from Table 6.1 the electrostatic repulsion can be written as the sum of two F integrals (note that $F^k = G^k$ for equivalent electrons),

$$E({}^3P) = F_0 - 2F_2 - 3G_2 = F_0 - 5F_2. \tag{9.23}$$

In accordance with convention the common denominators of the a^k and b^k coefficients were eliminated by redefining the F integrals: $F_0 \equiv F^0$, $F_2 \equiv F^2/25$.

Also the ${}^1D$ term represents a stretched angular momentum state (orbital angular momentum in this case),

$$E({}^1D) = \langle \dot{1}, \bar{1}| \frac{1}{\rho_{12}} |\dot{1}, \bar{1}\rangle_{np^2}. \tag{9.24}$$

This case is even simpler than the ${}^3P$ term because the exchange term vanishes; as the orbital part is stretched it is manifestly symmetric - note also the antisymmetric spin part. Looking up $a^k(p_1, p_1)$ for $k = 0, 2$ from Table 6.1 the expression for the electrostatic repulsion takes the form

$$E({}^1D) = \sum_{k=0,2} a^k(p_1, p_1) F^k(np^2) = F_0 + F_2. \tag{9.25}$$

Note that $E({}^1D) > E({}^3P)$, in accordance with the first Hund rule.

9.4.2.4 Slater sum rule

The electrostatic repulsion energy is more difficult to calculate for the ${}^1S$ term because it does not represent a stretched angular momentum state. In such cases the direct and exchange integrals do not suffice to calculate the energy with the straightforward way used above. Slater formulated a sum rule that enables to work around this complication. It is based on the invariance of the trace of the matrix of an operator $\mathcal{H}$ under a unitary transformation, $\mathcal{H}' = U^{-1}\mathcal{H}U$, from the coupled to the uncoupled representation,

$$\mathrm{tr}\,(\mathcal{H}') = \mathrm{tr}\,(U^{-1}\mathcal{H}U) = \mathrm{tr}\,(UU^{-1}\mathcal{H}) = \mathrm{tr}\,(\mathcal{H}). \tag{9.26}$$

This invariance follows from property (M.26) of square matrices. Secondly, the sum rule is based on the degeneracy of all magnetic sublevels M_S and M_L of a given LS term in zero field. Together with the states $|{}^3P; 0, 0\rangle$ and $|{}^1D; 0, 0\rangle$ the state $|{}^1S; 0, 0\rangle$ spans a 3 dimensional subspace. The electrostatic

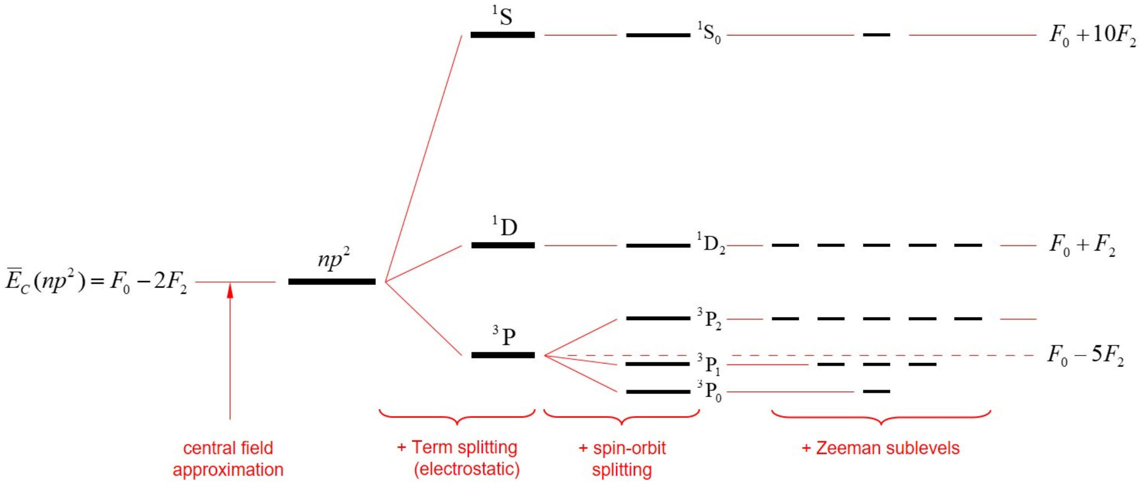


Figure 9.4: Fine-structure diagram of atoms from the carbon group (group 14) as an example of a np^2 valence configuration.

repulsion is diagonal in this basis and summing the diagonal matrix elements we obtain for the trace of the 3 × 3 matrix

$$\operatorname{tr}(3\times 3) = E(^3P) + E(^1D) + E(^1S) = 2F_0 - F_2 - 3G_2 + E(^1S), \tag{9.27}$$

where we used the M_S and M_L independence of the LS term energies. The same subspace is spanned by the determinantal states $|\dot{0},\bar{0}\rangle_{np^2}$, $|\dot{1},-\bar{1}\rangle_{np^2}$ and $|\bar{1},-\dot{1}\rangle_{np^2}$. The trace of the corresponding matrix is given by[1]

$$\begin{aligned}\operatorname{tr}(3\times 3) &= \langle\dot{0},\bar{0}|\frac{1}{\rho_{12}}|\dot{0},\bar{0}\rangle_{np^2} + \langle\dot{1},-\bar{1}|\frac{1}{\rho_{12}}|\dot{1},-\bar{1}\rangle_{np^2} + \langle\bar{1},-\dot{1}|\frac{1}{\rho_{12}}|\bar{1},-\dot{1}\rangle_{np^2} \\ &= (F_0+4F_2) + (F_0+F_2) + (F_0+F_2) = 3F_0 + 6F_2. \end{aligned} \tag{9.28}$$

Here the three matrix elements were calculated by looking up the a^k terms with the aid of Table 6.1. Comparing Eqs. (9.27) and (9.28), we find

$$E(^1S) = F_0 + 7F_2 + 3G_2 = F_0 + 10F_2. \tag{9.29}$$

Note that $E(^1S) > E(^3P)$, again in accordance with the first Hund rule.

9.4.3 Partially filled shells with nd^2 configuration

As a second example of a partially filled shell we consider the nd^2 configuration. This is the case of Ti, Zr, Hf and Rf. In the absence of the Pauli principle the electronic pair state $|\psi_{12}\rangle$ can be expressed as a linear combination of the $(2l+1)^2(2s+1)^2 = 100$ basis states of the *uncoupled* representation $\{|nlm_{l_1}sm_{s_1}\rangle \otimes |nlm_{l_2}sm_{s_2}\rangle\}$, with $l=2$ and $s=\frac{1}{2}$. In this case we obtain

$$\binom{10}{2} = 45 \tag{9.30}$$

linearly independent antisymmetric states consistent with the Pauli principle. In Table 9.5 they are grouped by the values of $M_L = m_{l_1} + m_{l_2}$ and $M_S = m_{s_1} + m_{s_2}$.

---

[1] For diagonal matrix elements of the type $\langle\dot{u},\bar{v}|\rho_{12}^{-1}|\dot{u},\bar{v}\rangle$ the exchange terms do not contribute - see Problems 7.2 and 7.5.

Table 9.5: Example nd^2 configuration: (top) uncoupled representation - pair states consistent with the Pauli principle; (bottom) coupled representation - all LS terms.

nd^2		M_S			deg.
uncoupled		1	0	-1	45
M_L	4		$\vert\dot{2},\bar{2}\rangle$		1
	3	$\vert\dot{2},\dot{1}\rangle$	$\vert\dot{2},\bar{1}\rangle\vert\bar{2},\dot{1}\rangle$	$\vert\bar{2},\bar{1}\rangle$	4
	2	$\vert\dot{2},\dot{0}\rangle$	$\vert\dot{2},\bar{0}\rangle\vert\dot{1},\bar{1}\rangle\vert\bar{2},\dot{0}\rangle$	$\vert\bar{2},\bar{0}\rangle$	5
	1	$\vert\dot{2},-\dot{1}\rangle\vert\dot{1},\dot{0}\rangle$	$\vert\dot{2},-\bar{1}\rangle\vert\dot{1},\bar{0}\rangle\vert\bar{1},\dot{0}\rangle\vert\bar{2},-\dot{1}\rangle$	$\vert\bar{2},-\bar{1}\rangle\vert\bar{1},\bar{0}\rangle$	8
	0	$\vert\dot{2},-\dot{2}\rangle\vert\dot{1},-\dot{1}\rangle$	$\vert\dot{2},-\bar{2}\rangle\vert\dot{1},-\bar{1}\rangle\vert\dot{0},\bar{0}\rangle\vert\bar{1},-\dot{1}\rangle\vert\bar{2},-\dot{2}\rangle$	$\vert\bar{2},-\bar{2}\rangle\vert\bar{1},-\bar{1}\rangle$	9
	-1	$\vert\dot{1},-\dot{2}\rangle\vert\dot{0},-\dot{1}\rangle$	$\vert\bar{1},-\dot{2}\rangle\vert\bar{0},-\dot{1}\rangle\vert\dot{0},-\bar{1}\rangle\vert\dot{1},-\bar{2}\rangle$	$\vert\bar{1},-\bar{2}\rangle\vert\bar{0},-\bar{1}\rangle$	8
	-2	$\vert\dot{0},-\dot{2}\rangle$	$\vert\bar{0},-\dot{2}\rangle\vert-\dot{1},-\bar{1}\rangle\vert\dot{0},-\bar{2}\rangle$	$\vert\bar{0},-\bar{2}\rangle$	5
	-3	$\vert-\dot{1},-\dot{2}\rangle$	$\vert-\bar{1},-\dot{2}\rangle\vert\dot{1},-\bar{2}\rangle$	$\vert-\bar{1},-\bar{2}\rangle$	4
	-4		$\vert-\dot{2},-\bar{2}\rangle$		1

nd^2		L					deg.
coupled		0	1	2	3	4	100
S	0	${}^1S_0$	${}^1P_1$	${}^1D_2$	${}^1F_3$	${}^1G_4$	25
	1	${}^3S_1$	${}^3P_2, {}^3P_1, {}^3P_0$	${}^3D_3, {}^3D_2, {}^3D_1$	${}^3F_4, {}^3F_3, {}^3F_2$	${}^3G_5, {}^3G_4, {}^3G_3$	75

Like in the case of the np^2 configuration we proceed by analyzing the nd^2 configuration in the *coupled representation* of the total orbital angular momentum $\mathbf{L} = \boldsymbol{l}_1 + \boldsymbol{l}_2$ coupled with the total spin $\mathbf{S} = \boldsymbol{s}_1 + \boldsymbol{s}_2$ into the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$. In a potential field of central symmetry J is a good quantum number. For the nd^2 configuration the orbital angular quantum number can take the values $L = 0, 1, 2, 3, 4$ with possible spin $S = 0, 1$. In the coupled representation $\{|JM_J\rangle\}$ again a total of 100 atomic states can be identified. In the LS-term notation these 100 states are also given in Table 9.5.

9.4.3.1 Determination of LS terms consistent with Pauli principle

Next we have to select the LS terms consistent with the Pauli principle. We start again with the terms of the highest multiplicity; i.e., with ${}^3G$. Since $L = 4$ and $S = 1$ the ${}^3G$ manifold must include the Slater determinant $|\dot{2}, \dot{2}\rangle$, which corresponds to $M_L = 4$ and $M_S = 1$. But this determinant violates the Pauli principle and therefore has to be excluded. Continuing with the triplet manifolds, we turn to the ${}^3F$ terms, where $L = 3$ and $S = 1$. Therefore, the ${}^3F$ manifold must include the Slater determinant $|2, 1\rangle$, which corresponds to $M_L = 3$ and $M_S = 1$. Because this determinant is consistent with the Pauli principle it must be within the fully anti-symmetric part of the Hilbert space. Further, because the ${}^3F$ terms represent a 21-fold degenerate manifold, also 21 linearly independent determinants are required to represent the ${}^3F$ sector of Hilbert space for all combinations of $M_S = 1, 0, -1$ and $M_L = 2, 1, 0, -1, -2$.

In Table 9.6 the determinants are assigned. Next we proceed with the ${}^3D$ term, where $L = 2$ and $S = 1$. It has to be excluded because there is only one linearly independent Slater determinant with $M_S = 1$ and $M_L = 2$ consistent with the Pauli principle and this one has to be assigned to the ${}^3F$ manifold. The next manifold is ${}^3P$, which can be accepted because it must include the Slater determinant $|\dot{1}, \dot{0}\rangle$ and we have two determinants with $M_S = 1$ and $M_L = 1$ consistent with the Pauli principle, one was already assigned to the ${}^3F$ manifold, but the other serves our purpose. To represent the full ${}^3P$ manifold 9 linearly independent determinants are required. The ${}^3S$ has to

Table 9.6: Decision table using the known LS-term degeneracy to determine the spin-orbit Terms consistent with the Pauli principle for two equivalent d-electrons .

M_S	M_L	Slater determinants [†]		*LS* terms		
		Pauli allowed:		options	accept	excl.
1	4	$\upharpoonleft\!\upharpoonleft - - - -$	0	^3G	-	^3G
	3	$\upharpoonleft\,\upharpoonleft - - -$	1	^3F	^3F	-
	2	$\upharpoonleft - \upharpoonleft - -, - \upharpoonleft\!\upharpoonleft - - -$	1	$^3F, {}^3D$	-	^3D
	1	$\upharpoonleft - - \upharpoonleft -, - \upharpoonleft\,\upharpoonleft - -$	2	$^3F, {}^3P$	^3P	-
	0	$\upharpoonleft - - - \upharpoonleft, - \upharpoonleft - \upharpoonleft -, - - \upharpoonleft\!\upharpoonleft - -$	2	$^3F, {}^3P, {}^3S$	-	^3S
0	4	$\upharpoonleft\!\downharpoonright - - - -$	1	^1G	^1G	-
	3	$\upharpoonleft\,\downharpoonright - - -\ (2\times)$	2	$^3F, {}^1G, {}^1F$	-	^1F
	2	$\upharpoonleft - \downharpoonright - -\ (2\times), - \upharpoonleft\!\downharpoonright - - -$	3	$^3F, {}^1G, {}^1D$	^1D	-
	1	$\upharpoonleft - - \downharpoonright -\ (2\times), - \upharpoonleft\,\downharpoonright - -\ (2\times)$	4	$^3F, {}^3P, {}^1G, {}^1D, {}^1P$	-	^1P
	0	$\upharpoonleft - - - \downharpoonright\ (2\times), - \upharpoonleft - \downharpoonright -\ (2\times), - - \upharpoonleft\!\downharpoonright - -$	5	$^3F, {}^3P, {}^1G, {}^1D, {}^1S$	^1S	-

[†] The sliced-arrow notation has no significance other than being compact.

be excluded because only two Slater determinants with $M_S = 1$ and $M_L = 0$ are consistent with the Pauli principle and these have to be assigned to the ^3F and the ^3P manifolds. We proceed by turning to the singlet manifolds, first the ^1G term ($L = 4; S = 0$). This 9-fold degenerate manifold must include the Slater determinant $|\dot{2}, \bar{2}\rangle$, which is allowed by the Pauli principle. Thus the ^1G sector must be part of the anti-symmetric part of Hilbert space spanned by 9 linearly independent Slater determinants corresponding to all combinations of $M_S = 0$ and $M_L = 2, 1, 0, -1, -2$. This procedure is continued until we decided on all LS terms. As illustrated in Table 9.6 only the terms $^3F_{4,3,2}$, $^3P_{2,1,0}$, $^1G_4$, $^1D_2$ and $^1S_0$ are consistent with the Pauli principle. Adding their degeneracies we recover the 45-dimensional Hilbert space of antisymmetric states identified above.

9.4.3.2 The second Hund rule

The relative energies of the LS terms of the nd^2 configuration may be estimated using the guidelines given for the np^2 configuration. All terms identified in Table 9.2 have the same energy in the screened Coulomb field of the nucleus

$$E_c(nd^2) = -2 \times hcR_\infty Z_{nd}^2/n^2, \tag{9.31}$$

where Z_{nd} is to be determined experimentally, or theoretically by numerical integration of the radial Schrödinger equation for a d-electron ($l = 2$) in the screened Coulomb field of the nucleus. The electrostatic repulsion between the electrons is lowest for the state with the largest average separation. In view of Pauli exclusion of close approach this is the case for terms of maximum S (*Hund's first rule*).

The decision procedure of Table 9.6 resulted for the nd^2 configuration in *two* terms with $S = 1$: ^3F and ^3P. These terms differ in the coupling of the orbital angular momenta of the d electrons, parallel for the $^3F_{4,3,2}$ terms and antiparallel for the $^3P_{2,1,0}$ terms. If the angular momenta are in the same direction the electrons meet less than if they rotate in opposite direction. Therefore, the electrostatic repulsion is less in the state of maximum L. This exposes the physics behind the *second Hund rule*:

> The *Coulomb repulsion* between the valence electrons of an atomic ground state configuration (of maximal Pauli-allowed S) is *minimal for the maximal (Pauli-allowed) value of the total orbital angular momentum*, L.

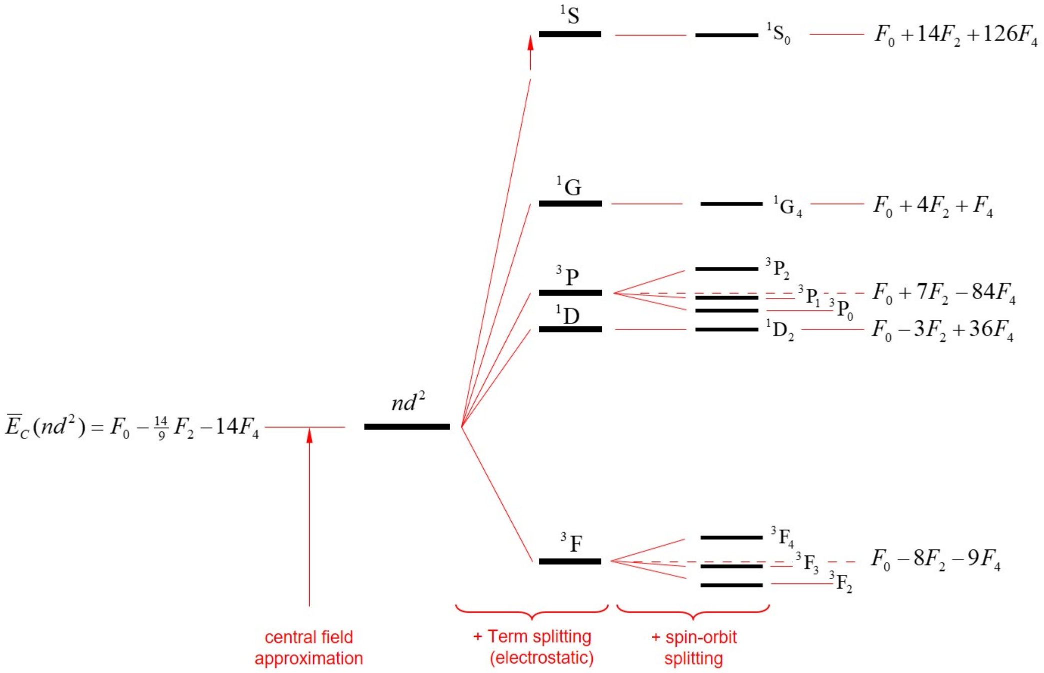


Figure 9.5: Fine-structure diagram of atoms from the titanium group (group 4) as an example of a nd^2 valence configuration.

Thus, the symmetric spin state with the largest angular momentum has the lowest energy. For the nd^2 configuration this corresponds to one of the triplet terms $^3F_{4,3,2}$. The ground state of Ti is indeed a ^3F term: $^3F_2$ (cf. Appendix B). To decide among the terms $^3F_4$, $^3F_3$ and $^3F_2$ we need *Hund's third rule*, which will be discussed in Section 9.5.

Calculation of Term energies: To conclude this section we verify that also the second Hund rule is confirmed by explicit calculation of the energy. Because in zero field the energy is independent of the magnetic quantum numbers we calculate the energy starting from the stretched spin states with the largest L, which can be written as a single nonzero Slater determinant, $|^3F;3,1\rangle = |\dot{2},\dot{1}\rangle_{nd^2}$. In Hartree atomic units the energy shift caused by electrostatic repulsion of this ^3F term is given by

$$E(^3F) = \langle\dot{2},\dot{1}|\frac{1}{\rho_{12}}|\dot{2},\dot{1}\rangle_{nd^2}. \tag{9.32}$$

This is a diagonal matrix element of a two-body operator and using Eq. (7.53) we find

$$E(^3F) = (d2,d1|\rho_{12}^{-1}|d2,d1) - (d1,d2|\rho_{12}^{-1}|d2,d1) = \mathcal{J} - \mathcal{K}, \tag{9.33}$$

where $\mathcal{J} = (d2,d1|\rho_{12}^{-1}|d2,d1)$ and $\mathcal{K} = (d1,d2|\rho_{12}^{-1}|d2,d1)$ are the Coulomb integrals defined by Eqs. (6.46). As we are dealing with equivalent electrons we have $F^k = G^k$ and Eq. (9.32) reduces to

$$E(^3F) = \sum_{k=0,2,4}\left[a^k(d2,d1)F^k(nd^2) - b^k(d2,d1)G^k(nd^2)\right]. \tag{9.34}$$

Substituting the values for $a^k(d2,d1)$ and $b^k(d2,d1)$ for $k=0,2,4$ from Table 6.1 the electrostatic repulsion can be written as the sum of three F integrals (note that $F^k = G^k$ for equivalent electrons),

$$E(^3F) = F_0 - 2F_2 - 4F_4 - 6G_2 - 5G_4 = F_0 - 8F_2 - 9F_4. \tag{9.35}$$

Here, the common denominators of the a^k and b^k coefficients where eliminated by redefining the F integrals: $F_0 \equiv F^0$, $F_2 \equiv F^2/49$ and $F_4 \equiv F^4/441$. The $|^3P;1,1\rangle$ term cannot be written as a single determinant. Thus we turn to the Slater sum rule and using the degeneracy of all magnetic sublevels M_S and M_L of a given LS term in zero field we obtain

$$\left.\begin{aligned} \operatorname{tr}(2\times 2) &= E(^3F) + E(^3P) = F_0 - 8F_2 - 9F_4 + E(^3P) \\ \operatorname{tr}(2\times 2) &= \langle \dot{2}, -\dot{1}|\rho_{12}^{-1}|\dot{2}, -\dot{1}\rangle_{nd^2} + \langle \dot{1}, \dot{0}|\rho_{12}^{-1}|\dot{1}, \dot{0}\rangle_{nd^2} = 2F_0 - F_2 - 93F_4 \end{aligned}\right\} \rightarrow$$

$$\rightarrow E(^3P) = F_0 + 7F_2 - 84F_4. \quad (9.36)$$

Thus $E(^3P) - E(^3F) = 15F_2 - 75F_4$ and the second Hund rule is satisfied if

$$E(^3P) > E(^3F) \;\Leftrightarrow\; 5F_4/F_2 = (245/441)\,F^4/F^2 < 1. \quad (9.37)$$

This is the case because $F^4 < F^2$ for a given screened potential - see Eq. (6.70).

9.4.4 Metastable excited state configurations

Before proceeding to the Hund's rule 3 we first have a look at some examples of excited states. Although the focus of the present chapter is on the determination of the electronic ground states the general formalism introduced in Chapter 7 can be applied to calculate the LS term energies of any electron configuration; i.e., also the energies of excited state configurations. In this section we shall demonstrate this for the $nsn's$ and $nsn'p$ configurations. In this way we can practice the procedure of Section 9.4 to calculate the Coulomb integrals for many-electron configurations. Thus, we consider the low-lying states of helium-like atoms; in particular, the $^1S_0$ and $^3S_1$ states of helium and the $^1P_1$ and $^3P_{0,1,2}$ states of the alkaline-earth atoms. The triplets terms are of special significance because they can represent metastable atomic states.

9.4.4.1 Metastable helium (He*)

Let us first convince ourselves that the general formalism leads to the same results as derived in Chapter 6 for the excited states of helium. The configuration of the ground state is $1s^2$. Exciting one of the $1s$ electrons to the $2s$ level we obtain the $1s2s$ configuration, which is compatible with $^1S_0$ (non-degenerate) and $^3S_1$ (three-fold degenerate) angular momentum terms, which are both Pauli allowed (because the electrons are non-equivalent). To calculate the electrostatic repulsion energy we turn to the uncoupled representation for which the $1s2s$ subspace can represented by the Slater determinants given in Table 9.7. As all $^3S_1$ levels are degenerate we can calculate the electrostatic repulsion energy in the $^3S_1$ term by by choosing the determinant of the spin-stretched state,

$$E(^3S) = \langle 1\dot{s}, 2\dot{s}|\rho_{12}^{-1}|1\dot{s}, 2\dot{s}\rangle = F_0 - G_0, \quad (9.38)$$

Table 9.7: Example $nsn's$ configuration: (top) uncoupled representation - the 4 linearly independent pair states consistent with the Pauli principle; (bottom) coupled representation.

$nsn's$		M_S			deg.
uncoupled		1	0	-1	
M_L	0	$\lvert n\dot{s}, n'\dot{s}\rangle$	$\lvert n\dot{s}, n'\bar{s}\rangle$ $\lvert n\bar{s}, n'\dot{s}\rangle$	$\lvert n\bar{s}, n'\bar{s}\rangle$	4

$\lvert ^{2S+1}L; M_L, M_S\rangle$		M_S			deg.
coupled		1	0	-1	
M_L	0	$\lvert ^3S;1,0\rangle$	$\lvert ^3S;0,0\rangle$ $\lvert ^1S;0,0\rangle$	$\lvert ^3S;-1,0\rangle$	4

Table 9.8: Example *nsnp* configuration: in the uncoupled representation we have 12 linearly independent pair states, all consistent with the Pauli principle

nsnp		M_S			deg.
uncoupled		1	0	-1	12
M_L	1	$\lvert n\dot{s}, n\dot{p}_1\rangle$	$\lvert n\dot{s}, n\bar{p}_1\rangle$ $\lvert n\bar{s}, n\dot{p}_1\rangle$	$\lvert n\bar{s}, n\bar{p}_1\rangle$	4
	0	$\lvert n\dot{s}, n\dot{p}_0\rangle$	$\lvert n\dot{s}, n\bar{p}_0\rangle$ $\lvert n\bar{s}, n\dot{p}_0\rangle$	$\lvert n\bar{s}, n\bar{p}_0\rangle$	4
	-1	$\lvert n\dot{s}, n\dot{p}_{-1}\rangle$	$\lvert n\dot{s}, n\bar{p}_{-1}\rangle$ $\lvert n\bar{s}, n\dot{p}_{-1}\rangle$	$\lvert n\bar{s}, n\bar{p}_{-1}\rangle$	4

with $F_0 = F^0$ and $G_0 = G^0/3$. The electrostatic repulsion energy in the ${}^1S_0$ term follows most conveniently with the Slater sum rule,

$$E({}^3S) + E({}^1S) = \mathrm{tr}\,(2\times 2) = \langle 1\dot{s}, 2\bar{s}|\rho_{12}^{-1}|1\dot{s}, 2\bar{s}\rangle + \langle 1\bar{s}, 2\dot{s}|\rho_{12}^{-1}|1\bar{s}, 2\dot{s}\rangle = 2F_0. \tag{9.39}$$

Thus, the electrostatic repulsion energy in the ${}^1S_0$ term is

$$E({}^1S) = \mathrm{tr}\,(2\times 2) - E({}^3S) = F_0 + G_0. \tag{9.40}$$

Comparing this result with the result of Section 6.3 we identify $\mathcal{J} = F_0$ and $\mathcal{K} = G_0$ (see Fig. 6.4). The average energy of the ${}^1S_0$ - ${}^3S_1$ manifold is

$$\bar{E}(1s, 2s) = [3(F_0 - G_0) + (F_0 + G_0)]/4 = F_0 - G_0/2. \tag{9.41}$$

9.4.4.2 Alkaline earth elements - lowest energy P manifolds

To discuss the low-lying excited states of the alkaline-earth elements we consider in particular the group 2 atoms, which all have a ground state configuration [inertgas]ns^2. The lowest-exited state configuration is [inertgas]$nsnp$, which is compatible with ${}^1P_1$ (three-fold degenerate) and ${}^3P_{0,1,2}$ (nine-fold degenerate) angular momentum terms, which are all Pauli allowed. The calculation of the term energies goes analogously to the case of helium. First we do the decomposition into Slater determinants, see Table 9.8. As all ${}^3P$ levels are degenerate we can calculate the electrostatic repulsion energy in the ${}^3P$ term by choosing the determinant of the fully stretched (spin and orbit) state,

$$E({}^3P) = \langle n\dot{s}, n\dot{p}_1|\rho_{12}^{-1}|n\dot{s}, n\dot{p}_1\rangle = F_0 - G_1, \tag{9.42}$$

with $F_0 = F^0$ and $G_1 = G^1/3$. Using the Slater sum rule,

$$E({}^3P) + E({}^1P) = \mathrm{tr}\,(2\times 2) = \langle n\dot{s}, n\bar{p}_0|\rho_{12}^{-1}|n\dot{s}, n\bar{p}_0\rangle + \langle n\bar{s}, n\dot{p}_0|\rho_{12}^{-1}|n\bar{s}, n\dot{p}_0\rangle = 2F_0, \tag{9.43}$$

we find for the electrostatic repulsion in the ${}^1S$ term

$$E({}^1P) = \mathrm{tr}\,(2\times 2) - E({}^3P) = F_0 + G_1. \tag{9.44}$$

Hence, the structure of the shift diagram is like that of metastable helium but the exchange integral is a different one, G_1 rather than G_0 (see Fig. 9.6). The average energy of the ${}^1P_1$ - ${}^3P_{0,1,2}$ manifold is

$$\bar{E}(n\dot{p}, n\dot{s}) = [9(F_0 - G_1) + 3(F_0 + G_1)]/12 = F_0 - G_1/2. \tag{9.45}$$

Since the Coulomb splitting is $2G_1$, the value of G_1 can be readily determined by atomic spectroscopy. From the data used for the level diagram 6.8 we can immediately determine the exchange integrals for the alkaline earth elements: Be: $G_1(2s, 2p) = 1.28$ eV, Mg: $G_1(3s, 3p) = 0.82$ eV, Ca: $G_1(4s, 4p) = 0.52$ eV, Sr: $G_1(5s, 5p) = 0.43$ eV, Ba: $G_1(6s, 6p) = 0.31$ eV. Note that the value drops by a factor 4 along the series which reflects the decreasing importance of the electrostatic interaction between the outer orbitals for growing atomic size (see also Table 9.9).

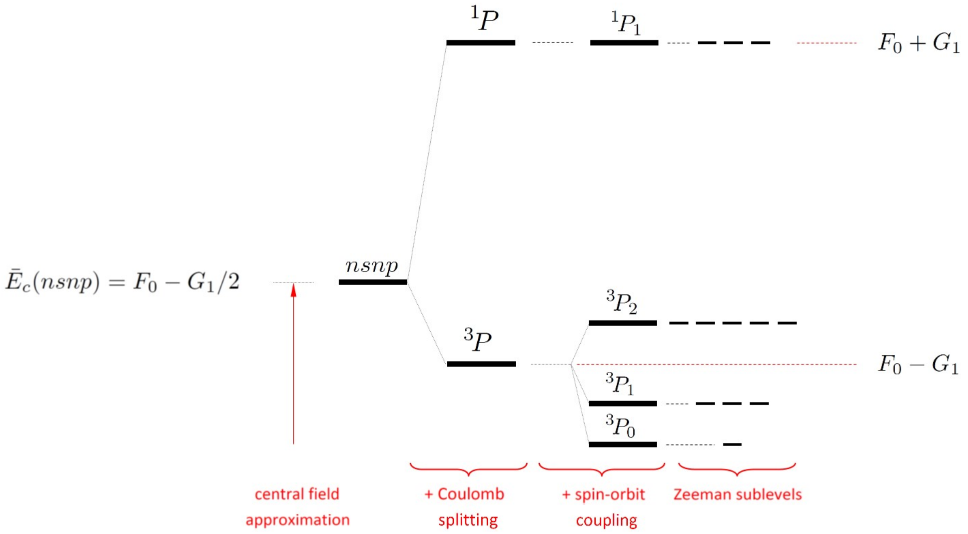


Figure 9.6: Coulomb splitting of $nsnp$ excited-state configuration of the alkaline-earth atoms (compare with Fig. 6.8).

Problem 9.1. Show that the central field contribution to the electrostatic energy of the $1s^22s2p$ excited state configuration of *beryllium* (see Fig. 9.6) is given by

$$\bar{E}_c(1s^22s2p) = 2I(1s) + I(2s) + I(2p) + F_0(1s,1s) + 2F_0(2s,1s) - G_0(2s,1s) \\ + 2F_0(2p,1s) - G_1(2p,1s) + \bar{E}(2p,2s).$$

Solution. The energy of the $1s^22s2p$ excited state configuration of *beryllium* can be written as

$$\bar{E}_c(1s^22s2p) = 2I(1s) + I(2s) + I(2p) + E_{1s^2} + E_{1s^2}(2s) + E_{1s^2}(2p) + \bar{E}(2p,2s),$$

where $\bar{E}(n\dot{p},n\dot{s})$ is energy average (9.45). Further we have $E_{1s^2} = F_0(1s^2)$, $E_{1s^2}(2s) = 2F_0(2s,1s) - G_0(2s,1s)$ and $E_{1s^2}(2p) = 2F_0(2p,1s) - G_1(2p,1s)$, where $F_0 = F^0$, $G_0 = G^0$ and $G_1 = G^1/3$. Collecting the terms we obtain the desired expression. □

9.5 Fine structure - Hund's rule 3

9.5.1 Zeeman interaction

The Zeeman Hamiltonian of a many-electron atom is given by the sum of the Zeeman contributions of the individual electrons,

$$\mathcal{H}_Z = \sum\nolimits_i [g_L(e/2m_e)\boldsymbol{l}_i + g_e(e/2m_e)\boldsymbol{s}_i] \cdot \mathbf{B}, \tag{9.46}$$

where $g_L \simeq 1$ is the effective g factor of the orbital motion and $g_e \simeq 2$ the anomalous g factor of the electron spin. In the absence of spin-orbit coupling the total orbital and spin angular momenta

$$\mathbf{L} = \sum_i \boldsymbol{l}_i \quad \text{and} \quad \mathbf{S} = \sum_i \boldsymbol{s}_i \tag{9.47}$$

are conserved quantities. Using the definition of the *Bohr magneton* $\mu_B = e\hbar/2m_e = 9.274\,0154(31) \times 10^{-24}$ J/T the Zeeman Hamiltonian can be written in the form

$$\mathcal{H}_Z = [g_L\mu_B\,(\mathbf{L}/\hbar) + g_e\mu_B\,(\mathbf{S}/\hbar)] \cdot \mathbf{B}. \tag{9.48}$$

Choosing the quantization axis (z axis) along the magnetic field direction this expression becomes

$$\mathcal{H}_Z = [g_L L_z + g_e S_z]\mu_B B/\hbar. \tag{9.49}$$

This Hamiltonian is diagonal in the uncoupled representation $\{|nLM_LSM_S\rangle\}$. The Zeeman shift is given by

$$\Delta E_Z = \langle nLM_LSM_S|\mathcal{H}_Z|nLM_LSM_S\rangle = [g_L M_L + g_e M_S]\mu_B B. \tag{9.50}$$

9.5.2 Spin-orbit interaction

In this section we return to spin-orbit coupling, introduced in Section 4.4 for hydrogen-like atoms. Now it is our task is to generalize this subject to the case of many-electron atoms. The analysis starts with the presence of a velocity-induced magnetic field whenever a particles is moving transversely to an electric field. In particular, electrons orbiting at velocity $\mathbf{v}$ in the electric field $\mathbf{E}$ of the partially shielded nucleus experience such a velocity-induced magnetic field, $\mathbf{B} = (\mathbf{E}\times\mathbf{v})/c^2$, where c is the speed of light (cf. Appendix E). The *spin-orbit coupling* is the coupling of the spin magnetic moment of the electron to this magnetic field. In the central-field approximation the electric field at the position of the electron is given by

$$\mathbf{E} = -\boldsymbol{\nabla}\varphi_{\mathrm{CF}}(r) = -\frac{\mathbf{r}}{r}\frac{d\varphi_{\mathrm{CF}}(r)}{dr}, \tag{9.51}$$

where $\varphi_{\mathrm{CF}}(r)$ is the central-field scalar potential. Including the *Thomas precession* the spin-orbit field takes the form

$$\mathbf{B}_{LS} \simeq -\frac{1}{2m_e c^2}\frac{1}{r}\frac{d\varphi_{\mathrm{CF}}(r)}{dr}\boldsymbol{l}, \tag{9.52}$$

where $\boldsymbol{l} = m_e(\mathbf{r}\times\mathbf{v})$ is the orbital angular momentum operator of the electron. The Hamiltonian describing the *spin-orbit interaction* of an individual electron is

$$\mathcal{H}_{ls} = g_e(e/2m_e)\boldsymbol{s}\cdot\mathbf{B}_{LS} = \xi(r)\boldsymbol{l}\cdot\boldsymbol{s}. \tag{9.53}$$

The function $\xi(r)$ is the *coupling strength*,

$$\xi(r) \simeq -\frac{1}{2m_e^2c^2}\frac{e}{r}\frac{d\varphi_{\mathrm{CF}}(r)}{dr}. \tag{9.54}$$

As long as the screened Coulomb potential, $\varphi_{\mathrm{CF}}(r)$, is a monotonically decreasing function of r, the coupling strength has to be positive,

$$\xi(\rho) \simeq -\frac{1}{2m_e^2c^2}\frac{e^2}{a^2}\frac{1}{\rho}\frac{d[Z_{\mathrm{scr}}(\rho)/\rho]}{d\rho} \equiv \frac{1}{2m_e^2c^2}\frac{e^2}{a^3}\frac{Z(\rho)}{\rho^3} > 0\,, \tag{9.55}$$

where $Z(\rho) \equiv Z_{\mathrm{scr}}(\rho) - \rho dZ_{\mathrm{scr}}(\rho)/d\rho$ is the spin-orbit screening function introduced in Section 4.4.2.

Summing the contributions of individual electrons we arrive at the Hamiltonian for the total spin-orbit interaction of the atom

$$\mathcal{H}_{LS} = \sum_{i=1}^{N}\xi_i(r)\boldsymbol{l}_i\cdot\boldsymbol{s}_i, \tag{9.56}$$

where $i \in \{1,\cdots N\}$ is the particle index of a system of N electrons.

9.5.3 Coupling schemes: LS coupling versus jj coupling

Adding the spin-orbit interaction to the central field Hamiltonian (8.7) we obtain the Hamiltonian

$$\mathcal{H} = \mathcal{H}_{\mathrm{CF}} + \mathcal{H}_{LS},$$

which suffices to calculate the ground states of most many-electron atoms. Recalling that closed subshells of equivalent electrons contribute zero angular momentum to the atom, the spin-orbit contribution can be calculated by restricting the particle-index summation in Eq. (9.56) to the valence electrons. In one-electron atoms this is particularly simple. The spin-orbit coupling results from the coupling of the spin $\boldsymbol{s}$ of the only valence electron with its own orbital angular momentum, $\boldsymbol{l}$. In Section 4.5.3 we found that the strength of this coupling scales with Z^4 for hydrogenic atoms in the weakly relativistic limit; for the alkalies we found it to increase by more than three orders of magnitude (in spite of the screening) when comparing *cesium* with *lithium* (see Section 4.6, in particular Table 4.3). It is straightforward to extend this analysis to one-electron atoms with one p electron (group III) or one d electron (group 3) in the ground state. Matters change when turning to two-electron atoms, two p electrons (group IV) or two d electrons (group 4) or to three-electron atoms, three d electrons (group 5) or three f electrons (*praseodymium*). In these cases not only the coupling strength but also the coupling scheme depends on Z but, whatever the coupling scheme, the total electronic angular momentum $\mathbf{J}$ is conserved (neglecting hyperfine structure).

- For small Z the electrostatic splitting between the LS terms is larger than the fine-structure splitting within the terms. In this case the various $\boldsymbol{l}_i$ couple into a total orbital angular momentum $\mathbf{L}$ and the $\boldsymbol{s}_i$ into a total spin $\mathbf{S}$, and these couple into the total electronic angular momentum
$$\mathbf{J} = \mathbf{L} + \mathbf{S}. \tag{9.57}$$
This coupling is known as LS coupling or *Russell-Saunders coupling.* In the presence of strong Russell-Saunders coupling the degeneracy of the LS terms is lifted by the spin-orbit interaction into levels differing in J. As long as this splitting remains small it can be calculated by perturbation theory with respect to the desired LS term. The Russell-Saunders coupling scheme applies typically to the atoms in the upper part of the periodic system.

- For large Z the spin-orbit interaction of the individual electrons dominates over the electrostatic coupling. This gives rise to a primary coupling into total angular momenta of the individual electrons, $\boldsymbol{j}_i = \boldsymbol{l}_i + \boldsymbol{s}_i$, and a subsequent coupling into the total electronic angular momentum of the atom,
$$\mathbf{J} = \sum_i \boldsymbol{j}_i.$$
This coupling scheme is known as jj coupling. The jj coupled levels differ in spin-orbit coupling energy. As long as the electrostatic interaction energy is small as compared to the splitting of the J levels it can be calculated by perturbation theory with respect to the jj coupled level. This coupling tends to reduce the spin-orbit splitting until the Russell-Saunders limit is reached in which all J levels are degenerate. The jj coupling scheme applies to the heaviest atoms of the periodic table.

- For the middle part of the periodic system the coupling is said to be *intermediate* between LS and jj coupling.

9.5.4 Russell-Saunders coupling

In this course we mostly restrict ourselves to LS coupling and calculate the spin-orbit shift in the *coupled representation* with first-order perturbation theory. This is allowed as long as the spin-orbit splitting is much smaller than the LS-term splitting. This approximation is valid except in cases of *configuration interaction.* For a given LS term, $|LSJM_J\rangle$, the *spin-orbit shift* is given by

$$\Delta E_{n,J}^{LS} = \langle nLSJM_J|\mathcal{H}_{LS}|nLSJM_J\rangle, \tag{9.58}$$

where it remains to be shown that the matrix is diagonal. Decomposing the LS term to the *uncoupled representation* we obtain

$$\Delta E_{n,J}^{LS} = \sum_{M_L,M_L'} \sum_{M_S,M_S'} \langle nLSJM_J|nLM_L'SM_S'\rangle \langle nLM_L'SM_S'|\mathcal{H}_{LS}|nLM_LSM_S\rangle\langle nLM_LSM_S|nLSJM_J\rangle. \tag{9.59}$$

Because the operators $\boldsymbol{l}_i$ operate on the orbital states and $\boldsymbol{s}_i$ on the spin states we can factor out the spin part from the orbital part, and the matrix elements take the form

$$\langle nLM_L'SM_S'|\mathcal{H}_{LS}|nLM_LSM_S\rangle = \sum_i \langle nLM_L'SM_S'|\xi_i(r)\boldsymbol{l}_i\cdot\boldsymbol{s}_i|nLM_LSM_S\rangle \tag{9.60}$$

$$= (\zeta_{nl}/\hbar^2)\sum_i \langle LM_L'|\boldsymbol{l}_i|LM_L\rangle\cdot\langle SM_S'|\boldsymbol{s}_i|SM_S\rangle, \tag{9.61}$$

where the summation runs over all equivalent electrons in the partially filled shell and ζ_{nl} is the *spin-orbit coupling constant* of the electrons in the shell under consideration. Recalling the Wigner-Eckart theorem we know that the components of $\boldsymbol{l}_i$ and $\boldsymbol{s}_i$ are proportional to the corresponding components of $\mathbf{L}$ and $\mathbf{S}$,

$$\langle LM_L'|\boldsymbol{l}_i|LM_L\rangle = \langle L||l_i||L\rangle\langle LM_L'|\mathbf{L}|LM_L\rangle \tag{9.62a}$$

$$\langle SM_S'|\boldsymbol{s}_i|SM_S\rangle = \langle S||s_i||S\rangle\langle SM_S'|\mathbf{S}|SM_S\rangle. \tag{9.62b}$$

Introducing the *Russell-Saunders coupling constant*,

$$\zeta_{nLS} = \zeta_{nl}\sum_i \langle L||l_i||L\rangle\langle S||s_i||S\rangle, \tag{9.63}$$

the Russell-Saunders matrix element becomes

$$\langle nLM_L'SM_S'|\mathcal{H}_{LS}|nLM_LSM_S\rangle = \zeta_{nLS}/\hbar^2\langle LM_L'SM_S'|\mathbf{L}\cdot\mathbf{S}|LM_LSM_S\rangle. \tag{9.64}$$

With the aid of the projection theorem the reduced matrix element $\langle L||l_i||L\rangle$ can be written in the form

$$\langle L||l_i||L\rangle = \frac{\langle LM_L|\boldsymbol{l}_i\cdot\mathbf{L}|LM_L\rangle}{\langle LM_L|\mathbf{L}^2|LM_L\rangle}. \tag{9.65}$$

A similar expression for $\langle S||s_i||S\rangle$ follows analogously. Returning to Eq. (9.64) and substituting this expression into Eq. (9.59) we obtain (after transforming back to the *coupled representation*) the following expression for the spin-orbit shift

$$\Delta E_{n,J}^{LS} = \left(\zeta_{nLS}/\hbar^2\right)\langle LSJM_J|\mathbf{L}\cdot\mathbf{S}|LSJM_J\rangle. \tag{9.66}$$

Here we used the inner product rule (3.109a) to establish that, in Russell-Saunders coupling, the spin-orbit interaction is indeed *diagonal* in the coupled representation. For the spin-orbit shift we find,

$$\Delta E_{n,J}^{LS} = \tfrac{1}{2}\zeta_{nLS}\left[J(J+1) - L(L+1) - S(S+1)\right]. \tag{9.67}$$

For positive coupling constant ($\zeta_{nLS} > 0$) the LS term with the lowest J has the lowest energy. Likewise, for negative coupling constant ($\zeta_{nLS} < 0$) the lowest energy is obtained for the highest value of J. With Eq. (9.67) we have obtained for the case of many-electron atoms the same expression for the spin-orbit shift as derived in Chapter 4 for hydrogen-like atoms. In the present context this expression is valid for the case of LS coupling of arbitrary L and S.

Table 9.9: Fine-structure splitting ratio $\Delta W_2/\Delta W_1$ of ${}^3P$ terms for increasing value of Z. For pure Russell-Saunders coupling this ratio is 2. Deviations from this value point to the presence of other coupling mechanisms. In the absence of nuclear spin ($I = 0$) and for large Z this points to a crossover to jj-coupling. This is confirmed by comparing the decrease in Coulomb integral G_1 (see Section) with the increase in fine-structure splitting ΔW_1.

element	Be	Mg	Ca	Sr	Ba	Ra	unit
Z	4	12	20	38	56	88	
G_1	1.28	0.82	0.52	0.43	0.31	0.33	eV
ΔW_1	8×10^{-5}	0.002	0.006	0.023	0.046	0.114	eV
$\Delta W_2/\Delta W_1$	3.6357	2.0297	2.0298	2.1100	2.3695	2.9200	

9.5.4.1 Shift rules for Russell-Saunders coupled angular momenta

As introduced in Section 4.5.2 expressions of the type (9.67) imply *three shift rules* for the spin-orbit manifolds of given J, L and S:

- the shift of the *stretched state* ($J = J_{\max} = L + S$) is given by

$$\Delta E^{LS}_{n,J_{\max}} = \zeta_{nLS} LS. \tag{9.68}$$

- the relative shift of two adjacent LS-coupled manifolds follows the *Landé interval rule*:

$$\Delta W_J = \Delta E^{LS}_{n,J} - \Delta E^{LS}_{n,J-1} = \zeta_{nLS} J, \tag{9.69}$$

  where J is the quantum number of the manifold with the highest multiplicity *of the pair.*

- the weighted mean of the energies of *all manifolds* of a given LS coupling scheme satisfies the *center-of-gravity rule*:

$$\frac{1}{(2L+1)(2S+1)} \sum_{J=|L-S|}^{L+S} (2J+1)\,\Delta E^{LS}_{n,J} = 0, \tag{9.70}$$

  where $(2L+1)(2S+1)$ is the degeneracy of the unperturbed level.

The shift rules are illustrated for the case $L = 1$, $S = 1$ in Fig. 9.7. In the absence of hyperfine interaction ($I = 0$), breakdown of the interval rule is an indicator for the breakdown of pure LS coupling. For ${}^3P$ terms this shows up as a deviation of the pure LS splitting ratio $\Delta W_2/\Delta W_1 = 2$. In Table 9.9 this is illustrated for the Z dependence of $\Delta W_2/\Delta W_1$ for the lowest ${}^3P$ terms of the alkaline-earth elements (see also Fig. 6.8).

9.5.5 Equivalence of electrons and holes

Let us consider a closed nl shell and remove one of the electrons. This leaves us with a partially filled shell occupied by $2(2l+1)-1$ electrons, total orbital angular momentum $\mathbf{L}$ and total spin $\mathbf{S}$. As closed shells have *zero* angular momentum, we know that the angular momenta of the partially filled shell and that of the vacant electron have to add up to zero, $\mathbf{L} + \boldsymbol{l} = 0$ and $\mathbf{S} + \boldsymbol{s} = 0$. In other words, the angular momenta of the partially filled shell are opposite to those of the electron removed,

$$\mathbf{L} = -\boldsymbol{l} \quad \text{and} \quad \mathbf{S} = -\boldsymbol{s}. \tag{9.71}$$

This simple result of what could have been a serious calculation (see Section 9.4) suggests to assign properties to a lacking electron. The configuration behaves like a closed shell in which an imaginary

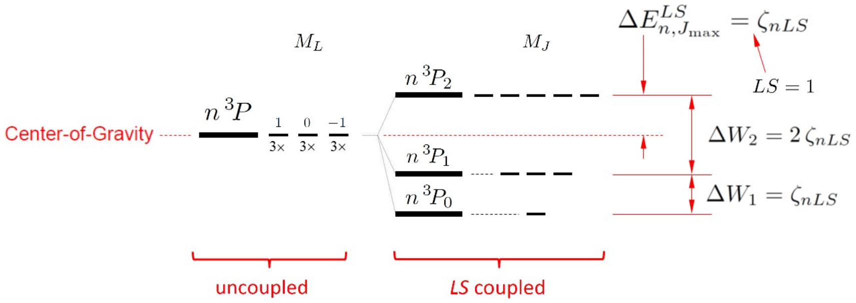


Figure 9.7: The shift rules for spin-orbit coupling in the Russell-Saunders limit plotted for the example $L = 1$, $S = 1$. Triplets of this type are observed in the ground state of low-Z group-12 atoms like *carbon* and in $ns \rightarrow np$ excited states of low-Z helium-like atoms like *calcium*.

particle, called the *hole*, cancels all properties of the electron filling the vacancy. So, rather than *removing* an electron from the shell we *add* a hole. The hole has to be always at the position of the vacancy filling electron (i.e., move at the same velocity) and have opposite charge ($q = +e$) and mass ($m_h = -m_e$). Apparently the hole has orbital angular momentum $\boldsymbol{l}_h \equiv \mathbf{L} = -\boldsymbol{l}$ and spin $\boldsymbol{s}_h \equiv \mathbf{S} = -\boldsymbol{s}$. In search for more properties of the hole we return to the nl shell with one vacancy. Rewriting Eq. (9.66) with the aid of Eq. (9.63) we obtain for the spin-orbit shift

$$\Delta E^{LS}_{n,J} = \zeta_{nl} \sum_{i=1}^{2(2l+1)-1} \langle L||l_i||L\rangle\langle S||s_i||S\rangle\langle nLSJM_J|\mathbf{L}\cdot\mathbf{S}|nLSJM_J\rangle, \tag{9.72}$$

Adding and subtracting the contribution of the lacking electron and using Eqs. (9.71) we find for the spin-orbit shift of a shell with a single hole

$$\Delta E^{LS}_{n,J} = -\left(\zeta_{nl}/\hbar^2\right)\langle nLSJM_J|\boldsymbol{l}\cdot\boldsymbol{s}|nLSJM_J\rangle = -\left(\zeta_{nl}/\hbar^2\right)\langle nLSJM_J|\mathbf{L}\cdot\mathbf{S}|nLSJM_J\rangle, \tag{9.73}$$

where

$$\zeta_{nl} = \langle nl||\xi(r)||nl\rangle\hbar^2 > 0 \tag{9.74}$$

is the spin-orbit coupling constant of a single nl electron moving in the shielded nuclear potential. Comparing Eqs. (9.73) and (9.66) we find

$$\zeta_{nLS} = -\zeta_{nl} = -\langle nl||\xi(r)||nl\rangle\hbar^2. \tag{9.75}$$

In view of Eq. (9.54) this expression represents the coupling constant of a single *positive* charge $q = +e$ moving in a nl orbital through the shielded field of the nuclear charge. For the intrinsic magnetic moment we require

$$\boldsymbol{\mu}_h = g_h(e/2m_h)\boldsymbol{s}_h = g_e(e/2m_e)\boldsymbol{s} = -\boldsymbol{\mu}_s, \tag{9.76}$$

which implies $g_h = g_e$. Since the electron and hole are at the same position $\mathbf{r}$ and move with the same velocity $\mathbf{v}$ (i.e., have opposite momentum) we find

$$\boldsymbol{l}_h = -\boldsymbol{l} = -m_e(\mathbf{r}\times\mathbf{v}) = m_h(\mathbf{r}\times\mathbf{v}). \tag{9.77}$$

9.5.6 Third Hund rule

Eq. (9.75) may be generalized to hold for the *ground state* of a partially filled shell of configuration nl^N. Let us suppose that we are interested in the Russell-Saunders coupling constant ζ_{nLS} of a given LS term of this configuration. Since the coupling constant is independent of J we are free to select the most convenient J manifold to find a relation between ζ_{nLS} and ζ_{nl}. This turns out to be the manifold of the *stretched state*, $J = J_{\max} = L + S$, because in this case the shift rule is particularly simple - see Eq. (9.68),

$$\Delta E^{LS}_{n,J_{\max}} = \langle nLSJ_{\max}M_J|\mathcal{H}_{LS}|nLSJ_{\max}M_J\rangle = \zeta_{nLS}LS. \tag{9.78}$$

Moreover, this is the manifold which contains the pure state, which is the state with $M_J = J$. It is called *pure* because it maps uniquely onto the $M_L = L$, $M_S = S$ sublevel of the uncoupled representation (see Section 4.7.3). In other words, we can equally well calculate the matrix element

$$\Delta E^{LS}_{n,J_{\max}} = \langle nLLSS|\mathcal{H}_{LS}|nLLSS\rangle. \tag{9.79}$$

As this matrix element is diagonal we only have to evaluate the diagonal part of $\mathcal{H}_{LS}$ as is confirmed by the actual calculation,

$$\Delta E^{LS}_{n,J_{\max}} = \langle nLLSS|\zeta_{nLS}L_zS_z|nLLSS\rangle = \zeta_{nLS}LS. \tag{9.80}$$

Next we make the connection with the spin-orbit terms of the individual electrons. For the *ground state* we know that the first two Hund rules are satisfied. This means that the LS term of the ground state carries the largest values of S and L consistent with the Pauli principle. For such terms the subspace with $M_S = S$ and $M_L = L$ can be represented by a *single* Slater determinant (corresponding to the stretched electron - or hole - spin state)

$$|nLLSS\rangle = |m_l^1 m_s^1, \cdots m_l^N m_s^N\rangle_{nl}, \tag{9.81}$$

where

$$M_S = \sum_{\kappa=1}^{N} m_s^\kappa = S \quad \text{and} \quad M_L = \sum_{\kappa=1}^{N} m_l^\kappa = L. \tag{9.82}$$

In terms of the Slater determinants the spin-orbit shift takes the form

$$\begin{aligned}\Delta E^{LS}_{n,J_{\max}} &= \langle m_l^1 m_s^1, \cdots m_l^N m_s^N|\mathcal{H}_{LS}|m_l^1 m_s^1, \cdots m_l^N m_s^N\rangle_{nl} \\ &= \sum_{i=1}^{N}\langle m_l^1 m_s^1, \cdots m_l^N m_s^N|\xi_i(r)\boldsymbol{l}_i\cdot\boldsymbol{s}_i|m_l^1 m_s^1, \cdots m_l^N m_s^N\rangle_{nl}.\end{aligned} \tag{9.83}$$

Hence, we have to evaluate the diagonal matrix elements of the one-body operators $\xi_i(r)\boldsymbol{l}_i\cdot\boldsymbol{s}_i$. With the aid of Eq. (7.45) we obtain

$$\Delta E^{LS}_{n,J_{\max}} = \zeta_{nl}/\hbar^2\sum_{\kappa=1}^{N}\langle lm_l^\kappa sm_s^\kappa|\boldsymbol{l}\cdot\boldsymbol{s}|lm_l^\kappa sm_s^\kappa\rangle = \zeta_{nl}\sum_{\kappa=1}^{N} m_l^\kappa m_s^\kappa. \tag{9.84}$$

Comparing the expressions (9.78) and (9.84) we find the following expression for the coupling constant,

$$\zeta_{nLS} = \frac{\zeta_{nl}}{LS}\sum_{\kappa=1}^{N} m_l^\kappa m_s^\kappa. \tag{9.85}$$

At this point we distinguish between two cases:

a.) *Shells less than half full.* In this case we have $M_S = S = N/2$ (Hund's rule 1), $M_L = L$ (Hund's rule 2), with $m_s^\kappa = \frac{1}{2}$ for $\kappa \in \{1, \cdots N\}$. Substituting these quantum numbers into Eq. (9.85) we obtain

$$\zeta_{nLS} = \frac{1}{2}\frac{\zeta_{nl}}{LS}\sum_{\kappa=1}^{N} m_l^\kappa = \frac{1}{2}\frac{\zeta_{nl}}{LS}M_L = +\frac{\zeta_{nl}}{2S}. \tag{9.86}$$

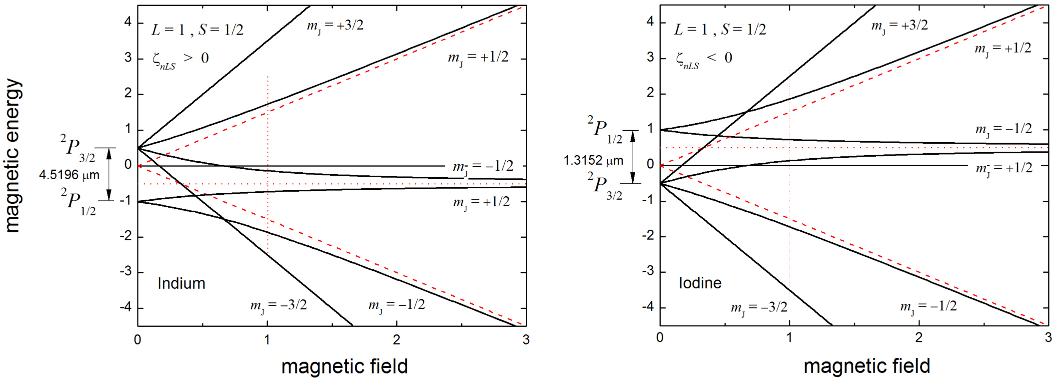


Figure 9.8: Magnetic energy versus magnetic field in units of the fine-structure field B_{fs}. Left: positive spin-orbit shift (one p electron - ground state of Indium). Right: negative spin-orbit shift (one p hole - ground state of Iodine).

b.) *Shells more than half full.* In this case we have $M_S = S = (2l+1) - N/2$ (Hund's rule 1), $M_L = L$ (Hund's rule 2), with $m_s^\kappa = \frac{1}{2}$ for $\kappa \in \{1, \cdots 2l+1\}$ and $m_s^\kappa = -\frac{1}{2}$ for $\kappa \in \{2l+2, \cdots N\}$

$$\zeta_{nLS} = \frac{1}{2}\frac{\zeta_{nl}}{LS}\sum_{\kappa=1}^{2l+1} m_l^\kappa - \frac{1}{2}\frac{\zeta_{nl}}{LS}\sum_{\kappa=2l+2}^{N} m_l^\kappa = -\frac{1}{2}\frac{\zeta_{nl}}{LS}M_L = -\frac{\zeta_{nl}}{2S}. \tag{9.87}$$

The first summation yields zero because $L = 0$ for half-filled shells. The second summation yields $M_L = L$.

Together with Eq. (9.67) the expressions (9.86) and (9.87) provide the scientific underpinning of the *third Hund rule*:

> The *most negative spin-orbit shift* is obtained for the valence electrons of an atomic ground state configuration (of maximal Pauli-allowed S and L) if the *total electronic angular momentum*, J, *is (a) minimal for less-than-half-filled shells*; *(b) maximal for more-than-half-filled shells.*

In Fig. 9.8 the two cases are illustrated by the ground-state fine-structure diagrams of Indium and Iodine. Note that the third Hund rule plays no role in the case of half-filled shells because in this case only a single J value ($J = S$) needs to be considered.

9.6 Zeeman effect in the presence of spin-orbit interaction

9.6.1 Hamiltonian

The Hamiltonian describing the fine-structure of many-electron atoms is given by

$$\mathcal{H} = \mathcal{H}_{\text{CF}} + \mathcal{H}_{LS} + \mathcal{H}_Z, \tag{9.88}$$

where $\mathcal{H}_{\text{CF}}$ is the central-field Hamiltonian (8.7), $\mathcal{H}_Z$ the Zeeman Hamiltonian (9.46) and $\mathcal{H}_{LS}$ the spin-orbit Hamiltonian (9.56). We write this as $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}'$, where

$$\mathcal{H}' = \mathcal{H}_{LS} + \mathcal{H}_Z \tag{9.89}$$

is the perturbation to be considered. Because in weak magnetic fields $\langle\mathcal{H}'\rangle \ll \langle\mathcal{H}_{\text{CF}}\rangle$ we look for a perturbative solution using perturbation theory for degenerate levels. The Hamiltonian $\mathcal{H}_{LS}$ is diagonal in the $\{|nLSJM_J\rangle\}$ (*coupled*) representation and $\mathcal{H}_Z$ in the $\{|nLM_LSM_S\rangle\}$ (*uncoupled*) representation. Physically, there is competition between the mutual coupling of **L** and **S** and the coupling of **L** and **S** individually with the external **B** field. In low fields the spin-orbit coupling is dominant, whereas in the high field limit it may be neglected. Thus, also for many electron atoms we have a *Paschen-Back effect.* In intermediate fields the Hamiltonian is neither diagonal in the coupled nor in the uncoupled representation and we have to solve the secular equation (I.24),

$$|\mathcal{H}'_{\mu,\nu} - W_n| = 0, \tag{9.90}$$

where $\mathcal{H}'_{\mu,\nu}$ are the matrix elements of the perturbation in the representation of choice. Note that the overlap integral in Eq. (I.24) vanishes when using orthonormal basis states.

9.6.2 Description in the uncoupled basis

To solve the secular equation in the *uncoupled* representation within a given LS term we use one of the inner product rules to write the perturbation in the form

$$\mathcal{H}' = \zeta_{nLS}\left[L_zS_z + \tfrac{1}{2}\left(L_+S_- + L_-S_+\right)\right] + (g_L\mu_BL_z + g_S\mu_BS_z)B/\hbar. \tag{9.91}$$

Note that this Hamiltonian conserves the total angular momentum along the z axis; i.e., $M_l + M_S$ is a "good" quantum number (J_z commutes with the Hamiltonian) for all values of the magnetic field. This quantum number will be denoted by M_J. The diagonal matrix elements for given values of $M_J = M_l + M_S$ are given by

$$\begin{aligned}\mathcal{H}'_{M_JM_SM_S} &= \langle L\,(M_J - M_S)\,SM_S|\zeta_{nLS}L_zS_z + (g_L\mu_BL_z + g_S\mu_BS_z)B/\hbar|L\,(M_J - M_S)\,SM_S\rangle \\ &= \zeta_{nLS}\,(M_J - M_S)\,M_S + \left[g_LM_J + (g_S - g_L)\,M_S\right]\mu_BB,\end{aligned} \tag{9.92}$$

where ζ_{nLS} is the spin-orbit energy given by Eq. (9.86) or (9.86). The off-diagonal elements are found using the shift operators (1.59) and the conservation of M_J,

$$\begin{aligned}\mathcal{H}'_{M_J(M_S\mp1)M_S} &= \tfrac{1}{2}\zeta_{nLS}\langle L\,(M_J - M_S \pm 1)\,S\,(M_S \mp 1)\,|L_\pm S_\mp|L\,(M_J - M_S)\,SM_S\rangle \\ &= \tfrac{1}{2}\zeta_{nLS}\sqrt{L\,(L+1) - (M_J - M_S)\,(M_J - M_S \pm 1)}\sqrt{S\,(S+1) - M_S(M_S \mp 1)}.\end{aligned} \tag{9.93}$$

As demonstrated in Section 4.7 for hydrogen-like atoms, the solution of the secular equation (9.90) can be obtained analytically for arbitrary magnetic fields because the electronic states are *electron spin doublets* $S = s = \frac{1}{2}$) and $J = j = l \pm \frac{1}{2}$. In this case, the characteristic equation is of the asymmetric type (*cf.* Appendix H.3) and factorizes into a product of two (1×1) and $l+1$ (2×2) determinants, each characterized by its own value of $M_J = M_l + M_S$. In Fig. 9.8 the magnetic field dependence based on these solutions is plotted for both positive and negative coupling constants.

For many-electron atoms the factorization of the secular equation into $(d\times d)$ determinants will in general also contain determinants with $d > 2$, for which the diagonalization is best done numerically with *Mathematica.* For the

9.6.3 High-field limit - Paschen-Back effect

In high magnetic fields the spin-orbit coupling is weak as compared to the Zeeman interaction. This is known as the Paschen-Back effect and best analyzed in the *uncoupled* representation $\{|nLM_LSM_S\rangle\}$ in which case the atomic Hamiltonian including the Zeeman interaction, $\mathcal{H}_0 + \mathcal{H}_Z$, is diagonal. In this basis the spin-orbit matrix elements are given by

$$\mathcal{H}^{LS}_{M'_LM'_S;M_LM_S} = \zeta_{nLS}\langle LM'_LSM'_S|L_zS_z + \tfrac{1}{2}\left(L_+S_- + L_-S_+\right)|LM_LSM_S\rangle\,. \tag{9.94}$$

As only the diagonal matrix elements of a weak perturbation contribute to leading order in perturbation theory, in high fields the spin-orbit shift is given by $\Delta E^{LS}_{n;M_L M_S} = \zeta_{nLS} M_L M_S$ and the field dependence can be expressed as

$$\Delta E_{n;M_L M_S} = \zeta_{nLS} M_L M_S + (g_L M_L + g_S M_S)\,\mu_B B. \tag{9.95}$$

Note that this expression is valid for *arbitrary* values of L and S.

9.6.4 Low-field limit - Landé factor

For magnetic fields much lower than the fine-structure field ($B \ll B_{\rm fs}$) the angular momenta $\mathbf{L}$ and $\mathbf{S}$ are coupled into the total electronic angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$. Under these conditions the atomic Hamiltonian

$$\mathcal{H} = (\mathcal{H}_{\rm CF} + \mathcal{H}_{LS}) + \mathcal{H}_Z \tag{9.96}$$

is best analyzed in the *coupled* representation $\{|nLSJM_J\rangle\}$, where the first two terms of Eq. (9.96) are diagonal and $\mathcal{H}_Z$ acts as the perturbation. In this basis and for given values of n, L, S, J the perturbation gives rise to an energy shift

$$\Delta E^Z_{J,M_J} = \langle JM_J| g_L L_z + g_e S_z \, |JM_J\rangle \, \mu_B B/\hbar \tag{9.97}$$

Note that the operators L_Z and S_Z are diagonal in the $\{|nLSJM_J\rangle\}$ basis because they conserve M_L and M_S and, in view of the M_J *selection rule* (3.90), also $M_L + M_S = M_J$. Hence, in sufficiently low fields J and M_J are "good" quantum numbers for the atomic Hamiltonian (9.96).

Intuitively this situation is also clear: for sufficiently low fields the magnetic moment of the atom will scale with the total electronic angular momentum $\mathbf{J}$,

$$\boldsymbol{\mu}_J = -g_J \mu_B \, (\mathbf{J}/\hbar) \qquad \text{for } B \to 0. \tag{9.98}$$

Hence, the Zeeman energy can be described by the effective Hamiltonian $\mathcal{H}_Z = -\boldsymbol{\mu}_J \cdot \mathbf{B}$. This Hamiltonian yields a Zeeman shift on top of the zero-field spin-orbit shift ΔE^{LS}_{nJ} (for given values of n, L, S, J):

$$\Delta E^Z_{J,M_J} = g_J \mu_B B \langle J m_J| (J_z/\hbar) \, |JM_J\rangle = g_J \mu_B M_J B \qquad \text{for } B \to 0. \tag{9.99}$$

As was to be expected, also this effective Hamiltonian is diagonal in the $\{|nLSJM_J\rangle\}$ basis if the quantization axis is chosen along the direction of the $\mathbf{B}$ field. Our task is of course to determine the value of g_J.

9.6.4.1 Wigner-Eckart theorem

To understand the relation between Eqs. (9.97) and (9.99) we recall the vector diagram in Fig. 4.6. The slow precession of $\mathbf{J} = \mathbf{L} + \mathbf{S}$ about the direction of a weak probe field (just serving to define a quantization axis) does not affect the projections of the angular momentum operators L_z, S_z and J_z. In other words these operators are invariant under rotation about the z axis. As all three are diagonal in the $\{|nLSJM_J\rangle\}$ basis, the matrix elements of L_z and S_z can be written in the form

$$\langle nLSJM_J| \, L_z \, |nLSJM_J\rangle = \langle lsJ \, \|L\| \, lsJ\rangle \, \langle JM_J| \, J_z \, |JM_J\rangle \tag{9.100a}$$

$$\langle nLSJM_J| \, S_z \, |nLSJM_J\rangle = \langle lsJ \, \|S\| \, lsJ\rangle \, \langle JM_J| \, J_z \, |JM_J\rangle \,, \tag{9.100b}$$

where the proportionality constants $\langle LSJ \, \|L\| \, LSJ\rangle$ and $\langle LSJ \, \|S\| \, LSJ\rangle$ are known as *reduced matrix elements* and are *independent of* M_J. Since $J_z = L_z + S_z$ we find by adding Eqs. (9.100a) and (9.100b) that the reduced matrix elements add up to unity,

$$\langle LSJ \, \|L\| \, LSJ\rangle + \langle LSJ \, \|S\| \, LSJ\rangle = 1 = \langle LSJ \, \|J\| \, LSJ\rangle. \tag{9.101}$$

The projections of $\mathbf{L}$ and $\mathbf{S}$ along the total angular momentum vector $\mathbf{J}$, i.e. $\mathbf{L}\cdot\mathbf{J}$ and $\mathbf{S}\cdot\mathbf{J}$, are conserved while $\mathbf{J}$ precesses about $\mathbf{B}$,

$$\langle nLSJM_J|\,L_z\,|nLSJM_J\rangle=\frac{\langle LSJM_J|\,(\mathbf{L}\cdot\mathbf{J})\,J_z\,|LSJM_J\rangle}{\langle LSJM_J|\,\mathbf{J}^2\,|LSJM_J\rangle}. \tag{9.102}$$

comparing with Eq. (9.100a) we find

$$\langle LSJ\,\|L\|\,LSJ\rangle=\frac{\langle LSJM_J|\,\mathbf{L}\cdot\mathbf{J}\,|LSJM_J\rangle}{\langle LSJM_J|\,\mathbf{J}^2\,|LSJM_J\rangle}=\frac{J(J+1)+L(L+1)-S(S+1)}{2J(J+1)}, \tag{9.103}$$

which is indeed independent of M_J. Analogously we find

$$\langle LSJ\,\|S\|\,LSJ\rangle=\frac{\langle LSJM_J|\,\mathbf{S}\cdot\mathbf{J}\,|LSJM_J\rangle}{\langle LSJM_J|\,\mathbf{J}^2\,|LSJM_J\rangle}=\frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}. \tag{9.104}$$

9.6.4.2 Landé factor g_J

In search for g_J we return to Eqs. (9.97) and (9.99). With the aid of the Wigner-Eckart theorem (9.100) as well as the relation (9.101) we obtain the expression

$$\begin{aligned} g_J&=g_L\langle LSJ\,\|L\|\,LSJ\rangle+g_e\langle LSJ\,\|S\|\,LSJ\rangle\\ &=\tfrac{1}{2}\left(g_L+g_e\right)+\tfrac{1}{2}\left(g_L-g_e\right)\left[\langle LSJ\,\|L\|\,LSJ\rangle-\langle LSJ\,\|S\|\,LSJ\rangle\right]. \end{aligned} \tag{9.105}$$

Substituting the expressions for the reduced matrix elements (9.103) and (9.104) we obtain for the g_J factor,

$$g_J=\tfrac{1}{2}\left(g_L+g_e\right)+\tfrac{1}{2}\left(g_L-g_e\right)\frac{L(L+1)-S(S+1)}{J(J+1)}. \tag{9.106}$$

In the important case of S terms $g_J=g_e$ as is intuitively clear for the absence of an orbital moment. Approximating $g_L=1$ and $g_e=2$ we find that g_J takes the form of the *Landé factor*

$$g_J=1+\frac{[J(J+1)+S(S+1)-L(L+1)]}{2J(J+1)}. \tag{9.107}$$

Note that just like the expressions for the high-field limit also the expressions for the low-field limit are valid for *arbitrary* values of L and S.

9.6.5 Low-field limit - quadratic Zeeman shift

For arbitrary value of S the quadratic shift can be calculated with the aid of second-order perturbation theory,

$$\Delta E^{(2)Z}_{J,M}\left(n\,{}^{2s+1}L_J\right)=\sum_{J'}\frac{|\langle LSJ'M|g_L\mu_BL_z+g_S\mu_BS_z|LSJM\rangle B/\hbar|^2}{E_{J,M}-E_{J',M}}. \tag{9.108}$$

Note that this contribution indeed scales like B^2. As L_z and S_z conserve M, the Zeeman operator couples states of different $J\in\{|J-S|,\cdots,J+S\}$ for given value of M.

We demonstrate the procedure for the special case in which *two* fine-structure levels, J' and $J=J'-1$, are coupled by the spin-orbit interaction. At sufficiently low fields, the energy splitting may be approximated by its zero-field value; i.e., by the interval rule (4.128), $\Delta W_J=\zeta_{nLS}J'$. The energy splitting also has a field dependence but if $g_J=g_{J'}$ the associated shift contributes in third or higher order. In any case, this field dependence is not included here. The quadratic shifts of the upper, J', and lower, $J=J'-1$, fine-structure levels (of given M) only differ in the sign, with the quadratic shifts of the upper manifold positive and those of the lower manifold negative. For the upper/lower manifold the quadratic shift becomes

$$\Delta E^{(2)Z}_{J'/J,M}\left(n\,^{2s+1}L_{J'/J}\right) = \pm\frac{|\langle LSJ'M|g_L L_z + g_S S_z|LSJM\rangle\mu_B B/\hbar|^2}{\zeta_{nLS}J'}. \tag{9.109}$$

Adding and subtracting $g_L\mu_B S_z$ and using the orthogonality relation $\langle LSJ'M|LSJM\rangle = 0$ as well as the definitions (4.174a) and (4.180) the Eq. (9.109) can be reduced to

$$\Delta E^{(2)Z}_{J'/J,M}\left(n\,^{2s+1}L_j\right) = \pm\zeta_{nLS}J'\left(B/B_{\mathrm{fs}}\right)^2|\langle LSJ'M|S_z/\hbar|LSJM\rangle|^2, \tag{9.110}$$

where

$$B_{\mathrm{fs}} \equiv \zeta_{nLS}J'/\mu'_B, \tag{9.111}$$

with $\mu'_B \equiv (g_S - g_L)\mu_B$, is the fine-structure crossover field. Evaluating the matrix element we find (see Problem 9.2)

$$\Delta E^{(2)Z}_{J'/J,M}\left(n\,^{2s+1}L_j\right) = \pm\tfrac{1}{4}\zeta_{nLS}J'\left(B/B_{\mathrm{fs}}\right)^2\left(1-\widetilde{M}^2\right)R(LSJ') \tag{9.112}$$

where $\widetilde{M} = M/J'$ and

$$R(LSJ') \equiv \frac{(L+S+1+J')(J'-L+S)(J'+L-S)(L+S+1-J')}{(2J'-1)(2J'+1)}. \tag{9.113}$$

For $L = 1$ and $s = \frac{1}{2}$ with $J' = L+S = \frac{3}{2}$ and $J = L-S = \frac{1}{2}$ we calculate $R(LSJ') = 1$ and we regain the result from Eq. (4.184). Note that we indeed regained the quadratic term of Eq. (5.102). Since $\widetilde{M}^2 \leq 1$ we note that the quadratic Zeeman shift is *largest* for the state of *smallest* $|M|$.

Problem 9.2. Derive the expression for the matrix element

$$\langle LSJ'M|S_z/\hbar|LSJM\rangle = \sqrt{\left(1-\widetilde{M}^2\right)\frac{(L+S+1+J')(J'-L+S)(J'+L-S)(L+S+1-J')}{4(2J'-1)(2J'+1)}}.$$

Solution. The matrix element is evaluated for arbitrary J' and $J = J'-1$ using the Wigner-Eckart theorem - see Eq. (K.24),

$$\langle LSJ'M|S_z/\hbar|LSJM\rangle = (-1)^{J'-M}\langle LSJ'||S||LSJ\rangle\begin{pmatrix} J' & 1 & J \\ -M & 0 & M\end{pmatrix},$$

where, in analogy with Eq. (K.62),

$$\langle LSJ'||S||LSJ\rangle = (-1)^{J+L+S}\langle S||S||S\rangle\sqrt{(2J+1)(2J'+1)}\begin{Bmatrix} S & J' & L \\ J & S & 1\end{Bmatrix},$$

with $\langle S||S||S\rangle = \sqrt{S(S+1)(2S+1)}$ - see Eq. (K.50). Combining these expressions we find

$$\langle LSJ'M|S_z/\hbar|LSJM\rangle = \\ (-1)^{J+J'+L+S-M}\sqrt{S(S+1)(2S+1)(2J+1)(2J'+1)}\begin{pmatrix} J' & 1 & J \\ -M & 0 & M\end{pmatrix}\begin{Bmatrix} S & J' & L \\ J & S & 1\end{Bmatrix}.$$

Since $J' = J+1$ we can use the special case (J.13) to evaluate the $3j$-symbol,

$$\begin{pmatrix} J' & 1 & J'-1 \\ -M & 0 & M\end{pmatrix} = (-1)^{-J'+M}\sqrt{\frac{(J'+M)(J'-M)}{J'(2J'+1)(2J'-1)}},$$

and special case (J.96b) to evaluate the $6j$-symbol,

$$\begin{Bmatrix} S & J' & L \\ J'-1 & S & 1\end{Bmatrix} = (-1)^{J+L+S}\sqrt{\frac{2(L+S+1+J')(J'-L+S)(J'+L-S)(L+S+1-J')}{2S(2S-1)(2S+1)(2J'-1)2J'(2J'+1)}}.$$

Substituting the Wigner symbols we obtain the desired expression. □

A

Periodic table of the elements

Table A.1: Eight primary groups cover the filling of s and p shells. The secondary groups 3-13 are the transition metals and correspond to the filling of d shells. The rare-earth elements (lanthanides and actinides) correspond to filling of the f shells.

I	II													III	IV	V	VI	VII	VIII
1	2		3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H															(np shells)			2 He
3 Li	4 Be												5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg						transition elements (nd shells)						13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca		21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr		39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	57−70 *	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	89−102 **	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg							

rare-earth elements (nf shells)

*lanthanides ($4f$)	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
**actinides ($5f$)	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

B

Properties of the elements

Table B.1: Configuration, Ground term, Ionization Energy (IE) and Electron Affinity (EA) of the first three rows of the periodic table. Note that all half-filled and fully-filled shells have an S term ground state and that the electron affinity of fully-filled shells vanishes. The table is based on the NIST Atomic Spectra Database [66]. The Electron Affinities (EA) are from T. Andersen, H.K. Haugen, H. Hotop [3].

				ground	IE in eV						EA in eV
Z	Element		configuration	term	1s	2s	2p	3s	3p	3d	
1	Hydrogen	H	1s	$^2S_{1/2}$	13.60						0.75
2	Helium	He	1s^2	$^1S_0$	24.59						< 0
3	Lithium	Li	[He]2s	$^2S_{1/2}$		5.39					0.62
4	Beryllium	Be	[He]2s^2	$^1S_0$		9.32					< 0
5	Boron	B	[He]2s^{2}2p	$^2P_{1/2}$			8.30				0.28
6	Carbon	C	[He]2s^{2}2p^2	$^3P_0$			11.26				1.26
7	Nitrogen	N	[He]2s^{2}2p^3	$^4S_{3/2}$			14.53				-0.07
8	Oxygen	O	[He]2s^{2}2p^4	$^3P_2$			13.62				1.46
9	Fluorine	F	[He]2s^{2}2p^5	$^2P_{3/2}$			17.42				3.40
10	Ne	Ne	[He]2s^{2}2p^6	$^1S_0$			21.56				< 0
11	Sodium	Na	[Ne]3s	$^2S_{1/2}$				5.14			0.55
12	Magnesium	Mg	[Ne]3s^2	$^1S_0$				7.65			< 0
13	Aluminium	Al	[Ne]3s^{2}3p	$^2P_{1/2}$					5.99		0.43
14	Silicon	Si	[Ne]3s^{2}3p^2	$^3P_0$					8.15		1.39
15	Phosphorus	P	[Ne]3s^{2}3p^3	$^4S_{3/2}$					10.49		0.75
16	Sulphur	S	[Ne]3s^{2}3p^4	$^3P_2$					10.36		2.08
17	Chlorine	Cl	[Ne]3s^{2}3p^5	$^2P_{3/2}$					12.97		3.61
18	Argon	Ar	[Ne]3s^{2}3p^6	$^1S_0$					15.76		< 0

Table B.2: Configuration, Ground term, Ionization energy (IE) and Electron affinity (EA) of the fourth and fifth row of the periodic table. This includes the filling of the 3d and 4d shells (transition metals). The table is based on the NIST Atomic Spectra Database [66]. The Electron Affinities (EA) are from T. Andersen, H.K. Haugen, H. Hotop [3].

Z	Element		configuration	ground term	IE in eV						EA in eV
					3d	4s	4p	4d	5s	5p	
19	Potassium	K	[Ar]$4s^1$	$^2S_{1/2}$		4.34					0.50
20	Calcium	Ca	[Ar]$4s^2$	$^1S_0$		6.11					0.02
21	Scandium	Sc	[Ar]$3d4s^2$	$^2D_{3/2}$	6.56						0.19
22	Titanium	Ti	[Ar]$3d^24s^2$	$^3F_2$	6.83						0.08
23	Vanadium	Va	[Ar]$3d^34s^2$	$^4F_{3/2}$	6.75						0.53
24	Chromium	Cr	[Ar]$3d^54s^1$	$^7S_3$	6.77						0.67
25	Manganese	Mn	[Ar]$3d^54s^2$	$^6S_{5/2}$	7.43						< 0
26	Iron	Fe	[Ar]$3d^64s^2$	$^5D_4$	7.90						0.16
27	Cobalt	Co	[Ar]$3d^74s^2$	$^4F_{9/2}$	7.88						0.66
28	Nickel	Ni	[Ar]$3d^84s^2$	$^3F_4$	7.64						1.16
29	Copper	Cu	[Ar]$3d^{10}4s^1$	$^2S_{1/2}$	7.73						1.23
30	Zinc	Zn	[Ar]$3d^{10}4s^2$	$^1S_0$	9.39						< 0
31	Gallium	Ga	[Ar]$3d^{10}4s^24p$	$^2P_{1/2}$			6.00				0.30
32	Germanium	Ge	[Ar]$3d^{10}4s^24p^2$	$^3P_0$			7.90				1.22
33	Arsenic	As	[Ar]$3d^{10}4s^24p^3$	$^4S_{3/2}$			9.79				0.81
34	Selenium	Se	[Ar]$3d^{10}4s^24p^4$	$^3P_2$			9.75				2.02
35	Bromine	Br	[Ar]$3d^{10}4s^24p^5$	$^2P_{3/2}$			11.81				3.37
36	Krypton	Kr	[Ar]$3d^{10}4s^24p^6$	$^1S_0$			14.00				< 0
37	Rubidium	Rb	[Kr]$5s$	$^2S_{1/2}$					4.18		0.49
38	Strontium	Sr	[Kr]$5s^2$	$^1S_0$					5.69		0.05
39	Yttrium	Y	[Kr]$4d5s^2$	$^2D_{3/2}$				6.22			0.31
40	Zirconium	Zr	[Kr]$4d^25s^2$	$^3F_2$				6.63			0.43
41	Niobium	Nb	[Kr]$4d^45s^1$	$^6D_{1/2}$				6.76			0.67
42	Molybdenium	Mo	[Kr]$4d^55s^1$	$^7S_3$				7.09			0.75
43	Technetium	Tc	[Kr]$4d^55s^2$	$^6S_{5/2}$				7.28			0.55
44	Ruthenium	Ru	[Kr]$4d^75s^1$	$^5F_5$				7.36			1.05
45	Rhodium	Rh	[Kr]$4d^85s^1$	$^4F_{9/2}$				7.46			1.14
46	Palladium	Pd	[Kr]$4d^{10}$	$^1S_0$				8.34			0.56
47	Silver	Ag	[Kr]$4d^{10}5s^1$	$^2S_{1/2}$					7.58		1.30
48	Cadmium	Cd	[Kr]$4d^{10}5s^2$	$^1S_0$					8.99		< 0
49	Indium	In	[Kr]$4d^{10}5s^25p$	$^2P_{1/2}$						5.79	0.4
50	Tin	Sn	[Kr]$4d^{10}5s^25p^2$	$^3P_0$						7.34	1.11
51	Antimony	Sb	[Kr]$4d^{10}5s^25p^3$	$^4S_{3/2}$						8.61	1.05
52	Tellurium	Te	[Kr]$4d^{10}5s^25p^4$	$^3P_2$						9.01	1.97
53	Iodine	I	[Kr]$4d^{10}5s^25p^5$	$^2P_{3/2}$						10.45	3.06
54	Xenon	Xe	[Kr]$4d^{10}5s^25p^6$	$^1S_0$						12.13	< 0

Table B.3: Configuration, Ground term, Ionization energy (IE) and Electron affinity (EA) of the sixth row of the periodic table. This includes the filling of the 4f (rare-earths lanthanides) and 5d (transition metals) shells. The table is based on the NIST Atomic Spectra Database [66]. The Electron Affinities (EA) are from T. Andersen, H.K. Haugen, H. Hotop [3].

				ground	IE in eV					EA in eV
Z	Element		configuration	term	4f	5d	6s	6p	7s	
55	Cesium	Cs	[Xe]6s	$^2S_{1/2}$			3.89			0.47
56	Barium	Ba	[Xe]$6s^2$	$^1S_0$			5.21			0.14
57	Lanthanum	La	[Xe]$5d\,6s^2$	$^2D_{3/2}$		5.58				0.47
58	Cerium	Ce	[Xe]$4f\,5d\,6s^2$	$^1G_4$	5.54					
59	Praseodymium	Pr	[Xe]$4f^3 6s^2$	$^4I_{9/2}$	5.47					
60	Neodymium	Nd	[Xe]$4f^4 6s^2$	$^5I_4$	5.53					
61	Promethium	Pm	[Xe]$4f^5 6s^2$	$^6H_{5/2}$	5.58					
62	Samarium	Sm	[Xe]$4f^6 6s^2$	$^7F_0$	5.64					
63	Europium	Eu	[Xe]$4f^7 6s^2$	$^8S_{7/2}$	5.67					
64	Gadolinium	Gd	[Xe]$4f^7 5d\,6s^2$	$^9D_2$	6.15					
65	Terbium	Tb	[Xe]$4f^9 6s^2$	$^6H_{15/2}$	5.86					
66	Dysprosium	Dy	[Xe]$4f^{10} 6s^2$	$^5I_8$	5.94					
67	Holmium	Ho	[Xe]$4f^{11} 6s^2$	$^4I_{15/2}$	6.02					
68	Erbium	Er	[Xe]$4f^{12} 6s^2$	$^3H_6$	6.11					
69	Thulium	Tm	[Xe]$4f^{13} 6s^2$	$^2F_{7/2}$	6.18					
70	Ytterbium	Yb	[Xe]$4f^{14} 6s^2$	$^1S_0$	6.25					
71	Lutetium	Lu	[Xe]$4f^{14} 5d\,6s^2$	$^2D_{3/2}$	5.43					
72	Hafnium	Hf	[Xe]$4f^{14} 5d^2 6s^2$	$^3F_2$		6.83				~0
73	Tantalum	Ta	[Xe]$4f^{14} 5d^3 6s^2$	$^4F_{3/2}$		7.55				0.32
74	Tungsten	W	[Xe]$4f^{14} 5d^4 6s^2$	$^5D_0$		7.86				0.82
75	Rhenium	Re	[Xe]$4f^{14} 5d^5 6s^2$	$^6S_{5/2}$		7.83				0.15
76	Osmium	Os	[Xe]$4f^{14} 5d^6 6s^2$	$^5D_4$		8.44				1.08
77	Iridium	Ir	[Xe]$4f^{14} 5d^7 6s^2$	$^4F_{9/2}$		8.97				1.56
78	Platinum	Pt	[Xe]$4f^{14} 5d^9 6s^1$	$^3D_3$		8.96				2.13
79	Gold	Au	[Xe]$4f^{14} 5d^{10} 6s^1$	$^2S_{1/2}$			9.23			2.31
80	Mercury	Hg	[Xe]$4f^{14} 5d^{10} 6s^2$	$^1S_0$			10.44			< 0
81	Thallium	Tl	[Xe]$4f^{14} 5d^{10} 6s^2 6p$	$^2P_{1/2}$				6.11		0.38
82	Lead	Pb	[Xe]$4f^{14} 5d^{10} 6s^2 6p^2$	$^3P_0$				7.42		0.36
83	Bismuth	Bi	[Xe]$4f^{14} 5d^{10} 6s^2 6p^3$	$^4S_{3/2}$				7.29		0.94
84	Polonium	Po	[Xe]$4f^{14} 5d^{10} 6s^2 6p^4$	$^3P_2$				8.14		1.93
85	Astatine	At	[Xe]$4f^{14} 5d^{10} 6s^2 6p^5$	$^2P_{3/2}$				-		2.82
86	Radon	Rn	[Xe]$4f^{14} 5d^{10} 6s^2 6p^6$	$^1S_0$				10.75		< 0

Table B.4: Configuration, Ground term, Ionization energy (IE) and Electron affinity (EA) of the seventh row of the periodic table up through the 5f shell (rare-earth actinides). The table is based on the NIST Atomic Spectra Database [66].

				ground	IE in eV			EA in eV
Z	Element		configuration	term	5f	6d	7s	
87	Francium	Fr	[Rn]7s	$^2S_{1/2}$			4.07	
88	Radium	Ra	[Ra]$7s^2$	$^1S_0$			5.28	
89	Actinium	Ac	[Ra]6d $7s^2$	$^2D_{3/2}$		5.38		
90	Thorium	Th	[Ra]$6d^2 7s^2$	$^3F_2$		6.31		
91	Protactinium	Pa	[Ra]$5f^2 6d\ 7s^2$	$^4K_{11/2}$	5.89			
92	Uranium	U	[Ra]$5f^3 6d\ 7s^2$	$^5L_6$	6.19			
93	Neptunium	Np	[Ra]$5f^4 6d\ 7s^2$	$^6L_{11/2}$	6.27			
94	Plutonium	Pu	[Ra]$5f^6 7s^2$	$^7F_0$	6.03			
95	Americium	Am	[Ra]$5f^7 7s^2$	$^8S_{7/2}$	5.97			
96	Curium	Cm	[Ra]$5f^7 6d\ 7s^2$	$^9D_2$	5.99			
97	Berkelium	Bk	[Ra]$5f^9 7s^2$	$^6H_{15/2}$	6.20			
98	Californium	Cf	[Ra]$5f^{10} 7s^2$	$^5I_8$	6.28			
99	Einsteinium	Es	[Ra]$5f^{11} 7s^2$	$^4I_{15/2}$	6.37			
100	Fermium	Fm	[Ra]$5f^{12} 7s^2$	$^3H_6$	6.50			
101	Mendelevium	Md	[Ra]$5f^{13} 7s^2$	$^2F_{7/2}$	6.58			
102	Nobelium	No	[Ra]$5f^{14} 7s^2$	$^1S_0$	6.65			

C

Properties of atomic isotopes

Table C.1: Properties of selected atomic isotopes. The Table is based on three databases: (a) AME2020 (atomic mass evaluation); (b) IAEA-INDC(NDS)-794 (magnetic dipole moments); (c) IAEA-INDC(NDS)-833 (electric quadrupole moments).

Z	element	symbol	A	N	radius (fm)	atomic mass (m_u)	I^{π} ($\hbar$)	μ_I (μ_N)	Q (barn)	RA (%)
0	neutron	n	1	1	-	1.008664915	$1/2^+$	-1.913043	0.0	-
1	hydrogen	$^{1}\mathrm{H}$	1	0	0.8783	1.007825032	$1/2^+$	2.792847351	0.0	99.9855
	deuterium	$^{2}\mathrm{D}$	2	1	2.1421	2.014101778	1^+	0.857438231	0.0028578	0.0145
	tritium	$^{3}\mathrm{T}$	3*	2	1.7591	3.016049281	$1/2^+$	2.97896246	0.0	trace
2	helium	$^{3}\mathrm{He}$	3	1	1.9661	3.016029322	$1/2^+$	-2.12762531	0.0	0.0002
		$^{4}\mathrm{He}$	4	2	1.6755	4.002603254	0^+	0.0	0.0	99.9998
3	lithium	$^{6}\mathrm{Li}$	6	3	2.589	6.015122887	1^+	0.822043	-0.000806	4.85
		$^{7}\mathrm{Li}$	7	4	2.444	7.016003434	$3/2^-$	3.256407	-0.04	95.15
4	beryllium	$^{9}\mathrm{Be}$	9	5	2.519	9.012183062	$3/2^-$	-1.17743	0.0529	100.0
		$^{10}\mathrm{Be}$	10*	6	2.355	10.01353469	0^+	0.0	0.0	trace
5	boron	$^{10}\mathrm{B}$	10	5	2.4277	10.01293686	3^+	1.8004636	0.0846	19.65
		$^{11}\mathrm{B}$	11	6	2.406	11.00930517	$3/2^-$	2.688378	0.04059	80.35
6	carbon	$^{12}\mathrm{C}$	12	6	2.4702	12.0	0^+	0.0	0.0	98.94
		$^{13}\mathrm{C}$	13	7	2.4614	13.00335484	$1/2^-$	0.702369	0.0	1.06
		$^{14}\mathrm{C}$	14*	8	2.5025	14.00324199	0^+	0.0	0.0	trace
7	nitrogen	$^{14}\mathrm{N}$	14	7	2.5582	14.003074	1^+	0.403573	0.02044	99.6205
		$^{15}\mathrm{N}$	15	8	2.6058	15.0001089	$1/2^-$	-0.2830569	0.0	0.3795
8	oxygen	$^{16}\mathrm{O}$	16	8	2.6991	15.99491462	0^+	0.0	0.0	99.757
		$^{17}\mathrm{O}$	17	9	2.6932	16.99913176	$5/2^+$	-1.893543	-0.0256	0.03835
		$^{18}\mathrm{O}$	18	10	2.7726	17.99915961	0^+	0.0	0.0	0.2045
9	fluorine	$^{19}\mathrm{F}$	19	10	2.8976	18.99840316	$1/2^+$	2.628321	0.0	100.0
10	neon	$^{20}\mathrm{Ne}$	20	10	3.0055	19.99244018	0^+	0.0	0.0	90.48
		$^{21}\mathrm{Ne}$	21	11	2.9695	20.99384669	$3/2^+$	-0.6617	0.1016	0.27
		$^{22}\mathrm{Ne}$	22	12	2.9525	21.99138511	0^+	0.0	0.0	9.25

*radioactive

Table 1.1 (continuation)

Z	element	symbol	A	N	radius (fm)	atomic mass (m_u)	I^{π} ($\hbar$)	μ_I (μ_N)	Q (barn)	RA (%)
11	sodium	^{22}Na	22*	11	2.9852	21.99443755	3^+	1.746	0.18	trace
		^{23}Na	23	12	2.9936	22.98976928	$3/2^+$	2.2175	0.104	100.0
12	magnesium	^{24}Mg	24	12	3.057	23.98504169	0^+	0.0	0.0	78.965
		^{25}Mg	25	13	3.0284	24.98583697	$5/2^+$	-0.85533	0.199	10.011
		^{26}Mg	26	14	3.0337	25.98259297	0^+	0.0	0.0	11.025
13	aluminium	^{26}Al	26*	13	3.05	25.98689188	5^+	2.803	0.26	trace
		^{27}Al	27	14	3.061	26.98153841	$5/2^+$	3.6407	0.1466	100.0
14	silicon	^{28}Si	28	14	3.1224	27.97692653	0^+	0.0	0.0	92.2545
		^{29}Si	29	15	3.1176	28.97649466	$1/2^+$	-0.555052	0.0	4.672
		^{30}Si	30	16	3.1336	29.97377014	0^+	0.0	0.0	3.0735
15	phosphorus	^{31}P	31	16	3.1889	30.973762	$1/2^+$	1.130925	0.0	100.0
16	sulfur	^{32}S	32	16	3.2611	31.97207117	0^+	0.0	0.0	94.85
		^{33}S	33	17	3.3	32.97145891	$3/2^+$	0.64325	-0.0678	0.763
		^{34}S	34	18	3.2847	33.96786701	0^+	0.0	0.0	4.365
		^{36}S	36	20	3.2985	35.96708069	0^+	0.0	0.0	0.0158
17	chlorine	^{35}Cl	35	18	3.3654	34.96885269	$3/2^+$	0.8217	-0.817	75.8
		^{36}Cl	36*	19	3.37	35.96830682	2^+	1.2849	-0.178	trace
		^{37}Cl	37	20	3.384	36.96590257	$3/2^+$	0.684	-0.0644	24.2
18	argon	^{36}Ar	36	18	3.3905	35.96754511	0^+	0.0	0.0	0.3336
		^{38}Ar	38	20	3.4028	37.9627321	0^+	0.0	0.0	0.0629
		^{39}Ar	39*	21	3.4093	38.96431304	$7/2^-$	-1.59	-0.116	trace
		^{40}Ar	40	22	3.4274	39.96238312	0^+	0.0	0.0	99.6035
19	potassium	^{39}K	39	20	3.4349	38.96370648	$3/2^+$	0.39147	0.0603	93.2581
		^{40}K	40*	21	3.4381	39.96399817	4^-	-1.29797	-0.075	0.0117
		^{41}K	41	22	3.4518	40.96182526	$3/2^+$	0.21482	-0.0734	6.7302
20	calcium	^{40}Ca	40	20	3.4776	39.96259085	0^+	0.0	0.0	96.941
		^{41}Ca	41*	21	3.478	40.96227791	$7/2^-$	-1.5942	-0.665	trace
		^{42}Ca	42	22	3.5081	41.95861778	0^+	0.0	0.0	0.647
		^{43}Ca	43	23	3.4954	42.95876638	$7/2^-$	-1.31733	-0.0408	0.135
		^{44}Ca	44	24	3.5179	43.95548149	0^+	0.0	0.0	2.086
		^{46}Ca	46	26	3.4953	45.95368773	0^+	0.0	0.0	0.004
		^{48}Ca	48*	28	3.4771	47.95252265	0^+	0.0	0.0	0.187
21	scandium	^{45}Sc	45*	24	3.5459	44.95590705	$7/2^-$	4.754	-0.22	100.0
22	titanium	^{46}Ti	46	24	3.607	45.95262636	0^+	0.0	0.0	8.25
		^{47}Ti	47	25	3.5962	46.95175749	$5/2^-$	-0.78848	0.302	7.44
		^{48}Ti	48	26	3.5921	47.94794068	0^+	0.0	0.0	73.72
		^{49}Ti	49	27	3.5733	48.94786439	$7/2^-$	-1.10417	0.247	5.41
		^{50}Ti	50	28	3.5704	49.94478562	0^+	0.0	0.0	5.18

Table 1.1 (continuation)

Z	element	symbol	A	N	radius (fm)	atomic mass (m_u)	I^π ($\hbar$)	μ_I (μ_N)	Q (barn)	RA (%)
23	vanadium	^{50}V	50*	27	3.58	49.94715668	6^+	3.3456889	0.21	0.25
		^{51}V	51	28	3.6002	50.94395766	$7/2^-$	5.1464	-0.52	99.75
24	chromium	^{50}Cr	50*	26	3.6588	49.94604221	0^+	0.0	0.0	4.345
		^{52}Cr	52	28	3.6452	51.94050471	0^+	0.0	0.0	83.789
		^{53}Cr	53	29	3.6511	52.9406463	$3/2^-$	-0.47431	-0.15	9.501
		^{54}Cr	54	30	3.6885	53.93887736	0^+	0.0	0.0	2.365
25	manganese	^{53}Mn	53*	28	3.6662	52.9412875	$7/2^-$	5.033	0.17	trace
		^{55}Mn	55	30	3.7057	54.93804304	$5/2^-$	3.4669	0.33	100.0
26	iron	^{54}Fe	54	28	3.6933	53.93960819	0^+	0.0	0.0	5.845
		^{56}Fe	56	30	3.7377	55.93493554	0^+	0.0	0.0	91.754
		^{57}Fe	57	31	3.7532	56.93539195	$1/2^-$	0.09064	0.0	2.119
		^{58}Fe	58	32	3.7745	57.93327358	0^+	0.0	0.0	0.282
		^{60}Fe	60*	34	3.78	59.93407025	0^+	0.0	0.0	trace
27	cobalt	^{59}Co	59	32	3.7875	58.93319352	$7/2^-$	4.615	0.42	100.0
		^{60}Co	60*	33	3.79	59.93381554	5^+	3.799	0.46	trace
28	nickel	^{58}Ni	58	30	3.7757	57.93534165	0^+	0.0	0.0	68.0769
		^{60}Ni	60	32	3.8118	59.93078513	0^+	0.0	0.0	26.2231
		^{61}Ni	61	33	3.8225	60.93105482	$3/2^-$	-0.74965	0.162	1.1399
		^{62}Ni	62	34	3.8399	61.92834475	0^+	0.0	0.0	3.6345
		^{64}Ni	64	36	3.8572	63.92796623	0^+	0.0	0.0	0.9256
29	copper	^{63}Cu	63	34	3.8823	62.92959712	$3/2^-$	2.2259	-0.22	69.15
		^{65}Cu	65	36	3.9022	64.92778948	$3/2^-$	2.3844	-0.204	30.85
30	zinc	^{64}Zn	64	34	3.9283	63.92914178	0^+	0.0	0.0	49.17
		^{66}Zn	66	36	3.9491	65.92603364	0^+	0.0	0.0	27.73
		^{67}Zn	67	37	3.953	66.92712742	$5/2^-$	0.87485	0.122	4.04
		^{68}Zn	68	38	3.9658	67.92484423	0^+	0.0	0.0	18.45
		^{70}Zn	70*	40	3.9845	69.92531918	0^+	0.0	0.0	0.61
31	gallium	^{69}Ga	69	38	3.9973	68.92557353	$3/2^-$	2.01502	0.171	60.108
		^{71}Ga	71	40	4.0118	70.92470255	$3/2^-$	2.56033	0.107	39.892
32	germanium	^{70}Ge	70	38	4.0414	69.92424854	0^+	0.0	0.0	20.52
		^{72}Ge	72	40	4.0576	71.92207582	0^+	0.0	0.0	27.45
		^{73}Ge	73	41	4.0632	72.92345895	$9/2^+$	-0.87824	-0.196	7.76
		^{74}Ge	74	42	4.0742	73.92117776	0^+	0.0	0.0	36.52
		^{76}Ge	76	44	4.0811	75.92140273	0^+	0.0	0.0	7.75
33	arsenic	^{75}As	75	42	4.0968	74.92159456	$3/2^-$	1.4383	0.311	100.0

Table 1.1 (continuation)

Z	element	symbol	A	N	radius	atomic mass	I^{π}	μ_I	Q	RA
					(fm)	(m_u)	$(\hbar)$	(μ_N)	(barn)	(%)
34	selenium	^{74}Se	74	40	4.07	73.92247593	0^+	0.0	0.0	0.86
		^{76}Se	76	42	4.1395	75.9192137	0^+	0.0	0.0	9.23
		^{77}Se	77	43	4.1395	76.91991415	$1/2^-$	0.53356	0.0	7.6
		^{78}Se	78	44	4.1406	77.91730924	0^+	0.0	0.0	23.69
		^{79}Se	79*	45	4.14	78.91849925	$7/2^+$	-1.018	nothing	trace
		^{80}Se	80	46	4.14	79.91652176	0^+	0.0	0.0	49.8
35	bromine	^{79}Br	79	44	4.1629	78.91833757	$3/2^-$	2.1046	0.3087	50.65
		^{81}Br	81	46	4.1599	80.9162882	$3/2^-$	2.2686	0.2579	49.35
36	krypton	^{78}Kr	78	42	4.2038	77.92036634	0^+	0.0	0.0	0.355
		^{80}Kr	80	44	4.197	79.91637794	0^+	0.0	0.0	2.286
		^{81}Kr	81*	45	4.1952	80.9165897	$7/2^+$	-0.907	0.644	trace
		^{82}Kr	82	46	4.1919	81.91348115	0^+	0.0	0.0	11.593
		^{83}Kr	83	47	4.1871	82.91412652	$9/2^+$	-0.97073	0.259	11.5
		^{84}Kr	84	48	4.1884	83.91149773	0^+	0.0	0.0	56.987
		^{85}Kr	85*	49	4.1846	84.91252726	$9/2^+$	-1.0055	0.433	trace
		^{86}Kr	86	50	4.1835	85.91061062	0^+	0.0	0.0	17.279
37	rubidium	^{85}Rb	85	48	4.2036	84.91178974	$5/2^-$	1.35306	0.276	72.17
		^{87}Rb	87*	50	4.1989	86.90918053	$3/2^-$	2.75129	0.1335	27.83
38	strontium	^{84}Sr	84	46	4.2394	83.91341912	0^+	0.0	0.0	0.56
		^{86}Sr	86	48	4.2307	85.90926072	0^+	0.0	0.0	9.86
		^{87}Sr	87	49	4.2249	86.90887749	$9/2^+$	-1.09316	0.305	7.0
		^{88}Sr	88	50	4.224	87.90561225	0^+	0.0	0.0	82.58
39	yttrium	^{89}Y	89	50	4.243	88.90583816	$1/2^-$	-0.137298	0.0	100.0
40	zirconium	^{90}Zr	90	50	4.2694	89.90469876	0^+	0.0	0.0	51.45
		^{91}Zr	91	51	4.2845	90.90564021	$5/2^+$	-1.3022	-0.176	11.22
		^{92}Zr	92	52	4.3057	91.90503534	0^+	0.0	0.0	17.15
		^{94}Zr	94	54	4.332	93.90631252	0^+	0.0	0.0	17.38
		^{96}Zr	96*	56	4.3512	95.90827762	0^+	0.0	0.0	2.8
41	niobium	^{91}Nb	91*	50	4.28	90.90699026	$9/2^+$	6.513	-0.25	trace
		^{92}Nb	92*	51	4.3026	91.90718858	7^+	5.13	-0.35	trace
		^{93}Nb	93	52	4.324	92.90637317	$9/2^+$	6.163	-0.32	100.0
42	molybdenum	^{92}Mo	92	50	4.3151	91.90680715	0^+	0.0	0.0	14.649
		^{94}Mo	94	52	4.3529	93.90508359	0^+	0.0	0.0	9.187
		^{95}Mo	95	53	4.3628	94.90583744	$5/2^+$	-0.9132	-0.022	15.873
		^{96}Mo	96	54	4.3847	95.90467477	0^+	0.0	0.0	16.673
		^{97}Mo	97	55	4.388	96.9060169	$5/2^+$	-0.9324	0.255	9.582
		^{98}Mo	98	56	4.4091	97.90540361	0^+	0.0	0.0	24.292
		^{100}Mo	100*	58	4.4468	99.90746798	0^+	0.0	0.0	9.744
43	technetium	^{99}Tc	99*	56	4.4	98.90624968	$9/2^+$	5.678	-0.129	trace

Table 1.1 (continuation)

Z	element	symbol	A	N	radius (fm)	atomic mass (m_u)	I^{π} ($\hbar$)	μ_I (μ_N)	Q (barn)	RA (%)
44	ruthenium	^{96}Ru	96	52	4.3908	95.90758891	0^+	0.0	0.0	5.54
		^{98}Ru	98	54	4.4229	97.90528671	0^+	0.0	0.0	1.87
		^{99}Ru	99	55	4.4338	98.90593028	$5/2^+$	-0.641	0.079	12.76
		^{100}Ru	100	56	4.4531	99.90421046	0^+	0.0	0.0	12.6
		^{101}Ru	101	57	4.4606	100.9055731	$5/2^+$	-0.718	0.46	17.06
		^{102}Ru	102	58	4.4809	101.9043403	0^+	0.0	0.0	31.55
		^{104}Ru	104	60	4.5098	103.9054253	0^+	0.0	0.0	18.62
45	rhodium	^{103}Rh	103	58	4.4945	102.9054941	$1/2^-$	-0.08829	0.0	100.0
46	palladium	^{102}Pd	102	56	4.4827	101.9056323	0^+	0.0	0.0	1.02
		^{104}Pd	104	58	4.5078	103.9040304	0^+	0.0	0.0	11.14
		^{105}Pd	105	59	4.515	104.9050795	$5/2^+$	-0.642	0.66	22.33
		^{106}Pd	106	60	4.5318	105.9034803	0^+	0.0	0.0	27.33
		^{108}Pd	108	62	4.5563	107.9038918	0^+	0.0	0.0	26.46
		^{110}Pd	110	64	4.5782	109.9051729	0^+	0.0	0.0	11.72
47	silver	^{107}Ag	107	60	4.5454	106.9050915	$1/2^-$	-0.11352	0.0	51.839
		^{109}Ag	109	62	4.5638	108.9047558	$1/2^-$	0.13051	0.0	48.161
48	cadmium	^{106}Cd	106	58	4.5383	105.9064598	0^+	0.0	0.0	1.245
		^{108}Cd	108	60	4.5577	107.9041836	0^+	0.0	0.0	0.888
		^{110}Cd	110	62	4.5765	109.9030075	0^+	0.0	0.0	12.47
		^{111}Cd	111	63	4.5845	110.9041838	$1/2^+$	-0.594	0.0	12.795
		^{112}Cd	112	64	4.5944	111.9027639	0^+	0.0	0.0	24.109
		^{113}Cd	113*	65	4.6012	112.9044081	$1/2^+$	-0.6213	0.0	12.227
		^{114}Cd	114	66	4.6087	113.903365	0^+	0.0	0.0	28.754
		^{116}Cd	116*	68	4.6203	115.9047632	0^+	0.0	0.0	7.512
49	indium	^{113}In	113	64	4.601	112.9040605	$9/2^+$	5.5208	0.761	4.281
		^{115}In	115*	66	4.6156	114.9038788	$9/2^+$	5.5326	0.772	95.719
50	tin	^{112}Sn	112	62	4.5948	111.9048249	0^+	0.0	0.0	0.97
		^{114}Sn	114	64	4.6099	113.9027801	0^+	0.0	0.0	0.66
		^{115}Sn	115	65	4.6148	114.9033447	$1/2^+$	-0.9174	0.0	0.34
		^{116}Sn	116	66	4.625	115.9017428	0^+	0.0	0.0	14.54
		^{117}Sn	117	67	4.6302	116.902954	$1/2^+$	-0.9995	0.0	7.68
		^{118}Sn	118	68	4.6393	117.9016066	0^+	0.0	0.0	24.22
		^{119}Sn	119	69	4.6438	118.9033113	$1/2^+$	-1.0459	0.0	8.59
		^{120}Sn	120	70	4.6519	119.9022026	0^+	0.0	0.0	32.58
		^{122}Sn	122	72	4.6634	121.9034455	0^+	0.0	0.0	4.63
		^{124}Sn	124	74	4.6735	123.9052796	0^+	0.0	0.0	5.79

Table 1.1 (continuation)

Z	element	symbol	A	N	radius (fm)	atomic mass (m_u)	I^{π} ($\hbar$)	μ_I (μ_N)	Q (barn)	RA (%)
51	antimony	^{121}Sb	121	70	4.6802	120.9038114	$5/2^+$	3.358	-0.543	57.21
		^{123}Sb	123	72	4.6879	122.9042153	$7/2^+$	2.5457	-0.692	42.79
52	tellurium	^{120}Te	120	68	4.7038	119.9040658	0^+	0.0	0.0	0.09
		^{122}Te	122	70	4.7095	121.9030447	0^+	0.0	0.0	2.55
		^{123}Te	123*	71	4.7117	122.904271	$1/2^+$	-0.7358	0.0	0.89
		^{124}Te	124	72	4.7183	123.9028183	0^+	0.0	0.0	4.74
		^{125}Te	125	73	4.7204	124.9044312	$1/2^+$	-0.887	0.0	7.07
		^{126}Te	126	74	4.7266	125.9033121	0^+	0.0	0.0	8.84
		^{128}Te	128*	76	4.7346	127.9044612	0^+	0.0	0.0	31.74
		^{130}Te	130*	78	4.7423	129.9062227	0^+	0.0	0.0	34.08
53	iodine	^{127}I	127	74	4.75	126.9044726	$5/2^+$	2.8087	-0.688	100.0
		^{129}I	129*	76	4.8	128.9049836	$7/2^+$	2.6165	-0.483	trace
54	xenon	^{124}Xe	124*	70	4.7661	123.9058852	0^+	0.0	0.0	0.095
		^{126}Xe	126	72	4.7722	125.9042974	0^+	0.0	0.0	0.089
		^{128}Xe	128	74	4.7774	127.9035308	0^+	0.0	0.0	1.91
		^{129}Xe	129	75	4.7775	128.9047809	$1/2^+$	-0.777961	0.0	26.401
		^{130}Xe	130	76	4.7818	129.9035093	0^+	0.0	0.0	4.071
		^{131}Xe	131	77	4.7808	130.9050841	$3/2^+$	0.691845	-0.114	21.232
		^{132}Xe	132	78	4.7859	131.9041551	0^+	0.0	0.0	26.909
		^{134}Xe	134*	80	4.7899	133.905393	0^+	0.0	0.0	10.436
		^{136}Xe	136*	82	4.7964	135.9072145	0^+	0.0	0.0	8.857
55	caesium	^{133}Cs	133	78	4.8041	132.905452	$7/2^+$	2.5778	-0.00343	100.0
		^{134}Cs	134*	79	4.805	133.9286804	4^+	2.9893	0.37	trace
		^{135}Cs	135*	80	4.8067	134.9059769	$7/2^+$	2.7283	0.048	trace
		^{137}Cs	137*	82	4.807	136.946162	$7/2^+$	2.8374	0.048	trace
56	barium	^{130}Ba	130	74	4.8283	129.906326	0^+	0.0	0.0	0.11
		^{132}Ba	132*	76	4.8303	131.9050612	0^+	0.0	0.0	0.1
		^{134}Ba	134	78	4.8322	133.9045082	0^+	0.0	0.0	2.42
		^{135}Ba	135	79	4.8294	134.9056884	$3/2^+$	0.8381	0.153	6.59
		^{136}Ba	136	80	4.8334	135.9045758	0^+	0.0	0.0	7.85
		^{137}Ba	137	81	4.8314	136.9058272	$3/2^+$	0.9375	0.236	11.23
		^{138}Ba	138	82	4.8378	137.9052471	0^+	0.0	0.0	71.7
57	lanthanum	^{137}La	137*	80	4.8496	136.9064504	$7/2^+$	2.696	0.21	trace
		^{138}La	138*	81	4.8473	137.907124	5^+	3.7084	0.39	0.08881
		^{139}La	139	82	4.855	138.9063629	$7/2^+$	2.7791	0.206	99.91119

Table 1.1 (continuation)

Z	element	symbol	A	N	radius (fm)	atomic mass (m_u)	I^{π} ($\hbar$)	μ_I (μ_N)	Q (barn)	RA (%)
58	cerium	^{136}Ce	136	78	4.8739	135.9071293	0^+	0.0	0.0	0.186
		^{138}Ce	138*	80	4.8737	137.9059942	0^+	0.0	0.0	0.251
		^{140}Ce	140	82	4.8771	139.9054484	0^+	0.0	0.0	88.449
		^{142}Ce	142*	84	4.9063	141.9092502	0^+	0.0	0.0	11.114
59	praseodymium	^{141}Pr	141	82	4.8919	140.9076596	$5/2^+$	4.266	-0.077	100.0
60	neodymium	^{142}Nd	142	82	4.9123	141.9077288	0^+	0.0	0.0	27.153
		^{143}Nd	143	83	4.9254	142.9098198	$7/2^-$	-1.065	-0.61	12.173
		^{144}Nd	144*	84	4.9421	143.9100928	0^+	0.0	0.0	23.798
		^{145}Nd	145	85	4.9535	144.9125792	$7/2^-$	-0.656	-0.314	8.293
		^{146}Nd	146	86	4.9696	145.9131225	0^+	0.0	0.0	17.189
		^{148}Nd	148	88	4.9999	147.916899	0^+	0.0	0.0	5.756
		^{150}Nd	150*	90	5.04	149.9209013	0^+	0.0	0.0	5.638
61	promethium	^{147}Pm	147*	86	5.0	146.9151449	$7/2^+$	2.58	0.74	trace
62	samarium	^{144}Sm	144	82	4.9524	143.9120063	0^+	0.0	0.0	3.08
		^{146}Sm	146*	84	4.9808	145.9130468	0^+	0.0	0.0	trace
		^{147}Sm	147*	85	4.9892	146.9149044	$7/2^-$	-0.809	-0.26	15.0
		^{148}Sm	148*	86	5.0042	147.9148292	0^+	0.0	0.0	11.25
		^{149}Sm	149	87	5.0134	148.9171912	$7/2^-$	-0.6677	0.078	13.82
		^{150}Sm	150	88	5.0387	149.917282	0^+	0.0	0.0	7.37
		^{151}Sm	151*	89	5.055	150.9199389	$5/2^-$	-0.3605	0.71	0.0057
		^{152}Sm	152	90	5.0819	151.9197386	0^+	0.0	0.0	26.74
		^{154}Sm	154	92	5.1053	153.9222158	0^+	0.0	0.0	22.74
63	europium	^{150}Eu	150*	87	5.0296	149.9197071	5^-	2.702	1.13	0.0108
		^{151}Eu	151*	88	5.0522	150.9198566	$5/2^+$	3.4635	0.903	47.81
		^{153}Eu	153	90	5.1115	152.9212368	$5/2^+$	1.5294	2.41	0.0062
		^{154}Eu	154*	91	5.1239	153.9229857	3^-	-2.0	2.85	0.0079
		^{155}Eu	155*	92	5.1221	154.9228998	$5/2^+$	1.516	2.5	0.0069
64	gadolinium	^{152}Gd	152*	88	5.0774	151.9197984	0^+	0.0	0.0	0.2
		^{154}Gd	154	90	5.1223	153.920873	0^+	0.0	0.0	2.18
		^{155}Gd	155	91	5.1319	154.9226294	$3/2^-$	-0.2591	1.27	14.8
		^{156}Gd	156	92	5.142	155.9221301	0^+	0.0	0.0	20.47
		^{157}Gd	157	93	5.1449	156.9239674	$3/2^-$	-0.3398	1.36	15.65
		^{158}Gd	158	94	5.1569	157.9241112	0^+	0.0	0.0	24.84
		^{160}Gd	160	96	5.1734	159.9270612	0^+	0.0	0.0	21.86
65	terbium	^{157}Tb	157*	92	5.0489	156.9240319	$3/2^+$	2.03	1.4	0.15
		^{159}Tb	159	94	5.06	158.9253537	$3/2^+$	2.009	1.432	100.0

Table 1.1 (continuation)

Z	element	symbol	A	N	radius (fm)	atomic mass (m_u)	I^{π} ($\hbar$)	μ_I (μ_N)	Q (barn)	RA (%)
66	dysprosium	^{156}Dy	156	90	5.1622	155.9242836	0^+	0.0	0.0	0.056
		^{158}Dy	158	92	5.1815	157.9244148	0^+	0.0	0.0	0.095
		^{160}Dy	160	94	5.1951	159.9252036	0^+	0.0	0.0	2.329
		^{161}Dy	161	95	5.1962	160.9269394	$5/2^+$	-0.479	2.51	18.889
		^{162}Dy	162	96	5.2074	161.9268045	0^+	0.0	0.0	25.475
		^{163}Dy	163	97	5.2099	162.9287372	$5/2^-$	0.671	2.65	24.896
		^{164}Dy	164	98	5.2218	163.9291808	0^+	0.0	0.0	28.26
67	holmium	^{163}Ho	163*	96	5.1907	162.9287403	$7/2^-$	4.22	3.7	trace
		^{165}Ho	165	98	5.2022	164.9303291	$7/2^-$	4.16	3.58	100.0
68	erbium	^{162}Er	162	94	5.2246	161.9287873	0^+	0.0	0.0	0.139
		^{164}Er	164	96	5.2389	163.9292077	0^+	0.0	0.0	1.601
		^{166}Er	166	98	5.2516	165.9303011	0^+	0.0	0.0	33.503
		^{167}Er	167	99	5.256	166.9320562	$7/2^+$	-0.5623	3.57	22.869
		^{168}Er	168	100	5.2644	167.9323783	0^+	0.0	0.0	26.978
		^{170}Er	170	102	5.2789	169.9354719	0^+	0.0	0.0	14.91
69	thulium	^{169}Tm	169	100	5.2256	168.934219	$1/2^+$	-0.231	0.0	100.0
		^{171}Tm	171*	102	5.2388	170.9364352	$1/2^+$	-0.23	0.0	trace
70	ytterbium	^{168}Yb	168	98	5.2702	167.9338913	0^+	0.0	0.0	0.123
		^{170}Yb	170	100	5.2853	169.9347672	0^+	0.0	0.0	2.982
		^{171}Yb	171	101	5.2906	170.9363315	$1/2^-$	0.4923	0.0	14.086
		^{172}Yb	172	102	5.2995	171.9363867	0^+	0.0	0.0	21.686
		^{173}Yb	173	103	5.3046	172.9382162	$5/2^-$	-0.678	2.8	16.103
		^{174}Yb	174	104	5.3108	173.9388675	0^+	0.0	0.0	32.025
		^{176}Yb	176	106	5.3215	175.9425747	0^+	0.0	0.0	12.995
71	lutetium	^{173}Lu	173*	102	5.3577	172.9389357	$7/2^+$	2.273	3.53	trace
		^{174}Lu	174*	103	5.3634	173.9403428	1^-	1.982	0.773	trace
		^{175}Lu	175	104	5.37	174.9407772	$7/2^+$	2.2257	3.49	97.401
		^{176}Lu	176*	105	5.3739	175.9426917	7^-	3.16	4.92	2.599
72	hafnium	^{174}Hf	174*	102	5.3201	173.9400484	0^+	0.0	0.0	0.16
		^{176}Hf	176	104	5.3286	175.9414098	0^+	0.0	0.0	5.26
		^{177}Hf	177	105	5.3309	176.9432302	$7/2^-$	0.791	3.37	18.6
		^{178}Hf	178	106	5.3371	177.9437083	0^+	0.0	0.0	27.28
		^{179}Hf	179	107	5.3408	178.9458257	$9/2^+$	-0.6389	3.79	13.62
		^{180}Hf	180	108	5.347	179.9465595	0^+	0.0	0.0	35.08
		^{182}Hf	182*	110	5.3516	181.9505637	0^+	0.0	0.0	trace
73	tantalum	^{179}Ta	179*	106	5.3	178.9459391	$7/2^+$	2.286	3.27	trace
		^{181}Ta	181	108	5.3507	180.9479985	$7/2^+$	2.365	3.17	99.98799

Table 1.1 (continuation)

Z	element	symbol	A	N	radius	atomic mass	I^{π}	μ_I	Q	RA
					(fm)	(m_u)	($\hbar$)	(μ_N)	(barn)	(%)
74	tungsten	^{180}W	180*	106	5.3491	179.9467133	0^+	0.0	0.0	0.12
		^{182}W	182	108	5.3559	181.9482056	0^+	0.0	0.0	26.5
		^{183}W	183*	109	5.3611	182.9502244	$1/2^-$	0.11739	0.0	14.31
		^{184}W	184	110	5.3658	183.9509332	0^+	0.0	0.0	30.64
		^{186}W	186	112	5.3743	185.9543651	0^+	0.0	0.0	28.43
75	rhenium	^{185}Re	185	110	5.3596	184.9529583	$5/2^+$	3.176	2.07	37.4
		^{187}Re	187*	112	5.3698	186.9557522	$5/2^+$	3.209	1.94	62.6
76	osmium	^{184}Os	184*	108	5.3823	183.9524929	0^+	0.0	0.0	0.02
		^{186}Os	186*	110	5.3909	185.9538376	0^+	0.0	0.0	1.59
		^{187}Os	187	111	5.3933	186.9557496	$1/2^-$	0.06442	0.0	1.96
		^{188}Os	188	112	5.3993	187.9558373	0^+	0.0	0.0	13.24
		^{189}Os	189	113	5.4016	188.9581459	$3/2^-$	0.6576	0.86	16.15
		^{190}Os	190	114	5.4062	189.9584454	0^+	0.0	0.0	26.26
		^{192}Os	192	116	5.4126	191.9614788	0^+	0.0	0.0	40.78
77	iridium	^{191}Ir	191	114	5.3968	190.9605915	$3/2^+$	0.1502	0.8	37.3
		^{193}Ir	193	116	5.4032	192.9629238	$3/2^+$	0.163	0.7	62.7
78	platinum	^{190}Pt	190*	112	5.4108	189.9599498	0^+	0.0	0.0	0.012
		^{192}Pt	192	114	5.4169	191.9610427	0^+	0.0	0.0	0.782
		^{193}Pt	193*	115	5.4191	192.9629845	$1/2^-$	0.601	0.0	trace
		^{194}Pt	194	116	5.4236	193.9626835	0^+	0.0	0.0	32.864
		^{195}Pt	195	117	5.427	194.9647943	$1/2^-$	0.6073	0.0	33.775
		^{196}Pt	196	118	5.4307	195.9649546	0^+	0.0	0.0	25.211
		^{198}Pt	198	120	5.4383	197.9678967	0^+	0.0	0.0	7.356
79	gold	^{197}Au	197	118	5.4371	196.9665701	$3/2^+$	0.1452	0.547	100.0
80	mercury	^{196}Hg	196	116	5.4385	195.9658334	0^+	0.0	0.0	0.15
		^{198}Hg	198	118	5.4463	197.9667692	0^+	0.0	0.0	10.04
		^{199}Hg	199	119	5.4474	198.968281	$1/2^-$	0.5039	0.0	16.94
		^{200}Hg	200	120	5.4551	199.9683269	0^+	0.0	0.0	23.14
		^{201}Hg	201	121	5.4581	200.9703031	$3/2^-$	-0.558	0.387	13.17
		^{202}Hg	202	122	5.4648	201.9706436	0^+	0.0	0.0	9.74
		^{204}Hg	204	124	5.4744	203.973494	0^+	0.0	0.0	6.82
81	thallium	^{203}Tl	203	122	5.4666	202.9723441	$1/2^+$	1.616	0.0	29.515
		^{205}Tl	205	124	5.4759	204.9744273	$1/2^+$	1.632	0.0	70.485
82	lead	^{202}Pb	202*	120	5.4705	201.9721516	0^+	0.0	0.0	trace
		^{204}Pb	204*	122	5.4803	203.9730435	0^+	0.0	0.0	1.4
		^{205}Pb	205*	123	5.4828	204.9744817	$5/2^-$	0.7094	0.23	trace
		^{206}Pb	206	124	5.4902	205.9744652	0^+	0.0	0.0	24.1
		^{207}Pb	207	125	5.4943	206.9758968	$1/2^-$	0.5906	0.0	22.1
		^{208}Pb	208	126	5.5012	207.976652	0^+	0.0	0.0	52.4

Table 1.1 (continuation)

Z	element	symbol	A	N	radius (fm)	atomic mass (m_u)	I^{π} ($\hbar$)	μ_I (μ_N)	Q (barn)	RA (%)
82	lead	^{202}Pb	202*	120	5.4705	201.9721516	0^+	0.0	0.0	trace
		^{204}Pb	204*	122	5.4803	203.9730435	0^+	0.0	0.0	1.4
		^{205}Pb	205*	123	5.4828	204.9744817	$5/2^-$	0.7094	0.23	trace
		^{206}Pb	206	124	5.4902	205.9744652	0^+	0.0	0.0	24.1
		^{207}Pb	207	125	5.4943	206.9758968	$1/2^-$	0.5906	0.0	22.1
		^{208}Pb	208	126	5.5012	207.976652	0^+	0.0	0.0	52.4
83	bismuth	^{207}Bi	207*	124	5.5103	206.9784706	$9/2^-$	4.0729	-0.76	trace
		^{208}Bi	208*	125	5.5147	207.9797421	5^+	4.572	-0.7	trace
		^{209}Bi	209*	126	5.5211	208.9803986	$9/2^-$	4.092	-0.516	100.0
84	polonium	^{209}Po	209*	125	5.5628	208.9824304	$1/2^-$	0.68	0.0	trace
89	actinium	^{227}Ac	227*	138	5.6	227.0277506	$3/2^-$	1.1	1.74	trace
90	thorium	^{229}Th	229*	139	5.7557	229.0317614	$5/2^+$	0.46	3.11	trace
		^{230}Th	230*	140	5.767	230.0331323	0^+	0.0	0.0	0.02
		^{232}Th	232*	142	5.7848	232.0380536	0^+	0.0	0.0	99.98
91	protactinium	^{231}Pa	231*	140	5.8	231.0358825	$3/2^-$	1.99	-1.72	100.0
92	uranium	^{233}U	233*	141	5.8203	233.0396343	$5/2^+$	-0.59	3.663	trace
		^{234}U	234*	142	5.8291	234.0409503	0^+	0.0	0.0	0.0054
		^{235}U	235*	143	5.8337	235.0439281	$7/2^-$	-0.38	4.936	0.7204
		^{236}U	236*	144	5.8431	236.0455661	0^+	0.0	0.0	trace
		^{238}U	238*	146	5.8571	238.0507869	0^+	0.0	0.0	99.2742
93	neptunium	^{237}Np	237*	144	5.85	237.0481716	$5/2^+$	3.16	3.886	trace
94	plutonium	^{239}Pu	239*	145	5.8601	239.0521616	$1/2^+$	0.202	0.0	trace
		^{240}Pu	240*	146	5.8701	240.0538117	0^+	0.0	0.0	trace
		^{241}Pu	241*	147	5.8748	241.0568497	$5/2^+$	-0.678	6.0	trace
		^{242}Pu	242*	148	5.8823	242.058741	0^+	0.0	0.0	trace
		^{244}Pu	244*	150	5.8948	244.0642044	0^+	0.0	0.0	trace
95	americium	^{241}Am	241*	146	5.8928	241.0568273	$5/2^-$	1.6	4.34	trace
		^{243}Am	243*	148	5.9048	243.0613799	$5/2^-$	1.52	4.32	trace

D

Classical Mechanics

D.1 Introduction

The goal of *classical mechanics* is to describe the time evolution of systems of point-like particles, called *classical particles*, subject to Newton's laws of motion. To introduce the subject we consider a mechanical *system* of N classical particles labeled by the *particle index* α. The *position* of particle α is denoted by the vector $\mathbf{r}_\alpha$. The first derivative with respect to time t, $\dot{\mathbf{r}}_\alpha \equiv d\mathbf{r}_\alpha/dt$, is called the *velocity* of the particle and the second derivative, $\ddot{\mathbf{r}}_\alpha \equiv d^2\mathbf{r}_\alpha/dt^2$, its *acceleration.*

We suppose that for a given state of the system the evolution does not depend on the instant that we choose to follow it in time; i.e., *time is homogeneous*. Space is taken to be both *homogeneous and isotropic*; i.e., it is possible to choose a frame of reference in which the evolution of the system is independent of position and orientation. Such a reference frame is called an *inertial* frame. In an inertial frame a *free* particle either stays at rest or moves at constant velocity. This is *Newton's first law*: the law of inertia.

A system of N classical particles has at most $3N$ *degrees of freedom*; i.e., $3N$ independent ways to change in time. For *free* particles this motion can be described by Newton's equations of motion using $3N$ coordinates, for instance the cartesian coordinates x_α, y_α and z_α, with $\alpha \in \{1, \cdots N\}$. However, in many cases the motion is not free (not described by) but subject to constraints. These can be time independent (*scleronomous*) or have an explicit time dependence (*rheonomous*). An example of a scleronomous constraint is the fixed distance between two particles in a rotating rigid body. A ball rolling on the deck of a ship is an example of a system (the ball) with a rheonomous constraint (confinement to the deck of the ship). The system is called *holonomic* if the constraining relations reduce the $3N$ degrees of freedom the system (which can greatly simplify calculations). This is the case for constraint *equalities*, for instance the above mentioned fixed distance between two particles in a rotating rigid body. If the constraints do *not* affect the $3N$ degrees of freedom the system is called *nonholonomic.* An example is a system of particles confined to a box, where the confinement is accounted for by constraint *inequalities.* Nonholonomic constraints can sometimes be avoided by idealizations of the theory.

The field of classical mechanics is subdivided into three subjects: (a) *statics*, which is the theory of mechanical equilibrium between forces; (b) *kinematics*, which is the theory of motion without entering into the origin of this motion; (c) *dynamics*, which is the theory of motion under the influence of forces. This appendix cannot be more than a summary. It is partially based on unpublished lecture notes by Sybren De Groot and Leendert Suttorp at the University of Amsterdam. For a comprehensive introduction the reader is referred to volume 1 of the Landau-Lifshitz series [69] and the book by Herbert Goldstein [52].

D.2 Dynamical evolution

D.2.1 Newton's equation of motion

The time development of a system of N particles under the influence of external forces is called the *dynamical evolution* of that system. In classical mechanics the dynamical evolution of a single particle is described by *Newton's second law*, which states that the total force $\mathbf{F}_\alpha$ acting on particle α is proportional to the acceleration of that particle, with the proportionality constant m_α being its *inertial mass*,

$$\mathbf{F}_\alpha = m_\alpha \ddot{\mathbf{r}}_\alpha. \tag{D.1}$$

For a system of N particles the force $\mathbf{F}_\alpha$ is given by the resultant of all external forces, $\mathbf{F}_\alpha^{\mathrm{ext}}$, and the sum of the interaction forces $\mathbf{F}_{\alpha\beta}$ of particle α with all other particles of the system,

$$\mathbf{F}_\alpha^{\mathrm{ext}} + \sum_{\beta=1}^{N}{}' \mathbf{F}_{\alpha\beta} = m_\alpha \ddot{\mathbf{r}}_\alpha, \tag{D.2}$$

where the prime indicates the condition $\beta \neq \alpha$. This expression represents a set of N coupled differential equations which is impossible to solve for a macroscopic number of particles. However, as we shall see below, in many cases excellent approximations can be obtained by replacing the interparticle forces by one or more relations acting as constraints on the motion of particle α.

D.2.2 Kinematic evolution of holonomous systems

For a *holonomic* system of N particles at positions $\mathbf{r}_\alpha$ with l constraints we can introduce $f = 3N - l$ *independent* coordinates $q_i(t)$ known as the *generalized coordinates*, with $i \in \{1, \cdots f\}$ being the *coordinate index.* Together these define the evolution of the system as a function of time in the f-dimensional *configuration space* of the system. The time derivatives $\dot{q}_i(t)$ are called the *generalized velocities.*[1]

- The *rheonomous* constraints of a N-particle system with $f = 3N - l$ degrees of freedom constitute a set of l *time-dependent* relations $f_k(\mathbf{r}_1, \cdots, \mathbf{r}_N, t) = 0$, where $k \in \{1, \cdots l\}$ is called the *constraint index.* In this case the position of particle α may be expressed as a function of the f coordinates q_i plus time,

  $$\mathbf{r}_\alpha = \mathbf{r}_\alpha(q_1, \cdots, q_f, t). \tag{D.3}$$

  The velocities follow with

  $$\dot{\mathbf{r}}_\alpha = \sum_{i=1}^{f} \frac{\partial \mathbf{r}_\alpha}{\partial q_i} \dot{q}_i + \frac{\partial \mathbf{r}_\alpha}{\partial t}, \tag{D.4}$$

  which shows that the velocity of particle α is (in general) a function of $q_i, \dot{q}_i$ and t,

  $$\dot{\mathbf{r}}_\alpha = \dot{\mathbf{r}}_\alpha(q_1, \cdots, q_f, \dot{q}_1, \cdots, \dot{q}_f, t). \tag{D.5}$$

  Hence, in *rheonomous* systems $\mathbf{r}_\alpha$ and $\dot{\mathbf{r}}_\alpha$ depend *explicitly* on time.

- For *scleronomous* constraints the l constraining relations do *not* contain an explicit time dependence: $f_k(\mathbf{r}_1, \cdots, \mathbf{r}_N) = 0$, with $k \in \{1, \cdots l\}$. This makes it possible to *choose* the f coordinates q_i in such a way that the position of the particles does not depend explicitly on time. For instance, a rigid body has $f = 6$ degrees of freedom, because the position of an arbitrary particle α is specified by $\mathbf{r}_\alpha = \mathbf{r}_\alpha(X, Y, Z, r, \theta, \phi)$, where $\mathbf{R} = (X, Y, Z)$ is the position of the center of mass and $\mathbf{r} = (r, \theta, \phi)$ to position relative to $\mathbf{R}$ inside the body. For a homogeneous rigid sphere on a static surface this reduces to $f = 4$ and $\mathbf{r}_\alpha = \mathbf{r}_\alpha(X, Y, \theta, \phi)$.

---

[1] In this appendix we adopt the following index convention: $\alpha \in \{1, \cdots N\}$ for the *particle index*, $i \in \{1, \cdots f\}$ for the *coordinate index* and $k \in \{1, \cdots l\}$. for the *constraint index.*

D.2.3 Virtual displacements - principle of d'Alembert

The force $\mathbf{F}_\alpha$ acting on particle α is the superposition of all forces acting on that particle. These forces can be separated into *known* forces $\mathbf{F}_\alpha^{\text{ext}}$, like gravity or externally applied forces, and *unknown* forces $\mathbf{K}_\alpha$,

$$\mathbf{F}_\alpha = \mathbf{F}_\alpha^{\text{ext}} + \mathbf{K}_\alpha. \tag{D.6}$$

The unknown forces are called *forces of constraint.* These unknown forces are *reaction forces* that reduce the freedom of motion of the particles. Their actual value is unknown but always such that the constraints remain satisfied, for instance to assure that a train stays in the rails.

The explicit appearance of the reaction forces can be eliminated from the formalism. To demonstrate this we first consider *mechanical equilibrium.* This is the state in which all particles are at rest in their equilibrium position; i.e., $\mathbf{F}_\alpha = 0$ and $\mathbf{K}_\alpha = -\mathbf{F}_\alpha^{\text{ext}}$ for all particles. By pulling on one of the particles we expect to perform work against the restoring force of the system but as long as the displacements $\delta\mathbf{r}_\alpha$ remain small (i.e., close to the equilibrium position) we find that this work vanishes because the forces $\mathbf{F}_\alpha$ remain vanishingly small,

$$\sum_\alpha \mathbf{F}_\alpha \cdot \delta\mathbf{r}_\alpha = \sum_\alpha (\mathbf{F}_\alpha^{\text{ext}} + \mathbf{K}_\alpha) \cdot \delta\mathbf{r}_\alpha = 0. \tag{D.7}$$

In other words, in equilibrium the restoring force is zero. This implies that in the absence of external forces the reaction forces have to be zero, $\mathbf{F}_\alpha = \mathbf{K}_\alpha = 0$.

Now suppose that the system is pulled *out of equilibrium.* In this case we have forces acting on (in general) all the particles; i.e., $\mathbf{F}_\alpha = m_\alpha \ddot{\mathbf{r}}_\alpha \neq 0$ and $\mathbf{K}_\alpha = -\mathbf{F}_\alpha^{\text{ext}} + \mathbf{F}_\alpha \neq 0$. Shifting the external forces to the l.h.s. of (D.7) this equation takes the form

$$\sum_\alpha (m_\alpha \ddot{\mathbf{r}}_\alpha - \mathbf{F}_\alpha^{\text{ext}}) \cdot \delta\mathbf{r}_\alpha = \sum_\alpha \mathbf{K}_\alpha \cdot \delta\mathbf{r}_\alpha. \tag{D.8}$$

The dot product leads us to distinguish between the component of $\delta\mathbf{r}_\alpha$ along $\mathbf{K}_\alpha$ and the component perpendicular to $\mathbf{K}_\alpha$. A parallel displacement would require work but is not in accordance with the constraints; a perpendicular displacement is in accordance with the constraints but does not result in work (as long as the displacements remain sufficiently small). This insight brings us to the *principle of d'Alembert*: under *virtual displacements* $\delta\mathbf{r}_\alpha$ (in accordance with the constraints) no work is done by the reaction forces,

$$\delta W = \sum_\alpha \mathbf{K}_\alpha \cdot \delta\mathbf{r}_\alpha = 0. \tag{D.9}$$

By analyzing virtual displacement we can eliminate the reaction forces from (D.8) and the principle of d'Alembert takes the form

$$\delta W = \sum_\alpha (m_\alpha \ddot{\mathbf{r}}_\alpha - \mathbf{F}_\alpha^{\text{ext}}) \cdot \delta\mathbf{r}_\alpha = 0. \tag{D.10}$$

As this expression also holds outside equilibrium, we have obtained a condition for the *dynamical* evolution of a mechanical system in which the constraints are satisfied implicitly. In Section D.3 the d'Alembert principle will be used to derive the central equations of classical mechanics: the equations of Lagrange.

Let us have a closer look at the virtual displacements. These can be thought of as instantly being present at any desired point in time just to analyze how they would affect the system; i.e., unlike real displacements, which are based on (D.3), virtual displacements have *no* (explicit) time dependence. To be relevant, they have to be in accordance with the constraints; i.e., $\delta\mathbf{r}_\alpha$ has to be perpendicular to $\mathbf{K}_\alpha$. This leads to the practical difficulty that (in principle) the virtual displacement of one particle affects the virtual reaction forces of all other particles; i.e., the $\delta\mathbf{r}_\alpha$ cannot be chosen independently. This difficulty can be eliminated by decomposing the virtual displacements in terms of the generalized coordinates defined by the constraints,

$$\delta\mathbf{r}_\alpha = \sum_{i=1}^{f} \frac{\partial \mathbf{r}_\alpha}{\partial q_i}\delta q_i. \tag{D.11}$$

Note the absence of the time variable. Substituting this expression into (D.10) we obtain

$$\delta W = \sum_{i=1}^{f} \left[\sum_{\alpha=1}^{N} (m_\alpha \ddot{\mathbf{r}}_\alpha - \mathbf{F}_\alpha^{\mathrm{ext}}) \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} \right] \delta q_i = 0 \tag{D.12}$$

and since the q_i are independent variables the principle of d'Alembert takes the form of a set of f coupled differential equations

$$\sum_\alpha (m_\alpha \ddot{\mathbf{r}}_\alpha - \mathbf{F}_\alpha^{\mathrm{ext}}) \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} = 0, \quad \text{with } i \in \{1, \cdots, f\}. \tag{D.13}$$

Inversely, it is easily verified that (D.10) is always valid provided the equations (D.13) are simultaneously satisfied.

Bernoulli's principle

In the absence of dynamical evolution (the *static* case - $\ddot{\mathbf{r}}_\alpha = 0$) the principle of d'Alembert reduces to the *Bernoulli principle* of virtual displacements :

$$\delta W = \sum_\alpha \mathbf{F}_\alpha^{\mathrm{ext}} \cdot \delta\mathbf{r}_\alpha = 0. \tag{D.14}$$

As an example of *Bernoulli's principle* without constraints we consider a system of $N = 2$ identical particles subject to a force $\mathbf{F}_\alpha^{\mathrm{ext}} = -\boldsymbol{\nabla}_\alpha U(\mathbf{r}_1, \mathbf{r}_2)$, where $U(\mathbf{r}_1, \mathbf{r}_2)$ is a potential function which depends on the position of both particles. This is an example of a conservative force (see Section D.5.2). The expression for the virtual work is $\delta W = \mathbf{F}_1^{\mathrm{ext}} \cdot \delta\mathbf{r}_1 + \mathbf{F}_2^{\mathrm{ext}} \cdot \delta\mathbf{r}_2 = 0$. For two free particles confined by a harmonic potential the condition $\delta W = 0$ is satisfied at the minimum of the potential. For two repulsive particles we have $\mathbf{F}_1^{\mathrm{ext}} = -\mathbf{F}_2^{\mathrm{ext}}$ and the condition $\delta W = 0$ is satisfied for $\delta\mathbf{r}_1 = -\delta\mathbf{r}_2$ perpendicular to $\mathbf{F}_1^{\mathrm{ext}}$. Note the correlation between $\delta\mathbf{r}_1$ and $\delta\mathbf{r}_2$ (these are *not* independent). The virtual displacements correspond to a rotation of the pair about the potential center. The latter case shows that the virtual variation does not automatically correspond to minimum of δW but can also represent a degenerate case.

Kinetic energy relations

Introducing the *kinetic energy*,

$$T \equiv \tfrac{1}{2} \sum_\alpha m_\alpha \dot{\mathbf{r}}_\alpha^2, \tag{D.15}$$

we note that, in view of (D.5), the kinetic energy of particle α is in the most general case a function of $q_i(t), \dot{q}_i(t)$ and t,

$$T = T(q_1, \cdots, q_f, \dot{q}_1, \cdots, \dot{q}_f, t). \tag{D.16}$$

For future convenience of use we derive the partial derivatives with respect to q_i and $\dot{q}_i$,

$$\frac{\partial T}{\partial q_i} = \sum_\alpha m_\alpha \dot{\mathbf{r}}_\alpha \cdot \frac{\partial \dot{\mathbf{r}}_\alpha}{\partial q_i} \tag{D.17a}$$

$$\frac{\partial T}{\partial \dot{q}_i} = \sum_\alpha m_\alpha \dot{\mathbf{r}}_\alpha \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i}. \tag{D.17b}$$

Here we used the relation $\partial \dot{\mathbf{r}}_\alpha / \partial \dot{q}_i = \partial \mathbf{r}_\alpha / \partial q_i$, which follows immediately from (D.4) and (D.3).

Problem D.1. Consider the kinematics of a single particle moving in a stationary circular orbit of radius $R = 1$. Derive the equations of motion and show that motion is in accordance with the principle of d'Alembert.

Solution. To describe the orbit we choose a two-dimensional cartesian coordinate system with the position vector $\mathbf{r}$ given by

$$x = r\cos\phi \qquad y = r\sin\phi,$$

where (x, y) are the cartesian coordinates and (r, ϕ) the polar coordinates of $\mathbf{r}$. In this problem we have one (skeronomic) constraint: the radius of the orbit is constant. Choosing the origin of the coordinate system at the center of the circle the constraint takes the form $r = 1$ and the motion can be described using ϕ as the only generalized coordinate, $\mathbf{r} = \mathbf{r}(\phi)$. The time evolution is given by

$$\begin{aligned} x(t) &= \cos\omega t, \quad & \dot{x}(t) &= -\omega\sin\omega t, \quad & \ddot{x}(t) &= -\omega^2\cos\omega t \\ y(t) &= \sin\omega t, \quad & \dot{y}(t) &= \omega\cos\omega t, \quad & \ddot{y}(t) &= -\omega^2\sin\omega t, \end{aligned}$$

where $\phi(t) \equiv \omega t$, with $\omega = \dot{\phi}$ the angular velocity about the origin. These expressions show that the velocity is always tangential to the circle with the acceleration pointing to the center. Note that the velocity is proportional to ω and the acceleration scales with ω^2. As there are no "known" forces acting on the particle, the d'Alembert principle tells us that the force must a reaction force, $\mathbf{K} = m\ddot{\mathbf{r}}$, which constrains the orbit of the particle to the circle. Furthermore, as no work is done by reaction forces under virtual displacements,

$$\delta W = m\ddot{\mathbf{r}} \cdot \delta\mathbf{r} = 0,$$

the virtual displacements have to be orthogonal to the force; i.e., tangential to the circle (in accordance with the constraint $r = 1$). Note that the x and y motion only differ by a constant phase shift. Hence, the circular motion maps on that of a one-dimensional harmonic oscillator, which also is a mechanical system with one degree of freedom ($f = 1$). □

D.3 Lagrange equations

In this section we derive the equations of Lagrange, which are the equations of motion of classical mechanics. The Lagrange equations offer important advantages over Newton's equations of motion because forces of constraint are eliminated from the formalism. In this way the *state* of a classical system of f degrees of freedom can be uniquely specified by a set of $2f$ generalized coordinates,

$$(\mathbf{q}, \dot{\mathbf{q}}) \equiv (q_1, \cdots, q_f; \dot{q}_1, \cdots, \dot{q}_f). \tag{D.18}$$

This set is known as the *classical phase* of the system. As the generalized coordinates are linearly independent, the classical phase corresponds to a point in a $2f$-dimensional vector space, called the (generalized) *phase space* of the system. Each point in the phase space corresponds to a different *classical state.* In the Lagrange equations the vectors $\mathbf{q} \equiv (q_1, \cdots, q_f)$ and $\dot{\mathbf{q}} \equiv (\dot{q}_1, \cdots, \dot{q}_f)$ serve as *dynamical variables* because the evolution of the classical phase in time represents the dynamical evolution of the system. Separately, $\mathbf{q}$ and $\dot{\mathbf{q}}$ evolve as vectors in the (generalized) *configuration space* and (generalized) *velocity space*, respectively. The Lagrange equations are equally valid for conservative and non-conservative forces. Examples of conservative forces are the *gravitational* force and the *Lorentz* force. The friction force is an example of a non-conservative force. In the present section the Lagrange equations are obtained from the principle of d'Alembert. In Section D.4 they will be derived from the principle of least action of Hamilton.

D.3.1 Absence of constraining forces

First we show how the Lagrange equations are obtained form Newton's second law. In the absence of constraining forces the particles are called *free* and the equations of motion (D.1) suffice to describe the motion of the particles. Using Eq. (D.15) for the kinetic energy, we find that a force acting on particle α can be written in the form

$$\mathbf{F}_\alpha = \frac{d}{dt}\frac{\partial T}{\partial \dot{\mathbf{r}}_\alpha}. \tag{D.19}$$

Now we restrict ourselves to forces that can be expressed as the negative gradient of a *potential function* $U = U(\mathbf{r}_1, \cdots, \mathbf{r}_N, t)$,

$$\mathbf{F}_\alpha^{\text{ext}} = -\frac{\partial U}{\partial \mathbf{r}_\alpha}. \tag{D.20}$$

In the special case of a *time-independent* potential function the force field is called *conservative* and potential function can be interpreted as the *potential energy* $U = V(\mathbf{r}_1, \cdots, \mathbf{r}_N)$. We return to the concept of potential energy in the context of the conservation laws (see Section D.5.2). Combining (D.19) and (D.20) we find

$$-\frac{\partial U}{\partial \mathbf{r}_\alpha} = \frac{d}{dt}\frac{\partial T}{\partial \dot{\mathbf{r}}_\alpha}. \tag{D.21}$$

At this point we introduce the *Lagrangian* $\mathcal{L}(\mathbf{r}_\alpha, \dot{\mathbf{r}}_\alpha, t)$ as the difference between the kinetic energy and the potential function,

$$\mathcal{L}(\mathbf{r}_\alpha, \dot{\mathbf{r}}_\alpha, t) \equiv T(\dot{\mathbf{r}}_\alpha) - U(\mathbf{r}_\alpha, t). \tag{D.22}$$

Because the potential function $U(\mathbf{r}_1, \cdots, \mathbf{r}_N, t)$ does not depend explicitly on $\dot{\mathbf{r}}_\alpha$, and $T(\dot{\mathbf{r}}_1, \cdots, \dot{\mathbf{r}}_N)$ not explicitly on $\mathbf{r}_\alpha$, (D.21) we arrive at the so-called *Lagrange equations* for particle α,

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_\alpha} - \frac{\partial \mathcal{L}}{\partial \mathbf{r}_\alpha} = 0, \quad \text{with } \alpha \in \{1, \cdots, N\}. \tag{D.23}$$

D.3.2 Presence of constraining forces

Lagrange equations can also be derived for systems with *holonomic* constraints. So let us turn to a system of N particles in which *holonomic* constraints add up to the reaction forces $\mathbf{K}_\alpha$. How do these constraints affect the evolution of the system? To answer this question we apply the principle of d'Alembert. Rewriting (D.13) in the form

$$\sum_\alpha \mathbf{F}_\alpha^{\text{ext}} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} = \sum_\alpha m_\alpha \ddot{\mathbf{r}}_\alpha \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i}, \quad \text{with } i \in \{1, \cdots, f\}, \tag{D.24}$$

we can introduce a *generalized force* defined by f components, one for every degree of freedom of the system,

$$Q_i \equiv \sum_\alpha \mathbf{F}_\alpha^{\text{ext}} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i}, \quad \text{with } i \in \{1, \cdots, f\}. \tag{D.25}$$

Using the r.h.s. of (D.24) the Q_i can be expressed as

$$Q_i = \frac{d}{dt}\left(\sum_\alpha m_\alpha \dot{\mathbf{r}}_\alpha \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i}\right) - \sum_\alpha m_\alpha \dot{\mathbf{r}}_\alpha \cdot \frac{\partial \dot{\mathbf{r}}_\alpha}{\partial q_i}, \tag{D.26}$$

and after substitution of (D.17) this becomes

$$Q_i = \frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \frac{\partial T}{\partial q_i}, \quad \text{with } i \in \{1, \cdots, f\}. \tag{D.27}$$

This expression holds for any holonomic system. It has the structure of a Lagrange equation but the forces remain to be specified. Below we derive the Lagrange equation for three types of forces: (a) conservative or non-conservative potential forces; (b) generalized potential forces; (c) non-potential forces.

(a) If the components of the generalized force can be expressed as the gradients of the potential function $U = U(\mathbf{r}_1, \cdots, \mathbf{r}_N, t)$ we can substitute (D.20) into (D.25) and obtain

$$Q_i \equiv -\sum_\alpha \frac{\partial U}{\partial \mathbf{r}_\alpha} \cdot \frac{\partial \mathbf{r}_\alpha}{\partial q_i} = -\frac{\partial U}{\partial q_i}. \tag{D.28}$$

Substituting this equation into (D.27) and taking into account $\partial U/\partial \dot{q}_i = 0$ (because U does not depend explicitly on the $\dot{q}_i$), the equations of motion takes the form

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = 0, \quad \text{with } i \in \{1, \cdots, f\}, \tag{D.29}$$

where

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) \equiv T(\mathbf{q}, \dot{\mathbf{q}}, t) - U(\mathbf{q}, t). \tag{D.30}$$

The set of equations (D.29) will be referred to as the *Lagrange equations*. Note that in the non-holonomic case we have $f = 3N$ and the set of generalized coordinates $\{q_i\}$ coincides with the full set of cartesian coordinates $\{\mathbf{r}_\alpha\}$; i.e., (D.29) coincides with (D.23).

(b) The derivation of Eq. (D.29) breaks down in cases where $\partial U/\partial \dot{q}_i \neq 0$; i.e., for velocity-dependent forces. However, by construction, the Lagrange equations will still be obtained if the generalized forces can be written the form

$$Q_i = -\frac{\partial U}{\partial q_i} + \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_i}\right), \quad \text{with } i \in \{1, \cdots, f\}. \tag{D.31}$$

Substituting this expression into (D.27) we find

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = 0, \quad \text{with } i \in \{1, \cdots, f\}. \tag{D.32}$$

Here the Lagrangian is defined as.

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) \equiv T(\mathbf{q}, \dot{\mathbf{q}}, t) - U(\mathbf{q}, \dot{\mathbf{q}}, t). \tag{D.33}$$

In the presence of a velocity dependence the function $U = U(\mathbf{q}, \dot{\mathbf{q}}, t)$ is called a *generalized potential function.*

(c) Not all forces can be derived from a potential function. However, also for non-potential forces (for instance friction forces) the equation (D.27) is valid. In the most general case, when both potential and non-potential forces are present, the Lagrange equations can be written in the form

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = Q_i, \quad \text{with } i \in \{1, \cdots, f\}. \tag{D.34}$$

Here Q_i are the components of the non-potential force. Forces that can be derived from a potential function are presumed to be contained in the Lagrangian.

D.3.2.1 Motion in a central potential field

As a first example we consider the motion of a particle of mass m freely moving in a *central* potential, $\mathcal{U}(r)$. This is a *conservative* potential which only depends on the distance of the particle to the center of the potential. As the force is central, the orbit will be confined to the plane defined by the position and velocity of the particle at $t = 0$. The Lagrangian is given by

$$\mathcal{L} = \tfrac{1}{2} m \, \dot{\mathbf{r}}^2 - \mathcal{U}(r). \tag{D.35}$$

Choosing the xy plane of a cartesian coordinate system with its origin at the potential center, the position vector given by

$$\mathbf{r} = \hat{\mathbf{x}}\, r\cos\phi + \hat{\mathbf{y}}\, r\sin\phi \tag{D.36}$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are the unit vectors along the x and y direction, and (r,ϕ) are the polar coordinates of $\mathbf{r}$. Differentiating Eq. (D.36) with respect to time we find for the velocity

$$\dot{\mathbf{r}} = \hat{\mathbf{x}}\,(\dot{r}\cos\phi - r\sin\phi\,\dot{\phi}) + \hat{\mathbf{y}}\,(\dot{r}\sin\phi + r\cos\phi\,\dot{\phi}). \tag{D.37}$$

Then, using the orthogonality of the unit vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ we obtain

$$\dot{\mathbf{r}}^2 = (\dot{r}\cos\phi - r\sin\phi\,\dot{\phi})^2 + (\dot{r}\sin\phi + r\cos\phi\,\dot{\phi})^2 = \dot{r}^2 + r^2\dot{\phi}^2 \tag{D.38}$$

and the Lagrangian becomes

$$\mathcal{L} = \tfrac{1}{2}m\,(\dot{r}^2 + r^2\dot{\phi}^2) - \mathcal{U}(r). \tag{D.39}$$

Since the polar coordinates are obtained from the cartesian coordinates by an orthogonal transformation, both sets of coordinates can serve as a complete set of generalized coordinates on which the action principle is based. In other words, the Lagrange equations also apply to the case of polar coordinates.

For the ϕ motion the Lagrange equation becomes

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\phi}} = \frac{\partial\mathcal{L}}{\partial\phi} \quad\Leftrightarrow\quad \frac{d}{dt}\left(mr^2\dot{\phi}\right) = 0. \tag{D.40}$$

This shows that the quantity $mr^2\dot{\phi}$ is *conserved.* In this quantity we recognize the angular momentum about the potential center, $m\,(\mathbf{r}\times\mathbf{v}) = mr^2\dot{\phi}$. Note that this conservation law implies that the area enclosed by the orbit in a given time is conserved in time

$$dA = \tfrac{1}{2}r^2 d\phi = c\,dt. \tag{D.41}$$

This Kepler's second law. In particular, for a *circular* orbit the angular velocity is conserved: $\dot{\phi} \equiv \omega$ and $\phi = \omega t$ (with the initial phase of the circular motion chosen to be zero at $t = 0$).

For the r motion the Lagrange equation becomes

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{r}} = \frac{\partial\mathcal{L}}{\partial r} \quad\Leftrightarrow\quad F = \frac{d}{dt}\,(m\dot{r}) = mr\dot{\phi}^2 - \frac{\partial\mathcal{U}(r)}{\partial r}. \tag{D.42}$$

For a *circular* orbit this reduces to the well-known condition for *stationary* motion,

$$m\omega^2 r - \partial\mathcal{U}(r)/\partial r = 0 \tag{D.43}$$

D.3.2.2 Friction force

In this example we ask for the steady-state velocity of particle falling in a gravitational field in the presence of friction. For a gravitational field along the z axis the potential is $V = \alpha z$. This is a conservative potential field. The friction force is a velocity dependent force,

$$F = -\beta\dot{z}. \tag{D.44}$$

This is a *non-conservative* and therefore *non-Newtonian* force. The Lagrangian is determined only by the kinetic energy and the potential fields (in this case only the gravitational potential),

$$\mathcal{L} = \tfrac{1}{2}\dot{z}^2 - \alpha z. \tag{D.45}$$

Substituting the friction force into the r.h.s. of the Lagrange equation (D.29) we obtain

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{z}}-\frac{\partial\mathcal{L}}{\partial z}=-\beta\dot{z}. \tag{D.46}$$

Substituting the Lagrangian (D.45) we obtain

$$\ddot{z}+\beta\dot{z}-\alpha=0. \tag{D.47}$$

The steady-state condition is $\ddot{z}=0$; hence, we find for the steady-state velocity

$$\dot{z}=-\alpha/\beta. \tag{D.48}$$

D.3.2.3 Lorentz force

The celebrated example of a velocity-dependent force for which a generalized potential can be defined is the *Lorentz* force. This force governs the motion of a charged particle moving through an electromagnetic field in vacuum. The *Lorentz* force is given by

$$\mathbf{F}=q(\mathbf{E}+\mathbf{v}\times\mathbf{B}), \tag{D.49}$$

where q is the electric charge and $\mathbf{v}=\dot{\mathbf{r}}$ the velocity of the particle. The *generalized electromagnetic potential* from which the Lorentz force may be derived has the form,

$$U(\mathbf{r},\mathbf{v},t)=q(\varphi-\mathbf{v}\cdot\mathbf{A}), \tag{D.50}$$

where $\varphi(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$ are the scalar and vector potentials of the electromagnetic field. To demonstrate this assertion we start from (D.31) and find in the vector notation [1]

$$\mathbf{F}(\mathbf{r},\dot{\mathbf{r}},t)=-\partial_{\mathbf{r}}U+\frac{d}{dt}\partial_{\dot{\mathbf{r}}}U=-q[\partial_{\mathbf{r}}\varphi-\partial_{\mathbf{r}}(\dot{\mathbf{r}}\cdot\mathbf{A})+\frac{d}{dt}\partial_{\dot{\mathbf{r}}}(\dot{\mathbf{r}}\cdot\mathbf{A})]. \tag{D.51}$$

Because $\mathbf{A}(\mathbf{r},t)$ does not depend explicitly on $\dot{\mathbf{r}}$ and in view of the vector rule (N.17) the last term of (D.51) can be rewritten as

$$\frac{d}{dt}\partial_{\dot{\mathbf{r}}}(\dot{\mathbf{r}}\cdot\mathbf{A})=\frac{d\mathbf{A}}{dt}=\partial_t\mathbf{A}+(\dot{\mathbf{r}}\cdot\partial_{\mathbf{r}})\mathbf{A}. \tag{D.52}$$

Substitution of this expression into (D.51) we obtain

$$\mathbf{F}=-q[\partial_{\mathbf{r}}\varphi+\partial_t\mathbf{A}+(\dot{\mathbf{r}}\cdot\partial_{\mathbf{r}})\mathbf{A}-\partial_{\mathbf{r}}(\dot{\mathbf{r}}\cdot\mathbf{A})], \tag{D.53}$$

which yields with (N.16) the expression for the Lorentz force

$$\mathbf{F}=-q(\partial_{\mathbf{r}}\varphi+\partial_t\mathbf{A}-\dot{\mathbf{r}}\times\partial_{\mathbf{r}}\times\mathbf{A})=q(\mathbf{E}+\mathbf{v}\times\mathbf{B}). \tag{D.54}$$

D.4 The Lagrange formalism

In Section D.3 the Lagrange equations were derived from the Newton equations of motion together with the principle of d'Alembert. An alternative approach is to derive the Lagrange equations from the *principle of Hamilton.* In this approach we start by noting that the time evolution of the system corresponds to a unique *path* of the vector $\mathbf{q}\equiv(q_1,\cdots,q_f)$ in the f-dimensional configuration space of the generalized coordinates. As these coordinates are mutually independent, this path can be predicted over an infinitesimal time interval from t to $t+\delta t$ by the relation

$$\mathbf{q}(t+\delta t)=\mathbf{q}(t)+\dot{\mathbf{q}}(t)\delta t. \tag{D.55}$$

Hence, to predict the continuation of the path at time t, *both* $\mathbf{q}(t)$ and $\dot{\mathbf{q}}(t)$ have to be known. In other words, together $\mathbf{q}$ and $\dot{\mathbf{q}}$ fully determine the *classical state* of the system (often represented by a point in the so-called *phase space* of position and momentum).

---

[1] In this example we use the short-hand notation $\partial_{\mathbf{q}}\equiv\partial/\partial\mathbf{q}\equiv(\partial/\partial q_x,\partial/\partial q_y,\partial/\partial q_z)$ and $\partial_t\equiv\partial/\partial t$.

D.4.1 Principle of Hamilton

Let us consider the case where the *actual* path $\mathbf{q}(t)$ runs from the starting point $\mathbf{q}(t_0)$ to the end point $\mathbf{q}(t_1)$ during the time interval (t_0, t_1), and let $\mathbf{q}'(t)$ be some *adjacent* path, slightly differing from the actual path but with the same starting point and end point. The key idea behind the principle of Hamilton is that any continuous deviation from the actual path,

$$\delta\mathbf{q}(t) = \mathbf{q}'(t) - \mathbf{q}(t), \tag{D.56}$$

with boundary conditions

$$\delta\mathbf{q}(t_0) = 0 \quad \text{and} \quad \delta\mathbf{q}(t_1) = 0, \tag{D.57}$$

has to be unfavorable in some respect. Thus we search for a functional of $\mathbf{q}$ and $\dot{\mathbf{q}}$ (the *cost functional* of the variational problem) which is minimal for the actual path from $\mathbf{q}(t_0)$ to $\mathbf{q}(t_1)$. Minimizing this functional with respect to the path provides us with criteria that have to be satisfied along the actual path (for instance conservation laws). This cost functional is called the *action* and is given by the integral of a lagrangian function $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$ over the time interval (t_0, t_1),

$$S = \int_{t_0}^{t_1} \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)dt. \tag{D.58}$$

Whatever this “Lagrangian” may turn out to be, the expression for the action is plausible; at each point in time t within the interval (t_0, t_1), knowledge of $\mathbf{q}$ and $\dot{\mathbf{q}}$ suffice to determine the actual path from $\mathbf{q}(t_0)$ to $\mathbf{q}(t_1)$. So if the Lagrangian is to be deterministic for the actual path, it can only be a function of the variables $\mathbf{q}$, $\dot{\mathbf{q}}$ and t. This minimization procedure is called the *action principle*[1]or the *principle of Hamilton.*

The principle of Hamilton is a variational principle that enables us to derive properties of the Lagrangian for systems in which the function $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$ is at least twice differentiable; i.e., *smooth* with respect to the variables $\mathbf{q}$, $\dot{\mathbf{q}}$ and t. Fortunately, many systems fall in this class and we shall study some examples. Let us consider the change in action of such a system over the time interval (t_0, t_1) when deforming the path defined by $\mathbf{q}(t)$ and $\dot{\mathbf{q}}(t)$ into an adjacent smooth continuous path $\mathbf{q}'(t)$ and $\dot{\mathbf{q}}'(t)$ given by $\mathbf{q}' = \mathbf{q} + \delta\mathbf{q}$ and $\dot{\mathbf{q}}' = \dot{\mathbf{q}} + \delta\dot{\mathbf{q}}$,

$$\delta S = \int_{t_0}^{t_1} [\mathcal{L}(\mathbf{q} + \delta\mathbf{q}, \dot{\mathbf{q}} + \delta\dot{\mathbf{q}}, t) - \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)]dt. \tag{D.59}$$

As the Lagrangian is smooth we can use a Taylor expansion with respect to $\delta\mathbf{q}$ and $\delta\dot{\mathbf{q}}$,[2]

$$\mathcal{L}(\mathbf{q} + \delta\mathbf{q}, \dot{\mathbf{q}} + \delta\dot{\mathbf{q}}, t) = \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) + \frac{\partial\mathcal{L}}{\partial\mathbf{q}} \cdot \delta\mathbf{q} + \frac{\partial\mathcal{L}}{\partial\dot{\mathbf{q}}} \cdot \delta\dot{\mathbf{q}} + \cdots, \tag{D.60}$$

and the change of action takes the form

$$\delta S = \int_{t_0}^{t_1} \sum_i \left(\frac{\partial\mathcal{L}}{\partial q_i}\delta q_i + \frac{\partial\mathcal{L}}{\partial\dot{q}_i}\delta\dot{q}_i \right) dt. \tag{D.61}$$

By partial integration of the second term this integral becomes

$$\delta S = \sum_i \left[\frac{\partial\mathcal{L}}{\partial\dot{q}_i}\delta q_i \right]_{t_0}^{t_1} + \int_{t_0}^{t_1} \sum_i \left(\frac{\partial\mathcal{L}}{\partial q_i} - \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{q}_i} \right) \delta q_i dt. \tag{D.62}$$

---

[1] Traditionally the action principle is called the *principle of least action.* This is an unfortunate name because the variational method provides stationary points rather than minima.

[2] Note the following equivalent notations for the gradient operator, $\partial_{\mathbf{q}} \equiv \partial/\partial\mathbf{q} \equiv (\partial/\partial q_1, \cdots, \partial/\partial q_f)$.

Since the variation is zero at the boundaries of the path, see (D.57), this expression reduces to

$$\delta S = \int_{t_0}^{t_1} \sum_i \left(\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \delta q_i dt. \tag{D.63}$$

At this point we can formulate a necessary condition for minimal action: the functional must be stationary ($\delta S = 0$) under *arbitrary* variations of the path,[1]

$$\int_{t_0}^{t_1} \sum_i \left(\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \delta q_i dt = 0. \tag{D.64}$$

Since the variations δq_i are independent this condition is only satisfied if all terms of the summation vanish simultaneously,

$$\frac{\partial \mathcal{L}}{\partial q_i} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} = 0, \quad \text{with } i \in \{1, \cdots, f\}. \tag{D.65}$$

These equations constitute the *Euler-Lagrange equations* of the action variational principle in which we recognize the *Lagrange equations* of classical mechanics. In Section D.3 these were obtained starting from the principle of d'Alembert.

The action principle has revealed the Lagrangian as the optimal functional from which the laws of physics of a mechanical of the system can be derived. As the Lagrangian is optimized by variation of the path in phase space it is independent of the choice of coordinates. The latter also holds for the form of the Lagrange equations.

Properties of the Lagrangian

Above the Lagrangian $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$ is introduced as a smooth (at least twice differentiable) function of the variables, $\mathbf{q}$, $\dot{\mathbf{q}}$ and t for which the action is stationary under the condition that the Lagrange equations (D.65) are satisfied. Lagrangians have the following properties:

(a) *Additive property*: The Lagrangian $\mathcal{L}$ of a system consisting of two non-interacting subsystems can be written as the sum of the Lagrangians $\mathcal{L}_1$ and $\mathcal{L}_2$ of the two subsystems,

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2. \tag{D.66}$$

This has to be the case because the paths of the subsystems in configuration space can be varied independently. Hence, the least action of the total system must coincide with the sum of the least actions of the two subsystems.

(b) *Definition up to a multiplicative factor*: The path of least action is invariant under multiplication of the Lagrangian by an arbitrary constant. For $\mathcal{L}' = \alpha \mathcal{L}$ we find

$$S' = \alpha \int_{t_0}^{t_1} \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) dt = \alpha S. \tag{D.67}$$

Hence, $\delta S' = \alpha \delta S$, which implies that the condition $\delta S' = 0$ is equivalent to $\delta S = 0$. This means that the multiplication by α does not affect the stationarity condition and leads to the same Euler-Lagrange equation.

(c) *Definition up to a total time derivative*: The path of least action is invariant under addition of a total time derivative to the Lagrangian. For $\mathcal{L}' = \mathcal{L} + d\mathrm{f}/dt$ we find

$$S' = \int_{t_0}^{t_1} \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) dt + \int_{t_0}^{t_1} \frac{d\mathrm{f}}{dt} dt = S + \mathrm{f}(t_1) - \mathrm{f}(t_0). \tag{D.68}$$

---

[1] Note that the inverse is not generally true: being stationary does not mean that the action is minimal.

Hence, S' and S differ by a constant and this does not affect the variation, $\delta S' = \delta S$. Thus, the addition of $d\alpha/dt$ does not change the stationarity condition and leads to the same Euler-Lagrange equation.

D.4.2 Lagrangian of a free particle

To gain insight in the functional form of the Lagrangian we consider a free particle. In classical mechanics space and time are postulated to be homogeneous, so the Lagrangian of the particle (which determines the time evolution of the system) must be independent of position $\mathbf{r}$ and time t. This leaves the velocity $\mathbf{v} = \dot{\mathbf{r}}$ as the only variable, $\mathcal{L} = \mathcal{L}(\mathbf{v})$. Since $\mathcal{L}$ is independent of $\mathbf{r}$ we have $\partial\mathcal{L}/\partial\mathbf{r} = 0$ and the Lagrange equation (D.65) becomes

$$\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\mathbf{v}} = 0. \tag{D.69}$$

Hence, $\partial\mathcal{L}/\partial\mathbf{v}$ must be a constant and since $\mathbf{v}$ is the only variable of $\mathcal{L}$ the velocity must be constant, $\mathbf{v}(t) = \mathbf{v}_0$. Thus we found *Newton's law of inertia.*

Space is postulated to be isotropic. Thus, the Lagrangian may only depend on the absolute value of the velocity (the *speed* v); i.e., it has to be some function of the velocity squared,

$$\mathcal{L} = \mathcal{L}(v^2). \tag{D.70}$$

This observation immediately implies that the Lagrangian is invariant under time reversal; i.e., time is isotropic in newtonian mechanics. To get more precise information about the functional dependence of $\mathcal{L}$ on v^2 we turn to *Galileo's principle of relativity*, which is also based on the homogeneity and isotropy of space and time. According to this principle the equations of motion (hence also the Lagrangian) must be *invariant under a Galilean transformation*, which is a transformation between two inertial frames, $\mathcal{S}$ and $\mathcal{S}'$, in which one ($\mathcal{S}'$) moves at constant velocity $\mathbf{V}$ with respect to the other ($\mathcal{S}$),

$$\mathbf{r}' = \mathbf{r} + \mathbf{V}t; \qquad t' = t, \tag{D.71}$$

Thus we are led to compare two Lagrangians, $\mathcal{L}(v'^2)$ and $\mathcal{L}(v^2)$, defined with respect to two inertial frames which move with an infinitesimally small but *constant* velocity $\mathbf{w} = \mathbf{v}' - \mathbf{v}$ with respect to each other. Using the relation $v'^2 = v^2 + 2\mathbf{v}\cdot\mathbf{w} + w^2$ we find

$$\mathcal{L}(v'^2) = \mathcal{L}(v^2) + \frac{\partial\mathcal{L}}{\partial v^2}(2\mathbf{v}+\mathbf{w})\cdot\mathbf{w} + \frac{\partial^2\mathcal{L}}{\partial(v^2)^2}\left[(2\mathbf{v}+\mathbf{w})\cdot\mathbf{w}\right]^2 + \cdots. \tag{D.72}$$

As the time evolution of the system must be independent of the choice of inertial frame, the two Lagrangians must be equivalent. In view of property (D.68) this means that they cannot differ by more than a total time derivative. Since $\mathbf{v}$ is by definition a total time derivative and $\mathbf{w}$ a constant, $\partial\mathcal{L}/\partial v^2$ has to be a constant; i.e., the function $\mathcal{L}$ has to be a linear function of v^2,

$$\mathcal{L} = \mathcal{L}_0 + \alpha v^2. \tag{D.73}$$

For any other functional dependence on v^2 the Lagrangian is either trivially zero or differs from $\mathcal{L}$ by more than a total time derivative. Once we accept that $\mathcal{L}$ has to be a linear function of v^2 we find that all higher order terms of the expansion vanish. As the constant $\mathcal{L}_0$ can also be regarded as a total time derivative we may write the Lagrangian as simply proportional to v^2,

$$\mathcal{L} = \tfrac{1}{2}mv^2 = T. \tag{D.74}$$

The proportionality constant m is called the *mass* of the particle and serves to calibrate the Lagrangian. This calibration is possible in view of property (D.67). The chosen calibration yields the *kinetic energy* of the particle, $T = \frac{1}{2}mv^2$.

For a free particle the principle of least action is very intuitive. The particle has to follow a straight path in configuration space (which coincides in this case with real space) because for a constant Lagrangian any deviation from a straight path leads to an increase of the action integral. The example also shows that the mass must be positive ($m > 0$) because otherwise the action could be reduced by deviating from the straight path, which contradicts experimental observation.

D.4.3 Lagrangian of a single particle in a potential field

Let us now add to the kinetic energy of the particle a (smooth) function of position and time (i.e., a time-dependent field),

$$\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) = \tfrac{1}{2}mv^2 - U(\mathbf{r}, t). \tag{D.75}$$

Being a function of $\mathbf{r}$, $\dot{\mathbf{r}}$ and t, the new expression is again a Lagrangian if the Lagrange equations are satisfied:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{r}} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \mathbf{v}}. \tag{D.76}$$

When will this be the case? Since $U(\mathbf{r}, t)$ is independent of $\mathbf{v}$ and $T = \frac{1}{2}m\mathbf{v}^2$ is independent of $\mathbf{r}$ and t, substitution of (D.75) into the Lagrange equations (D.76) leads to the following condition

$$-\frac{\partial U}{\partial \mathbf{r}} = \frac{d}{dt}\frac{\partial T}{\partial \mathbf{v}} = m\frac{d\mathbf{v}}{dt}. \tag{D.77}$$

Interpreting the negative gradient of $U(\mathbf{r}, t)$ as a force,

$$\mathbf{F} = -\frac{\partial U}{\partial \mathbf{r}} \tag{D.78}$$

we recognize in (D.77) *Newton's second law*, $\mathbf{F} = m\ddot{\mathbf{r}}$. Hence, the Lagrange equations are satisfied if the particle moves in accordance with the Newton equations of motion; i.e., under non-relativistic conditions. The function $U(\mathbf{r}, t)$ is called the *potential field* of the system. If this field does not depend on position, $U(\mathbf{r}, t) = U(t)$, it is called *uniform.* If it does not depend on time, $U(\mathbf{r}, t) = U(\mathbf{r})$, the field is called *conservative.*

D.4.3.1 Generalized potential functions

The above discussion can be generalized to include velocity-dependent potentials, called *generalized potential functions.* For this purpose we add to the kinetic energy a (smooth) function of $\mathbf{r}$, $\dot{\mathbf{r}}$ and t,

$$\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) = \tfrac{1}{2}mv^2 - U(\mathbf{r}, \dot{\mathbf{r}}, t). \tag{D.79}$$

This function is a Lagrangian if it satisfies the Lagrange equations. Substituting $\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t)$ into (D.76), we find that the condition (D.77) is replaced by

$$-\frac{\partial U}{\partial \mathbf{r}} + \frac{d}{dt}\frac{\partial U}{\partial \mathbf{v}} = \frac{d}{dt}\frac{\partial T}{\partial \mathbf{v}}. \tag{D.80}$$

In other words, the Lagrange equations are satisfied for *generalized forces* of the type

$$\mathbf{Q} = -\frac{\partial U}{\partial \mathbf{r}} + \frac{d}{dt}\frac{\partial U}{\partial \mathbf{v}}. \tag{D.81}$$

Note that (D.81) reduces to (D.78) for velocity-independent potentials; hence, the class of generalized forces includes gradient forces as a subclass. Although any generalized potential function gives rise to a generalized force, the inverse is *not* true. For instance, friction forces, $\mathbf{F} = -\alpha\mathbf{v}$, cannot be expressed in the form (D.81). On the other hand, the Lorentz force is an example of a generalized force because we can define a generalized potential function for which (D.81) is satisfied (see Problem D.2).

Problem D.2. Show, using the Einstein notation, that the Lorentz force, $F_i = q[E_i + (\mathbf{v}\times\mathbf{B})_i]$, is the generalized force resulting from the generalized potential function $U = q(\varphi - v_j \cdot A_j)$, where $\varphi(\mathbf{r},t)$ is the scalar and $\mathbf{A}(\mathbf{r},t)$ the vector potential. Further, $E_i = -\partial_i\varphi - \partial_t A_i$ and $B_i = \varepsilon_{ijk}\partial_j A_k$ are the cartesian components of the (generally time dependent) **E** and **B** fields, respectively. Use the short-hand notation $\partial_i \equiv \partial/\partial r_i$ and $\partial_t \equiv \partial/\partial t$.

Solution. The first term of (D.81) yields

$$-\partial_i U = q\left[-\partial_i\varphi - v_j\partial_i A_j(\mathbf{r},t)\right] \tag{a}$$

and the second term can be written as

$$\frac{d}{dt}\frac{\partial U}{\partial v_i} = q\frac{d}{dt}\left(\frac{\partial\varphi(\mathbf{r},t)}{\partial v_i} - \frac{\partial v_j A_j(\mathbf{r},t)}{\partial v_i}\right) = -q\frac{d}{dt}A_i(\mathbf{r},t) = q\left[-\partial_j A_i(\mathbf{r},t)\,\dot{r}_j - \partial_t A_i(\mathbf{r},t)\right]. \tag{b}$$

Combining (a) and (b) we obtain for the generalized force

$$Q_i = q\left[-\partial_i\varphi - \partial_t A_i + v_j\partial_i A_j - v_j\partial_j A_i\right]. \tag{c}$$

In the first two terms within the brackets we recognize the E_i-component of the **E** field. Likewise, recalling $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$, we recognize in the last two terms the component $(\mathbf{v}\times\mathbf{B})_i$ of the $\mathbf{v}\times\mathbf{B}$ field[1]

$$(\mathbf{v}\times\mathbf{B})_i = \varepsilon_{ijk}v_j(\boldsymbol{\nabla}\times\mathbf{A})_k = \varepsilon_{ijk}\varepsilon_{klm}v_j\partial_l A_m = v_j\partial_i A_j - v_j\partial_j A_i.$$

Hence, Eq. (c) can be rewritten as $Q_i = q\left[E_i + (\mathbf{v}\times\mathbf{B})_i\right]$, which is indeed the expression for the Lorentz force. □

D.5 Many-particle systems

D.5.1 Lagrangian

At this point we return to many-particle systems. In view of property (D.66) we know that the Lagrangian of a system of N *non*-interacting particles is given by

$$\mathcal{L} = \tfrac{1}{2}\sum_\alpha m_\alpha v_\alpha^2, \tag{D.82}$$

where $\alpha \in \{1,\cdots N\}$. In search for the Lagrangian in the presence of interactions between the particles as well as interactions with the environment we use a procedure similar to the one introduced for the single particle: we add to the Lagrangian (D.82) a generalized potential function

$$\mathcal{L}(\cdots,\mathbf{r}_\alpha,\dot{\mathbf{r}}_\alpha,\cdots,t) = \tfrac{1}{2}\sum_\alpha m_\alpha\dot{\mathbf{r}}_\alpha^2 - U(\cdots,\mathbf{r}_\alpha,\dot{\mathbf{r}}_\alpha,\cdots,t). \tag{D.83}$$

The potential function depends in principle on the position and velocities of *all* the particles as well as on time

$$U(\cdots,\mathbf{r}_\alpha,\dot{\mathbf{r}}_\alpha,\cdots,t) = U^{int}(\cdots,\mathbf{r}_\alpha,\dot{\mathbf{r}}_\alpha,\cdots) + U^{ext}(\cdots,\mathbf{r}_\alpha,\dot{\mathbf{r}}_\alpha,\cdots,t). \tag{D.84}$$

The function U^{int} is called the *internal* generalized potential function and represents the interaction between all particles of the system; *e.g.*, the Van der Waals interaction between the atoms in a classical gas. The function U^{ext} is called the *external* generalized potential function and represents externally applied fields; *e.g.*, the electromagnetic potential of a light field.

The function (D.83) is a Lagrangian if it satisfies the Lagrange equations,

$$\frac{\partial\mathcal{L}}{\partial\mathbf{r}_\alpha} = \frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\mathbf{v}_\alpha}, \quad \text{with } \alpha \in \{1,\cdots N\}. \tag{D.85}$$

---

[1] In the Einstein notation the contraction of the Levi-Civita tensor is given by $\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$.

Since the first term on the r.h.s. of (D.83) is independent of $\mathbf{r}_\alpha$ we find that the condition (D.85) can be written in the form

$$-\frac{\partial U}{\partial \mathbf{r}_\alpha} + \frac{d}{dt}\frac{\partial U}{\partial \mathbf{v}_\alpha} = \frac{d}{dt}\frac{\partial T}{\partial \mathbf{v}_\alpha} = m_\alpha \ddot{\mathbf{r}}_\alpha, \quad \text{with } \alpha \in \{1, \cdots N\}. \tag{D.86}$$

Hence, the quantity

$$\mathbf{Q}_\alpha = -\frac{\partial U}{\partial \mathbf{r}_\alpha} + \frac{d}{dt}\frac{\partial U}{\partial \mathbf{v}_\alpha}, \quad \text{with } \alpha \in \{1, \cdots N\}, \tag{D.87}$$

can be interpreted as the generalized force acting on particle α as the result of its interaction with all other particles in the system.

Generalization to the presence of scleronomous constraints

It is straightforward to generalize the formalism to holonomic systems with *scleronomous* constraints. In this case the positions $\mathbf{r}_\alpha$ and velocities $\dot{\mathbf{r}}_\alpha$ of the particles can be expressed in terms of $f = 3N - l$ generalized coordinates $q_1, \cdots, q_f$ and velocities $\dot{q}_1, \cdots, \dot{q}_f$,

$$\mathbf{r}_\alpha = \mathbf{r}_\alpha(q_1, \cdots, q_f) \tag{D.88a}$$

$$\dot{\mathbf{r}}_\alpha = \sum_{i=1}^{f} \frac{\partial \mathbf{r}_\alpha}{\partial q_i}\dot{q}_i. \tag{D.88b}$$

Substituting these expressions into (D.83) we obtain

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}) = \tfrac{1}{2}\sum_{i,j=1}^{f} a_{ij}(\mathbf{q})\dot{q}_i\dot{q}_j - U(\mathbf{q}, \dot{\mathbf{q}}). \tag{D.89}$$

Here we use the vector notation (D.18); the coefficient $a_{ij}(\mathbf{q})$ is given by

$$a_{ij}(\mathbf{q}) = \sum_\alpha m_\alpha \frac{\partial \mathbf{r}_\alpha}{\partial q_i}\frac{\partial \mathbf{r}_\alpha}{\partial q_j}. \tag{D.90}$$

In terms of the generalized coordinates and velocities the Lagrange equation take the form

$$\frac{\partial \mathcal{L}}{\partial q_i} = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_i}, \quad \text{with } i \in \{1, \cdots f\}, \tag{D.91}$$

and the components of the generalized force become

$$Q_i = -\frac{\partial U}{\partial q_i} + \frac{d}{dt}\frac{\partial U}{\partial \dot{q}_i}, \quad \text{with } i \in \{1, \cdots f\}. \tag{D.92}$$

Note that in the absence of constraints we have $f = 3N$ and the generalized coordinates turn into the cartesian coordinates of the particles, $(q_1, \cdots, q_f) \to (x_1, y_1, z_1, \cdots, x_N, y_N, z_N)$, and we regain (D.85) and (D.87).

Open and closed systems

A mechanical system is called *closed* if it does not interact with its environment; i.e., the potential function can be written as $U = U^{int}(\cdots, \mathbf{r}_\alpha, \dot{\mathbf{r}}_\alpha, \cdots)$. Hence, for closed systems both the potential function and the Lagrangian are time independent. An example of a closed mechanical system is an atom in field-free space. If the system is not closed it is called open. An example of an *open* system is a cloud of interacting particles confined by an external potential.

D.5.2 Energy conservation

In the absence of *time-dependent* external forces the evolution of a mechanical system only depends on the state of that system and not on the instant that this state is created in time. In view of this *homogeneity of time* the Lagrangian of such a system cannot depend explicitly on time; i.e., the total time derivative of (D.83) can be written in the form

$$\frac{d\mathcal{L}}{dt} = \sum_\alpha \frac{\partial \mathcal{L}}{\partial \mathbf{r}_\alpha} \cdot \dot{\mathbf{r}}_\alpha + \sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_\alpha} \cdot \ddot{\mathbf{r}}_\alpha. \tag{D.93}$$

With the aid of the Lagrange equation (D.85) this becomes

$$\frac{d\mathcal{L}}{dt} = \sum_\alpha \dot{\mathbf{r}}_\alpha \cdot \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_\alpha} + \sum_\alpha \ddot{\mathbf{r}}_\alpha \cdot \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_\alpha} = \sum_\alpha \frac{d}{dt}\left(\dot{\mathbf{r}}_\alpha \cdot \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_\alpha} \right). \tag{D.94}$$

Rewriting this expression in the form

$$\frac{d}{dt}\left(\sum_\alpha \dot{\mathbf{r}}_\alpha \cdot \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_\alpha} - \mathcal{L} \right) = 0 \tag{D.95}$$

we find that the quantity

$$E = \sum_\alpha \dot{\mathbf{r}}_\alpha \cdot \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_\alpha} - \mathcal{L} \tag{D.96}$$

is a *constant of the motion*; i.e., it is *conserved* along the path from t_0 to t_1. This quantity is called the *energy* of the system. Mechanical systems for which the energy is conserved are called *conservative* systems. It is straightforward to generalize (D.96) to holonomic systems with *scleronomous* constraints,

$$E(\mathbf{q}, \dot{\mathbf{q}}) = \sum_i \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{r}_i} - \mathcal{L}. \tag{D.97}$$

D.5.2.1 Potential energy

The energy of a *conservative* system can always be written as the sum of a *kinetic energy* and a *potential energy* contribution. This can be seen as follows. The kinetic energy is uniquely determined by the velocities of the particles,

$$T(\dot{\mathbf{r}}_1, \cdots, \dot{\mathbf{r}}_N) = \tfrac{1}{2} \sum_\alpha m_\alpha v_\alpha^2.$$

In general this quantity will change in time under the influence of inter-particle forces but (since the system is conservative) the (total) energy has to remain constant, $E(\mathbf{r}_1, \cdots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \cdots, \dot{\mathbf{r}}_N) = E$. Since the same kinetic energy can be obtained for different distributions of velocity over the particles the difference of E and T has to be a function of the positions *only.* This difference is called the *potential energy* $V(\mathbf{r}_1, \cdots, \mathbf{r}_N)$ of the system,

$$V(\mathbf{r}_1, \cdots, \mathbf{r}_N) = E(\mathbf{r}_1, \cdots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \cdots, \dot{\mathbf{r}}_N) - T(\dot{\mathbf{r}}_1, \cdots, \dot{\mathbf{r}}_N). \tag{D.98}$$

Beware of the difference between the potential energy $V(\mathbf{r}_1, \cdots, \mathbf{r}_N)$ and the generalized potential function $U(\mathbf{r}_1, \cdots, \mathbf{r}_N, \dot{\mathbf{r}}_1, \cdots, \dot{\mathbf{r}}_N)$; the former determines the kinetic energy of a conservative system (also in the presence of generalized forces between the particles); the latter determines the force on all particles, both for conservative and non-conservative systems.

An *open* mechanical system is conservative if the external potential is constant in time.

D.5.3 Momentum conservation in closed systems

The evolution of a *closed* mechanical system only depends on the internal state of that system and not on the absolute position where this state is created in space. In view of this *homogeneity of space* the Lagrangian of a closed system must be independent of the absolute position of that system; i.e., (D.83) must be invariant under a translation of the system over a distance $\mathbf{R}$,

$$\mathbf{r}'_\alpha = \mathbf{r}_\alpha + \mathbf{R}. \tag{D.99}$$

Thus we are led to compare two Lagrangians, defined with respect to two inertial frames shifted with respect to each other over an infinitesimally small distance $\delta\mathbf{r}_\alpha = \mathbf{r}'_\alpha - \mathbf{r}_\alpha = \delta\mathbf{R}$. For this displacement the change change in Lagrangian is given by

$$\delta\mathcal{L} = \sum_\alpha \frac{\partial\mathcal{L}}{\partial\mathbf{r}_\alpha} \cdot \delta\mathbf{R} + \cdots . \tag{D.100}$$

Hence, a necessary condition that the Lagrangian be invariant under an *arbitrary* small shift $\delta\mathbf{R}$ is

$$\sum_\alpha \frac{\partial\mathcal{L}}{\partial\mathbf{r}_\alpha} = 0. \tag{D.101}$$

Note that also the sum over all higher order terms has to vanish but this has no relevance in the present context. Using the Lagrange equation (D.85) the condition (D.101) becomes

$$\sum_\alpha \frac{\partial\mathcal{L}}{\partial\mathbf{r}_\alpha} = \frac{d}{dt}\sum_\alpha \frac{\partial\mathcal{L}}{\partial\mathbf{v}_\alpha} = 0. \tag{D.102}$$

Apparently, in a closed system the quantity

$$\mathbf{P} = \sum_\alpha \mathbf{p}_\alpha \tag{D.103}$$

with

$$\mathbf{p}_\alpha = \frac{\partial\mathcal{L}}{\partial\dot{\mathbf{r}}_\alpha} \tag{D.104}$$

is a *constant of the motion*. The quantity $\mathbf{P}$ is called the *canonical momentum* of the system and $\mathbf{p}_\alpha$ is the canonical momentum of particle α. Hence, for closed systems the canonical momenta of the individual particles always add up to the total momentum of the system, irrespective of the absence or presence of generalized forces between the particles. With (D.85) it immediately follows that

$$\dot{\mathbf{p}}_\alpha = \frac{\partial\mathcal{L}}{\partial\mathbf{r}_\alpha}. \tag{D.105}$$

Example: velocity-independent potential functions

In the special case of systems with only *velocity-independent* interactions between the particles (e.g., gravitational forces) the canonical momentum only depends on the kinetic energy and we find the well-known expression

$$\mathbf{p}_\alpha = m_\alpha\mathbf{v}_\alpha. \tag{D.106}$$

The Hamiltonian is given by

$$E = \sum_\alpha m_\alpha\mathbf{v}_\alpha^2 - \mathcal{L}, \tag{D.107}$$

which becomes after substitution of the Lagrangian (D.83)

$$E = T + V. \tag{D.108}$$

D.5.3.1 Generalization to holonomic systems with scleronomous constraints - symmetry

It is straightforward to generalize (D.104) and (D.105) to *holonomic* systems with *scleronomous* constraints,

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \tag{D.109a}$$

$$\dot{p}_i = \frac{\partial \mathcal{L}}{\partial q_i} \tag{D.109b}$$

with $i \in \{1, \cdots f\}$.

D.5.4 Conservation of angular momentum in closed systems

In view of the *isotropy of space* the Lagrangian of a *closed* system must be independent of the absolute orientation of that system; i.e., (D.83) must be invariant under rotation of the system over an angle $\boldsymbol{\Omega}$ about an axis in direction $\hat{\boldsymbol{\Omega}}$ through the origin of the inertial frame of observation,

$$\mathbf{r}'_\alpha = \mathbf{r}_\alpha + \boldsymbol{\Omega} \times \mathbf{r}_\alpha. \tag{D.110}$$

Under such a rotation also the velocity vector changes direction

$$\mathbf{v}'_\alpha = \mathbf{v}_\alpha + \boldsymbol{\Omega} \times \mathbf{v}_\alpha. \tag{D.111}$$

Thus we are led to compare two Lagrangians, defined with respect to two inertial frames rotated with respect to each other over an infinitesimally small angle $\delta\boldsymbol{\Omega}$ so that $\delta\mathbf{r}_\alpha = \mathbf{r}'_\alpha - \mathbf{r}_\alpha = \delta\boldsymbol{\Omega} \times \mathbf{r}_\alpha$ and $\delta\mathbf{v}_\alpha = \mathbf{v}'_\alpha - \mathbf{v}_\alpha = \delta\boldsymbol{\Omega} \times \mathbf{v}_\alpha$. For this rotation we have

$$\delta\mathcal{L} = \sum_\alpha \frac{\partial \mathcal{L}}{\partial \mathbf{r}_\alpha} \cdot \delta\boldsymbol{\Omega} \times \mathbf{r}_\alpha + \sum_\alpha \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_\alpha} \cdot \delta\boldsymbol{\Omega} \times \dot{\mathbf{r}}_\alpha + \cdots . \tag{D.112}$$

Hence, a necessary condition that the Lagrangian be invariant under an *arbitrary* rotation (about an axis through the origin) is

$$\sum_\alpha \left(\mathbf{r}_\alpha \times \frac{\partial \mathcal{L}}{\partial \mathbf{r}_\alpha} + \dot{\mathbf{r}}_\alpha \times \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_\alpha} \right) = 0. \tag{D.113}$$

Here we used the permutation rules for the vector products (see Appendix N). Turning to canonical variables the condition becomes

$$\sum_\alpha (\mathbf{r}_\alpha \times \dot{\mathbf{p}}_\alpha + \dot{\mathbf{r}}_\alpha \times \mathbf{p}_\alpha) = \frac{d}{dt} \sum_\alpha (\mathbf{r}_\alpha \times \mathbf{p}_\alpha) = 0. \tag{D.114}$$

Apparently, in a closed system the quantity

$$\mathbf{L} = \sum_\alpha (\mathbf{r}_\alpha \times \mathbf{p}_\alpha), \tag{D.115}$$

with

$$\mathbf{L}_\alpha \equiv \mathbf{r}_\alpha \times \mathbf{p}_\alpha, \tag{D.116}$$

is a constant of the motion. The quantity $\mathbf{L}$ is called the *angular momentum* of the system and $\mathbf{L}_\alpha$ is the angular momentum of particle α, both with respect to the origin of the inertial system. Although the value of $\mathbf{L}$ depends on the choice of origin, the conservation holds irrespective of this choice.

D.6 The Hamilton formalism

In the Lagrange formalism the dynamical evolution of mechanical systems is expressed in terms of the dynamical variables $\mathbf{q} \equiv (q_1, \cdots, q_f)$ and $\dot{\mathbf{q}} \equiv (\dot{q}_1, \cdots, \dot{q}_f)$ with the aid of the Lagrangian $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$ and the Lagrange equations. With this approach we identified the sum over the canonical momenta $\mathbf{p} \equiv (p_1, \cdots, p_f)$ as a conserved quantity of a closed mechanical system. In the Hamilton formalism the dynamical evolution is described in terms of the dynamical variables $\mathbf{q}$ and $\mathbf{p}$; i.e., the dependence on $\dot{\mathbf{q}}$ is eliminated from the formalism and replaced by a dependence on $\mathbf{p}$.

D.6.1 Legendre transformation of Lagrangian - Hamiltonian

To replace the dependence on $\dot{\mathbf{q}}$ by a dependence on $\mathbf{p}$ we use a Legendre transformation of the function $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$ into the function

$$H(\mathbf{q}, \mathbf{p}, t) = \mathbf{p} \cdot \dot{\mathbf{q}} - \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t). \tag{D.117}$$

The function $H(\mathbf{q}, \mathbf{p}, t)$ is called the *function of Hamilton* or *Hamiltonian* of the system. Note that with the sign convention of (D.117) the Hamiltonian represents the energy of the system

$$E = H(\mathbf{q}, \mathbf{p}, t). \tag{D.118}$$

To convince ourselves that the Hamiltonian has the desired properties we consider the total differential

$$dH = -d\mathcal{L} + d(\sum_i p_i \dot{q}_i). \tag{D.119}$$

For the total differential of the Lagrangian $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t)$ we have with the aid of (D.109a) and (D.109b)

$$d\mathcal{L} = \sum_i \frac{\partial \mathcal{L}}{\partial q_i} dq_i + \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} d\dot{q}_i + \frac{\partial \mathcal{L}}{\partial t} dt = \sum_i \dot{p}_i dq_i + \sum_i p_i d\dot{q}_i + \frac{\partial \mathcal{L}}{\partial t} dt. \tag{D.120}$$

The total differential of the transformation term is

$$d(\sum_i p_i \dot{q}_i) = \sum_i \dot{q}_i dp_i + \sum_i p_i d\dot{q}_i. \tag{D.121}$$

Substituting (D.120) and (D.121) into (D.119) we obtain

$$dH = -\sum_i \dot{p}_i dq_i + \sum_i \dot{q}_i dp_i + \frac{\partial \mathcal{L}}{\partial t} dt. \tag{D.122}$$

Thus we have verified that H is indeed a function of $\mathbf{q}$ and $\mathbf{p}$. From the total differential we obtain

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{D.123a}$$

$$\dot{q}_i = +\frac{\partial H}{\partial p_i} \tag{D.123b}$$

with $i \in \{1, \cdots f\}$. These equations are called the *Hamilton equations* of motion. In view of their symmetrical form they are called the *canonical equations* of motion.

D.6.1.1 Energy conservation

From (D.122) we further infer

$$\frac{\partial H}{\partial t} = \frac{\partial \mathcal{L}}{\partial t}. \tag{D.124}$$

Together with (D.118) this shows that the energy is conserved if the Lagrangian (hence, also the Hamiltonian) do not depend explicitly on time; i.e., for $\partial H / \partial t = 0$.

D.6.1.2 Example: charged particle in an electromagnetic field

As an example of a system with a *time-independent* generalized potential function $U(\mathbf{r},\dot{\mathbf{r}})$ we consider the motion of a charged particle in a static electromagnetic field. As we showed in Example D.3.2.3 the force on a charged particle can be written as the gradient of a generalized potential function

$$U(\mathbf{r},\dot{\mathbf{r}}) = q[\varphi(\mathbf{r}) - \mathbf{v}\cdot\mathbf{A}(\mathbf{r})] \tag{D.125}$$

and the Lagrangian is of the form (D.33). As the kinetic energy is given by

$$T(\dot{\mathbf{r}}) = \tfrac{1}{2}m\mathbf{v}^2, \tag{D.126}$$

the Lagrangian becomes

$$\mathcal{L}(\mathbf{r},\dot{\mathbf{r}}) \equiv \tfrac{1}{2}m\mathbf{v}^2 - q[\varphi - \mathbf{v}\cdot\mathbf{A}]. \tag{D.127}$$

From this we calculate the canonical momentum

$$\mathbf{p} \equiv \frac{\partial\mathcal{L}}{\partial\mathbf{v}} = m\mathbf{v} + q\mathbf{A}. \tag{D.128}$$

Thus we found that the canonical momentum consists in this case of two contributions; the first term is called the *kinetic momentum* and the second term the *electromagnetic momentum.* The Hamiltonian is given by

$$H = \mathbf{v}\cdot\mathbf{p} - \mathcal{L} = m\mathbf{v}^2 + q\mathbf{A}\cdot\mathbf{v} - \tfrac{1}{2}m\mathbf{v}^2 + q(\varphi - \mathbf{v}\cdot\mathbf{A}) = \tfrac{1}{2}m\mathbf{v}^2 + q\varphi. \tag{D.129}$$

The potential energy is given by

$$V(\mathbf{r}) = H - T = q\varphi(\mathbf{r}). \tag{D.130}$$

Expressing (D.129) in terms of the canonical variables $\mathbf{r}$ and $\mathbf{p}$ we obtain with the aid of (D.128) the Hamiltonian of the system,

$$H = \frac{1}{2m}(\mathbf{p} - q\mathbf{A})^2 + q\varphi. \tag{D.131}$$

D.7 Center of mass and relative coordinates

D.7.1 Center of mass motion of a closed system

In Section D.5.3 we found that the momentum of a *closed* mechanical system with *velocity-independent* interparticle forces,

$$\mathbf{P} = \sum_\alpha m_\alpha\dot{\mathbf{r}}_\alpha, \tag{D.132}$$

is conserved in time. However, this quantity is not conserved when changing from one inertial frame to another. Obviously, the velocity of the particles depends on the inertial system in which the velocity is measured. Hence, also the kinetic momentum depends on the inertial system. To analyze this dependence we change from the inertial frame $\mathcal{S}$ to a new inertial frame $\mathcal{S}'$ in which the origin of old frame is moving at velocity $\mathbf{V}$. This change of inertial system is described by a *galilean transformation*,

$$\mathbf{r}'_\alpha = \mathbf{r}_\alpha - \mathbf{V}t \tag{D.133a}$$

$$\dot{\mathbf{r}}'_\alpha = \dot{\mathbf{r}}_\alpha - \mathbf{V}. \tag{D.133b}$$

The observed difference in momentum between both inertial systems is given by

$$\mathbf{P}' = \mathbf{P} - \sum_\alpha m_\alpha\mathbf{V}. \tag{D.134}$$

This expression reveals the existence of an inertial frame in which the momentum is zero, $\mathbf{P}' = 0$. Thus we find for the momentum in an inertial system moving with velocity $\mathbf{V}$ with respect to the zero-momentum frame,

$$\mathbf{P} = \sum_\alpha m_\alpha \mathbf{V}. \tag{D.135}$$

In other words, the momentum of the entire system behaves like that of a single particle with mass

$$M = \sum_\alpha m_\alpha. \tag{D.136}$$

This quantity is called the *total mass* of the system. Introducing the position vector

$$\mathbf{R} = \sum_\alpha m_\alpha \mathbf{r}_\alpha / \sum_\alpha m_\alpha \tag{D.137}$$

we find the relation

$$\mathbf{P} = M\dot{\mathbf{R}} = M\mathbf{V}. \tag{D.138}$$

The vector $\mathbf{R}$ is called the *center of mass* of the system. Hence, in the zero-momentum frame the center of mass is at rest.

It is also instructive to compare the total angular momentum in the frame $\mathcal{S}'$ with that in the frame $\mathcal{S}$. Using Eq. (D.133a) the total angular momentum in the frame $\mathcal{S}'$ can be expressed as

$$\mathbf{L}' = \sum_\alpha (\mathbf{r}_\alpha \times \mathbf{p}'_\alpha) - \mathbf{V}t \times \sum_\alpha \mathbf{p}'_\alpha, \tag{D.139}$$

where $\mathbf{V}t$ is the position of the origin of $\mathcal{S}$ as observed in $\mathcal{S}'$. The second term vanishes if we choose for $\mathcal{S}'$ the zero-momentum frame. To proceed we multiply both sides of Eq. (D.133b) by the particles mass,

$$\mathbf{p}'_\alpha = \mathbf{p}_\alpha - m_\alpha \mathbf{V}. \tag{D.140}$$

Substituting this expression into Eq. (D.139) we find

$$\mathbf{L}' = \mathbf{L} - \sum_\alpha (m_\alpha \mathbf{r}_\alpha \times \mathbf{V}). \tag{D.141}$$

In terms of the position and momentum of the center of mass this becomes

$$\mathbf{L} = \mathbf{L}' + \mathbf{R} \times \mathbf{P}. \tag{D.142}$$

Thus we recognize two contributions to the angular momentum of the system: $\mathbf{L}'$ is called the *intrinsic* angular momentum, which is the angular momentum with the center of mass at rest, and $\mathbf{R} \times \mathbf{P}$ is the angular momentum resulting from the center of mass motion in the inertial frame of reference.

D.7.2 Relative motion in a closed system of two atoms

To deal with interatomic interactions and collisions between particles one introduces *relative coordinates*. The position of particle 1 relative to particle 2 is given by

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2. \tag{D.143}$$

Taking the derivative with respect to time we find for the relative velocity of particle 1 with respect to particle 2

$$\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2. \tag{D.144}$$

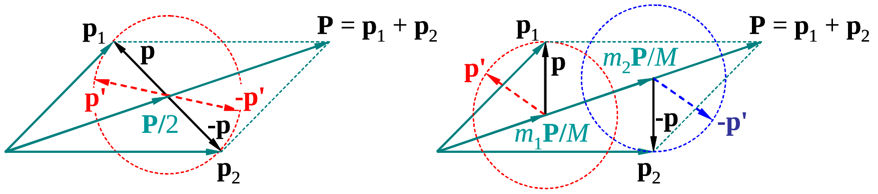


Figure D.1: Center of mass and relative momenta for two colliding atoms: *left*: equal mass; *right*: unequal mass with $m_1/m_2 = 1/3$.

Let us analyze in some detail the case of a system of two particles. The total momentum of the pair (the center of mass momentum) is a conserved quantity and given by

$$\mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2 = m_1\mathbf{v}_1 + m_2\mathbf{v}_2 = m_1\dot{\mathbf{r}}_1 + m_2\dot{\mathbf{r}}_2. \tag{D.145}$$

The total mass is given by the sum of the two particles masses, $M = m_1 + m_2$. With the relation $\mathbf{P} = M\mathbf{V}$, where $\mathbf{V} = \dot{\mathbf{R}}$ is the center of mass velocity, we find for the position of the center of mass

$$\mathbf{R} = (m_1\mathbf{r}_1 + m_2\mathbf{r}_2)/(m_1 + m_2). \tag{D.146}$$

Adding and subtracting (D.145) and (D.144) allows us to express $\mathbf{v}_1$ and $\mathbf{v}_2$ in terms of $\mathbf{P}$ and $\mathbf{v}$,

$$\mathbf{P} + m_2\mathbf{v} = (m_1 + m_2)\,\mathbf{v}_1 \tag{D.147a}$$

$$\mathbf{P} - m_1\mathbf{v} = (m_1 + m_2)\,\mathbf{v}_2. \tag{D.147b}$$

With these expressions the total kinetic energy of the pair, $E = \epsilon_1 + \epsilon_2$, can be split in a contribution of the *center of mass* and a contribution of the relative motion

$$E = \frac{1}{2}m_1\mathbf{v}_1^2 + \frac{1}{2}m_2\mathbf{v}_2^2 = \frac{1}{2}m_1\frac{(\mathbf{P} + m_2\mathbf{v})^2}{(m_1 + m_2)^2} + \frac{1}{2}m_2\frac{(\mathbf{P} - m_1\mathbf{v})^2}{(m_1 + m_2)^2} = \frac{\mathbf{P}^2}{2M} + \frac{\mathbf{p}^2}{2m_r}, \tag{D.148}$$

where

$$\mathbf{p} = m_r\mathbf{v} = m_r\dot{\mathbf{r}} = \frac{m_r}{m_1}\mathbf{p}_1 - \frac{m_r}{m_2}\mathbf{p}_2 \tag{D.149}$$

is the relative momentum with

$$m_r = m_1m_2/(m_1 + m_2) \tag{D.150}$$

representing the *reduced mass* of the pair. Because both $\mathbf{P}$ and E are conserved quantities in elastic collisions also $\mathbf{p}^2$ must be conserved. This implies that in elastic collisions the relative momentum $p = |\mathbf{p}|$ is conserved, but *not* its direction. Adding and subtracting (D.143) and (D.146) we can express $\mathbf{r}_1$ and $\mathbf{r}_2$ in terms of $\mathbf{R}$ and $\mathbf{r}$,

$$\mathbf{r}_1 = \mathbf{R} + \frac{m_2}{M}\mathbf{r} \text{ and } \mathbf{r}_2 = \mathbf{R} - \frac{m_1}{M}\mathbf{r}. \tag{D.151}$$

Likewise combining (D.145) and (D.149) we can express $\mathbf{p}_1$ and $\mathbf{p}_2$ in terms of $\mathbf{P}$ and $\mathbf{p}$,

$$\mathbf{p}_1 = \frac{m_1}{M}\mathbf{P} + \mathbf{p} \text{ and } \mathbf{p}_2 = \frac{m_2}{M}\mathbf{P} - \mathbf{p}. \tag{D.152}$$

The vector diagram is shown in Fig. D.1.

Problem D.3. Show that the Jacobian of the transformation $d\mathbf{r}_1 d\mathbf{r}_2 = \left|\frac{\partial(\mathbf{r}_1,\mathbf{r}_2)}{\partial(\mathbf{R},\mathbf{r})}\right| d\mathbf{R}d\mathbf{r}$ is -1.

Solution. Because the x, y and z directions separate we can write the Jacobian as the product of three 1D Jacobians.

$$\left|\frac{\partial(\mathbf{r}_1,\mathbf{r}_2)}{\partial(\mathbf{R},\mathbf{r})}\right| = \prod_{i=x,y,z}\left|\frac{\partial(r_{1i},r_{2i})}{\partial(R_i,r_i)}\right| = \prod_{i=x,y,z}\begin{vmatrix} 1 & m_1/M \\ 1 & -m_2/M \end{vmatrix} = -1. \qquad \square$$

Problem D.4. Show that the Jacobian of the transformation $d\mathbf{p}_1 d\mathbf{p}_2 = \left|\frac{\partial(\mathbf{p}_1,\mathbf{p}_2)}{\partial(\mathbf{P},\mathbf{p})}\right| d\mathbf{P}d\mathbf{p}$ is -1.

Solution. Because the x, y and z directions separate we can write the Jacobian as the product of three 1D Jacobians.

$$\left|\frac{\partial(\mathbf{p}_1,\mathbf{p}_2)}{\partial(\mathbf{P},\mathbf{p})}\right| = \prod_{i=x,y,z}\left|\frac{\partial(p_{1i},p_{2i})}{\partial(P_i,p_i)}\right| = \prod_{i=x,y,z}\begin{vmatrix} m_1/M & 1 \\ m_2/M & -1 \end{vmatrix} = -1. \qquad \square$$

D.7.3 Kinematics of scattering

In any collision the momentum $\mathbf{P}$ is conserved. Thus, also the center of mass energy $\mathbf{P}^2/2M$ is conserved and since also the total energy must be conserved also the relative kinetic energy $\mathbf{p}^2/2m_r$ is conserved in *elastic* collisions, be it in general not *during* the collision. In this section we consider the consequence of the conservation laws for the momentum transfer between particles in elastic collisions in which the relative momentum changes from $\mathbf{p}$ to $\mathbf{p}'$, with $\mathbf{q} = \mathbf{p}' - \mathbf{p}$. Because the relative energy is conserved, also the modulus of the relative momentum will be conserved, $|\mathbf{p}| = |\mathbf{p}'|$, and the only effect of the collision is to change the direction of the relative momentum over an angle ϑ. Hence, the *scattering angle* ϑ fully determines the energy and momentum transfer in the collision. Using (D.147) the momenta of the particles before and after the collision (see Fig. D.1) are given by

$$\mathbf{p}_1 = m_1\mathbf{P}/M + \mathbf{p} \longrightarrow \mathbf{p}_1' = m_1\mathbf{P}/M + \mathbf{p}' \tag{D.153a}$$

$$\mathbf{p}_2 = m_2\mathbf{P}/M - \mathbf{p} \longrightarrow \mathbf{p}_2' = m_2\mathbf{P}/M - \mathbf{p}'. \tag{D.153b}$$

Hence, the momentum transfer is

$$\Delta\mathbf{p}_1 = \mathbf{p}_1' - \mathbf{p}_1 = \mathbf{p}' - \mathbf{p} = \mathbf{q} \tag{D.154a}$$

$$\Delta\mathbf{p}_2 = \mathbf{p}_2' - \mathbf{p}_2 = \mathbf{p} - \mathbf{p}' = -\mathbf{q}. \tag{D.154b}$$

The energy transfer is

$$\Delta E_1 = \frac{\mathbf{p}_1'^2}{2m_1} - \frac{\mathbf{p}_1^2}{2m_1} = \frac{(m_1\mathbf{P}/M + \mathbf{p}')^2}{2m_1} - \frac{(m_1\mathbf{P}/M + \mathbf{p})^2}{2m_1} = \frac{\mathbf{P}\cdot\mathbf{q}}{M} \tag{D.155a}$$

$$\Delta E_2 = \frac{\mathbf{p}_2'^2}{2m_2} - \frac{\mathbf{p}_2^2}{2m_2} = \frac{(m_2\mathbf{P}/M - \mathbf{p}')^2}{2m_2} - \frac{(m_2\mathbf{P}/M - \mathbf{p})^2}{2m_2} = -\frac{\mathbf{P}\cdot\mathbf{q}}{M}. \tag{D.155b}$$

In the special case $\mathbf{p}_1 = 0$ we have

$$\mathbf{P} = \mathbf{p}_2 = \frac{-\mathbf{p}}{1 - m_2/M} = -\frac{M}{m_1}\mathbf{p} \tag{D.156}$$

or

$$\mathbf{p} = -m_r\mathbf{v}_2. \tag{D.157}$$

The momentum transfer becomes

$$q = \sqrt{\mathbf{q}^2} = \sqrt{(\mathbf{p}' - \mathbf{p})^2} = \sqrt{2\mathbf{p}^2 - 2\mathbf{p}'\cdot\mathbf{p}} = p\sqrt{2 - 2\cos\vartheta}. \tag{D.158}$$

For small angles this implies

$$\vartheta = q/p. \tag{D.159}$$

The energy transfer becomes

$$\Delta E_1 = \frac{\mathbf{P}\cdot\mathbf{q}}{M} = -\frac{\mathbf{p}\cdot(\mathbf{p}' - \mathbf{p})}{m_1} = \frac{\mathbf{p}^2}{m_1}(1 - \cos\vartheta) = \frac{m_r^2}{m_1}\mathbf{v}_2^2(1 - \cos\vartheta), \tag{D.160}$$

where ϑ is the scattering angle. This can be written in the form

$$\Delta E_1 = \tfrac{1}{4}\xi m_2 \mathbf{v}_2^2 \left(1 - \cos\vartheta\right), \tag{D.161}$$

where

$$\xi = \frac{4m_r^2}{m_1 m_2} = \frac{4m_1 m_2}{\left(m_1 + m_2\right)^2} = 4\frac{m_r}{M}. \tag{D.162}$$

is the thermalization efficiency parameter. For $m_1 = m_2$ this parameter reaches its maximum value ($\xi = 1$) and we obtain

$$\Delta E_1 = \tfrac{1}{2} E_2 \left(1 - \cos\vartheta\right). \tag{D.163}$$

For $m_1 \ll m_2$ the efficiency parameter is given by $\xi \simeq 4m_1/m_2$.

E

Classical electrodynamics

E.1 Maxwell equations

The central equations of the classical electrodynamics of continuous media are the four equations of Maxwell (the *macroscopic* Maxwell equations). In the presence of a *charge-density distribution* ϱ and a *current-density-distribution* $\mathbf{J}$ these equations are [1]

$$\boldsymbol{\nabla}\cdot\mathbf{D}=\varrho \qquad \text{(Gauss)} \tag{E.1a}$$

$$\boldsymbol{\nabla}\times\mathbf{H}-\frac{\partial\mathbf{D}}{\partial t}=\mathbf{J} \qquad \text{(Ampère/Maxwell)} \tag{E.1b}$$

$$\boldsymbol{\nabla}\cdot\mathbf{B}=0 \tag{E.1c}$$

$$\boldsymbol{\nabla}\times\mathbf{E}+\frac{\partial\mathbf{B}}{\partial t}=\mathbf{0} \qquad \text{(Faraday).} \tag{E.1d}$$

Here $\mathbf{E}$ is called the *electric field strength* and $\mathbf{B}$ the *magnetic induction.* These are the fields that determine the forces acting on charges and currents, respectively. $\mathbf{H}$ is the *magnetic field strength* and $\mathbf{D}$ the *dielectric displacement.*; These fields are introduced to account for the material-specific properties of matter. The relation between the two pairs of fields, $\mathbf{B}$ versus $\mathbf{H}$ and $\mathbf{E}$ versus $\mathbf{D}$, is given by the *materials equations*

$$\mathbf{B}=\mu_0(\mathbf{H}+\mathbf{M}) \qquad \mathbf{E}=\varepsilon_0^{-1}\left(\mathbf{D}-\mathbf{P}\right), \tag{E.2}$$

where $\mathbf{M}$ is called the *magnetization* and $\mathbf{P}$ the *polarization* of the material under consideration; $\mu_0\approx 1.256\times10^{-6}\,\mathrm{N\,A^{-2}}$ is the *magnetic permeability of vacuum* and $\varepsilon_0=1/\mu_0c^2\approx 8.854\times10^{-12}\,\mathrm{F\,m^{-1}}$ the *electric permittivity of vacuum.* Outside materials both $\mathbf{M}$ and $\mathbf{P}$ vanish; i.e., there is no essential difference between $\mathbf{B}$ and $\mathbf{H}$ as well as between $\mathbf{E}$ and $\mathbf{D}$. Inside materials, the material-specific relations between $\mathbf{D}$ and $\mathbf{P}$ determine the primary field $\mathbf{E}$. Likewise, $\mathbf{H}$ and $\mathbf{M}$ determine $\mathbf{B}$.

E.1.1 Linear media

The simplest materials are *linear media.* In such materials, the magnetization $\mathbf{M}$ (the magnetic moment per unit volume) and polarization $\mathbf{P}$ (the electric dipole moment per unit volume) are given by

$$\mathbf{M}=\chi_m\mathbf{H} \qquad \mathbf{P}=\varepsilon_0\chi_e\mathbf{E}, \tag{E.3}$$

where χ_m is the *magnetic susceptibility* and χ_e the *electric susceptibility* of the material. Dividing the susceptibilities by the atomic density N/V one obtains (for dilute systems) the magnetic and

---

[1] See for instance: J.D. Jackson, *Classical electrodynamics*, John Wiley & Sons, 3rd edition, New York 1998.

electric *polarizabilities* α_M and α_E,

$$\chi_m = \mu_0\,(N/V)\,\alpha_M \tag{E.4a}$$

$$\chi_e = (N/V)\,\alpha_E. \tag{E.4b}$$

Substituting these expressions into the materials equations they become

$$\mathbf{B} = \mu_0(1+\chi_m)\mathbf{H} = \mu\mathbf{H} \tag{E.5a}$$

$$\mathbf{D} = \varepsilon_0(1+\chi_e)\mathbf{E} = \varepsilon\mathbf{E}, \tag{E.5b}$$

where $\mu = \mu_0(1+\chi_m) = \mu_0\mu_{rel}$ is the *permeability* and $\varepsilon = \varepsilon_0(1+\chi_e) = \varepsilon_0\varepsilon_{rel}$ the *permittivity* of the material. The quantity $\mu_{rel} = (1+\chi_m)$ is called the *relative* permeability with respect to the vacuum. Likewise, $\varepsilon_{rel} = (1+\chi_e)$ is called the permittivity relative to vacuum.

In linear media the **D** and **H** fields can be eliminated and the Maxwell equations can be written in the form

$$\boldsymbol{\nabla}\cdot\mathbf{E} = \varrho/\varepsilon \quad \text{(Gauss)} \tag{E.6a}$$

$$\boldsymbol{\nabla}\times\mathbf{B} - \frac{\partial\mathbf{E}}{c^2\partial t} = \mu\mathbf{J} \quad \text{(Ampère/Maxwell)} \tag{E.6b}$$

$$\boldsymbol{\nabla}\cdot\mathbf{B} = 0 \tag{E.6c}$$

$$\boldsymbol{\nabla}\times\mathbf{E} + \frac{\partial\mathbf{B}}{\partial t} = 0 \quad \text{(Faraday),} \tag{E.6d}$$

where $c = (\mu\varepsilon)^{-1/2}$ is the speed of light in the medium. Combining equations we obtain, with the aid of Eq. (N.24), for the **E** and **B** fields separately

$$-\boldsymbol{\nabla}^2\mathbf{B} + \frac{\partial^2\mathbf{B}}{c^2\partial t^2} = \mu\boldsymbol{\nabla}\times\mathbf{J} \tag{E.7}$$

$$-\boldsymbol{\nabla}^2\mathbf{E} + \frac{\partial^2\mathbf{E}}{c^2\partial t^2} = -\mu\mathbf{J} - \frac{1}{\varepsilon}\boldsymbol{\nabla}\varrho. \tag{E.8}$$

E.1.2 Maxwell-Lorentz equations

In search of a microscopic foundation of the Maxwell equations, Hendrik Antoon Lorentz postulated a microscopic theory for the electromagnetic interaction between elementary electric charges in vacuum. In this theory, the relation between the microscopic currrent-charge distribution, given by $\varrho(\mathbf{r},t)$ and $\mathbf{J}(\mathbf{r},t)$, and the microscopic electric and magnetic fields, $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{B}(\mathbf{r},t)$, is given by the *Maxwell-Lorentz equations*,

$$\boldsymbol{\nabla}\cdot\mathbf{E} = \varrho/\varepsilon_0 \quad \text{(Gauss)} \tag{E.9a}$$

$$\boldsymbol{\nabla}\times\mathbf{B} - \frac{\partial\mathbf{E}}{c^2\partial t} = \mu_0\mathbf{J} \quad \text{(Ampère/Maxwell)} \tag{E.9b}$$

$$\boldsymbol{\nabla}\cdot\mathbf{B} = 0 \tag{E.9c}$$

$$\boldsymbol{\nabla}\times\mathbf{E} + \frac{\partial\mathbf{B}}{\partial t} = 0 \quad \text{(Faraday).} \tag{E.9d}$$

These equations correspond to Eqs. (E.6 treating the vacuum as a linear medium with $\mathbf{P} = \mathbf{M} = 0$. The goal of Lorentz was to obtain expressions for **P** and **M** starting from the microscopic expressions for condensed matter composed of atoms and ions in vacuum. For this purpose the Maxwell-Lorentz equations have to be completed with the equation of motion of the charges under the influence of **E** and **B**. This is provided by the *Lorentz force* - see Section D.3.2.3. In the context of atomic physics the Maxwell-Lorentz equations serve in particular to describe the response of single atoms on an externally applied electromagnetic radiation field; e.g., a laser.

So let us consider a system of elementary charges labeled with particle index i at positions $\{\mathbf{r}_i\}$ moving at velocities $\{\mathbf{v}_i\}$ in vacuum. The microscopic expressions for the charge density $\varrho(\mathbf{r},t)$ and current density $\mathbf{J}(\mathbf{r},t)$ are given by

$$\varrho(\mathbf{r},t) = \sum_i q_i \delta[\mathbf{r} - \mathbf{r}_i(t)] \tag{E.10a}$$

$$\mathbf{J}(\mathbf{r},t) = \sum_i q_i \mathbf{v}_i \delta[\mathbf{r} - \mathbf{r}_i(t)]. \tag{E.10b}$$

The equations of motion take the form

$$m_i \ddot{\mathbf{r}}_i = q(\mathbf{E}(\mathbf{r}_i,t) + \dot{\mathbf{r}} \times \mathbf{B}(\mathbf{r}_i,t)). \tag{E.11}$$

It is important to realize that the Lorentz theory is not without issues. What is holding together a point-like charge and how to deal with its diverging selfenergy?

E.1.2.1 Charge conservation

Combining the Gauss law with Ampère's law we find the continuity equation

$$\frac{\partial}{\partial t}\varrho(\mathbf{r},t) + \boldsymbol{\nabla}\cdot\mathbf{J}(\mathbf{r},t) = 0, \tag{E.12}$$

which expresses the local conservation of charge. The integral form of this conservation law is obtained by integration over a volume V,[1]

$$-\frac{d}{dt}\int_V d\mathbf{r}\, \varrho(\mathbf{r},t) = \oint_A \mathbf{J}(\mathbf{r},t)\cdot d\mathbf{S}. \tag{E.13}$$

In other words, the rate of change of the charge inside the volume V is equal to charge per unit time crossing the surface of this volume, A. For current-charge distributions strongly localized in space about a nucleus (such as in atoms or ions) the surface integral vanishes far from the nucleus. Hence, the total charge Q is conserved, $dQ/dt = 0$, and given by

$$Q = \int_V d\mathbf{r}\, \varrho(\mathbf{r},t). \tag{E.14}$$

E.1.3 Scalar and vector potentials and gauge freedom

Using the vector identity (N.23) and the Maxwell equation (E.6c) the $\mathbf{B}$ field can be derived from a vector potential $\mathbf{A}$ by requiring

$$\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}. \tag{E.15}$$

Substituting this expression into the Maxwell equation (E.6d) we obtain

$$\boldsymbol{\nabla}\times(\mathbf{E} + \partial\mathbf{A}/\partial t) = 0. \tag{E.16}$$

Thus, in view of the vector identity (N.22) we infer that the $\mathbf{E}$ field can be derived from the equation

$$\mathbf{E} = -\boldsymbol{\nabla}\varphi - \partial\mathbf{A}/\partial t. \tag{E.17}$$

In view of the Helmholtz theorem (N.5) the vector potential $\mathbf{A}$ can be separated in a *solenoidal* or *divergence-free* part, $\mathbf{A}_\perp$, and a *irrotational* or *rotation-free* part, $\mathbf{A}_\parallel$,

$$\mathbf{A} = \mathbf{A}_\perp + \mathbf{A}_\parallel \text{ with } \boldsymbol{\nabla}\cdot\mathbf{A}_\perp = 0 \text{ and } \boldsymbol{\nabla}\times\mathbf{A}_\parallel = 0. \tag{E.18}$$

---

[1] Here we use the Gauss law, $\int_V \mathrm{div}\mathbf{J}\, d\mathbf{r} = \oint_A \mathbf{J}\cdot d\mathbf{S}$.

The potentials $\mathbf{A}$ and φ are not uniquely defined. Eq. (E.15) remains satisfied under the transformation $\mathbf{A}' \to \mathbf{A} = \mathbf{A}' + \boldsymbol{\nabla}\chi$, where $\chi(\mathbf{r}, t)$ is an arbitrary scalar field because $\boldsymbol{\nabla} \times \boldsymbol{\nabla}\chi = 0$ for all scalar fields χ. Furthermore, under the simultaneous transformation $\varphi = \varphi' - \partial\chi/\partial t$ also Eq. (E.17) remains satisfied. This is known as *gauge freedom*. The transformation $(\mathbf{A}', \varphi') \to (\mathbf{A}, \varphi)$ is called a *gauge transformation* and the field $\chi(\mathbf{r}, t)$ the *gauge field*. The invariance of Eqs. (E.15) and (E.17) under gauge transformations is called *gauge invariance*. Also the fields $\mathbf{B}$ and $\mathbf{E}$ are *gauge invariant*. In view of the Helmholtz theorem (E.18), the gauge invariance of Eq. (E.15) always allows us to freely choose $\mathbf{A}_{\|}$.

For the electromagnetic field we mention two important gauges:

- In the *Coulomb gauge* the gauge freedom is used to make the vector potential *divergence-free*,

  $$\boldsymbol{\nabla} \cdot \mathbf{A} = \boldsymbol{\nabla} \cdot (\mathbf{A}' + \boldsymbol{\nabla}\chi) = 0. \tag{E.19}$$

  This condition is satisfied if the gauge field satisfies the relation $\boldsymbol{\nabla}^2\chi = -\boldsymbol{\nabla} \cdot \mathbf{A}'$. This gauge implies $\mathbf{A}_{\|} = 0$ as follows by using the Helmholtz relation $\boldsymbol{\nabla} \cdot \mathbf{A}_{\perp} = 0$. In this sense the vector potential can be called *transverse*. More precisely, by spatial Fourier transformation

  $$\mathbf{A}(\mathbf{r}, t) = \frac{1}{(2\pi)^{3/2}} \int \vec{\mathcal{A}}(\mathbf{k}, t) e^{i\mathbf{k}\cdot\mathbf{r}} d\mathbf{k} \tag{E.20}$$

  the Eq. (E.19) turns into

  $$\mathbf{k} \cdot \vec{\mathcal{A}}(\mathbf{k}, t) = 0. \tag{E.21}$$

  Hence, the Fourier components of the vector potential are transverse with respect to the vector $\mathbf{k}$ in reciprocal space.

- In the *Lorenz gauge* the gauge freedom is used to satisfy the Lorenz condition,

  $$\begin{aligned}\boldsymbol{\nabla} \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial\varphi}{\partial t} &= \boldsymbol{\nabla} \cdot (\mathbf{A}' + \boldsymbol{\nabla}\chi) + \frac{1}{c^2}\frac{\partial}{\partial t}(\varphi' - \partial\chi/\partial t) \\ &= \boldsymbol{\nabla} \cdot \mathbf{A}' + \boldsymbol{\nabla}^2\chi + \frac{1}{c^2}\frac{\partial\varphi'}{\partial t} - \frac{1}{c^2}\frac{\partial^2\chi}{\partial t^2} = 0.\end{aligned}$$

  This condition is satisfied if the gauge field satisfies the relation $\boldsymbol{\nabla}^2\chi = -\boldsymbol{\nabla} \cdot \mathbf{A}' - c^{-2}\partial\varphi/\partial t + c^{-2}\partial^2\chi/\partial^2 t$.

E.2 Transformation formulas for moving objects

An observer moving in an static electromagnetic field experiences a *velocity-induced* electromagnetic field given by

$$\begin{array}{ll} \mathbf{E}'_{\perp} = \gamma\left(\mathbf{E}_{\perp} + \mathbf{v} \times \mathbf{B}_{\perp}\right) & \mathbf{B}'_{\perp} = \gamma\left(\mathbf{B}_{\perp} - \dfrac{\mathbf{v}}{c^2} \times \mathbf{E}_{\perp}\right) \\ \mathbf{E}'_{\|} = \mathbf{E}_{\|} & \mathbf{B}'_{\|} = \mathbf{B}_{\|}, \end{array} \tag{E.22}$$

where $\gamma = (1 - v^2/c^2)^{-1/2}$ is the *Lorentz contraction factor*. For an observer moving at a non-relativistic velocity in a static magnetic field $\mathbf{B}$ this reduces to

$$\begin{array}{ll} \mathbf{E}'_{\perp} \simeq \mathbf{v} \times \mathbf{B}_{\perp} & \mathbf{B}'_{\perp} \simeq \mathbf{B}_{\perp} \\ \mathbf{E}'_{\|} = 0 & \mathbf{B}'_{\|} = \mathbf{B}_{\|}. \end{array} \tag{E.23}$$

Likewise, for an observer moving at a non-relativistic velocity in a static electric field $\mathbf{E}$ we have

$$\begin{array}{ll} \mathbf{B}'_{\perp} \simeq -\dfrac{\mathbf{v}}{c^2} \times \mathbf{E}_{\perp} & \mathbf{E}'_{\perp} \simeq \mathbf{E}_{\perp} \\ \mathbf{B}'_{\|} = 0 & \mathbf{E}'_{\|} = \mathbf{E}_{\|}. \end{array} \tag{E.24}$$

For an observer moving orthogonal to the direction of a static electric or magnetic field this is illustrated in Fig. E.1.

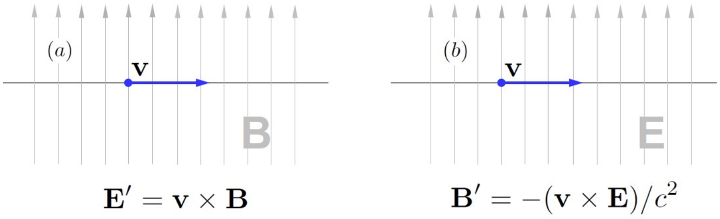


Figure E.1: (*a*) An observer moving perpendicularly to the direction of a static magnetic field experiences a velocity-induced electric field; (*b*) an observer moving orthogonal to the direction of a static electric field experiences a velocity-induced magnetic field. The formulas are valid in the weakly-relativistic limit ($v \ll c$).

E.3 Current-charge distributions in the quasi-static approximation

E.3.1 Introduction

Let us consider a *current-charge distribution* in vacuum, localized at the origin of a coordinate system as illustrated in Fig. E.1. The current-charge distribution is defined by the *charge-density distribution* $\varrho_e(\mathbf{r},t)$ and the *current-density distribution* $\mathbf{J}(\mathbf{r},t)$, which are coupled through the *continuity equation*

$$\frac{\partial \varrho_e}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{J} = 0. \tag{E.25}$$

The charge-density distribution is normalized on the total free charge q of the system under consideration,

$$\int \varrho_e(\mathbf{r},t) d\mathbf{r} = q. \tag{E.26}$$

Rather than using charge densities and current densities we give preference to work with neutral densities n and neutral currents $\mathbf{j}$,

$$\varrho_e(\mathbf{r},t) = q\, n(\mathbf{r},t) \tag{E.27a}$$

$$\mathbf{J}(\mathbf{r},t) = q\, \mathbf{j}(\mathbf{r},t), \tag{E.27b}$$

where

$$\int n(\mathbf{r},t) d\mathbf{r} = 1. \tag{E.28}$$

Under quasi-static conditions (no retardation), the electromagnetic potentials of this current-charge distribution are given by

$$\varphi(\mathbf{r},t) = \frac{q}{4\pi\varepsilon_0} \int \frac{1}{|\mathbf{r}-\mathbf{r}'|} n(\mathbf{r}',t) d\mathbf{r}' \tag{E.29a}$$

$$\mathbf{A}(\mathbf{r},t) = \frac{\mu_0}{4\pi} q \int \frac{1}{|\mathbf{r}-\mathbf{r}'|} \mathbf{j}(\mathbf{r}',t) \cdot d\mathbf{r}'. \tag{E.29b}$$

Note that by using for $n(\mathbf{r},t)$ a delta-function distribution centered at the origin Eq. (E.29a) reduces to the well-known expression for the *Coulomb potential* of an electric point charge q,

$$\varphi(\mathbf{r},t) = \int \frac{q\delta^3(\mathbf{r}',t)}{4\pi\varepsilon_0|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' = \frac{q}{4\pi\varepsilon_0|\mathbf{r}|}. \tag{E.30}$$

Hence, the integrand of Eq. (E.29a) represents the contribution of a point-charge density at position $\mathbf{r}'$ to the scalar potential. Likewise, the integrand of Eq. (E.29a) represents the contribution of the current density at position $\mathbf{r}'$ to the vector potential.

For a point $\mathbf{r}$ far from the current-charge distribution we can expand $1/|\mathbf{r}-\mathbf{r}'|$ in terms of a Taylor series in $\mathbf{r}'$ about the origin. For $r'/r<1$ the expansion is given by,

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|}=\sum_{n=0}^{\infty}\frac{1}{n!}(\mathbf{r}'\cdot\boldsymbol{\nabla}')^n\left.\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right|_{\mathbf{r}'=0}. \tag{E.31}$$

Since $\boldsymbol{\nabla}'|\mathbf{r}-\mathbf{r}'|=-\boldsymbol{\nabla}|\mathbf{r}-\mathbf{r}'|$ (as is easily verified in cartesian coordinates) the Taylor series can be written in the form

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|}=\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}(\mathbf{r}'\cdot\boldsymbol{\nabla})^n\frac{1}{|\mathbf{r}|}. \tag{E.32}$$

The function $1/|\mathbf{r}-\mathbf{r}'|$ is called a generating function for the Legendre polynomials because the expansion can be written in the form

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|}=\frac{1}{r}\sum_l\left(\frac{r'}{r}\right)^l P_l(\cos\theta)\quad\text{for } r'/r<1, \tag{E.33}$$

where $r=|\mathbf{r}|$ and $u=\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}'=\cos\theta$, with θ the angle included by the unit vectors $\hat{\mathbf{r}}$ and $\hat{\mathbf{r}}'$. This is readily verified order for order,

- *zero* order

$$\frac{1}{r}P_0(\cos\theta)=\frac{1}{r}, \tag{E.34}$$

- in *first* order we obtain with the aid of Eq. (N.18)

$$\frac{r'}{r^2}P_1(\cos\theta)=-(\mathbf{r}'\cdot\boldsymbol{\nabla})\frac{1}{r}=\frac{1}{r^3}(\mathbf{r}'\cdot\mathbf{r})=\frac{r'}{r^2}\cos\theta, \tag{E.35}$$

- in *second* order we obtain with the aid of Eq. (N.18)

$$\begin{aligned}\frac{r'^2}{r^3}P_2(\cos\theta)&=\frac{1}{2}(\mathbf{r}'\cdot\boldsymbol{\nabla})(\mathbf{r}'\cdot\boldsymbol{\nabla})\frac{1}{r}\\&=-\frac{1}{2}(\mathbf{r}'\cdot\boldsymbol{\nabla})\frac{1}{r^3}(\mathbf{r}'\cdot\mathbf{r})\\&=-\frac{1}{2}\left[(\mathbf{r}'\cdot\mathbf{r})(\mathbf{r}'\cdot\boldsymbol{\nabla})\frac{1}{r^3}+\frac{1}{r^3}(\mathbf{r}'\cdot\boldsymbol{\nabla})(\mathbf{r}'\cdot\mathbf{r})\right]\\&=\frac{1}{2}\left[\frac{3}{r^5}(\mathbf{r}'\cdot\mathbf{r})(\mathbf{r}'\cdot\mathbf{r})-\frac{1}{r^3}(\mathbf{r}'\cdot\mathbf{r}')\right]\\&=\frac{r'^2}{r^3}\tfrac{1}{2}\left[3\cos^2\theta-1\right].\end{aligned} \tag{E.36}$$

E.3.2 Expansion of the scalar potential

Substituting Eq. (E.33) into Eq. (E.29a) we obtain for the scalar potential

$$\varphi(\mathbf{r},t)=\frac{q}{4\pi\varepsilon_0 r}\sum_l\int\left(\frac{r'}{r}\right)^l P_l(\cos\theta)n(\mathbf{r}',t)d\mathbf{r}'\quad\text{for } r'/r<1. \tag{E.37}$$

Limiting the expansion to power $n=2$,

$$\varphi(\mathbf{r},t)=\frac{q}{4\pi\varepsilon_0 r}\int\left[1+P_1(\cos\theta)\frac{r'}{r}+P_2(\cos\theta)\left(\frac{r'}{r}\right)^2+\cdots\right]n(\mathbf{r}',t)d\mathbf{r}'. \tag{E.38}$$

Evaluating these terms order for order we obtain:

- *zero* order

$$\varphi^{(0)}(\mathbf{r},t)=\frac{q}{4\pi\varepsilon_0 r}\int n(\mathbf{r}',t)d\mathbf{r}'=\frac{q}{4\pi\varepsilon_0 r},\tag{E.39}$$

- *first* order

$$\varphi^{(1)}(\mathbf{r},t)=\frac{q}{4\pi\varepsilon_0 r^2}\int P_1(\cos\theta)r'n(\mathbf{r}',t)d\mathbf{r}'=\frac{q}{4\pi\varepsilon_0}\frac{\mathbf{d}\cdot\mathbf{r}}{r^3},\tag{E.40}$$

where the dipole moment of the charge distribution is defined as

$$\mathbf{d}=q\int\mathbf{r}'n(\mathbf{r}',t)d\mathbf{r}'.\tag{E.41}$$

- *second* order

$$\begin{aligned}\varphi^{(2)}(\mathbf{r},t)&=\frac{q}{4\pi\varepsilon_0 r^3}\int P_2(\cos\theta)r'^2n(\mathbf{r}',t)d\mathbf{r}'\\&=\frac{q}{4\pi\varepsilon_0}\frac{1}{r^3}\int\tfrac{1}{2}\left[3(\mathbf{r}'\cdot\hat{\mathbf{r}})^2-r'^2\right]n(\mathbf{r}',t)d\mathbf{r}'\\&=\frac{q}{4\pi\varepsilon_0}\frac{1}{2r^5}\int\left[3(\textstyle\sum_i x_i'x_i)(\sum_j x_j'x_j)-r'^2\right]n(\mathbf{r}',t)d\mathbf{r}'\\&=\frac{q}{4\pi\varepsilon_0}\frac{1}{2}\sum_{i,j}Q_{ij}\frac{x_ix_j}{r^5},\end{aligned}\tag{E.42}$$

where x_i are the cartesian components of $\mathbf{r}$ and Q_{ij} the components of the quadrupole moment tensor

$$Q_{ij}=\int\left[3x_i'x_j'-r'^2\delta_{ij}\right]n(\mathbf{r}',t)d\mathbf{r}'.\tag{E.43}$$

E.3.3 Expansion of the Vector potential

Substituting Eq. (E.33) into Eq. (E.29b) we obtain for the vector potential

$$\mathbf{A}(\mathbf{r},t)=\frac{\mu_0}{4\pi r}q\sum_l\left(\frac{r'}{r}\right)^l P_l(\cos\theta)\mathbf{j}(\mathbf{r}',t)d\mathbf{r}'\quad\text{for } r'/r<1.\tag{E.44}$$

Limiting the expansion to power $n=2$,

$$\mathbf{A}(\mathbf{r},t)=\frac{\mu_0}{4\pi r}q\int\mathbf{j}(\mathbf{r}',t)d\mathbf{r}'+\frac{\mu_0}{4\pi r^2}q\int P_1(\cos\theta)r'\mathbf{j}(\mathbf{r}',t)d\mathbf{r}'+\frac{\mu_0}{4\pi r^3}\int P_2(\cos\theta)r'^2\mathbf{j}(\mathbf{r}',t)d\mathbf{r}'+\cdots.\tag{E.45}$$

Let us write the current density of a moving point charge in the form

$$\mathbf{j}(\mathbf{r}')=q\,\dot{\mathbf{r}}'\delta(\mathbf{r}')\tag{E.46}$$

and evaluate the terms of the expansion order for order we obtain:

- zero order

$$\mathbf{A}^{(0)}(\mathbf{r},t)=\frac{\mu_0}{4\pi}\frac{q}{r}\int\dot{\mathbf{r}}'\delta(\mathbf{r}')d\mathbf{r}'=\frac{\mu_0}{4\pi r}q\frac{\mathbf{v}}{r}\tag{E.47}$$

- first order

$$\mathbf{A}^{(1)}(\mathbf{r},t)=\frac{\mu_0}{4\pi}\frac{q}{r^2}\int P_1(\cos\theta)r'\dot{\mathbf{r}}'\delta(\mathbf{r}')d\mathbf{r}'\tag{E.48}$$

- second order

$$\begin{aligned}\mathbf{A}^{(2)}(\mathbf{r},t) &= \frac{\mu_0}{4\pi}\frac{q}{r^3}\int P_2(\cos\theta)r'^2\mathbf{j}(\mathbf{r}',t)d\mathbf{r}' \\ &= \frac{\mu_0}{4\pi}\frac{q}{2r^3}\int\left[3(\mathbf{r}'\cdot\hat{\mathbf{r}})^2-\mathbf{r}'\cdot\mathbf{r}'\right]\mathbf{j}(\mathbf{r}',t)d\mathbf{r}' \\ &= \frac{\mu_0}{4\pi}\frac{q}{2r^5}\int\left[3(\textstyle\sum_i x_i' x_i)(\sum_j x_j' x_j)-r'^2\right]\dot{\mathbf{r}}'\delta(\mathbf{r}')d\mathbf{r}'. \qquad \text{(E.49)}\end{aligned}$$

F

Various concepts from Quantum Mechanics

In this appendix we give a summary of formal quantum mechanics for purposes of internal reference. For a proper introduction the reader is referred to one of the major textbooks on quantum mechanics, such as those by Paul Dirac [36], Albert Messiah [75, 76], Gordon Baym [10], Claude Cohen-Tannoudji, Bernard Diu and Franck Laloë [26], Eugen Merzbacher [74], and Jun John Sakurai [97].

F.1 Dirac formalism

F.1.1 Introduction

In the formulation of Dirac, the dynamical state of a physical system is established by observation of its *dynamical variables* (e.g., the position of a particle, its translational momentum, orbital angular momentum with respect to a point of reference, spin, energy, ...). Each dynamical variable $\mathcal{A}$ corresponds to a *hermitian operator* A which defines a *complete orthonormal set of eigenstates* $\{|a_i\rangle\}$ called *kets*, representing the *basis vectors* of a complex vector space, the *Hilbert space* of the dynamical system. A hermitian operator with the mentioned properties is called an *observable.* The set of eigenstates $\{|a_i\rangle\}$ is called a *representation* of the Hilbert space, often referred to as the representation $\{A\}$. The representation $\{A\}$ serves to describe measurements of the observable A. The measurement is expressed by the eigenvalue relation

$$A|a_i\rangle = a_i|a_i\rangle, \tag{F.1}$$

where a_i is the eigenvalue corresponding to the eigenstate $|a_i\rangle$. Observables are characterized by a spectrum of *real* eigenvalues. The kets are abstract state vectors which provide a *representation-free notation* for the *states.* In this introduction we focus on *discrete* representations (e.g., the standard representation $\{L^2, L_z\}$ of orbital angular momentum - cf. Appendix F.1.3). The discussion can be extended to *continuous* representations in which summations are replaced by integrations and in the orthonormality condition the Kronecker symbol becomes a Dirac delta function - cf. Appendix F.1.2.

By *completeness* we mean that an arbitrary state $|\psi\rangle$ of the system can be *represented* by the linear superposition,

$$|\psi\rangle = \sum_i |a_i\rangle\langle a_i|\psi\rangle, \tag{F.2}$$

where $\langle a_i|\psi\rangle$ is the *inner product* of the vectors $|\psi\rangle$ and $|a_i\rangle$; i.e., the *projection* of $|\psi\rangle$ onto $|a_i\rangle$. This is called the *superposition principle.* In matrix notation $\langle a_i|$ is written as a row vector and $|\psi\rangle$ as a column vector. In the Dirac formalism, the vectors $\langle a_i|$ are called *bras*. They are related to the *kets* by hermitian conjugation (antilinear transposition - in matrix notation the column vector is transformed into a row vector with the coordinates replaced by their complex conjugates;

i.e., $\langle\phi|\psi\rangle^* = \langle\psi|\phi\rangle$). The *orthonormality* of the basis is expressed by the property $\langle a_j|a_i\rangle = \delta_{ij}$. The inner product $\langle a_i|\psi\rangle$ represents the *amplitude*, short for *probability amplitude*, to observe the eigenstate $|a_i\rangle$ when the system prepared in the state $|\psi\rangle$. The corresponding *density*, short for *probability density* or *occupation density*, is given by

$$\mathcal{P}_i = |\langle a_i|\psi\rangle|^2. \tag{F.3}$$

The completeness of the representation $\{A\}$ is expressed by the *Parseval relation* (which is a probability sum rule),

$$\sum_i \mathcal{P}_i = \sum_i |\langle a_i|\psi\rangle|^2 = 1 \tag{F.4}$$

and the *closure relation* (which is a decomposition sum rule),

$$\mathbb{1} = \sum_i |a_i\rangle\langle a_i|. \tag{F.5}$$

Here $\mathbb{1}$ is the *identity operator*.

Presuming $|\psi\rangle$ to be normalized, $\langle\psi|\psi\rangle = 1$, the quantity

$$\langle A\rangle = \langle\psi|A|\psi\rangle \tag{F.6}$$

is called the *expectation value* of the operator A. For $|\phi\rangle = A|\psi\rangle$ we write $\langle\psi|A^\dagger = \langle\phi|$, where $A^\dagger$ is called the hermitian conjugate of A. This implies the relation

$$\langle\psi|A|\psi\rangle^* = \langle\psi|A^\dagger|\psi\rangle, \tag{F.7}$$

which is real for hermitian operators ($A^\dagger = A$). If the matrix elements $\langle b_i|A|b_j\rangle$ of the operator A are known in some representation $\{B\}$, the expectation value is given by

$$\langle A\rangle = \sum_{i,j}\langle\psi|b_j\rangle\langle b_j|A|b_i\rangle\langle b_i|\psi\rangle. \tag{F.8}$$

For the representation $\{A\}$, defined above, the matrix is diagonal, $\langle a_i|A|a_j\rangle = a_i\delta_{i,j}$, and Eq. (F.8) reduces to

$$\langle A\rangle = \sum_i\langle\psi|a_i\rangle a_i\langle a_i|\psi\rangle = \sum_i a_i\mathcal{P}_i. \tag{F.9}$$

This result is also intuitively correct: the expectation value corresponds to the sum of the eigenvalues weighted by the probability to have that eigenvalue.

Two dynamical variables $\mathcal{A}$ and $\mathcal{B}$ are called *compatible* if they can be measured simultaneously to arbitrary precision; i.e., share a complete set of eigenstates. This is the case if and only if the corresponding operators A and B commute. Operators that do not commute are called *incompatible*. In general, the shared basis is not uniquely defined, because two eigenstates can have the same eigenvalues. This is called *degeneracy*. To obtain a unique basis a *complete set of commuting observables* $A, B, C, \cdots$ is required, together providing all quantum numbers of the dynamical system. This is called the representation $\{A, B, C, \cdots\}$.

Problem F.1. Show that two commuting observables A and B share a complete set of eigenstates.

Solution. Consider the eigenvalue equations for the operators A and B,

$$A\,|\alpha,k\rangle = \alpha\,|\alpha,k\rangle \text{ and } B\,|\beta,l\rangle = \beta\,|\beta,l\rangle\,,$$

where $\{|\alpha,k\rangle\}$ is a complete set of k_α-fold degenerate states corresponding to the operator A and $\{|\beta,l\rangle\}$ a complete set of l_β-fold degenerate states corresponding to the operator B. In search for a joint basis we know that the eigenstates of A can be expressed in the eigenstates of B

$$|\alpha,k\rangle = \sum_\beta\sum_{l=1}^{l_\beta}|\beta,l\rangle\,\langle\beta,l|\alpha,k\rangle = \sum_\beta|\beta\rangle\,,$$

where

$$|\beta\rangle \equiv \sum_{l=1}^{l_\beta} |\beta, l\rangle \langle \beta, l|\alpha, k\rangle \,.$$

Note that $|\beta\rangle$ is a linear combination of *degenerate* eigenstates of B, all with eigenvalue β. This implies that $|\beta\rangle$ itself is an eigenstate of B with eigenvalue β as follows from

$$B\,|\beta\rangle \equiv \sum_{l=1}^{l_\beta} B\,|\beta, l\rangle \langle \beta, l|\alpha, k\rangle = \sum_{l=1}^{l_\beta} \beta\,|\beta, l\rangle \langle \beta, l|\alpha, k\rangle = \beta\,|\beta\rangle \,.$$

Next we show that $|\beta\rangle$ is also an eigenstate of A. Since $[A, B] = 0$ we have

$$B(A\,|\beta\rangle) = AB\,|\beta\rangle = A\beta\,|\beta\rangle = \beta(A\,|\beta\rangle).$$

Hence $A\,|\beta\rangle$ is also an eigenstate of B with eigenvalue β. Having the eigenvalue β this eigenstate has to be of the form, $A\,|\beta\rangle = \lambda\,|\beta\rangle$. This shows that $|\beta\rangle$ is also an eigenstate of A. The corresponding eigenvalue λ has to be equal to α as follows from

$$\sum_\beta A\,|\beta\rangle = A\,|\alpha, k\rangle = \alpha\,|\alpha, k\rangle = \sum_\beta \alpha\,|\beta\rangle \,. \qquad \square$$

F.1.2 Continuous bases - position and momentum representation

For a particle in the *arbitrary* state $|\psi\rangle$ the wavefunction $\psi(\mathbf{r})$ represents the probability amplitude to observe the particle at position $\mathbf{r}$. The $\psi(\mathbf{r})$ defines the state $|\psi\rangle$ in the *position representation* and $\mathbf{r}$ stands for the *position coordinates.* In the Dirac formalism the same probability amplitude is obtained by projecting the state vector $|\psi\rangle$ onto the eigenstate $|\mathbf{r}\rangle$ of the position operator $\underline{\mathbf{r}}$,

$$\psi(\mathbf{r}) = \langle \mathbf{r}|\psi\rangle. \tag{F.10}$$

The set $\{|\mathbf{r}\rangle\}$ forms a basis of the Hilbert space of the particle because $|\psi\rangle$ is completely defined by specifying the probability amplitude $\psi(\mathbf{r})$ for all values of $\mathbf{r}$,

$$|\psi\rangle = \int d\mathbf{r}\, |\mathbf{r}\rangle\langle \mathbf{r}|\psi\rangle. \tag{F.11}$$

This shows that the closure relation is given by

$$\mathbb{1} = \int d\mathbf{r}\, |\mathbf{r}\rangle\langle \mathbf{r}| \tag{F.12}$$

and the Parseval relation takes the form of the normalization integral,

$$1 = \int d\mathbf{r} |\langle \mathbf{r}|\psi\rangle|^2 = \int |\psi(\mathbf{r})|^2 d\mathbf{r}. \tag{F.13}$$

To find the orthogonality relation for the position representation we note that the probability amplitude to observe the particle at position $\mathbf{r}'$ is given by

$$\psi(\mathbf{r}') = \int d\mathbf{r}\, \langle \mathbf{r}'|\mathbf{r}\rangle \psi(\mathbf{r}). \tag{F.14}$$

Since $|\psi\rangle$ was chosen arbitrary this implies

$$\langle \mathbf{r}'|\mathbf{r}\rangle = \delta(\mathbf{r} - \mathbf{r}'), \tag{F.15}$$

which is the Dirac deltafunction in three dimensions.

In search for the *momentum representation* we consider the Fourier transform

$$\psi(\mathbf{p}) = \int d\mathbf{r}\, e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar}\psi(\mathbf{r}). \tag{F.16}$$

The inverse transform is given by

$$\psi(\mathbf{r}) = (2\pi\hbar)^{-3}\int d\mathbf{p}\, e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}\psi(\mathbf{p}). \tag{F.17}$$

In the plane wave $e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}$ we recognize the wavefunction of a free particle in state $|\mathbf{p}\rangle$,

$$\langle\mathbf{r}|\mathbf{p}\rangle = e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}. \tag{F.18}$$

Substituting this expression into the Fourier transform (F.16) we find with the aid of the closure relation (F.12)

$$\psi(\mathbf{p}) = \int d\mathbf{r}\, \langle\mathbf{p}|\mathbf{r}\rangle\langle\mathbf{r}|\psi\rangle = \langle\mathbf{p}|\psi\rangle. \tag{F.19}$$

Substituting this expression into the inverse transform we obtain

$$|\psi\rangle = (2\pi\hbar)^{-3}\int d\mathbf{p}\, |\mathbf{p}\rangle\langle\mathbf{p}|\psi\rangle. \tag{F.20}$$

Since $|\psi\rangle$ was chosen arbitrarily, this shows that the set $\{|\mathbf{p}\rangle\}$ forms the basis of the momentum representation for the Hilbert of the particle. The wavefunction $\psi(\mathbf{p}) = \langle\mathbf{p}|\psi\rangle$ represents the probability amplitude that the particle is measured with momentum $\mathbf{p}$. The closure relation is given by

$$\mathbb{1} = (2\pi\hbar)^{-3}\int d\mathbf{p}\, |\mathbf{p}\rangle\langle\mathbf{p}| \tag{F.21}$$

and the Parseval relation becomes,

$$1 = (2\pi\hbar)^{-3}\int d\mathbf{p}|\langle\mathbf{p}|\psi\rangle|^2 = (2\pi\hbar)^{-3}\int |\psi(\mathbf{p})|^2 d\mathbf{p}. \tag{F.22}$$

To find the orthogonality relation we note that the probability amplitude to find the particle with momentum $\mathbf{p}'$ is given by

$$\psi(\mathbf{p}') = (2\pi\hbar)^{-3}\int d\mathbf{p}\, \langle\mathbf{p}'|\mathbf{p}\rangle\psi(\mathbf{p}). \tag{F.23}$$

Since $|\psi\rangle$ was chosen arbitrary we obtain

$$\langle\mathbf{p}'|\mathbf{p}\rangle = (2\pi\hbar)^3\delta(\mathbf{p}-\mathbf{p}'). \tag{F.24}$$

Using the closure relation (F.12) as well as Eq. (F.18) we find the expression

$$\int d\mathbf{r}\langle\mathbf{p}'|\mathbf{r}\rangle\langle\mathbf{r}|\mathbf{p}\rangle = \int d\mathbf{r}\, e^{i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{r}/\hbar} = (2\pi\hbar)^3\delta(\mathbf{p}-\mathbf{p}'). \tag{F.25}$$

F.1.3 Discrete basis - orbital angular momentum

As an example of a discrete basis, we discuss the case of orbital angular momentum. The three cartesian components of the angular momentum operator $\mathbf{L}$ satisfy the commutation relations

$$[L_x, L_y] = i\hbar L_z,\ [L_y, L_z] = i\hbar L_x \text{ and } [L_z, L_x] = i\hbar L_y. \tag{F.26}$$

These are the generic commutation relations for any type of angular momentum. The Hermitian operators that uniquely define the state of orbital angular momentum are the operators $\mathbf{L}^2$ and L_z. Their shared basis $\{|l,m_l\rangle\}$ is defined by

$$\mathbf{L}^2\,|l,m_l\rangle = l(l+1)\hbar^2\,|l,m_l\rangle \tag{F.27a}$$

$$L_z\,|l,m_l\rangle = m_l\hbar\,|l,m_l\rangle\,, \tag{F.27b}$$

where the $|l,m_l\rangle$ are abstract (*i.e.*, representation-free) state vectors in Hilbert space, with l and m_l the *rotational quantum numbers.* The basis $\{|l,m_l\rangle\}$ is called the *standard representation* $\{L^2,L_z\}$ of angular momentum. The spherical harmonics

$$Y_l^{m_l}(\theta,\phi) \equiv Y_l^{m_l}(\hat{\mathbf{r}}) = \langle\hat{\mathbf{r}}|l,m_l\rangle \tag{F.28}$$

are the corresponding wavefunctions in the position representation of spherical coordinates.

F.1.4 Spin coordinates, spinor states and spinorbitals

The “wavefunction” $\chi(\sigma)$ of the spin of a particle in the angular momentum state $|\chi\rangle$ is obtained by projection of $|\chi\rangle$ onto one of the eigenstates $\{|\sigma\rangle\}$ of the s_z operator,

$$\chi(\sigma) = \langle\sigma|\chi\rangle = \sum_{m_s=-s}^{s}\langle\sigma|s,m_s\rangle\langle s,m_s|\chi\rangle, \tag{F.29}$$

where σ assumes discrete values in the interval $-s\le\sigma\le s$. Hence, the function $\chi(\sigma)$ is a function of the discrete variable σ, the *spin coordinate.* For a particle in spin state χ the spinor $\chi(\sigma)$ represents the probability amplitude to detect the particle in the eigenstate σ. For s being *half-integral* it is called a $2s+1$ component *spinor* (to distinguish its rotation properties from scalar and vector functions). The basis states are denoted by the unit spinors

$$\chi_{m_s}(\sigma) \equiv \langle\sigma|s,m_s\rangle = \delta_{m_s,\sigma}, \tag{F.30}$$

where m_s assumes discrete values in the interval $s\le m_s\le s$. For the special case of $s=\frac{1}{2}$ the eigenstates are often written in the arrow notation, $\chi_{+1/2}(\sigma)\equiv\chi_\uparrow(\sigma)$ and $\chi_{-1/2}(\sigma)\equiv\chi_\downarrow(\sigma)$.

As an example we consider the case $s=\frac{1}{2}$, where the summation runs over the values $m_s\in\{-\frac{1}{2},+\frac{1}{2}\}$,

$$\chi(\sigma) = \chi_\uparrow(\sigma)\langle+\tfrac{1}{2}|\chi\rangle + \chi_\downarrow(\sigma)\langle-\tfrac{1}{2}|\chi\rangle. \tag{F.31}$$

Note that $\chi(\sigma)=\langle\pm\frac{1}{2}|\chi\rangle$ for $\sigma=\pm\frac{1}{2}$. In the *matrix representation* the $s=\frac{1}{2}$ spinors take the form of a column vector,

$$\chi(\sigma) = \begin{pmatrix}\langle+\frac{1}{2}|\chi\rangle\\ \langle-\frac{1}{2}|\chi\rangle\end{pmatrix} = \langle+\tfrac{1}{2}|\chi\rangle\begin{pmatrix}1\\0\end{pmatrix} + \langle-\tfrac{1}{2}|\chi\rangle\begin{pmatrix}0\\1\end{pmatrix}. \tag{F.32}$$

and the basis vectors take the form (up to a not observable phase factor)

$$\chi_\uparrow(\sigma) = \begin{pmatrix}1\\0\end{pmatrix} \quad\text{and}\quad \chi_\downarrow(\sigma) = \begin{pmatrix}0\\1\end{pmatrix}. \tag{F.33}$$

The spinor concept is readily extended from *discrete spinor* functions to *spinor fields.* Such an extension is essential to describe particles with both motional and spin degrees of freedom (for example the electron). Since $\boldsymbol{s}$ and $\underline{\mathbf{r}}$ are commuting observables they must share a complete set of basis functions, obtained as the *Kronecker product* of their eigenstates

$$|\mathbf{r},\sigma\rangle \equiv |\mathbf{r}\rangle\otimes|\sigma\rangle. \tag{F.34}$$

The Kronecker product assures that for every eigenvalue of one operator the state can assume all eigenvalues of the other operator. An arbitrary state can be written in the form

$$\psi\left(\mathbf{r},\sigma\right)=\langle\mathbf{r},\sigma|\psi\rangle=\varphi\left(\mathbf{r}\right)\chi(\sigma),\tag{F.35}$$

where $\varphi\left(\mathbf{r}\right)=\langle\mathbf{r}|\varphi\rangle$ is the wavefunction in the position representation and $\chi(\sigma)=\langle\sigma|\chi\rangle$ a discrete spinor.

For the case $s=\frac{1}{2}$ this implies

$$\psi\left(\mathbf{r},\sigma\right)=\varphi\left(\mathbf{r}\right)\langle+\tfrac{1}{2}|\chi\rangle\chi_{\uparrow}(\sigma)+\varphi\left(\mathbf{r}\right)\langle-\tfrac{1}{2}|\chi\rangle\chi_{\downarrow}(\sigma).\tag{F.36}$$

In matrix form this becomes

$$\psi\left(\mathbf{r},\sigma\right)=\begin{pmatrix}\psi\left(\mathbf{r},+1/2\right)\\ \psi\left(\mathbf{r},-1/2\right)\end{pmatrix}=\psi\left(\mathbf{r},+1/2\right)\begin{pmatrix}1\\0\end{pmatrix}+\psi\left(\mathbf{r},-1/2\right)\begin{pmatrix}0\\1\end{pmatrix},\tag{F.37}$$

where $\psi\left(\mathbf{r},\pm1/2\right)$ is a *double-valued* function representing the probability amplitude to observe the particle at position $\mathbf{r}$ in spin state $|\pm\frac{1}{2}\rangle$.

If a particle is prepared in the motional state $|\varphi\rangle=|\mathbf{k}\rangle$ and spin state $|\chi\rangle=|\uparrow\rangle$ the Eq. (F.36) takes the form of a *spinorbital*

$$\psi_{\mathbf{k}\uparrow}\left(\mathbf{r},\sigma\right)=\varphi_{\mathbf{k}}\left(\mathbf{r}\right)\chi_{\uparrow}(\sigma).\tag{F.38}$$

In terms of spinorbitals the spinor field becomes

$$\psi\left(\mathbf{r},\sigma\right)=\psi_{\mathbf{k}\uparrow}\left(\mathbf{r},\sigma\right)\chi(+\tfrac{1}{2})+\psi_{\mathbf{k}\downarrow}\left(\mathbf{r},\sigma\right)\chi(-\tfrac{1}{2}),\tag{F.39}$$

where $\chi(\pm\frac{1}{2})$ is the probability amplitude to observe the particle in spin state $\left|\pm\frac{1}{2}\right\rangle$.

F.2 The Schrödinger and Heisenberg pictures of time evolution

In this section we discuss two alternative views on the time evolution in non-relativistic quantum mechanics. These are known as the *Schrödinger* and the *Heisenberg picture.* To introduce these pictures, we start by recalling that the state of a quantum mechanical system is postulated to be *uniquely* defined by a state vector in Hilbert space (see Appendix F.1.1). At any point in time a quantum mechanical state can be written as a linear superposition of eigenstates. This superposition is (in general) not stationary but evolves in time. In the Schrödinger picture we postulate that the time evolution is *uniquely* defined by a *linear* transformation,

$$|\psi_S(t)\rangle=U(t,t_0)|\psi_S(t_0)\rangle,\tag{F.40}$$

where $|\psi_S(t)\rangle$ is the state of the system at time t. The linearity implies that $|\psi_S(t)\rangle$ is given by a linear superposition of *independently* evolving eigenstates; i.e., the superposition is conserved but the coefficients vary in time. The operator $U(t,t_0)$ is called the evolution operator of the system and accounts for the time dependence of the states over the interval $t_0\to t$. By definition $U(t_0,t_0)=1$. To assure the postulated uniqueness of the state evolution, the transformation has to be *norm conserving*,

$$\langle\psi_S(t)|\psi_S(t)\rangle=\langle\psi_S(t_0)|\psi_S(t_0)\rangle.\tag{F.41}$$

Only in this way we can be sure that a system, initially in state $|\psi_S(t_0)\rangle$, arrives with unit probability at time t in the state $|\psi_S(t)\rangle$. Substituting the transformation (F.40) into (F.41), we find that the norm is conserved if and only if $U^{\dagger}U=\mathbb{1}$; i.e., the evolution has to be *unitary*,

$$U^{\dagger}(t,t_0)=U^{-1}(t,t_0)=U(t_0,t).\tag{F.42}$$

In view of the uniqueness of the time evolution, $U(t,t_0)$ can be written as a product of unitary operators,

$$U(t,t_0) = U(t,t_n)U(t_n,t_{n-1})\cdots U(t_2,t_1)U(t_1,t_0). \tag{F.43}$$

This shows that the set $\{U(t,t')\}$ forms a group under multiplication, with unit element $U(t,t)$ and inverse $U(t',t)$. In particular, as t is a continuous variable, $U(t,t_0)$ can be written as an *infinite* product of *infinitesimal* evolution operators $U(t+\delta t,t)$. Infinitesimal unitary operators with the property $U(t,t)=1$ can be expressed in the form

$$U(t+\delta t,t) = \mathbb{1} - (i/\hbar)\mathcal{H}(t)\delta t = e^{-i\mathcal{H}(t)\delta t/\hbar} \quad \text{for } \delta t \to 0, \tag{F.44}$$

where $\mathcal{H}(t)$ has to be a *hermitian* operator to assure the unitarity of U. The factor $1/\hbar$ has been included for future convenience. Note that (for any point in time) the evolution operator commutes with the Hamiltonian,

$$[U(t+\delta t,t),\mathcal{H}(t)] = 0. \tag{F.45}$$

Using $U(t,t)=1$, we rewrite Eq. (F.44) in the form

$$\lim_{\delta t\to 0}\frac{U(t+\delta t,t)-U(t,t)}{\delta t} = -(i/\hbar)\mathcal{H}(t). \tag{F.46}$$

Multiplying both sides with $U(t,t_0)$ we obtain the differential equation [1]

$$i\hbar\frac{\partial}{\partial t}U(t,t_0) = \mathcal{H}(t)U(t,t_0), \tag{F.47}$$

subject to the boundary condition $U=1$ for $t=t_0$. This equation can be rewritten in the form of an integral equation by integration over the finite time interval $t_0 \to t$,

$$U(t,t_0) = \mathbb{1} - (i/\hbar)\int_{t_0}^{t} dt'\mathcal{H}(t')U(t',t_0). \tag{F.48}$$

This form for the evolution operator is particularly suited for generating iterative solutions for application in time-dependent perturbation theory.

As the time dependences of $U(t,t_0)$ and $|\psi_S(t)\rangle$ are related by the equation

$$i\hbar\frac{\partial}{\partial t}|\psi_S(t)\rangle = \left[i\hbar\frac{\partial}{\partial t}U(t,t_0)\right]|\psi_S(t_0)\rangle \tag{F.49}$$

we find for the time dependence of the state vector

$$i\hbar\frac{\partial}{\partial t}|\psi_S(t)\rangle = \mathcal{H}(t)|\psi_S(t)\rangle, \tag{F.50}$$

with the boundary condition that at $t=t_0$ the system is in the state $|\psi_S(t_0)\rangle$. In this differential equation we recognize the familiar form of the time-dependent Schrödinger equation. In view of this formal similarity the hermitian operator $\mathcal{H}$ is called the Hamiltonian of the quantum mechanical system. In hindsight this explains the factor $1/\hbar$ included in Eq. (F.44). This being said, it should be emphasized that Eq. (F.50) was obtained from the postulates without any reference to the Hamilton formalism of classical mechanics. It shows that the Hamiltonian can be defined, *also in the absence of any correspondence to a classical system*, as the hermitian operator that assures unitary evolution of the system.

Recalling that $U(t,t_0)$ can be written as an infinite product of infinitesimal operators we distinguish three cases

---

[1] The time derivative of the operator $U(t,t')$ is defined as

$$\frac{\partial}{\partial t}U(t,t') \equiv \lim_{\delta t\to 0}\frac{U(t+\delta t,t')-U(t,t')}{\delta t}.$$

- *Time-dependent Hamiltonian* with the property $[\mathcal{H}(t),\mathcal{H}(t')] \neq 0$.

$$U(t,t_0) \equiv \lim_{\delta t\to 0} e^{-i\mathcal{H}(t-\delta t)\delta t/\hbar} \cdots e^{-i\mathcal{H}(t_0)\delta t/\hbar} \tag{F.51}$$

- *Time-dependent Hamiltonian* with the property $[\mathcal{H}(t),\mathcal{H}(t')] = 0$. Using the properties of exponential operators (see Appendix L.8)

$$U(t,t_0) \equiv \lim_{\delta t\to 0} e^{-i\mathcal{H}(t-\delta t)\delta t/\hbar} \cdots e^{-i\mathcal{H}(t_0)\delta t/\hbar} = e^{-(i/\hbar)\int_{t_0}^{t}\mathcal{H}(t')dt'} \tag{F.52}$$

- *Time-independent Hamiltonian* $\mathcal{H} = \mathcal{H}_0$ (conservative system). In this case we write $\delta t \equiv (t-t_0)/n$ and obtain with the aid of Eq. (L.33a)

$$U(t,t_0) \equiv \lim_{n\to\infty} \left[e^{-(i/\hbar)\mathcal{H}_0(t-t_0)/n}\right]^n = e^{-i\mathcal{H}(t-t_0)/\hbar}. \tag{F.53}$$

F.2.1 Schrödinger picture

In this section we summarize the primary features of the *Schrödinger picture.* The evolution of the physical system is contained in the time dependence of the state vector,

$$|\psi_S(t)\rangle = U(t,t_0)|\psi_S(t_0)\rangle, \tag{F.54}$$

where $|\psi_S(t)\rangle$ is the *Schrödinger state* of the system at time t. In the Schrödinger picture we write A_S for the operator that represents the dynamical variable A. Likewise, we shall write A_H when turning to the Heisenberg picture. An exception is made for the Hamiltonian because, as we shall see, this operator is the same in both pictures. For this reason we simply write $\mathcal{H}$ rather than $\mathcal{H}_S$ or $\mathcal{H}_H$. The time dependence of the state vectors is governed by the Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}|\psi_S(t)\rangle = \mathcal{H}(t)|\psi_S(t)\rangle, \tag{F.55}$$

subject to the boundary condition that at $t = t_0$ the system is in the state $|\psi_S(t_0)\rangle$. According to the postulate of quantum measurement (see Appendix F.1.1), the expectation value of $A_S(t)$ is given by

$$\langle A_S(t)\rangle = \frac{\langle\psi_S(t)|A_S(t)|\psi_S(t)\rangle}{\langle\psi_S(t)|\psi_S(t)\rangle}. \tag{F.56}$$

As the evolution is norm conserving, the time derivative of the expectation value is given by - see Problem F.2

$$i\hbar\frac{d}{dt}\langle A_S(t)\rangle = \langle[A_S(t),\mathcal{H}(t)]\rangle + i\hbar\langle\frac{\partial}{\partial t}A_S(t)\rangle. \tag{F.57}$$

Here the first term arises from the time development of the state and the second one from that of the operator. Note that the latter is only nonzero if the operator depends explicitly on time.

Problem F.2. Show that the equation of motion for the expectation value of a (generally time-dependent) operator $A_S(t)$ is given by

$$i\hbar\frac{d}{dt}\langle A_S(t)\rangle = \langle[A_S(t),\mathcal{H}(t)]\rangle + i\hbar\langle\frac{\partial}{\partial t}A_S(t)\rangle.$$

Solution. The answer is obtained by differentiation of (F.56) using the Leibniz integral rule (L.28). Since the norm of the state is conserved we obtain the following three terms,

$$\frac{d}{dt}\langle A_S(t)\rangle = \left[\frac{\partial}{\partial t}\langle\psi_S(t)|\right] A_S(t)|\psi_S(t)\rangle + \langle\psi_S(t)|\left[\frac{\partial}{\partial t}A_S(t)\right]|\psi_S(t)\rangle + \langle\psi_S\,(t)\,|A_S(t)\left[\frac{\partial}{\partial t}|\psi_S(t)\rangle\right].$$

The first and the third term are evaluated using the Schrödinger equation (F.55) and its hermitian conjugate. The second term is zero unless the operator has an explicit time dependence. Thus we obtain

$$i\hbar\frac{d}{dt}\langle A_S(t)|\rangle = -\langle\psi_S(t)|\left[\mathcal{H}(t)A_S(t) - A_S(t)\mathcal{H}(t)\right]|\psi_S(t)\rangle + i\hbar\langle\psi_S(t)|\left[\frac{\partial}{\partial t}A_S(t)\right]|\psi_S(t)\rangle,$$

which can be rewritten in the form of the desired expression. □

F.2.2 Heisenberg picture

The *Heisenberg picture* is obtained by a unitary transformation of the Schrödinger states and operators in the Hilbert space. This unitary transformation is chosen such that it exactly removes the time dependence from the Schrödinger state $|\psi_S(t)\rangle$ by evolving it back to $t = t_0$,

$$|\psi_H\rangle \equiv U^\dagger(t,t_0)|\psi_S(t)\rangle = |\psi_S(t_0)\rangle. \tag{F.58}$$

The same unitary transformation puts a time dependence on the operators,

$$A_H(t) = U^\dagger(t,t_0)A_S(t)U(t,t_0). \tag{F.59}$$

Note that with this transformation the Heisenberg and Schrödinger pictures coincide at $t = t_0$,

$$A_H(t_0) = A_S(t_0) \quad \text{and} \quad |\psi_H\rangle = |\psi_S(t_0)\rangle. \tag{F.60}$$

Importantly, since $U(t,t_0)$ commutes with the Hamiltonian $\mathcal{H}(t)$, the transformation to the Heisenberg picture leaves the Hamiltonian invariant. For this reason we omit the subscript and simply write $\mathcal{H}(t)$. It is straightforward to show, using Eq. (F.56), that the expectation value $\langle A_H\rangle$ coincides with $\langle A_S\rangle$ at any time t and is given by

$$\langle A_H(t)\rangle = \frac{\langle\psi_H|A_H(t)|\psi_H\rangle}{\langle\psi_H|\psi_H\rangle} = \frac{\langle\psi_H|U^\dagger(t,t_0)A_S(t)U(t,t_0)|\psi_H\rangle}{\langle\psi_H|U^\dagger(t,t_0)U(t,t_0)|\psi_H\rangle} = \langle A_S(t)\rangle. \tag{F.61}$$

As, by construction, $|\psi_H\rangle$ does not depend on time, the time dependence of $\langle A_H\,(t)\rangle$ is fully determined by the time dependence of $A_H\,(t)$. The latter is known as the the *Heisenberg equation of motion.* By differentiation of (F.59) we obtain - see Problem F.3

$$i\hbar\frac{d}{dt}A_H(t) = [A_H(t),\mathcal{H}(t)] + i\hbar\frac{\partial}{\partial t}A_H(t), \tag{F.62}$$

where

$$\frac{\partial}{\partial t}A_H(t) \equiv \left(\frac{\partial}{\partial t}A_S(t)\right)_H = U^\dagger(t)\left[\frac{\partial}{\partial t}A_S(t)\right]U(t). \tag{F.63}$$

The correspondence (F.61) shows that the Heisenberg equation of motion is completely equivalent to the Schrödinger equation.

Problem F.3. Consider the Heisenberg operator $A_H(t) = U^\dagger(t)A_S(t)U(t)$. Show that the equation of motion of this operator is given by

$$i\hbar\frac{d}{dt}A_H(t) = [A_H(t),\mathcal{H}(t)] + i\hbar\frac{\partial}{\partial t}A_H(t).$$

Solution. By differentiation of $\langle A_H(t)\rangle$ we obtain three terms using the Leibniz integral rule (L.28)

$$\frac{d}{dt}\langle A_H(t)\rangle = \langle\left[\frac{\partial}{\partial t}U^\dagger(t)\right]A_S(t)U(t) + U^\dagger(t)\left[\frac{\partial}{\partial t}A_S(t)\right]U(t) + U^\dagger(t)A_S(t)\left[\frac{\partial}{\partial t}U(t)\right]\rangle.$$

The first and the third term are evaluated using the differential equation (F.47) and its hermitian conjugate. The second term is zero unless the operator has an explicit time dependence. As the Heisenberg states are time independent we can assign all time dependence to the Heisenberg operators,

$$i\hbar\frac{d}{dt}A_H(t) \equiv -U^\dagger(t)\left[\mathcal{H}(t)A_S(t) - A_S(t)\mathcal{H}(t)\right]U(t) + i\hbar U^\dagger(t)\frac{\partial}{\partial t}A_S(t)U(t).$$

As $\mathcal{H}(t)$ commutes with the evolution operator this expression simplifies to

$$i\hbar\frac{d}{dt}A_H(t) = [A_H(t),\mathcal{H}(t)] + i\hbar U^\dagger(t)\frac{\partial}{\partial t}A_S(t)U(t).$$

Defining

$$\frac{\partial}{\partial t}A_H(t) \equiv \left(\frac{\partial}{\partial t}A_S(t)\right)_H \equiv U^\dagger(t)\left[\frac{\partial}{\partial t}A_S(t)\right]U(t)$$

we arrive at the desired expression. □

F.2.3 Interaction picture

Rather than removing the full time dependence, we can also use a unitary transformation that removes only part of it. It may speak for itself that this can be done in many ways. A particularly valuable variant arises when the Hamiltonian of the system can be separated into two parts,

$$\mathcal{H}(t) = \mathcal{H}_0 + \mathcal{H}_1(t), \tag{F.64}$$

where $\mathcal{H}_0$ is the time-independent Hamiltonian of a closed system and $\mathcal{H}_1(t)$ is a (generally time-dependent) perturbation representing some interaction with an external field. To analyze this case, we consider the unitary transformations

$$|\psi_I(t)\rangle = U_0^\dagger(t)|\psi_S(t)\rangle, \tag{F.65}$$

and

$$A_I(t) = U_0^\dagger(t)A_S(t)U_0(t). \tag{F.66}$$

Here $U_0(t)$ is the evolution operator corresponding to $\mathcal{H}_0$,

$$U_0(t) = e^{-i\mathcal{H}_0(t-t_1)/\hbar}, \tag{F.67}$$

which is the solution of the differential equation

$$i\hbar\frac{\partial}{\partial t}U_0(t) = \mathcal{H}_0 U_0(t), \tag{F.68}$$

under the boundary condition $U_0(t) = 1$ for $t = t_1$. With this definition, the state $|\psi_I(t_1)\rangle$ coincides with the Schrödinger state $|\psi_S(t)\rangle$ for $t = t_1$. As we are free to choose the time t_1, we conveniently choose $t_0 = 0$.

The above transformations define the *interaction* picture and the states $|\psi_I(t)\rangle$ are called the *intermediate* states. The operator $A_I(t)$ is called the *intermediate* operator. Note that $U_0(t)$ removes all $\mathcal{H}_0$-related evolution from the state but leaves the part related to $\mathcal{H}_1(t)$. For small $\mathcal{H}_1(t)$ this means that $|\psi_I(t)\rangle$ is "almost" stationary. It follows with the aid of Eqs. (F.65) and (F.56) that the expectation values $\langle A_I\rangle$ and $\langle A_S\rangle$ coincide at all times,

$$\langle A_I\,(t)\rangle = \frac{\langle\psi_I(t)|A_I(t)|\psi_I(t)\rangle}{\langle\psi_I(t)|\psi_I(t)\rangle} = \frac{\langle\psi_I(t)|U_0^\dagger(t)A_S(t)U_0(t)|\psi_I(t)\rangle}{\langle\psi_I(t)|U_0^\dagger(t)U_0(t)|\psi_I(t)\rangle} = \langle A_S(t)\rangle. \tag{F.69}$$

The time dependence of the intermediate states can be obtained by differentiation of Eq. (F.65) - see Problem F.4,

$$i\hbar\frac{\partial}{\partial t}|\psi_I(t)\rangle = \mathcal{H}_I(t)|\psi_I(t)\rangle, \tag{F.70}$$

where $\mathcal{H}_I(t)$ is defined by

$$\mathcal{H}_I(t) \equiv U_0^\dagger(t)\mathcal{H}_1(t)U_0(t), \tag{F.71}$$

which is consistent with the general case (F.66). In complete analogy with the derivation of the Heisenberg equation of motion we obtain by differentiation of (F.66) the *intermediate* equation of motion,

$$i\hbar\frac{d}{dt}A_I(t) = [A_I(t), \mathcal{H}_0] + i\hbar\frac{\partial}{\partial t}A_I(t), \tag{F.72}$$

where

$$\frac{\partial}{\partial t}A_I(t) \equiv \left(\frac{\partial}{\partial t}A_S(t)\right)_I = U_0^\dagger(t)\left[\frac{\partial}{\partial t}A_S(t)\right]U_0(t). \tag{F.73}$$

Hence, just like the time evolution of A_H follows from the full Hamiltonian, the time dependence of A_I follows from the partial Hamiltonian $\mathcal{H}_0$.

As also the time evolution driven by $\mathcal{H}_I(t)$ has to be uniquely defined the evolution of the intermediate state over the interval $t_0 \to t$ can be written in the form

$$|\psi_I(t)\rangle = U_I(t,t_0)|\psi_I(t_0)\rangle, \tag{F.74}$$

where $U_I(t,t_0)$ is a *unitary* operator and the solution of the differential equation

$$i\hbar\frac{\partial}{\partial t}U_I(t,t_0) = \mathcal{H}_I(t)U_I(t,t_0), \tag{F.75}$$

under the boundary condition $U_I(t_0,t_0) = 1$. This differential equation can be rewritten in the form of an integral equation,

$$U_I(t,t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} dt_1 \mathcal{H}_I(t_1)U_I(t_1,t_0). \tag{F.76}$$

In search for a relation between $U_I(t,t_0)$ and $U(t,t_0)$ we write, with the aid of Eqs. (F.65) and (F.54), the intermediate state in the form

$$|\psi_I(t)\rangle = U_0^\dagger(t)|\psi_S(t)\rangle = U_0^\dagger(t)U(t,t_0)|\psi_S(t_0)\rangle = U_0^\dagger(t)U(t,t_0)U_0(t_0)|\psi_I(t_0)\rangle. \tag{F.77}$$

Comparing this expression with Eq. (F.65) we obtain the following relation between the full and the intermediate evolution operator,

$$U_I(t,t_0) = e^{i\mathcal{H}_0 t/\hbar}U(t,t_0)e^{-i\mathcal{H}_0 t/\hbar}. \tag{F.78}$$

Note that in the intermediate picture the time dependence is divided between the states and the operators. Its operators resemble the Heisenberg operators - compare Eqs. (F.62) and (F.72), whereas its states resemble the Schrödinger states - compare Eqs. (F.55) and (F.70).

Problem F.4. Show that the equation of motion for the expectation value of a (generally time-dependent) intermediate operator $A_I(t)$ is given by

$$i\hbar\frac{\partial}{\partial t}|\psi_I(t)\rangle = \mathcal{H}_I(t)|\psi_I(t)\rangle.$$

Solution. Differentiating (F.65) we obtain with the aid of Eqs. (F.68) and (F.55)

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\psi_I(t)\rangle &= \left[\frac{\partial}{\partial t}U_0^\dagger(t)\right]|\psi_S(t)\rangle + i\hbar U_0^\dagger(t)\frac{\partial}{\partial t}|\psi_S(t)\rangle \\ &= -\mathcal{H}_0 U_0^\dagger(t)|\psi_S(t)\rangle + U_0^\dagger(t)\mathcal{H}(t)|\psi_S(t)\rangle. \end{aligned}$$

Using Eq. (F.65) and its inverse, this becomes

$$i\hbar\frac{\partial}{\partial t}|\psi_I(t)\rangle = -\mathcal{H}_0|\psi_I(t)\rangle + U_0^\dagger(t)\mathcal{H}(t)U_0(t)|\psi_I(t)\rangle.$$

Since $\mathcal{H}_0 = U_0^\dagger(t)\mathcal{H}_0 U_0(t)$ and $\mathcal{H}_1(t) = \mathcal{H}(t) - \mathcal{H}_0$ this corresponds to

$$i\hbar\frac{\partial}{\partial t}|\psi_I(t)\rangle = U_0^\dagger(t)\mathcal{H}_1(t)U_0(t)|\psi_I(t)\rangle.$$

Defining $\mathcal{H}_I(t) \equiv U_0^\dagger(t)\mathcal{H}_1(t)U_0(t)$ we arrive at the desired form. □

F.2.4 Constants of the motion and "good" quantum numbers

The dynamical variable A of a quantum mechanical system is called a *constant of the motion* if the expectation value $\langle A_S\rangle$ is independent of time, whatever the state of the system. This implies that the Heisenberg operator A_H is stationary as follows by differentiation of (F.61),

$$\frac{d}{dt}\langle A_S\rangle = \langle \frac{d}{dt}A_H\rangle \equiv 0 \quad \Rightarrow \quad \frac{d}{dt}A_H = 0. \tag{F.79}$$

Furthermore, expressing the Heisenberg equation of motion for the operator A_H in terms of the corresponding Schrödinger operator A_S we find with the aid of Eqs. (F.62) and (F.45)

$$i\hbar\frac{d}{dt}A_H = U^{\dagger}(t,t_0)[A_S,\mathcal{H}]U(t,t_0) + i\hbar U^{\dagger}(t)\left[\partial A_S(t)/\partial t\right]U(t). \tag{F.80}$$

This shows that the dynamical variable A is a constant of the motion if A_S does not depend explicitly on time, $\partial A_S(t)/\partial t = 0$, and commutes with the Hamiltonian, $[A_S,\mathcal{H}] = 0$.

In particular, suppose that at $t = t_0$ the system is in the eigenstate $|a\rangle$ of A_S with eigenvalue a. In this case we have $\langle A_S\rangle = a$ and a is called a "good" quantum number if this expectation value is conserved in time. This invariance is satisfied if $|a\rangle$ is also an eigenstate of $\mathcal{H}$ (i.e., if $[A_S,\mathcal{H}] = 0$). If a is a good quantum number the time development of the state is given by

$$|\psi(t)\rangle = \exp[-(i/\hbar)\varepsilon_a(t-t_0)]\,|a\rangle\,, \tag{F.81}$$

where $\exp[-(i/\hbar)\varepsilon_a(t-t_0)]$ is called the *dynamical phase* of the state. This shows explicitly that eigenstates are stationary up to the development of the dynamical phase.

To illustrate these concepts we consider a quantum mechanical system at $t = 0$ in the state

$$|\psi_0\rangle = |a\rangle = \sqrt{1/2}\,(|\alpha\rangle + |\beta\rangle)\,, \tag{F.82}$$

which is itself *not* an eigenstate of $\mathcal{H}$ but a linear combination of the (normalized) eigenstates $|\alpha\rangle$ and $|\beta\rangle$ of $\mathcal{H}$ with eigenvalues ε_a and ε_b, respectively. By integrating the time-dependent Schrödinger equation we obtain for the time dependence of the state

$$|\psi(t)\rangle = e^{-(i/\hbar)\mathcal{H}t}\,|a\rangle = \sqrt{1/2}\left[e^{-(i/\hbar)\varepsilon_\alpha t}\,|\alpha\rangle + e^{-(i/\hbar)\varepsilon_\beta t}\,|\beta\rangle\right]. \tag{F.83}$$

Note that for $t = 0$ we regain Eq. (F.82). Projecting $|\psi(t)\rangle$ onto the state $|a\rangle$ we find that the occupation of $|a\rangle$ will oscillate in time between 1 and 0 at the frequency corresponding to the energy splitting $\varepsilon_a - \varepsilon_b$ between the levels,

$$|\langle a\,|\psi(t)\rangle\,|^2 = \tfrac{1}{2} + \tfrac{1}{2}\cos[(\varepsilon_\alpha - \varepsilon_\beta)\,t/\hbar]. \tag{F.84}$$

Note that the Hamiltonian is a constant of the motion,

$$\langle\psi(t)|\mathcal{H}\,|\psi(t)\rangle = \tfrac{1}{2}\left(\varepsilon_\alpha + \varepsilon_\beta\right). \tag{F.85}$$

Let us specialize to the case that the initial state $|a\rangle$ is an eigenstate of the operator A, with eigenvalue a,

$$A\,|a\rangle = a\,|a\rangle\,. \tag{F.86}$$

When is A is constant of the motion (and a a good quantum number)?

- If $[A,\mathcal{H}] = 0$ we find with the aid of Eq. (L.33b),

$$\langle\psi(t)|A\,|\psi(t)\rangle = \langle a|\,e^{(i/\hbar)\mathcal{H}t}Ae^{-(i/\hbar)\mathcal{H}t}\,|a\rangle = a. \tag{F.87}$$

Hence, in this case A is a constant of the motion (and a is a good quantum number).

- If $[A,\mathcal{H}] = -B \neq 0$ we find with the aid of Eq. (L.33b)

$$\begin{aligned}\langle\psi(t)|A\,|\psi(t)\rangle &= \langle a|\,e^{(i/\hbar)\mathcal{H}t}Ae^{-(i/\hbar)\mathcal{H}t}\,|a\rangle \\ &= a + (i/\hbar)\,\langle a|\,B\,|a\rangle\,t + \tfrac{1}{2!}(i/\hbar)^2\,\langle a|\,[\mathcal{H},B]\,|a\rangle\,t^2 + \cdots\,. \end{aligned} \tag{F.88}$$

Hence, in this case A is *not* a constant of the motion (and a *not* a good quantum number).

Problem F.5. Show by explicit calculation the equivalence of the Heisenberg equation of motion and the Schrödinger equation for the example of a particle of mass m moving in the conservative potential energy field $\mathcal{V}(\mathbf{r})$.

Solution. At $t = 0$ the particle is in the state $|\psi_H\rangle$. As the Hamiltonian does not depend on time and the position operator $\underline{\mathbf{r}}$ does not depend *explicitly* on time, the time development of the Heisenberg position operator $\mathbf{r}(t) = \mathbf{r}_H(t)$ is given by Eq. (F.80) in the form

$$i\hbar\frac{d}{dt}\mathbf{r}(t) = [\mathbf{r}(t),\mathcal{H}],\ \text{with}\ \mathcal{H} = -\frac{\underline{\mathbf{p}}^2}{2m} + \mathcal{V}(\underline{\mathbf{r}}).$$

Using $\mathbf{r}(t) = U^{\dagger}(t,t_0)\underline{\mathbf{r}}U(t,t_0)$ and recalling Eqs. (F.45) and (F.47) we have

$$i\hbar\frac{d}{dt}U^{\dagger}(t,t_0)\underline{\mathbf{r}}U(t,t_0)\,|\psi_H\rangle = U^{\dagger}(t,t_0)[\underline{\mathbf{r}},\mathcal{H}]U(t,t_0)\,|\psi_H\rangle\,.$$

Multiplying from the left with $U(t,t_0)$ and turning to Schrödinger states this becomes

$$i\hbar U(t,t_0)\frac{d}{dt}U^{\dagger}(t,t_0)\underline{\mathbf{r}}\,|\psi(t)\rangle = [\underline{\mathbf{r}},\mathcal{H}]\,|\psi(t)\rangle$$

With the aid of the differential equation (F.47) this becomes

$$-\mathcal{H}\underline{\mathbf{r}}\,|\psi(t)\rangle + i\hbar\frac{d}{dt}\underline{\mathbf{r}}\,|\psi(t)\rangle = [\underline{\mathbf{r}},\mathcal{H}]\,|\psi(t)\rangle\,.$$

Turning to the position representation this becomes the Schrödinger equation,

$$i\hbar\frac{d}{dt}\langle\mathbf{r}\,|\psi(t)\rangle = \mathcal{H}_0\langle\mathbf{r}\,|\psi(t)\rangle\,,\quad \text{with}\quad \mathcal{H}_0 = \langle\mathbf{r}|\mathcal{H}|\mathbf{r}\rangle = -\frac{\hbar^2}{2m}\Delta + \mathcal{V}(\mathbf{r}). \qquad \square$$

F.2.5 Translation as an example of a symmetry operation

In Appendix D.5.3 we introduced the *homogeneity of space* in relation to conservation laws. The evolution of a *closed* mechanical system only depends on the internal state of that system and not on the inertial frame in which the state is observed in space and time. Now let us suppose that the quantum mechanical state of the system is given by de wavefunction

$$\psi(\mathbf{r}) = \langle\mathbf{r}|\psi\rangle,$$

where $\mathbf{r}$ is the position with respect to a cartesian coordinate system $\mathcal{S}$. We now consider a *passive* translation T of the cartesian coordinate system $\mathcal{S}$ defined by the displacement vector $\mathbf{d}$. Denoting the translated coordinate system by $\mathcal{S}' = T\mathcal{S}$, the position $\mathbf{r} = (x,y,z)$ with respect to $\mathcal{S}$ changes into $\mathbf{r}' = (x',y',z')$ with respect to $\mathcal{S}'$, where

$$\mathbf{r}' = T\,\mathbf{r} = \mathbf{r} - \mathbf{d}. \tag{F.89}$$

As the translation is passive the physical state of the system is not affected by the translation, we require

$$\psi'(\mathbf{r}') \equiv \psi(\mathbf{r}), \tag{F.90}$$

or, equivalently,

$$\psi'(\mathbf{r}) = \psi(\mathbf{r} + \mathbf{d}). \tag{F.91}$$

The appearance of the new state ψ' does not come as a surprise. From general quantum mechanics we know that a new coordinate system comes with a new linear combination of eigenstates, $|\psi'\rangle$, which is related to the original linear combination of eigenstates, $|\psi\rangle$, by a *norm-conserving* basis transformation,

$$|\psi'\rangle = \underline{u}|\psi\rangle. \tag{F.92}$$

The transformation $\underline{u}$ has to be norm conserving because $|\psi\rangle$ and $|\psi'\rangle$ represent the *same* state with respect to two *different* bases,

$$\langle\psi|\psi\rangle = \langle\psi'|\psi'\rangle = \langle\psi|\underline{u}^{\dagger}\underline{u}|\psi\rangle \to \underline{u}^{\dagger}\underline{u} = \mathbb{1}. \tag{F.93}$$

This identity shows that the transformation must be unitary, $\underline{u}^{\dagger} = \underline{u}^{-1}$, which implies that the inverse transformation is given by

$$|\psi\rangle = \underline{u}^{\dagger}|\psi'\rangle. \tag{F.94}$$

So let us consider the wavefunction of a *free particle*,

$$\psi'(x) = \underline{u}(\Delta x)\psi(x)$$

$$\begin{aligned}\psi'(x) &= \psi(x+\Delta x)\\ &= \psi(x) + \Delta x\partial_x\psi(x) + \tfrac{1}{2!}\Delta x^2\partial_x^2\psi(x) + \cdots\\ &= e^{\Delta x\partial_x}\psi(x)\end{aligned}$$

F.3 Density matrix formalism

F.3.1 Density operators for pure states

The *density matrix formalism* provides an alternative for the familiar representation of quantum mechanical states by linear combinations of eigenvectors spanning a Hilbert space. The central quantity of the density matrix formalism is the *density operator.* For a quantum mechanical system in the state $|\psi(t)\rangle$ the density operator is defined as

$$\rho(t) \equiv |\psi(t)\rangle\langle\psi(t)|. \tag{F.95}$$

Note that this operator is *hermitian,* $\rho^{\dagger} = \rho$, as well as *idempotent,* $\rho^2 = \rho$. The operator ρ can be represented by a (in general *nondiagonal*) square matrix, $\rho_{ij} = \langle i|\rho|j\rangle$, which explains the use of the names *density operator* and *density matrix.* Being hermitian, ρ is free of the phase ambiguity of the state vectors; i.e., a quantum mechanical state is fully defined by its density matrix. As the state kets satisfy the Schrödinger equation (F.50) and the state bras its complex conjugate we have for a time-independent Hamiltonian

$$i\hbar\frac{\partial\rho(t)}{\partial t} = \mathcal{H}|\psi(t)\rangle\langle\psi(t)| - |\psi(t)\rangle\langle\psi(t)|\mathcal{H} = -[\rho(t),\mathcal{H}]. \tag{F.96}$$

Apparently, the density matrix is stationary, $\partial\rho(t)/\partial t = 0$, if it commutes with the Hamiltonian. This has the important consequence that *a stationary density matrix is diagonal in the representation of* $\mathcal{H}$. Furthermore, by diagonalization of a given density matrix we obtain the state of the system and its eigenvalue is unity; i.e., the state of the system is the solution of the eigenvalue equation

$$\rho|\psi\rangle = \rho_0|\psi\rangle \tag{F.97}$$

with eigenvalue $\rho_0 = 1$. In this state, the expectation value of the density operator is unity,

$$\langle\rho\rangle = \langle\psi|\rho|\psi\rangle = 1. \tag{F.98}$$

To demonstrate the calculation of properties of the system in state $|\psi\rangle$, we ask for the *expectation value* of an *arbitrary* operator A,

$$\langle A\rangle = \langle\psi|A|\psi\rangle. \tag{F.99}$$

In some *arbitrarily chosen* alternative representation $\{B\}$, this expectation value can be rewritten (using the closure relation) in the form

$$\langle A\rangle = \sum_i \langle\psi|A|b_i\rangle\langle b_i|\psi\rangle = \sum_i \langle b_i|\psi\rangle\langle\psi|A|b_i\rangle. \tag{F.100}$$

Recognizing the density operator $\rho = |\psi\rangle\langle\psi|$ we find that the expectation value of A for the system in the state $|\psi\rangle$ is given by

$$\langle A\rangle = \sum_i \langle b_i|\rho A|b_i\rangle = \operatorname{tr}\rho A. \tag{F.101}$$

Note that $\operatorname{tr}\rho A = \operatorname{tr}A\rho$. As the representation $\{B\}$ was chosen arbitrarily, the trace is invariant under unitary transformation (a property of square matrices - see Appendix M.2.1). This is expressed by the representation-free form of Eq. (F.101).

To compare the two formalisms a bit further, we reconsider the Parseval relation for the state $|\psi\rangle$ in the representation $\{B\}$,

$$1 = \sum_i |\langle b_i|\psi\rangle|^2 = \sum_i \mathcal{P}_i, \tag{F.102}$$

where

$$\mathbb{1} = \sum_i |b_i\rangle\langle b_i| = \sum_i \rho_i \tag{F.103}$$

is the closure relation in the representation $\{B\}$ and $\mathcal{P}_i$ represents the probability density to observe the system in the eigenstate $|b_i\rangle$. In the *state formalism* we write

$$\mathcal{P}_i = \langle\psi|b_i\rangle\langle b_i|\psi\rangle = \langle\psi|\rho_i|\psi\rangle, \tag{F.104}$$

which is interpreted as the expectation value of the projection operator $\rho_i = |b_i\rangle\langle b_i|$. In the *density matrix formalism*, $\mathcal{P}_i$ is interpreted as the *expectation value of the density operator*,[1]

$$\mathcal{P}_i = \langle b_i|\psi\rangle\langle\psi|b_i\rangle = \langle b_i|\rho|b_i\rangle = \rho_{ii}. \tag{F.105}$$

The representation $\{B\}$ being arbitrary, the Parseval relation takes the representation-free form

$$\operatorname{tr}\rho = 1. \tag{F.106}$$

F.3.2 Density operators for mixed ensembles of states

The importance of the density matrix formalism finds its origin in the probabilistic nature of quantum mechanical measurement. A *single* state-selective measurement (state *filtration*) on a quantum mechanical system is of little value. It tells us whether or not the selected state was observed but nothing more. Only by repeating the measurement many times we can determine the probability of observing that state. If the measurements are done on an ensemble of identically prepared systems, say a beam of particles all prepared in the state $|\psi\rangle$, the probability density of observing this state, $\mathcal{P}_s = |\langle s|\psi\rangle|^2$, is obtained by averaging over the set of measurements. This procedure is called the *ensemble-averaging*. As was shown at the end of the previous section, formally $\mathcal{P}_s$ can be expressed as the *expectation value of the density operator* for the system in state $|s\rangle$, $\mathcal{P}_s = \langle s|\rho|s\rangle$. Such an ensemble of identically prepared systems is called a *pure ensemble*, to distinguish it from *mixed*

---

[1] Notice the difference between the density operator ρ_i and the diagonal matrix element of the density matrix, ρ_{ii}. In our notation $\mathcal{P}_i = \langle\psi|\rho_i|\psi\rangle = \langle\psi|b_i\rangle\langle b_i|\psi\rangle = \langle b_i|\psi\rangle\langle\psi|b_i\rangle = \langle b_i|\rho|b_i\rangle = \rho_{ii}$.

ensembles, in which the state preparation may vary or is even absent.[1] It may speak for itself that whatever the ensemble we can always determine ensemble averages. Therefore, to refer to an arbitrary type of ensemble averaging we use a different notation for the density matrix, $\rho \to \varrho$, denoting the ensemble-averaged probability to detect the state $|s\rangle$ by $\langle s|\varrho|s\rangle$. If there is no state preparation whatsoever one speaks of a *random ensemble* and $\langle s|\varrho|s\rangle = 1/n$, where n is the order of the density matrix. The value of n cannot exceed the number of measurements but it can be smaller. For instance, measuring spin "up" or "down" on a beam of particles we can do many measurements but $n = 2$.

The relative probability of occurrence of one of the states of a mixed ensemble $\{|s\rangle\}$ is given by the statistical weight w_s, a number between zero and one $(0 \leq w_s \leq 1)$, and subject to the normalization condition

$$\sum_s w_s = 1. \tag{F.107}$$

Here, the probability is defined for a discrete ensemble but can be generalized to the continuous case. Note that orthogonality is neither mentioned nor required for the states of the ensemble $\{|s\rangle\}$.

Let us analyze ensemble averaging in more detail. Knowing the probabilities $\{w_s\}$, we can calculate the *quantum statistical average* for an arbitrary operator A,

$$\bar{A} \equiv \langle\!\langle A \rangle\!\rangle = \sum_s w_s \langle s|A|s\rangle. \tag{F.108}$$

Note that this is the *ensemble average* of the expectation values $\langle A\rangle_s = \langle s|A|s\rangle = A_{ss}$. Choosing an arbitrary representation $\{B\}$, we can use its closure relation to write the quantum-statistical average in the form

$$\bar{A} = \sum_s w_s \sum_i \langle s|A|b_i\rangle\langle b_i|s\rangle = \sum_i \sum_s w_s \langle b_i|s\rangle\langle s|A|b_i\rangle = \sum_i \langle b_i|\varrho A|b_i\rangle, \tag{F.109}$$

where

$$\varrho \equiv \sum_s w_s |s\rangle\langle s| = \sum_s w_s \rho_s \tag{F.110}$$

defines the *statistical operator* of the ensemble defined by the weights $\{w_s\}$, which represents the (by construction *diagonal*) density matrix of a *statistical mixture of states.* Importantly, as for a time-independent mixture the weights $\{w_s\}$ are conserved, it follows in the same way as demonstrated for Eq. F.96 that the time dependence of its density matrix is given by

$$i\hbar \frac{\partial \varrho(t)}{\partial t} = -[\varrho(t), \mathcal{H}]. \tag{F.111}$$

Thus we established that (also) for *mixed* ensembles a *stationary* density matrix is *diagonal* in the representation of $\mathcal{H}$.

Using the statistical operator the quantum statistical average can be written in the representation-free form

$$\bar{A} \equiv \langle\!\langle A \rangle\!\rangle = \mathrm{tr}\, \varrho A. \tag{F.112}$$

Substituting $A \to \mathbb{1}$ we find that for mixed ensembles the Parseval relation (F.106) is replaced by

$$\mathrm{tr}\, \varrho = 1. \tag{F.113}$$

---

[1] We distinguish between a *mixed ensemble*, which consists of a statistical mixture of states, and a *mixed state*, which is a coherent superposition of states. Beware that some authors use mixed state as an abbreviation for "statistical mixture of states".

Substituting $A \to \rho_s \equiv |s\rangle\langle s|$ we obtain the probability to observe the (pure) state $|s\rangle$ in a measurement on the ensemble,

$$\bar{\rho}_s = \operatorname{tr} \varrho\rho_s = \sum_i \langle b_i|\varrho|s\rangle\langle s|b_i\rangle = \langle s|\varrho|s\rangle = \varrho_{ss} = w_s. \tag{F.114}$$

Thus, using the density matrix for mixed ensembles of states, we can calculate quantum statistical averages using the same methods as used for the density matrix of pure states - see Eq. (F.101).[1] The density matrix of a pure state is in general nondiagonal, whereas that of mixed states is by construction diagonal. Furthermore, an important difference is revealed when calculating $\operatorname{tr} \varrho^2$. For *pure* states the density matrix is idempotent, $\rho^2 = \rho$, and we have

$$\operatorname{tr} \rho^2 = \operatorname{tr} \rho = 1. \tag{F.115}$$

For *mixed* states this is replaced by the inequalities.

$$0 \leq \operatorname{tr} \varrho^2 \leq (\operatorname{tr} \varrho)^2 = 1. \tag{F.116}$$

F.3.3 Density matrices for quantum-statistical mixtures

Thus far we emphasized ensemble averaging for a real set of measurements; e.g., of the spin states of particles in a beam. Such ensembles of experimental data are called *real* ensembles. Interestingly, density matrices based on *fictitious* ensembles are at least as important. In many-body systems it is simply impossible to measure all particles individually but, as is well known from statistical physics, often this is not necessary. For instance, to calculate the pressure of an ideal gas, all we need to know is the probability of occupation of the single-particle states and this information can be provided by a *fictitious* ensemble such as the canonical distribution over the energy eigenstates of a physical system.

To start the discussion of quantum statistical ensembles we define the entropy of a physical system in terms of its density matrix

$$S/k_B = -\operatorname{tr} \varrho \ln\varrho. \tag{F.117}$$

To explore the plausibility of this definition as a measure for the disorder in the system we choose the representation of $\mathcal{H}$ (in which ϱ is diagonal - see Section F.3.2). Denoting the basis of this representation by $\{|s\rangle\}$ the expression for the entropy can be written in the form

$$S/k_B = -\sum_{s,s'} \langle s|\varrho|s'\rangle\langle s'|\ln\varrho|s\rangle = -\sum_s \langle s|\varrho|s\rangle \ln\langle s|\varrho|s\rangle = -\sum_s \varrho_{ss} \ln \varrho_{ss}. \tag{F.118}$$

From classical statistical mechanics we recall that zero entropy corresponds to the system being in a single microstate. In quantum statistics this translates to the system being in a pure state. Denoting this pure state by $|\psi\rangle$, its density operator is given by $\varrho = \rho = |\psi\rangle\langle\psi|$ and

$$\varrho_{ss} = \langle s|\varrho|s\rangle = |\langle s|\psi\rangle|^2. \tag{F.119}$$

To avoid the complications of entanglement we restrict ourselves to a single eigenstate of the representation $\{|s\rangle\}$, $|\psi\rangle \to |t\rangle \in \{|s\rangle\}$. In this case we have $\varrho_{ss} = \delta_{s,t}$ and calculate zero entropy, $S = 0$, because either $\varrho_{ss} = 0$ or $\ln \varrho_{ss} = 0$. On the other hand, for a completely random ensemble (where all energy eigenstates have the same probability), the density matrix is given by $\varrho = 1/n$, where n is the size of the ensemble (the order of the density matrix),

$$S/k_B = -\operatorname{tr} \frac{1}{n} \ln\frac{1}{n} = \ln n. \tag{F.120}$$

---

[1] Many authors do not distinguish between $\langle\langle A\rangle\rangle$ and $\langle A\rangle$ as all relevant information is contained in the form of ϱ.

For example, the maximum entropy of a two-level system is $\ln 2$.

In general a statistical mixture will not have maximum entropy. We know from the *statistical principle* for ergodic systems that only states of *equal energy* have to be equally probable. This limits the entropy because, at finite temperature, high-energy states will be less probable. In statistical physics the entropy is determined by a maximization procedure in which one or more constraints are enforced. In a canonical ensemble we search for the maximum entropy under the constraints of given total total energy $E = \langle \mathcal{H} \rangle = \operatorname{tr} \varrho \mathcal{H}$ and unit normalization, $\mathcal{N} = \operatorname{tr} \varrho = 1$. To describe the procedure, we turn to the representation of $\mathcal{H}$, denoted by $\{|s\rangle\} = \{|1\rangle, \cdots, |n\rangle\}$, because in this representation the density matrix is diagonal. For the sake of the argument we restrict ourselves to discrete systems with a finite-dimensional Hilbert space. The lagrangian for the variation of the diagonal elements of the density matrix $\{\varrho_{ss}\}$, and the multipliers β (for total energy) and γ (for normalization) is stationary if the following conditions are satisfied - see Problem F.6

$$\delta L = \sum_{s=1}^{n} \delta \varrho_{ss} \left[-(\ln \varrho_{ss} + 1) - \beta E_s - \gamma \right] = 0, \tag{F.121a}$$

$$\delta E = \sum_{s=1}^{n} \varrho_{ss} E_s - E = 0, \tag{F.121b}$$

$$\delta \mathcal{N} = \sum_{s=1}^{n} \varrho_{ss} - 1 = 0, \tag{F.121c}$$

where $E_s = \langle s|\mathcal{H}|s\rangle$ is the energy of the eigenstate $|s\rangle$. To assure that Eq. (F.121a) be valid for *arbitrary* variations of $\{\delta \varrho_{ss}\}$, we require

$$[(\ln \varrho_{ss} + 1) + \beta E_s + \gamma] = 0. \tag{F.122}$$

Renaming the Lagrange multiplier $\gamma \to -(1 + \beta F)$ we obtain

$$\ln \varrho_{ss} = -\beta (E_s - F), \tag{F.123}$$

which can be written in the form

$$\varrho_{ss} = e^{-\beta(E_s - F)} = \frac{1}{Z} e^{-\beta E_s}. \tag{F.124}$$

Hence, the corresponding density matrix can be defined in the form of a statistical operator,

$$\varrho = \frac{1}{Z} e^{-\beta \mathcal{H}}. \tag{F.125}$$

Imposing the normalization condition (F.121c) we find

$$Z = e^{-\beta F} = \sum_{s=1}^{n} e^{-\beta E_s}. \tag{F.126}$$

By identifying $\beta \equiv 1/k_B T$, we recognize the well-known expression for the canonical partition sum. Note that in the high-temperature limit ($\beta \to 0$) we obtain $Z = n$ and $\varrho_{ss} = 1/n$, which shows that the limiting value for the entropy is that of a random ensemble, $S = \ln n$.

For finite temperature we find for the entropy

$$S/k_B = -\sum_{s=1}^{n} \varrho_{ss} \ln \varrho_{ss} = \beta \sum_{s=1}^{n} \varrho_{ss} (E_s - F) = \beta (E - F), \tag{F.127}$$

where we used Eq. (F.121b). Identifying F with the Helmholtz free energy and E with the internal energy U we obtain the thermodynamic relation

$$F = U - TS. \tag{F.128}$$

Problem F.6. Show that the deviation of the lagrangian of the canonical ensemble of size n vanishes under variation of the $\{\varrho_{ss}\}$, with $s \in \{1, \cdots n\}$ if

$$\delta L = \sum_{s=1}^{n} \delta\varrho_{ss}\, \left[-(\ln \varrho_{ss} + 1) - \beta E_s - \gamma\right] = 0.$$

Solution. The lagrangian for the variation of the $\{\varrho_{ss}\}$ of the canonical ensemble under the constraints of fixed total total energy $E = \langle \mathcal{H} \rangle = \operatorname{tr} \varrho\mathcal{H}$ and normalization, $\operatorname{tr} \varrho = 1$, is given by

$$L(\{\varrho_{ss}\}_n, \beta, \gamma) = \sum_{s=1}^{n} \left[-\varrho_{ss} \ln \varrho_{ss} + \beta(E - \varrho_{ss}E_s) + \gamma(1 - \varrho_{ss})\right].$$

Hence, L is stationary under partial variation of ϱ_{ss} provided

$$\frac{\delta L}{\delta \varrho_{ss}} = -(\ln \varrho_{ss} + 1) - \beta E_s - \gamma = 0$$

and also stationary under partial variation of β and γ provided

$$\frac{\delta L}{\delta \beta} = E - \sum_{s=1}^{n} \varrho_{ss}E_s = 0 \quad \text{and} \quad \frac{\delta L}{\delta \gamma} = 1 - \sum_{s=1}^{n} \varrho_{ss} = 0.$$

Hence, L is stationary under simultaneous variation of of $\{\varrho_{ss}\}_n$, β and γ if

$$\delta L = \sum_{s=1}^{n} \left[-(\ln \varrho_{ss} + 1) - \beta E_s - \gamma\right] \delta\varrho_{ss} = 0, \quad \sum_{s=1}^{n} \varrho_{ss}E_s = E \quad \text{and} \quad \sum_{s=1}^{n} \varrho_{ss} = 1. \qquad \square$$

F.4 Conservation of normalization and current density

The rate of change of normalization of a wave function can be written as a continuity equation

$$\frac{\partial}{\partial t} |\Psi(r,t)|^2 + \boldsymbol{\nabla} \cdot \mathbf{j} = 0, \tag{F.129}$$

which defines $\mathbf{j}$ as the probability current density of the wave function. With the time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi(r,t) = H\, \Psi(r,t) \tag{F.130a}$$

$$-i\hbar \frac{\partial}{\partial t} \Psi^*(r,t) = H\, \Psi^*(r,t) \tag{F.130b}$$

we find

$$\begin{aligned}\frac{\partial}{\partial t} |\Psi(r,t)|^2 &= \Psi^*(r,t) \frac{\partial}{\partial t} \Psi(r,t) + \Psi(r,t) \frac{\partial}{\partial t} \Psi^*(r,t) \\ &= \frac{1}{i\hbar} \left[\Psi^*(H\Psi) - \Psi(H\Psi^*)\right]. \end{aligned} \tag{F.131}$$

Hence,

$$\boldsymbol{\nabla} \cdot \mathbf{j} = \frac{i}{\hbar} \left[\Psi^*(H\Psi) - \Psi(H\Psi^*)\right]. \tag{F.132}$$

Hence, together with the continuity equation this equation shows that the normalization of a stationary state is conserved if the Hamiltonian is hermitian.

For a Hamiltonian of the type

$$H = -\frac{\hbar^2}{2m}\Delta + V(r) \tag{F.133}$$

the Eq. (F.132) takes the form

$$\boldsymbol{\nabla}\cdot\mathbf{j} = -\frac{i\hbar}{2m}\left[\psi^*(\boldsymbol{\nabla}^2\psi) - (\boldsymbol{\nabla}^2\psi^*)\psi\right] = -\frac{i\hbar}{2m}\boldsymbol{\nabla}\cdot\left[\psi^*(\boldsymbol{\nabla}\psi) - (\boldsymbol{\nabla}\psi^*)\psi\right]. \tag{F.134}$$

Hence, the continuity equation is satisfied by defining the probability current density as

$$\mathbf{j} = -\frac{i\hbar}{2m}\left(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*\right). \tag{F.135}$$

The probability current density is a real quantity,

$$\mathbf{j} = \mathrm{Re}\left[\frac{-i\hbar}{m}\psi^*\boldsymbol{\nabla}\psi\right] = \mathrm{Re}\left[\psi^*\mathbf{v}\psi\right], \tag{F.136}$$

where $\mathbf{v} = \mathbf{p}/m = (-i\hbar/m)\,\boldsymbol{\nabla}$ is the *velocity operator*, in which we recognize the well-known *correspondence rule* $\mathbf{p} \to -\,i\hbar\boldsymbol{\nabla}$.

Writing ψ as the product of a *position-independent* amplitude $|\psi|$ and a *position-dependent* phase factor, $\psi = |\psi|e^{i\phi}$, we have

$$\mathbf{j} = \frac{i\hbar}{2m}\left(\psi\boldsymbol{\nabla}\psi^* - \psi^*\boldsymbol{\nabla}\psi\right) = \frac{i\hbar}{2m}|\psi|^2\left(e^{i\phi}\boldsymbol{\nabla}e^{-i\phi} - e^{-i\phi}\boldsymbol{\nabla}e^{i\phi}\right) = |\psi|^2\,(\hbar/m)\,\boldsymbol{\nabla}\phi. \tag{F.137}$$

Keeping in mind the flux relation $\mathbf{j} = n\mathbf{v}$, we identify $n = |\psi|^2$ as the *probability density* and

$$\mathbf{v} = (\hbar/m)\,\boldsymbol{\nabla}\phi \tag{F.138}$$

as the *probability-flow velocity.*

G

Two-level systems

G.1 Introduction

In this appendix we review the unique properties of *two-level systems.* A two-level system is a discrete quantum mechanical system with exactly two energy levels. Along with the harmonic oscillator and the Bohr atom, two-level systems belong to the select group of quantum mechanical systems for which exact analytical solutions can be obtained. The mathematical form of the solutions to be obtained in the present appendix is independent of the physical origin of the two levels. This reveals the universal properties of two-level systems. The iconic example is the electron spin in an applied magnetic field. The two-level equations are widely used in physics, be it that usually an approximation is required before they can be applied.

G.2 Two-level systems - general properties

Let us consider a closed quantum mechanical system with exactly two energy levels and described by the *time-independent* Hamiltonian $\mathcal{H}_0$. The Schrödinger equation for the eigenstates of this system,

$$\mathcal{H}_0|\phi_n\rangle = E_n^0|\phi_n\rangle, \tag{G.1}$$

defines a two-dimensional Hilbert space spanned by a basis of two state vectors, ϕ_1 and ϕ_2, corresponding to the eigenvalues E_1^0 and E_2^0, respectively. Degeneracy ($E_1^0 = E_2^0$) is not excluded. The eigenvectors can be chosen to be normalized and orthogonal and we assume this to be the case,

$$\langle\phi_n|\phi_m\rangle = \delta_{nm}, \quad \text{with} \;\; n, m \in \{1, 2\}. \tag{G.2}$$

What happens to the time evolution when we apply a perturbing field? To describe this situation we append the Hamiltonian $\mathcal{H}_0$ with a hermitian term, $\mathcal{H}_1$, representing the perturbation,

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1. \tag{G.3}$$

In this context $\mathcal{H}$ is called the *full* Hamiltonian, $\mathcal{H}_0$ the *unperturbed* part and $\mathcal{H}_1$ the *perturbation.* Unlike the practice in perturbation theory the perturbation may be *strong* (i.e. dominate over $\mathcal{H}_0$) without compromising the validity of the equations. Furthermore, whereas $\mathcal{H}_0$ remains *constant* in time, the perturbation may (and often will) have an *explicit time dependence*, $\mathcal{H}_1 \to \mathcal{H}_1(t)$.

To understand how the perturbation affects the two-level system we solve the *time-dependent* Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \mathcal{H}(t)|\psi(t)\rangle \tag{G.4}$$

under the initial condition $|\psi(0)\rangle = |\psi_i\rangle$, where ψ_i is the state of the system at time $t = 0$. If $\mathcal{H}_1$ depends on time there will be energy exchange between the system and the perturbing field. There are methods to deal with this situation but these fall outside the scope of the present appendix. Leaving aside their energy, the states of the system can always be decomposed with respect to the unperturbed basis,

$$|\psi(t)\rangle = \sum_{n=1,2} |\phi_n\rangle \langle\phi_n|\psi(t)\rangle = a_1(t)|\phi_1\rangle + a_2(t)|\phi_2\rangle, \tag{G.5}$$

where the amplitudes $a_n(t) \equiv \langle\phi_n|\psi(t)\rangle$ are complex functions (depending *only on time*) and subject to the *Parseval relation*

$$|a_1(t)|^2 + |a_2(t)|^2 = 1. \tag{G.6}$$

The probability to observe the system at time $t > 0$ in the state ψ_f is given by the expression

$$W_{i\to f}(t) = |\langle\psi_f|\psi(t)\rangle|^2. \tag{G.7}$$

This is called the *transition probability* for the transition $\psi_i \to \psi_f$.

G.2.1 Solving the Schrödinger equation

To solve the Schrödinger equation we project both sides of Eq. (G.4) onto the eigenstates ϕ_n,

$$i\hbar\frac{d}{dt}a_n(t) = \langle\phi_n|\mathcal{H}_0 + \mathcal{H}_1(t)|\psi(t)\rangle, \tag{G.8}$$

with $n \in \{1,2\}$. Substituting Eq. (G.5) into this expression and using the orthonormality relations (G.2) as well as the hermiticity of $\mathcal{H}_0$, we find that the amplitudes $a_1(t)$ and $a_2(t)$ satisfy a set of two coupled linear differential equations

$$i\hbar\frac{d}{dt}a_n(t) = E_n^0 a_n(t) + \sum_{m=1,2} \langle\phi_n|\mathcal{H}_1(t)|\phi_m\rangle a_m(t), \tag{G.9}$$

with $n \in \{1,2\}$. Let us suppose for convenience that at $t = 0$ the system is in the eigenstate ϕ_1 (i.e., $\psi_i \to \phi_1$). In this case the initial condition for the amplitudes is

$$a_n(0) = \langle\phi_n|\psi(0)\rangle = \langle\phi_n|\phi_1\rangle = \delta_{n1}. \tag{G.10}$$

Before continuing we introduce a compact notation and derive some properties for the matrix elements:

$$H_{nm}(t) \equiv \langle\phi_n|\mathcal{H}(t)|\phi_m\rangle \quad \text{and} \quad H'_{nm}(t) \equiv \langle\phi_n|\mathcal{H}_1(t)|\phi_m\rangle, \tag{G.11}$$

with $n, m \in \{1,2\}$. In view of the orthogonality of ϕ_1 and ϕ_2 we have

$$H_{nm}(t) = \begin{cases} H'_{nm}(t) & \text{for } n \neq m \\ E_n^0 + H'_{nn}(t) & \text{for } n = m. \end{cases} \tag{G.12}$$

Given the hermiticity of $\mathcal{H}_1$ we find[1]

$$H'_{nm} = \langle\phi_n|\mathcal{H}_1|\phi_m\rangle = \langle\phi_m|\mathcal{H}_1|\phi_n\rangle^* = H'^*_{mn}. \tag{G.13}$$

In particular, for the diagonal matrix elements this becomes

$$H'_{nn} = H'^*_{nn}, \tag{G.14}$$

showing that the diagonal elements are all *real.* In this notation the coupled equations (G.9) take the form

$$i\hbar\dot{a}_1 = H_{11}a_1 + H'_{12}a_2 \tag{G.15a}$$

$$i\hbar\dot{a}_2 = H'^*_{12}a_1 + H_{22}a_2. \tag{G.15b}$$

---

[1] From here on we mostly suppress showing the time variable in the notation.

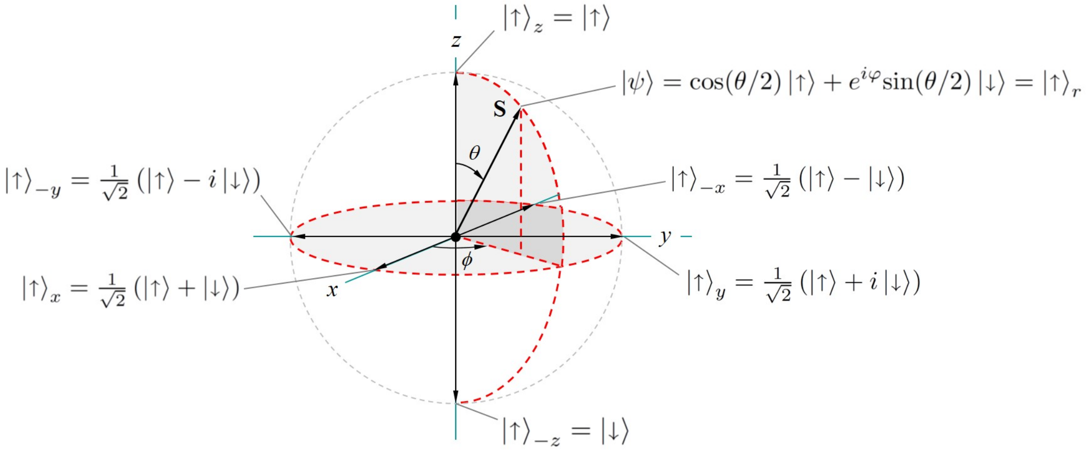


Figure G.1: A two-level system can be represented by a pseudospin **S** in 3D real space. The states of the system are defined by the angles θ and ϕ corresponding to a point on the surface of the Bloch sphere. Note that by changing the coupling angle (θ) over its range from 0 to π the mixing angle ($\alpha = \theta/2$) changes from 0 to $\pi/2$. Note further that the relative phase of the two components depends on the definition of the coordinate system (in the plot the two components are in phase in the x direction).

G.2.2 Pseudospin and Bloch sphere - formal correspondence with spin-1/2

The unique position of two-level systems among the discrete quantum systems finds its origin in the simplicity of the Parseval relation (G.6). As only two levels are involved, the occupation of one of the levels, $|a_1|^2$, fully determines that of the other, $|a_2|^2 = 1 - |a_1|^2$. This imposes a strong condition on the relation between the amplitudes a_1 and a_2. It is straightforward to show that the Parseval relation is satisfied for amplitudes of the following form:

$$a_1 = e^{-i\phi/2}\cos\alpha \quad \text{and} \quad a_2 = e^{i\phi/2}\sin\alpha. \tag{G.16}$$

In this form the quantum numbers $n \in \{1,2\}$ are replaced by two new quantities, the *mixing angle* α, with $0 \le \alpha < \pi/2$,[1] and the *relative phase* ϕ, with $0 \le \phi < 2\pi$ Actually, it is advantageous from a geometric point of view to write the mixing angle in the form

$$\alpha = \theta/2, \tag{G.17}$$

where θ is called the *coupling angle*. In Sections G.3.2 and G.3.3 we discuss why. Since θ is twice as large as α it is defined over the range $0 \le \theta < \pi$. In this notation the *states* of any two-level system are represented by two angles, θ and ϕ, which define the *polar* and *azimuthal* angle of a three-dimensional unit vector, $\hat{\mathbf{r}}$, in an *abstract real* space,

$$r_x = a_1 a_2^* + a_1^* a_2 = \sin\theta\cos\phi \tag{G.18a}$$

$$r_y = i(a_1 a_2^* - a_1^* a_2) = \sin\theta\sin\phi \tag{G.18b}$$

$$r_z = a_1 a_1^* - a_2 a_2^* = \cos\theta. \tag{G.18c}$$

Keeping in mind that a spin-$\frac{1}{2}$ system is a two-level system, Eqs. (G.18) suggest a relation between the unit vector $\hat{\mathbf{r}}$ and the direction of the magnetic moment of a spin-$\frac{1}{2}$ particle. For this reason $\hat{\mathbf{r}}$ is called the *pseudo* or *fictitious spin* of the two-level system. A first indication for the correctness of this conjecture is found by writing the decomposition for an *arbitrary* state in the form

$$|\psi\rangle = \cos(\theta/2)|\phi_1\rangle + e^{i\phi}\sin(\theta/2)|\phi_2\rangle. \tag{G.19}$$

---

[1] Any interval of $\pi/2$ could be chosen; in Section G.3.3 preference is given to $-\pi/4 \le \alpha \le \pi/4$.

In this form a_1 is chosen to be *real*, $a_1 \to \cos(\theta/2)$. This is allowed without loss of generality because we are free to choose the global phase. As is illustrated in Fig. G.1, the decomposition (G.19) relates the 2D Hilbert space of the two-level system to the surface of a sphere - the *Bloch sphere* - introduced by Felix Bloch to describe the evolution of a nuclear moment in a electromagnetic field [16]. In 1957 Richard Feynman, Frank Vernon and Robert Hellwarth showed that the evolution of *any* two-level system in a perturbing field can be mapped onto the equation of motion of a spin-$\frac{1}{2}$ system subject to a perturbation [42]. In this map the Schrödinger equation is written in the form of a real three-dimensional vector equation (see Problem G.1),

$$\frac{d\hat{\mathbf{r}}}{dt} = (\omega \times \hat{\mathbf{r}}), \tag{G.20}$$

where the components of the pseudospin $\hat{\mathbf{r}}$ uniquely refer to the state of the system and the components of the vector ω represent the perturbation,

$$\omega_x = (H'_{12} + H'^{*}_{12})/\hbar, \quad \omega_y = i(H'_{12} - H'^{*}_{12})/\hbar, \quad \omega_z = (H_{11} - H_{22})/\hbar. \tag{G.21}$$

In particular, for the magnetic interaction of a real spin-$\frac{1}{2}$ particle (e.g., electron spin) the 3D real space reduces to the physical space of observation. In optics the sphere is called the *Poincaré sphere* and the vector corresponds to an arbitrary state of polarized light (elliptical polarization). With Eqs. (G.20) and (G.21) we have obtained a universal relation for the time evolution of two-level systems.

Problem G.1. Derive the Feynman-Vernon-Hellwarth transformation by showing that the time evolution of two-level systems may be formulated in the universal form of a real three-dimensional vector equation for the pseudospin $\hat{\mathbf{r}}$,

$$\frac{d\hat{\mathbf{r}}}{dt} = (\omega \times \hat{\mathbf{r}}),$$

where the components of the $\hat{\mathbf{r}}$ uniquely refer to the state of the system and the components of the vector ω represent the perturbation.

Solution. We start by differentiating Eqs. (G.18) with respect to time,

$$\begin{aligned} \dot{r}_x &= (\dot{a}_1 a_2^* + a_1 \dot{a}_2^* + c.c.) \\ -i\dot{r}_y &= (\dot{a}_1 a_2^* + a_1 \dot{a}_2^* - c.c.) \\ \dot{r}_z &= (\dot{a}_1 a_1^* - \dot{a}_2 a_2^* + c.c). \end{aligned}$$

Each term in these equations follows from the Schrödinger equation using Eq. (G.15),

$$\begin{aligned} \dot{a}_1 a_2^* &= -i(H_{11} a_1 a_2^* + H'_{12} a_2 a_2^*)/\hbar \\ a_1 \dot{a}_2^* &= i(H'_{12} a_1 a_1^* + H_{22} a_1 a_2^*)/\hbar \\ \dot{a}_1 a_1^* &= -i(H_{11} a_1 a_1^* + H'_{12} a_1^* a_2)/\hbar \\ \dot{a}_2 a_2^* &= -i(H'^{*}_{12} a_1 a_2^* + H_{22} a_2 a_2^*)/\hbar. \end{aligned}$$

Then, substituting these terms and using the Parseval relation (G.6) we obtain after some rearrangement

$$\begin{aligned} \hbar\dot{r}_x &= -i(H_{11} - H_{22})(a_1 a_2^* - a_1^* a_2) + i(H'_{12} - H'^{*}_{12})(a_1 a_1^* - a_2 a_2^*) \\ -i\hbar\dot{r}_y &= -i(H_{11} - H_{22})(a_1 a_2^* + a_1^* a_2) + i(H'_{12} + H'^{*}_{12})(a_1 a_1^* - a_2 a_2^*) \\ \hbar\dot{r}_z &= +i(H'_{12} + H'^{*}_{12})(a_1 a_2^* - a_1^* a_2) - i(H'_{12} - H'^{*}_{12})(a_1 a_2^* + a_1^* a_2). \end{aligned}$$

Here we recognize the appearance of the components of $\hat{\mathbf{r}}$ as given in Eqs. (G.18) and write

$$\begin{aligned} i\hbar\dot{r}_x &= -i(H_{11} - H_{22})r_y - (H'_{12} - H'^{*}_{12})r_z \\ i\hbar\dot{r}_y &= i(H_{11} - H_{22})r_x - i(H'_{12} + H'^{*}_{12})r_z \\ i\hbar\dot{r}_z &= i(H'_{12} + H'^{*}_{12})r_y + (H'_{12} - H'^{*}_{12})r_x. \end{aligned}$$

In vector notation this becomes

$$\frac{d}{dt}\hat{\mathbf{r}} = -\begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ r_x & r_y & r_z \\ (H'_{12}+H'^{*}_{12})/\hbar & i(H'_{12}-H'^{*}_{12})/\hbar & (H_{11}-H_{22})/\hbar \end{vmatrix}.$$

This expression can be rewritten in the form of the vector equation (G.20), where the components of the vector $\boldsymbol{\omega}$ are defined in terms of the matrix elements of the Hamiltonian as given in Eq. (G.21). □

G.3 Two-level systems - static perturbations

Let us return to the coupled equations (G.15). These equations govern the evolution of two-level systems for a (generally time-dependent) Hamiltonian. They followed from the Schrödinger equation and were transformed to the vector equation (G.20) but *they were not solved.* The reason is simple: analytic solution of the differential equations is not possible for an *arbitrary* time dependence of the perturbation. This being said, analytic solutions can be given for special cases. In the present section we specialize to *static perturbations* (i.e., perturbations without an explicit time dependence). Time-dependent perturbations are not discussed in this Appendix.

Before proceeding we note that the coupled equations (G.15) can be written in the form of a matrix equation as is common practice in linear algebra,

$$i\hbar\frac{d}{dt}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} H_{11} & H'_{12} \\ H'^{*}_{12} & H_{22} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}. \tag{G.22}$$

Not surprisingly, this equation has the form of the time-dependent Schrödinger equation,

$$i\hbar\frac{d}{dt}\psi = \mathcal{H}\psi, \tag{G.23}$$

where, in matrix notation, the Hamiltonian and the state are given by

$$\mathcal{H} = \begin{pmatrix} H_{11} & H'_{12} \\ H'^{*}_{12} & H_{22} \end{pmatrix} \quad \text{and} \quad \psi = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}. \tag{G.24}$$

Specializing to the case of *static* perturbations turns $\mathcal{H}$ into an operator without an explicit time dependence. This makes it possible to integrate Eq. (G.23) by *separation of variables*,

$$\frac{1}{\psi}\frac{d\psi}{dt} = -\frac{i}{\hbar}\mathcal{H}. \tag{G.25}$$

Under the *initial condition* $\psi(0) = \psi_i$ we obtain

$$\psi(t) = e^{-(i/\hbar)\mathcal{H}t}\psi_i. \tag{G.26}$$

Note the appearance of the evolution operator. If ψ_i is an *eigenstate* of $\mathcal{H}$ (corresponding to the eigenvalue E_i) it satisfies the Schrödinger equation

$$\mathcal{H}\psi_i = E_i\psi_i, \tag{G.27}$$

and the evolution becomes stationary; i.e., there is only evolution of the *global* phase,

$$|\psi(t)\rangle = e^{-(i/\hbar)E_i t}|\psi_i\rangle. \tag{G.28}$$

The exponential prefactor is called the *dynamical phase* of the state.

In cases where ψ_i is *not* an eigenstate of $\mathcal{H}$ the time evolution is not stationary. Non-stationary evolution under *static* perturbations is the subject of Section G.4

G.3.1 Level shifts and mixing of the eigenstates

As is well known from linear algebra Eq. (G.27) represents an eigenvalue problem. Non-singular solutions exist if and only if the *characteristic determinant*, traditionally referred to as the *secular determinant*,[1] is zero,

$$\begin{vmatrix} H_{11}-E & H'_{12} \\ H'^{*}_{12} & H_{22}-E \end{vmatrix} = 0. \tag{G.29}$$

Eq. (G.29) is known as the *characteristic equation* (*secular equation*) of the problem. By solving this equation we find the eigenvalues of $\mathcal{H}$. As is evident from Eq. (G.29), for two-level systems the characteristic equation takes the form of a quadratic equation in E,

$$(H_{11}-E)(H_{22}-E) = |H'_{12}|^2 . \tag{G.30}$$

The eigenvalues are given by the roots of this equation.

- In the *absence of mixing* of the eigenstates ($H'_{12}=0$) the Hamiltonian is *diagonal* and we easily find $E_+ = H_{11}$ and $E_- = H_{22}$. Furthermore, we are free (without loss of generality) to choose the state labeling such that
$$H_{11} \geq H_{22}. \tag{G.31}$$
In this convention E_+ and E_- correspond to the upper and lower level, $E_+ \geq E_-$.

- Also in the *presence of mixing* ($H'_{12} \neq 0$) we have $E_+ > E_-$ under the convention (G.31) because the *off-diagonal* matrix elements always result in *level repulsion* (see Section G.3.2). The *general solution* of the quadratic equation is given by the "square root formula",
$$E_{\pm} = \tfrac{1}{2}(H_{11}+H_{22}) \pm \tfrac{1}{2}\sqrt{(H_{11}-H_{22})^2 + 4|H'_{12}|^2}. \tag{G.32}$$
Note that the discriminant is always positive or zero; i.e., the eigenvalues are real as required for the eigenvalues of a hermitian operator.

G.3.2 Eigenvalues and coupling angle

To study the effect of the perturbation we introduce the *coupling angle* θ and its tangent

$$\tan\theta \equiv \frac{2|H'_{12}|}{H_{11}-H_{22}}, \tag{G.33}$$

where θ is defined on the interval $-\pi/2 \leq \theta \leq \pi/2$ (under the convention (G.31) this is the interval $0 \leq \theta \leq \pi/2$). The name *coupling angle* refers to the coupling of the two differential equations in Eqs. (G.15) by the off-diagonal matrix elements H'_{12} and H'^{*}_{12}. In Section G.3.3 this coupling angle will be identified with the coupling angle introduced in Section G.2.2. We distinguish between *weak* and *strong* coupling (cf. Section G.3.4):

- The coupling is called *weak* if the off-diagonal matrix elements are "smaller than the level splitting", $|H'_{12}| \ll |H_{11}-H_{22}|$. In this case the coupling angle is *small* ($0 \leq \tan^2\theta \ll 1$).
- Likewise, the coupling is called *strong* for $|H'_{12}| \gg |H_{11}-H_{22}|$. In this case the coupling angle is *large*; i.e., close to $\pm\pi/2$ ($0 \leq \cot^2\theta \ll 1$). Note that in the *absence of asymmetry* ($H_{11}=H_{22}$) *any* coupling is strong.

---

[1] The adjective secular was introduced in astronomy to describe non-cyclic variations in the planetary motion caused by the interactions between planets.

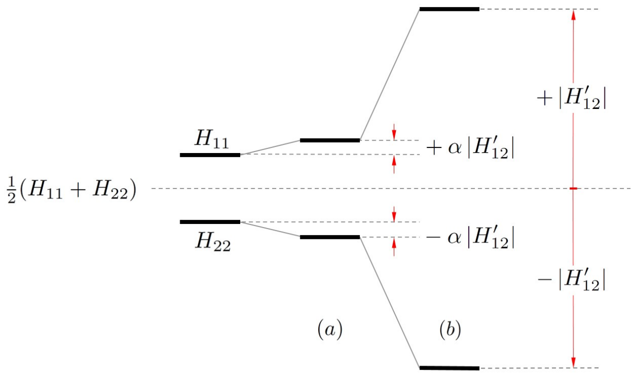


Figure G.2: Level shifts in two-level systems are symmetric with respect to the average value of the *diagonal* matrix elements, $\frac{1}{2}(H_{11}+H_{22})$: (a) *weak* repulsion by an *off-diagonal* matrix element ($|H'_{12}| \ll H_{11}-H_{22}$); (b) *strong* repulsion by an *off-diagonal* matrix element ($H_{11}-H_{22} \ll |H'_{12}|$).

Let us have a closer look at Eq. (G.32). Using the coupling angle this equation can be rewritten in one of the two following equivalent forms

$$E_\pm = \tfrac{1}{2}(H_{11}+H_{22}) \pm \tfrac{1}{2}(H_{11}-H_{22})\sqrt{1+\tan^2\theta} \tag{G.34a}$$

$$E_\pm = \tfrac{1}{2}(H_{11}+H_{22}) \pm |H'_{12}|\,\sqrt{1+\cot^2\theta}, \tag{G.34b}$$

where the labeling of the energy levels is chosen in accordance with the convention (G.31); i.e., $E_+ \to H_{11}$ and $E_- \to H_{22}$ in the limit of weak coupling ($\theta \to 0$). Note that there is no ambiguity in the sign of the square root,

$$\sqrt{1+\tan^2\theta} = 1/\cos\theta > 1 \quad \text{for} \quad -\pi/2 \le \theta \le \pi/2. \tag{G.35}$$

Also the level splitting $\Delta E = E_+ - E_-$ follows directly from Eq. (G.32),

$$\Delta E = \sqrt{(H_{11}-H_{22})^2 + 4\,|H'_{12}|^2}. \tag{G.36}$$

In terms of the coupling angle this becomes

$$\Delta E = \tfrac{1}{2}(H_{11}-H_{22})\sqrt{1+\tan^2\theta} \tag{G.37a}$$

$$\Delta E = 2\,|H'_{12}|\,\sqrt{1+\cot^2\theta}. \tag{G.37b}$$

The upper form of Eqs. (G.34) and (G.37) is best suited to analyze weak coupling ($\tan^2\theta \ll 1$); the lower form for strong coupling ($\cot^2\theta \ll 1$). The expressions shows that for weak coupling the splitting is dominated by the diagonal matrix elements whereas for strong coupling it is a direct measure of the coupling strength - see Fig. G.2.

Level repulsion

Note that the coupling gives rise to a shift of the energy levels which is symmetric about the average value $\frac{1}{2}(H_{11}+H_{22})$. More precisely, under the convention (G.31) the levels *repel* each other [118]. This is illustrated in Fig. G.2 for the degenerate case. The repulsion is easily verified by rewriting Eq. (G.34a) in the form

$$E_+ = H_{11} + \tfrac{1}{2}(H_{11}-H_{22})(\sqrt{1+\tan^2\theta} - 1) \tag{G.38a}$$

$$E_- = H_{22} - \tfrac{1}{2}(H_{11}-H_{22})(\sqrt{1+\tan^2\theta} - 1). \tag{G.38b}$$

The average value $\frac{1}{2}(H_{11}+H_{22})$ may be chosen as the zero of energy for most practical purposes.

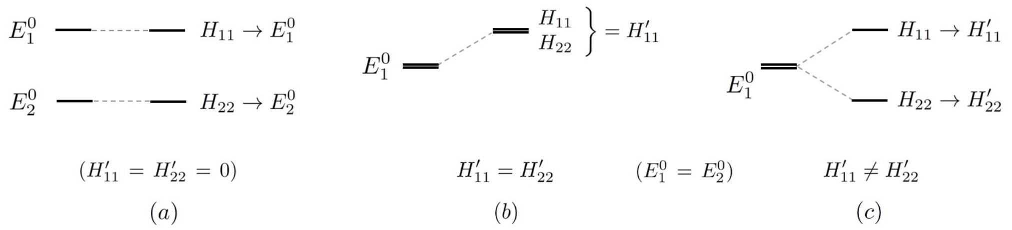


Figure G.3: Three special cases: (*a*) *absence* of diagonal shifts; (*b*) *degeneracy* with *symmetric* diagonal shift; (*c*) *degeneracy* with *asymmetric* diagonal shift.

Special cases

When dealing with energy levels and shifts induced by the perturbation it is good to mention a few special cases in which the equations simplify.

- *zero diagonal shift* ($H'_{11} = H'_{22} = 0$): in this case the diagonal matrix elements of the *full* Hamiltonian coincide with those of $\mathcal{H}_0$ - see Fig. G.3*a*

$$H_{11} \to E_1^0 \quad \text{and} \quad H_{22} \to E_2^0. \tag{G.39}$$

- *degeneracy* of the unperturbed levels ($E_1^0 = E_2^0$): in this case the diagonal matrix elements of the *full* Hamiltonian coincide with those of the perturbation,

$$H_{11} \to H'_{11} \quad \text{and} \quad H_{22} \to H'_{22}. \tag{G.40}$$

  We distinguish between two cases: (*a*) *symmetric* diagonal shifts ($H'_{11} = H'_{22}$); (*b*) *asymmetric* diagonal shifts ($H'_{11} \neq H'_{22}$) - see Fig. G.3*b,c.*

- *zero off-diagonal shift* ($H'_{12} = H'_{21} = 0$): in this case there is no coupling between the levels although the levels shift as a result of the diagonal matrix elements of the perturbation,

$$E_{\pm} = \tfrac{1}{2}(H_{11} + H_{22}) \pm \tfrac{1}{2}(H_{11} - H_{22}) = \begin{cases} H_{11} \\ H_{22} \end{cases}. \tag{G.41}$$

G.3.3 Eigenstates

Ones we have the eigenvalues $E_{\pm}$ we obtain the corresponding eigenstates $\psi_{\pm}$ by substituting the eigenvalues into the Schrödinger equation. In matrix notation this becomes

$$\begin{pmatrix} H_{11} & H'_{12} \\ H'^{*}_{12} & H_{22} \end{pmatrix} \begin{pmatrix} A_{\pm} \\ B_{\pm} \end{pmatrix} = E_{\pm} \begin{pmatrix} A_{\pm} \\ B_{\pm} \end{pmatrix}, \quad \text{with} \quad \psi_{\pm} \equiv \begin{pmatrix} A_{\pm} \\ B_{\pm} \end{pmatrix}. \tag{G.42}$$

As we recall from Section G.2.2, the eigenstates of a two-level system are fully defined by the mixing angle and the relative phase - see Eq. (G.16). In Problem G.2 we show that in terms of these angles the amplitudes of the eigenstates are given by

$$a_1 \to \begin{cases} A_{+} = +e^{-i\phi/2}\cos\alpha \\ A_{-} = -e^{-i\phi/2}\sin\alpha \end{cases} \quad \text{and} \quad a_2 \to \begin{cases} B_{+} = e^{i\phi/2}\sin\alpha = -A_{-}e^{i\phi} \\ B_{-} = e^{i\phi/2}\cos\alpha = A_{+}e^{i\phi} \end{cases}. \tag{G.43}$$

In this format the eigenstates become (in Dirac notation)

$$|\psi_+\rangle = +e^{-i\phi/2}\cos\alpha|\phi_1\rangle + e^{i\phi/2}\sin\alpha|\phi_2\rangle \tag{G.44a}$$

$$|\psi_-\rangle = -\mathrm{e}^{-\mathrm{i}\phi/2}\sin\alpha|\phi_1\rangle + e^{i\phi/2}\cos\alpha|\phi_2\rangle. \tag{G.44b}$$

As expected for a basis transformation Eqs. (G.44) define a unitary transformation between the eigenstates of $\mathcal{H}$ and those of $\mathcal{H}_0$ (cf. Problem 3.9). Importantly, as the *probabilities of occupation*,

$$|B_-|^2 = |A_+|^2 = \cos^2\alpha \quad \text{and} \quad |B_+|^2 = |A_-|^2 = \sin^2\alpha, \tag{G.45}$$

are independent of the phase we are free to choose the coordinate system such that the relative phase vanishes, $\phi = 0$. In this convention the amplitudes are *real* quantities and related by

$$B_+ = -A_- \quad \text{and} \quad B_- = A_+. \tag{G.46}$$

So the eigenstates of $\mathcal{H}$ take the convenient form

$$|\psi_+\rangle = +\cos\alpha|\phi_1\rangle + \sin\alpha|\phi_2\rangle \tag{G.47a}$$

$$|\psi_-\rangle = -\sin\alpha|\phi_1\rangle + \cos\alpha|\phi_2\rangle. \tag{G.47b}$$

The inverse transformation is given by

$$|\phi_1\rangle = \cos\alpha|\psi_+\rangle - \sin\alpha|\psi_-\rangle \tag{G.48a}$$

$$|\phi_2\rangle = \sin\alpha|\psi_+\rangle + \cos\alpha|\psi_-\rangle. \tag{G.48b}$$

What remains to be done is demonstrate that the coupling angle defined in Eq. (G.33) coincides with the coupling angle introduced for the pseudospin in Eq. (G.17). This follows in a few steps from Eqs (G.45). First we notice that $\tan^2\alpha$ may be expressed as

$$\tan^2\alpha = |B_+/A_+|^2 = |A_-/A_+|^2. \tag{G.49}$$

Substituting E_+ from Eq. (G.38a) into the Eq. (G.42) we obtain under the convention (G.31) [1]

$$\tan^2\alpha = \frac{(E_+ - H_{11})^2}{|H_{12}'|^2} = \frac{(\sqrt{1+\tan^2\theta} - 1)^2}{\tan^2\theta} = \frac{(1-\cos\theta)^2}{\sin^2\theta} = \frac{1-\cos\theta}{1+\cos\theta} = \tan^2(\theta/2). \tag{G.50}$$

Here we used Eq. (G.35) as well as the double-angle formulas for the cosine. With this result we established that the mixing angle α, introduced in Eq. (G.33), can be identified with *half* the coupling angle as defined in Eq. (G.33), $\alpha = \theta/2$. Since θ is defined over a range of π on the interval $-\pi/2 < \theta \le \pi/2$, the mixing angle is restricted to a range of $\pi/2$ on the interval $-\pi/4 < \alpha \le \pi/4$.[2] Alternatively, we can use the periodicity of $\tan\theta$ to redefine θ on the interval $0 \le \theta < \pi$ (i.e., α restricted to the interval $0 \le \alpha < \pi/2$). The latter convention was used to map the states of a pseudospin onto the Bloch sphere (see Section G.2.2).

An expression for the coupling coefficient $|A_+|^2$ in terms of the matrix elements is obtained using the trigonometric relation between $\cos^2\alpha$ and $\tan^2\alpha$,

$$|A_+|^2 = \cos^2\alpha = \frac{1}{1+\tan^2\alpha} = \frac{|H_{12}'|^2}{(E_+ - H_{11})^2 + |H_{12}'|^2}. \tag{G.51}$$

A nice consequence of the convention $-\pi/4 \le \alpha \le \pi/4$ is the inequality $\cos^2\alpha \ge \sin^2\alpha$. Note that this can serve as a convenient pointer to the dominant component of a mixed state (e.g., see the figures of magnetic fine- and hyperfine structure in Chapters 4 and 5).

---

[1] By substituting the solution (G.38b) into Eq. (G.42) we obtain the same result.

[2] Adopting the convention (G.31) we can restrict the angles to positive values, $0 \le \theta \le \pi/2$ and $0 \le \alpha \le \pi/4$.

Problem G.2. Derive Eqs. (G.43).

Solution. Introducing for the eigenstates $\psi_\pm$ the mixing angles $\alpha_\pm$ and phase angles $\theta_\pm$ we have in the notation of Section G.2.2

$$a_1 \to A_\pm = e^{-i\phi_\pm/2}\cos\alpha_\pm \quad \text{and} \quad a_2 \to B_\pm = e^{i\phi_\pm/2}\sin\alpha_\pm .$$

Since the eigenstates have to be orthogonal we require $\langle\psi_+|\psi_-\rangle = A_+^* A_- + B_+^* B_- = 0$. In terms of the angles this condition becomes

$$e^{i(\phi_+-\phi_-)/2}\cos\alpha_+\cos\alpha_- + e^{-i(\phi_+-\phi_-)/2}\sin\alpha_+\sin\alpha_- = 0.$$

This condition is satisfied for $\cos\alpha_- = -\sin\alpha_+$ and $\sin\alpha_- = \cos\alpha_+$ in combination with

$$e^{i(\phi_+-\phi_-)/2} - e^{-i(\phi_+-\phi_-)/2} = -2i\sin[(\phi_+ - \phi_-)/2] = 0.$$

In term of the angles this is satisfied for $\alpha_- = \alpha_+ + \pi/2$ and $\phi_- = \phi_+$. Changing notation $\alpha_+ \to \alpha$ and $\phi_+ \to \phi$ this results in the requested expressions. □

G.3.4 Weak versus strong coupling

In Section G.3.2 we introduced the concepts of weak and strong coupling by comparing the strength of the off-diagonal matrix elements, $|H'_{12}|$, with the difference of the diagonal ones, $H_{11} - H_{22}$, under the convention (G.31). We now analyze these cases in more detail using Fig. G.2 for purposes of illustration:

- *Weak coupling (strong asymmetry),* $|H'_{12}|^2 \ll (H_{11} - H_{22})^2 \to \alpha^2 \simeq \frac{1}{4}\tan^2\theta \ll 1$. In this regime Eq. (G.34a) reduces to

  $$E_+ = H_{11} + \alpha^2 (H_{11} - H_{22}) + \cdots = H_{11} + \alpha\,|H'_{12}| + \cdots , \tag{G.52a}$$

  $$E_- = H_{22} - \alpha^2 (H_{11} - H_{22}) + \cdots = H_{22} - \alpha\,|H'_{12}| + \cdots . \tag{G.52b}$$

  Here we used the expansion $\sqrt{1+\tan^2\theta} = 1 + \frac{1}{2}\tan^2\theta + \cdots = 1 + 2\alpha^2 + \cdots$. Note that the contribution of the off-diagonal matrix elements $|H'_{12}|$ is suppressed by a factor α with respect to the diagonal ones, H_{11} and H_{22}. In other words, the *diagonal elements contribute in first order*, cf. Eq. (H.56), whereas the *off-diagonal elements contribute in second-order*, cf. Eq. (H.63). The corresponding basis coefficients satisfy the relations

  $$A_\pm|^2 = \begin{cases} \cos^2\alpha = 1 - \alpha^2 + \cdots \\ \sin^2\alpha \;= \alpha^2 + \cdots . \end{cases} \tag{G.53}$$

  In the limit $|H'_{12}| \to 0$ we obtain the case of *no coupling* (zero off-diagonal shift), in which

  $$E_+ = H'_{11} \text{ and } E_- = H_{22}, \text{ with } |A_+|^2 = 1 \text{ and } |A_-|^2 = 0. \tag{G.54}$$

- *Strong coupling (weak asymmetry),* $(H_{11} - H_{22})^2 \ll |H'_{12}|^2 \Leftrightarrow \cot^2\theta \ll 1$. In this case Eq. (G.34b) reduces to

  $$E_\pm = \tfrac{1}{2}(H_{11} + H_{22}) \pm |H'_{12}|\,(1 + \tfrac{1}{2}\cot^2\theta + \cdots). \tag{G.55}$$

  Here we used the expansion $\sqrt{1+\cot^2\theta} = 1 + \frac{1}{2}\cot^2\theta + \cdots$. In this limit the mixing coefficients satisfy the relations

  $$A_\pm^2 = \tfrac{1}{2}(1 \pm \cot^2\theta + \cdots). \tag{G.56}$$

  In the limit $H_{11} \to H_{22}$ we obtain the *symmetric case*, in which

  $$E_\pm = H_{11} \pm |H'_{12}| \text{ and } A_\pm^2 = \tfrac{1}{2}. \tag{G.57}$$

- *Crossover regime.* In case we are neither weakly nor strongly coupled we are dealing with crossover behavior between the two limits. In this case the shifts cannot be approximated and we have to apply the exact expression; i.e., Eq. (G.32).

G.4 Non-stationary evolution under a static perturbation

Let us return to the time-dependent aspects of the two-level system. We start by assuming that the system is at time $t = 0$ in the eigenstate ϕ_1 of the unperturbed Hamiltonian, $\mathcal{H}_0$. To find the time evolution under the *full* Hamiltonian we solve the time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \mathcal{H} |\psi(t)\rangle \tag{G.58}$$

under the initial condition $|\psi(0)\rangle \equiv |\phi_1\rangle$. For a static perturbation, i.e., a perturbation without explicit time dependence, the full Hamiltonian is also independent of time and along the lines of Section G.3 we obtain for the time evolution of the initial state ϕ_1

$$|\psi(t)\rangle = e^{-(i/\hbar)\mathcal{H}t} |\phi_1\rangle. \tag{G.59}$$

In the absence of the perturbation, $\mathcal{H}_0 = 0$, the system remains in the state ϕ_1. This is the stationary evolution of Section G.3. In the present section we discuss how the state evolves for a non-vanishing static perturbation. In particular we are interested in finding out how the perturbation gives rise to a probability of observing the system at $t > 0$ in the eigenstate ϕ_2. As discussed in the introduction of Section G.2 this means that we are looking for the *transition probability*

$$W_{1\to 2}(t) = |\langle \phi_2 | \psi(t) \rangle|^2. \tag{G.60}$$

G.4.1 Special case

To introduce the subject we start with a special case. We consider the full Hamiltonian $\mathcal{H}$ with eigenvalues E_+ and E_- and suppose for convenience that the corresponding eigenstates ψ_+ and ψ_- are given by symmetric (+) and anti-symmetric (−) linear combinations of the eigenstates of $\mathcal{H}_0$),

$$|\psi_+\rangle = \sqrt{1/2}\,(|\phi_1\rangle + |\phi_2\rangle) \tag{G.61a}$$

$$|\psi_-\rangle = \sqrt{1/2}\,(|\phi_1\rangle - |\phi_2\rangle)\,. \tag{G.61b}$$

Note that this is a special case of Eqs. (G.47a) with $\sin^2\alpha = \cos^2\alpha = \frac{1}{2}$. This is called a 50/50 *coherent* mixture of the eigenstates ϕ_1 and ϕ_2 corresponding to a mixing angle of $\alpha = \pi/4$ and coupling angle $\theta = \pi/2$ ($\cot^2\theta \to 0$). Adding Eqs. (G.61) and solving for $|\phi_1\rangle$ (the initial state) we find

$$|\phi_1\rangle = \sqrt{1/2}\,(|\psi_+\rangle + |\psi_-\rangle)\,. \tag{G.62}$$

Next we substitute this expression into the time-dependent solution (G.59),

$$|\psi(t)\rangle = \sqrt{1/2}\left[e^{-(i/\hbar)E_+t}|\psi_+\rangle + e^{-(i/\hbar)E_-t}|\psi_-\rangle\right], \tag{G.63}$$

and project the result on the eigenstate ϕ_2. Using the orthonormality of the ϕ_1 and ϕ_2 as well as the decompositions (G.61) this leads to the *amplitude* for observing the system at time $t > 0$ in "level 2" (i.e., in the eigenstate ϕ_2),

$$a_2(t) = \tfrac{1}{2} e^{-(i/\hbar)E_+t}\left[1 - e^{(i/\hbar)(E_+ - E_-)t}\right]. \tag{G.64}$$

Introducing the angular frequency

$$\omega = (E_+ - E_-)/\hbar \tag{G.65}$$

the *probability* for observing the system in level 2 becomes

$$W_{1\to 2}(t) = \tfrac{1}{4}\left[2 - e^{i\omega t} - e^{-i\omega t}\right] = \tfrac{1}{2}(1 - \cos\omega t) = \sin^2 \tfrac{1}{2}\omega t. \tag{G.66}$$

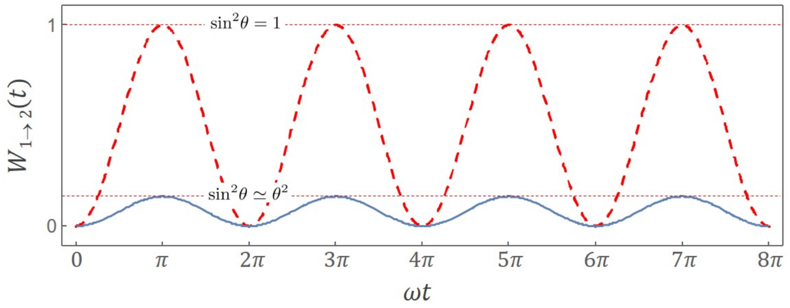


Figure G.4: Transition probability from level 1 to level 2 as a function of time after switching on a perturbation. Full transfer of the population can be achieved for the coupling angle $\theta = \pm\pi/2$ (dashed red line). In the limit of weak coupling ($\theta \ll 1$) the transfer efficiency is at best $\sin^2\theta \simeq \theta^2$ (solid blue line).

This result is plotted as the dashed red line in Fig. G.4. Note that at $t = 0$ the probability to observe the system in level 2 is zero, in accordance with the initial condition $|\psi(0)\rangle \equiv |\phi_1\rangle$. With increasing time the transition probability grows until at $\omega t = \pi$ the system is with unit probability in level 2. Continuing, the probability oscillates in time between zero and one at angular frequency ω. This has an important consequence. By applying the perturbation at $t = 0$ for the duration $\Delta t = \pi/\omega_0$ (better: $\Delta t = (\pi \pm 2n\pi)/\omega$, with $n = 0, \pm 1, \pm 2, \cdots$) we can transfer the entire population of a two-level system from level 1 to level 2. Such a timed pulse is called a *Rabi pulse* (better: a *zero-frequency* Rabi pulse). Since for the 50/50 mixture the coupling angle is $\theta = \pi/2$ ($\cot^2\theta \to 0$) we find for the angular frequency - see Eqs. (G.34b)

$$\omega = \frac{2}{\hbar}|H'_{12}|\sqrt{1+\cot^2\theta} \to \frac{2}{\hbar}|H'_{12}| \quad \text{for} \quad \theta \to \pi/2. \tag{G.67}$$

G.4.2 The general case - Rabi's formula

It is straightforward to generalize the example of the previous section to cases with an arbitrary mixing angle. For this purpose we recall Eqs. (G.47),

$$|\psi_+\rangle = +\cos\alpha|\phi_1\rangle + \sin\alpha|\phi_2\rangle \tag{G.68a}$$

$$|\psi_-\rangle = -\sin\alpha|\phi_1\rangle + \cos\alpha|\phi_2\rangle \tag{G.68b}$$

and its inverse, Eqs. (G.48). Sticking to the same initial condition as before the decomposition of the initial state is given by

$$|\psi(0)\rangle = |\phi_1\rangle = \cos\alpha|\psi_+\rangle - \sin\alpha|\psi_-\rangle. \tag{G.69}$$

Using Eq. (G.59) we find for the time evolution

$$\langle\phi_2|\psi(t)\rangle = \sin\alpha\cos\alpha\left[e^{-(i/\hbar)E_+t} - e^{-(i/\hbar)E_-t}\right]. \tag{G.70}$$

$$= 2\sin\alpha\cos\alpha e^{-(i/2\hbar)(E_+ + E_-)t} i\frac{\left[e^{-(i/2\hbar)(E_+ - E_-)t} - e^{(i/2\hbar)(E_+ - E_-)t}\right]}{2i} \tag{G.71}$$

$$= -i\sin 2\alpha \sin\tfrac{1}{2}\omega t, \tag{G.72}$$

where

$$\omega = (E_+ - E_-)/\hbar \tag{G.73}$$

is the *Rabi frequency* and $\frac{1}{2}(H_{11} + H_{22})$ was chosen as the zero of energy.

In terms of the Rabi frequency the *probability* to observe the system at time t in level 2 is given by

$$W_{1\to2}(t) = |\langle\phi_2|\psi(t)\rangle|^2 = \sin 2\alpha \sin^2 \tfrac{1}{2}\omega t. \tag{G.74}$$

In terms of the coupling angle this further simplifies to the form of *Rabi's formula* [85],

$$W_{1\to2}(t) = \sin^2\theta \sin^2 \tfrac{1}{2}\omega t. \tag{G.75}$$

Just as in the previous section, the probability oscillates at angular frequency ω. The difference is in coupling strength. For *strong* coupling, $\cot^2\theta \to 0$ ($\theta \to \pi/2$), we regain Eq. (G.66). The transition probability oscillates between zero and one. On the other hand, for *weak* coupling, $\tan^2\theta \to 0$, the transition probability oscillates between zero and the small value $\sin^2\theta$. In other words, aside from oscillating in time, the transition probability increases with the strength of the coupling until it saturates because a probability cannot become larger than one. This conclusion is satisfactory, also from the intuitive point of view. Using the trigonometric relation $\sin^2\theta = \tan^2\theta/(1+\tan^2\theta)$ we find by substitution of Eq. (G.33)

$$\sin^2\theta = \frac{4(H_{11}-H_{22})^2}{(H_{11}-H_{22})^2 + 4\left|H'_{12}\right|^2}. \tag{G.76}$$

For the level splitting we find

$$\omega = 2\left|H'_{12}\right|/\hbar \quad (\cot^2\theta \to 0); \qquad \omega = (H_{11}-H_{22})/\hbar \quad (\tan^2\theta \to 0). \tag{G.77}$$

Note that for $\cot^2\theta \to 0$ ($\theta \to \pi/2$) we regain Eq. (G.66). In the absence of a diagonal shift ($H'_{11} = H'_{22} = 0$) the angular frequency approaches in the limit of weak coupling ($\tan^2\theta \to 0$) the transition frequency of the unperturbed system.

$$\omega_0 \equiv (E_2^0 - E_1^0)/\hbar. \tag{G.78}$$

In atomic physics this quantity is known as the *Bohr frequency.* In Fig. G.4 we show the transition probability (G.75) for transitions from level 1 to level 2 as a function of time for both *strong* ($\cot^2\theta \to 0$ - dashed red line) and *weak* coupling ($\tan^2\theta \to 0$ - solid blue line).

H

Time-independent perturbation theory

H.1 Introduction

In most cases of practical importance physical systems cannot be studied without approximation and quantum mechanical systems are no exception to this rule. Unfortunately, exact solutions of the Schrödinger equation can only be obtained for a hand full of simple systems. Therefore, methods of approximation play a crucial role in the application of quantum mechanics. Obviously, in the computer era with massive calculational power at our disposal, approximate solutions of high precision are available using numerical methods. However, these methods constitute a field in itself and are better introduced in the context of mathematics. In this course we follow the tradition of focusing on analytical approaches as these offer structured insight in the physics of the quantum mechanical system under consideration. Importantly, although the approximate methods to be discussed are often intuitively convincing and appealing one should be aware that their foundation is generally non-trivial and forms an important subject of mathematical physics.[1] In this appendix we discuss analytical approximations for *time-independent* phenomena in systems with *discrete energy levels* resulting from a *time-independent* Hamiltonian. Time-dependent phenomena form a topic in itself and are best discussed separately.

In *time-independent* perturbation theory we consider physical systems in which the eigenstates and eigenvalues of the Hamiltonian $\mathcal{H}$ of the system are known, except for the effects of a *small* and *stationary* perturbing term $\mathcal{H}_1$. Typically this gives rise to level shifts and/or level splittings. To develop the theory the Hamiltonian $\mathcal{H}$ is broken up in two hermitian parts,

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1, \tag{H.1}$$

where $\mathcal{H}_0$ is called the *unperturbed part* and $\mathcal{H}_1$ the *perturbation.* We shall consider systems where the Schrödinger equation for the unperturbed Hamiltonian

$$\mathcal{H}_0 \left|\phi_n\right\rangle = E_n^0 \left|\phi_n\right\rangle . \tag{H.2}$$

defines a *discrete* set of (known) eigenvalues $\{E_n^0\}$ corresponding to an *orthonormal* set of (known) eigenstates $\{|\phi_n\rangle\}$, for instance corresponding to the *bound states* of a system. The presence of an additional continuum of states is not considered here but can be included in the closure of the Hilbert space. The Schrödinger equation for the full Hamiltonian

$$\mathcal{H}|\psi_n\rangle = E_n|\psi_n\rangle \tag{H.3}$$

defines a discrete set of (unknown) eigenvalues $\{E_n\}$ corresponding to an orthonormal set of (unknown) eigenstates $\{|\psi_n\rangle\}$. Because the perturbation is small we expect only a small shift of the

---

[1] See P.M. Morse and H. Feshbach, *Methods of Theoretical Physics*, McGraw-Hill Book Company, Boston 1953.

energy level E_n with respect to the value E_n^0. Projecting both sides of the Schrödinger equation (H.3) on an arbitrary unperturbed state $|\phi_\nu\rangle$ we obtain, after splitting off the perturbation and using the hermiticity of $\mathcal{H}_0$,

$$E_n\langle\phi_\nu|\psi_n\rangle = \langle\phi_\nu|\mathcal{H}_0 + \mathcal{H}_1|\psi_n\rangle = E_\nu^0\langle\phi_\nu|\psi_n\rangle + \langle\phi_\nu|\mathcal{H}_1|\psi_n\rangle. \tag{H.4}$$

Thus we found a set of expressions relating the energy of the perturbed level E_n with the energies E_ν^0 of all the eigenstates $\{|\phi_n\rangle\}$ of the unperturbed Hamiltonian $\mathcal{H}_0$,

$$E_n = E_\nu^0 + \frac{\langle\phi_\nu|\mathcal{H}_1|\psi_n\rangle}{\langle\phi_\nu|\psi_n\rangle}. \tag{H.5}$$

For the special case $\nu = n$ this relation corresponds to an expression for the energy shift,

$$\Delta E_n = E_n - E_n^0 = \frac{\langle\phi_n|\mathcal{H}_1|\psi_n\rangle}{\langle\phi_n|\psi_n\rangle}. \tag{H.6}$$

It will be our approach to derive an approximate solution for the level shift with a method of successive approximation starting from Eq. (H.6). This method is called *Rayleigh-Schrödinger perturbation theory* Schrödinger [100]. We distinguish between perturbation theory for non-degenerate levels (*ordinary* perturbation theory - see Section H.2) and that for degenerate levels (see Section H.3). It will be shown that the ordinary theory embodies a method of successive approximation describing the level shift (H.6); for the degenerate case the method amounts to the diagonalization of the degenerate subspace and describes the lifting of degeneracy by splitting into nondegenerate sublevels.

H.2 Perturbation theory for non-degenerate levels

We start with the case of a *non-degenerate* energy level E_n. To develop the perturbation theory we introduce a parameterized Hamiltonian

$$\mathcal{H}(\lambda) = \mathcal{H}_0 + \lambda\mathcal{H}_1, \tag{H.7}$$

with a set of (unknown) orthonormal eigenstates $\{|\psi_n(\lambda)\rangle\}$ satisfying the Schrödinger equation

$$\mathcal{H}(\lambda)\,|\psi_n(\lambda)\rangle = E_n(\lambda)\,|\psi_n(\lambda)\rangle\,. \tag{H.8}$$

The parameter λ enables us to vary the strength of the perturbation and is introduced solely to keep track of the order at which the perturbation contributes; for $\lambda \to 0$ the $\mathcal{H}(\lambda)$ reduces to the unperturbed Hamiltonian, $\lim_{\lambda\to 0}\mathcal{H}(\lambda) = \mathcal{H}_0$; for $\lambda \to 1$ we obtain the full Hamiltonian, $\lim_{\lambda\to 1}\mathcal{H}(\lambda) = \mathcal{H}_0 + \mathcal{H}_1$. Using the Hamiltonian (H.7) the expression (H.5) changes into

$$E_n(\lambda) = E_\nu^0 + \lambda\frac{\langle\phi_\nu|\mathcal{H}_1\,|\psi_n(\lambda)\rangle}{\langle\phi_\nu|\psi_n(\lambda)\rangle}. \tag{H.9}$$

The denominator of the second term plays an important role in the perturbation theory and will be denoted by

$$a_\nu(\lambda) \equiv \langle\phi_\nu|\psi_n(\lambda)\rangle. \tag{H.10}$$

Eq. (H.9) is the starting point to obtain approximate solutions for the levels $E_n(\lambda)$ and the coefficients $a_\nu(\lambda)$. This is done with an iterative procedure in which Eq. (H.9) is expressed in the form of an hierarchy of equations of increasing power of λ obtained from the Ansatz

$$|\psi_n(\lambda)\rangle \equiv |\phi_n\rangle + \lambda|\psi_n^1\rangle + \lambda^2|\psi_n^2\rangle + \cdots\,, \tag{H.11}$$

where $|\psi_n^p\rangle$ corresponds to the pth-order correction to the wavefunction, with $p \in \{0, 1, \cdots\}$ and $|\psi_n^0\rangle \equiv |\phi_n\rangle$. As will be shown below the corrections $|\psi_n^p\rangle$ can be chosen orthogonal to the unperturbed state, $\langle\phi_n|\psi_n^p\rangle = 0$, and we will do so in a later stage. The Ansatz (H.11) implies an expansion in powers of λ for the function $a_\nu(\lambda)$,

$$a_\nu(\lambda) = \langle\phi_\nu|\psi_n(\lambda)\rangle = \delta_{\nu n} + \lambda a_\nu^{(1)} + \lambda^2 a_\nu^{(2)} + \cdots , \tag{H.12}$$

where the coefficients are defined by $a_\nu^{(p)} = \langle\phi_\nu|\psi_n^p\rangle$. For weak perturbations the Ansatz assures that the coefficients $a_\nu(\lambda)$ will change regularly with λ over the full interval $0 \leq \lambda \leq 1$ (in particular at the end points) and differ little, $a_\nu(\lambda) - a_\nu^{(0)} \ll 1$, from the unperturbed value $a_\nu^{(0)}$ given by

$$a_\nu^{(0)} \equiv \lim_{\lambda\to 0} a_\nu(\lambda) = \lim_{\lambda\to 0} \langle\phi_\nu|\psi_n(\lambda)\rangle = \langle\phi_\nu|\phi_n\rangle = \delta_{\nu n}. \tag{H.13}$$

The norm of the parameterized wavefunction is given by

$$\langle\psi_n(\lambda)|\psi_n(\lambda)\rangle = 1 + \lambda(a_n^{(1)} + a_n^{(1)*}) + \lambda^2(a_n^{(2)} + a_n^{(2)*} + \langle\psi_n^1|\psi_n^1\rangle) + \cdots . \tag{H.14}$$

Substituting the Ansatz (H.11) and the expansion (H.12) into Eq. (H.9) we obtain for the special case $\nu = n$

$$E_n(\lambda) = E_n^0 + \frac{\lambda\langle\phi_n|\mathcal{H}_1|\phi_n\rangle + \lambda^2\langle\phi_n|\mathcal{H}_1|\psi_n^1\rangle + \cdots}{1 + \lambda a_n^{(1)} + \lambda^2 a_n^{(2)} + \cdots}. \tag{H.15}$$

Expanding this expression in powers of λ we obtain

$$\begin{aligned} E_n(\lambda) = E_n^0 &+ \lambda\langle\phi_n|\mathcal{H}_1|\phi_n\rangle \\ &+ \lambda^2[\langle\phi_n|\mathcal{H}_1|\psi_n^1\rangle - \langle\phi_n|\mathcal{H}_1|\phi_n\rangle a_n^{(1)}] \\ &+ \lambda^3[\langle\phi_n|\mathcal{H}_1|\psi_n^2\rangle - \langle\phi_n|\mathcal{H}_1|\psi_n^1\rangle a_n^{(1)} - \langle\phi_n|\mathcal{H}_1|\phi_n\rangle a_n^{(2)}] + \cdots . \end{aligned} \tag{H.16}$$

This expansion defines the energy shift order by order,

$$E_n(\lambda) \equiv E_n^0 + \lambda\Delta E_n^{(1)} + \lambda^2\Delta E_n^{(2)} + \lambda^3\Delta E_n^{(3)} + \cdots , \tag{H.17}$$

where $\Delta E_n^{(p)}$ represents the pth-order contribution to the energy shift. With regard to the eigenstate $|\psi_n(\lambda)\rangle$, the perturbation can be regarded as giving rise to an admixture of all *other* states $(m \neq n)$ as follows with the aid of the closure relation under the assumption $\langle\phi_n|\psi_n^1\rangle = 0$,

$$|\psi_n(\lambda)\rangle \equiv |\phi_n\rangle + \lambda\sum_m{}'|\phi_m\rangle\langle\phi_m|\psi_n^1\rangle + \lambda^2\sum_m{}'|\phi_m\rangle\langle\phi_m|\psi_n^2\rangle + \cdots . \tag{H.18}$$

Here the prime on the summation indicates $m \neq n$ (*i.e.*, the term $m = n$ is skipped in the summation). In terms of the coefficients $a_m^{(p)}$ the corrections to the wavefunction can be expressed as as

$$|\psi_n(\lambda)\rangle \equiv |\phi_n\rangle + \lambda\sum_m{}' a_m^{(1)}|\phi_m\rangle + \lambda^2\sum_m{}' a_m^{(2)}|\psi_m\rangle + \cdots . \tag{H.19}$$

Note that to calculate the energy shift to order $p \geq 1$ we need to know the wavefunction $|\psi_n^q\rangle$ to order $q \leq p - 1$. It is good to emphasize that there is no guarantee that the expansion (H.16) actually converges for $\lambda \to 1$. We will simply presume convergence, as is the case in many applications of practical interest. Convergence criteria and alternative perturbation methods can be found in advanced texts.

Before proceeding to the calculation of the energy shifts $\Delta E_n^{(p)}$ we return to Eq. (H.9) and derive a set of equations for the coefficients $a_\nu^{(p)}$, this time for the case $\nu = m \neq n$. Substituting Eqs. (H.11) and (H.12) for the case $\nu = m \neq n$ into Eq. (H.9) we obtain

$$E_n(\lambda) = E_m^0 + \frac{\lambda\langle\phi_m|\mathcal{H}_1|\phi_n\rangle + \lambda^2\langle\phi_m|\mathcal{H}_1|\psi_n^1\rangle + \cdots}{\lambda a_m^{(1)} + \lambda^2 a_m^{(2)} + \cdots}. \tag{H.20}$$

Replacing the l.h.s. by the expansion (H.17) and collecting the terms of equal power in λ we have

$$\begin{aligned} (E_n^0 - E_m^0)a_m^{(1)} &= \langle\phi_m|\mathcal{H}_1|\phi_n\rangle && \text{first order} && \text{(H.21a)}\\ \Delta E_n^{(1)}a_m^{(1)} + (E_n^0 - E_m^0)a_m^{(2)} &= \langle\phi_m|\mathcal{H}_1|\psi_n^1\rangle && \text{second order} && \text{(H.21b)}\\ \Delta E_n^{(2)}a_m^{(1)} + \Delta E_n^{(1)}a_m^{(2)} + (E_n^0 - E_m^0)a_m^{(3)} &= \langle\phi_m|\mathcal{H}_1|\psi_n^2\rangle && \text{third order} && \text{(H.21c)}\\ \cdots &= \cdots . \end{aligned}$$

These equations are valid for $m \neq n$. Extension to any order in λ is straightforward. Note that the parameter λ has dropped out of the expressions; it only served to identify the orders of the perturbation expansion. Note further that to calculate $a_m^{(p)}$ to order $p \geq 1$ we need to know the correction to the wavefunction $|\psi_n^q\rangle$ as well as the energy shift $\Delta E_n^{(q)}$ to order $q \leq p-1$.

Zeroth order

We are now prepared to calculate the energy shifts $\Delta E_n^{(p)}$. To zeroth order we have $E_n \simeq E_n^0$ and $a_m^{(0)} \simeq \langle\phi_m|\psi_n^0\rangle = \langle\phi_m|\phi_n\rangle = \delta_{mn}$. Note that the approximate wavefunction $|\psi_n\rangle \simeq |\phi_n\rangle$ is normalized.

H.2.0.1 First order

Proceeding to first order we compare Eqs. (H.17) and (H.16) to obtain

$$\Delta E_n^{(1)} = \langle\phi_n|\mathcal{H}_1|\phi_n\rangle. \tag{H.22}$$

With this result we expressed the first-order energy shifts in terms of the (known) unperturbed states $\{|\phi_n\rangle\}$. Eq. (H.22) is of enormous practical importance. It shows that in the presence of the perturbation $\mathcal{H}_1$ the first-order contribution to the energy is obtained by calculating the expectation value of $\mathcal{H}_1$ under the assumption that the wavefunctions remain unchanged. Also from the intuitive point of view this is plausible: we expect

$$\Delta E_n = \frac{\langle\phi_n|\mathcal{H}_1|\psi_n\rangle}{\langle\phi_n|\psi_n\rangle} \simeq \frac{\langle\phi_n|\mathcal{H}_1|\phi_n\rangle}{\langle\phi_n|\phi_n\rangle} = \langle\phi_n|\mathcal{H}_1|\phi_n\rangle \tag{H.23}$$

because replacing in Eq. (H.6) the exact (but unknown) state $|\psi_n\rangle$ by the only slightly different (but known) state $|\phi_n\rangle$ will hardly affect the result.

H.2.0.2 Second order

Comparing Eq. (H.16) with (H.17), and using the closure relation the second-order contribution to the energy shift takes the form

$$\Delta E_n^{(2)} = \sum_m \langle\phi_n|\mathcal{H}_1|\phi_m\rangle\langle\phi_m|\psi_n^1\rangle - \langle\phi_n|\mathcal{H}_1|\phi_n\rangle a_n^{(1)}. \tag{H.24}$$

Separating the term $m = n$ from the summation it cancels against the second term on the r.h.s. and Eq. (H.24) reduces to

$$\Delta E_n^{(2)} = \sum_m{}' a_m^{(1)}\langle\phi_n|\mathcal{H}_1|\phi_m\rangle, \tag{H.25}$$

where the prime indicates the constraint $m \neq n$. Interestingly, the result for $\Delta E_n^{(2)}$ is independent of $a_n^{(1)}$. This leaves us the freedom to choose the value of $a_n^{(1)}$ to best serve our purpose. The obvious choice is $a_n^{(1)} = 0$. This not only assures the normalization of the wavefunction to first order in λ, see Eq. (H.14), but also has the advantage that the expression for the third-order shift simplifies, see Eq. (H.16). Eq. (H.25) shows that the second-order correction to the energy follows from the first-order correction to the state - see Eq. (H.19),

$$|\psi_n(\lambda)\rangle \simeq |\phi_n\rangle + \lambda \sum_m{}' a_m^{(1)} |\phi_m\rangle. \tag{H.26}$$

To obtain the coefficient $a_m^{(1)}$ we rewrite Eq. (H.21a) in the form

$$a_m^{(1)} = \frac{\langle \phi_m | \mathcal{H}_1 | \phi_n \rangle}{E_n^0 - E_m^0} \quad \text{with } m \neq n. \tag{H.27}$$

As this expression is not defined for $E_n^0 = E_m^0$ it is *only valid in the absence of degeneracy.* Substituting Eq. (H.27) into Eq. (H.26) the *first-order* expression for the *state* is given by

$$|\psi_n\rangle \simeq |\phi_n\rangle + \sum_m{}' |\phi_m\rangle \frac{\langle \phi_m | \mathcal{H}_1 | \phi_n \rangle}{E_n^0 - E_m^0}. \tag{H.28}$$

Substituting Eq. (H.27) into Eq. (H.25) the *second-order* correction to the *energy* is found to be

$$\Delta E_n^{(2)} = \sum_m{}' \frac{|\langle \phi_n | \mathcal{H}_1 | \phi_m \rangle|^2}{E_n^0 - E_m^0}. \tag{H.29}$$

Note that this second-order contribution to the energy shift consists of a summation over all states $|\phi_m\rangle$ except the state $|\phi_n\rangle$ under consideration. This feature of second-order perturbation theory is called summation over *virtual excitations* or *virtual transitions to excited states.*

Closure approximation: It is possible to set limits on $|\Delta E_n^{(2)}|$ by over-estimating or under-estimating all terms of the summation (H.29). We illustrate this for the ground state ($n = 0$). Since $\left|E_1^0 - E_0^0\right| \leq \left|E_m^0 - E_0^0\right| \leq \left|E_0^0\right|$ the following inequality holds:

$$\frac{1}{|E_0^0|} \sum_m{}' |\langle \phi_0 | \mathcal{H}_1 | \phi_m \rangle|^2 \leq |\Delta E_0^{(2)}| \leq \frac{1}{|E_1^0 - E_0^0|} \sum_m{}' |\langle \phi_0 | \mathcal{H}_1 | \phi_m \rangle|^2 \tag{H.30}$$

Using the closure relation the summation can be rewritten in the form

$$\sum_m{}' |\phi_m\rangle \langle \phi_m| = \mathbb{1} - |\phi_n\rangle \langle \phi_n| \tag{H.31}$$

and the inequality becomes

$$\frac{\langle \phi_0 | \mathcal{H}_1^2 | \phi_0 \rangle - |\langle \phi_0 | \mathcal{H}_1 | \phi_0 \rangle|^2}{|E_0^0|} \leq |\Delta E_0^{(2)}| \leq \frac{\langle \phi_0 | \mathcal{H}_1^2 | \phi_0 \rangle - |\langle \phi_0 | \mathcal{H}_1 | \phi_0 \rangle|^2}{|E_1^0 - E_0^0|}, \tag{H.32}$$

The quantity

$$(\Delta \mathcal{H}_1)^2 = \langle \phi_0 | \mathcal{H}_1^2 | \phi_0 \rangle - |\langle \phi_0 | \mathcal{H}_1 | \phi_0 \rangle|^2$$

is the variance of the perturbing term $\mathcal{H}_1$ in the state $|\phi_0\rangle$. This approximation is known as the *closure approximation.*

H.2.0.3 Third order

In many cases the determination of the first-order contribution $\Delta E_n^{(1)}$ is already adequate for a good estimate of the level shift, with $\Delta E_n^{(2)}$ serving as a tool to get an impression of the convergence of the perturbation expansion. When the first-order shift happens to be zero, $\Delta E_n^{(1)} = 0$, the second order contribution may still be nonzero and provide the leading contribution to the shift of the energy levels. Rarely we will need the third-order contribution. The reason to proceed anyhow and calculate the third order is that a new element enters the discussion with consequences for the normalization of the wavefunction to second order in λ. Comparing Eqs. (H.17) and (H.16) and using the choice $a_n^{(1)} = 0$ the third-order contribution to the energy shift can be written in the form

$$\Delta E_n^{(3)} = \sum_m \langle \phi_n | \mathcal{H}_1 | \phi_m \rangle \langle \phi_m | \psi_n^2 \rangle - \langle \phi_n | \mathcal{H}_1 | \phi_n \rangle a_n^{(2)}. \tag{H.33}$$

As the term $m = n$ from the summation cancels against the second term on the r.h.s. the expression for $\Delta E_n^{(3)}$ becomes

$$\Delta E_n^{(3)} = \sum_m{}' a_m^{(2)} \langle \phi_n | \mathcal{H}_1 | \phi_m \rangle. \tag{H.34}$$

The result for $\Delta E_n^{(3)}$ is independent of $a_n^{(2)}$. Thus, like for $a_n^{(1)}$ we have the freedom to choose $a_n^{(2)}$. Choosing $a_n^{(2)} = 0$ has the advantage that the expression for the fourth-order term simplifies. However, unlike in the second-order case this choice does not conserve the normalization as is evident from Eq. (H.14). Normalization requires $a_n^{(2)} + a_n^{(2)*} + \langle \psi_n^1 | \psi_n^1 \rangle = 0$. Hence, we have the choice between unit normalization or a simplified hierarchy of equations of higher order. To keep the procedure simple we choose for the latter ($a_n^{(2)} = 0$) and correct for the loss of normalization by renormalizing the final result. We return to this renormalization below. To determine the coefficient $a_m^{(2)}$ for $m \neq n$ we use the closure relation to rewrite Eq. (H.21b) in the form

$$(E_n^0 - E_m^0) a_m^{(2)} = \sum_\nu{}' \langle \phi_m | \mathcal{H}_1 | \phi_\nu \rangle \langle \phi_\nu | \psi_n^{(1)} \rangle - \Delta E_n^{(1)} a_m^{(1)} \quad \text{with } m \neq n, \tag{H.35}$$

where the term $\nu = n$ could be excluded because $a_n^{(1)}$ was previously chosen to be zero. Substituting the first-order results (H.22) and (H.27) the expression for the second-order coefficient becomes

$$a_m^{(2)} = \sum_\nu{}' \frac{\langle \phi_m | \mathcal{H}_1 | \phi_\nu \rangle \langle \phi_\nu | \mathcal{H}_1 | \phi_n \rangle}{(E_n^0 - E_\nu^0)(E_n^0 - E_m^0)} - \frac{\langle \phi_m | \mathcal{H}_1 | \phi_n \rangle \langle \phi_n | \mathcal{H}_1 | \phi_n \rangle}{(E_n^0 - E_m^0)^2} \quad \text{with } m \neq n. \tag{H.36}$$

Substituting this result into Eq. (H.34) we find for the third-order shift

$$\Delta E_n^{(3)} = \sum_{m,\nu}{}' \frac{\langle \phi_n | \mathcal{H}_1 | \phi_m \rangle \langle \phi_m | \mathcal{H}_1 | \phi_\nu \rangle \langle \phi_\nu | \mathcal{H}_1 | \phi_n \rangle}{(E_n^0 - E_\nu^0)(E_n^0 - E_m^0)} - \langle \phi_n | \mathcal{H}_1 | \phi_n \rangle \sum_m{}' \frac{|\langle \phi_m | \mathcal{H}_1 | \phi_n \rangle|^2}{(E_n^0 - E_m^0)^2}. \tag{H.37}$$

H.2.1 Renormalization of the wavefunction

It is straightforward to show by induction that Eq. (H.25) can be generalized to arbitrary order by choosing $a_n^{(p)} = 0$ for all orders $p \geq 1$. With this convention we have $a_n(\lambda) = 1$ and the expression for the energy shifts (H.15) reduces to

$$\Delta E_n(\lambda) = \lambda \langle \phi_n | \mathcal{H}_1 | \phi_n \rangle + \lambda^2 \langle \phi_n | \mathcal{H}_1 | \psi_n^1 \rangle + \cdots = \lambda \langle \phi_n | \mathcal{H}_1 | \psi_n(\lambda) \rangle. \tag{H.38}$$

Here we recognize a conveniently simplified form of Eq. (H.6). However, as mentioned above, with this convention the normalization differs from unity,

$$\langle \psi_n(\lambda) | \psi_n(\lambda) \rangle = 1 + \lambda^2 \langle \psi_n^1 | \psi_n^1 \rangle + \lambda^3 [\langle \psi_n^1 | \psi_n^2 \rangle + \langle \psi_n^2 | \psi_n^1 \rangle] + \cdots . \tag{H.39}$$

As the result (H.38) depends on the norm of $|\psi_n(\lambda)\rangle$ we have to correct for the normalization error by multiplying with a renormalization factor,

$$\Delta E_n(\lambda) = \frac{\lambda\langle\phi_n|\mathcal{H}_1|\psi_n(\lambda)\rangle}{\langle\psi_n(\lambda)|\psi_n(\lambda)\rangle^{1/2}} = \lambda\langle\phi_n|\mathcal{H}_1|\psi_n(\lambda)\rangle Z^{1/2}. \tag{H.40}$$

The quantity

$$Z = \frac{1}{\langle\psi_n(\lambda)|\psi_n(\lambda)\rangle} = 1 - \lambda^2\langle\psi_n^{(1)}|\psi_n^{(1)}\rangle + \cdots \tag{H.41}$$

is called the *renormalization constant.* Note that the renormalization correction enters as a second-order correction in λ. Using the closure relation as well as Eq. (H.27) the renormalization constant becomes

$$Z = 1 - \lambda^2 \sum_m{}' \frac{|\,\langle\phi_n|\,\mathcal{H}_1\,|\phi_m\rangle\,|^2}{(E_n^0 - E_m^0)^2} + \cdots . \tag{H.42}$$

H.3 Perturbation theory for degenerate levels

In this section we turn to the perturbation theory for a g-fold *degenerate* energy level E_n^0 of the unperturbed Hamiltonian $\mathcal{H}_0$, corresponding to *bound states* $|\phi_{n,k}\rangle$, with $k \in \{1, \cdots, g\}$, where the value of g will generally depend on the primary quantum number n. In this case the Schrödinger equation of $\mathcal{H}_0$ takes the form

$$\mathcal{H}_0\,|\phi_{n,k}\rangle = E_n^0\,|\phi_{n,k}\rangle \;\text{ with } k \in \{1, \cdots, g\}, \tag{H.43}$$

where the manifold of normalized eigenstates $\{|\phi_{n,k}\rangle\}$, with $k \in \{1, \cdots, g\}$, span the degenerate subspace of level E_n^0 but are not *necessarily* orthogonal,

$$S_{ik} = \langle\phi_{n,i}|\phi_{n,k}\rangle \neq \delta_{ik}. \tag{H.44}$$

The quantity S_{ik} represents the overlap matrix for the eigenstates $|\phi_{n,k}\rangle$ of the degenerate subspace. The set $\{|\phi_{n,k}\rangle\}$, with $k \in \{1, \cdots, g\}$, can of course always be orthogonalized to provide a complete set of orthonormal eigenstates. Implicitly, *we suppose that this orthogonalization has been done*

$$\langle\phi_{n,i}|\phi_{n,k}\rangle = \delta_{ik}. \tag{H.45}$$

As we will see this diagonalization, although convenient, is not sufficient to deal with the degenerate case.

Ordinary perturbation theory fails in the degenerate case by divergence at the level of Eqs. (H.27) and (H.29) because $E_{n,k}^0 = E_{n,l}^0$ for $k \neq l \in \{1, \cdots, g\}$. As will appear, for a degenerate manifold of states, the effect of the perturbation is not so much an over-all shift of the manifold but a relative shift of the degenerate sublevels with respect to each other. This is called *lifting of the degeneracy* by level splitting. To approximate the perturbed basis states $\{|\psi_{n,k}\rangle\}$ we reconsider the Hamiltonian (H.7) with corresponding Schrödinger equation

$$\mathcal{H}(\lambda)\,|\psi_{n,k}(\lambda)\rangle = E_{n,k}(\lambda)\,|\psi_{n,k}(\lambda)\rangle\,. \tag{H.46}$$

For this case the expression (H.9) takes the form

$$E_{n,k}(\lambda) = E_\nu^0 + \lambda\frac{\langle\phi_{\nu,i}|\mathcal{H}_1\,|\psi_{n,k}(\lambda)\rangle}{\langle\phi_{\nu,i}|\psi_{n,k}(\lambda)\rangle}. \tag{H.47}$$

To obtain approximate expressions for the levels $E_{n,k}(\lambda)$ we start, in analogy with Eq. (H.11), with the Ansatz

$$|\psi_{n,k}(\lambda)\rangle \equiv |\phi_{n,k}\rangle + \lambda|\psi_{n,k}^1\rangle + \lambda^2|\psi_{n,k}^2\rangle + \cdots . \tag{H.48}$$

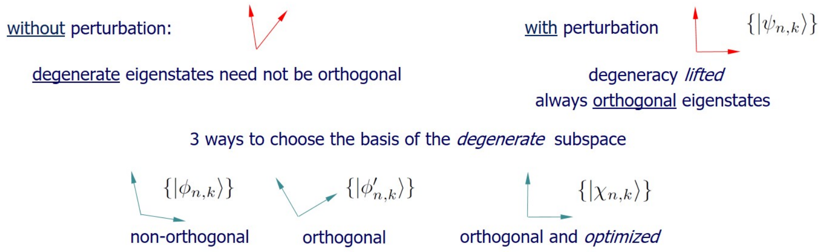


Figure H.1: Heuristic diagram for the selection of the optimal basis for perturbation theory of a degenerate level. The perturbation removes the ambiguity in the choice of basis vectors and this imposes a constraint in the choice of the unperturbed basis.

However, in the present case immediately an ambiguity arises. It is not obvious *which* of the unperturbed states from the degenerate manifold $\{|\phi_{n,k}\rangle\}$ of energy E_n^0 should be chosen in the expansion. Actually, *any* normalized linear combination of states from the degenerate the manifold $\{|\phi_{n,k}\rangle\}$ will also be an eigenstate of $\mathcal{H}_0$ of energy E_n^0. Therefore, it is not *a priori* clear how to choose an orthonormal set $\{|\chi_{n,k}\rangle\}$ within the degenerate subspace of the manifold n such that the $|\psi_{n,k}^p\rangle$ with $p \geq 1$ represent a small correction with respect to $|\chi_{n,k}\rangle$ for $\lambda \to 1$. This becomes clear only *after* we know how the symmetry of $\mathcal{H}_0$ is broken by the perturbation $\mathcal{H}_1$.

Let us presume that $\{|\chi_{n,i}\rangle\}$ is the (unknown) orthonormal set of the manifold n that we are looking for and $\{|\phi_{m,k}\rangle\}$ is the subset of the basis $\{|\phi_{n,k}\rangle\}$ for which $m \neq n$, (see Fig. H.1). The selection of the optimal basis is illustrated in Fig. H.1. For the optimal basis the following closure relation holds

$$\mathbb{1} = \sum_i |\chi_{n,i}\rangle\langle\chi_{n,i}| + \sum_{m,k}{}' |\phi_{m,k}\rangle\langle\phi_{m,k}| \equiv \sum_{\nu,\mu} |\psi_{\nu,\mu}^0\rangle\langle\psi_{\nu,\mu}^0|, \tag{H.49}$$

where the prime indicates $m \neq n$ and

$$\{|\psi_{\nu,\mu}^0\rangle\} = \begin{cases} \{|\chi_{n,\mu}\rangle\} & \text{for } \nu = n \\ \{|\phi_{\nu,\mu}\rangle\} & \text{for } \nu \neq n. \end{cases} \tag{H.50}$$

For this new set of eigenstates the Ansatz (H.48) takes the form

$$|\psi_{n,k}(\lambda)\rangle \equiv |\chi_{n,k}\rangle + \lambda|\psi_{n,k}^1\rangle + \lambda^2|\psi_{n,k}^2\rangle + \cdots . \tag{H.51}$$

This implies for the function $a_{\nu\mu,nk}(\lambda)$ the following expansion in powers of λ:

$$a_{\nu\mu,nk}(\lambda) = \langle\psi_{\nu,\mu}^0|\psi_{n,k}(\lambda)\rangle = \delta_{\nu\mu,nk} + \lambda a_{\nu\mu,nk}^{(1)} + \lambda^2 a_{\nu\mu,nk}^{(2)} + \cdots , \tag{H.52}$$

where the coefficients are defined by $a_{\nu\mu,nk}^{(p)} = \langle\psi_{\nu,\mu}^0|\psi_{n,k}^p\rangle$. Substituting this expansion into Eq. (H.47) we obtain for the special case $\nu = n, \mu = k$

$$E_{n,k}(\lambda) = E_n^0 + \frac{\lambda\langle\chi_{n,k}|\mathcal{H}_1|\chi_{n,k}\rangle + \lambda^2\langle\chi_{n,k}|\mathcal{H}_1|\psi_{n,k}^1\rangle + \cdots}{1 + \lambda a_{nk,nk}^{(1)} + \lambda^2 a_{nk,nk}^{(2)} + \cdots}. \tag{H.53}$$

Expanding Eq. (H.53) in powers of λ we obtain

$$E_{n,k}(\lambda) = E_n^0 + \lambda\langle\chi_{n,k}|\mathcal{H}_1|\chi_{n,k}\rangle + \lambda^2[\langle\chi_{n,k}|\mathcal{H}_1|\psi_{n,k}^1\rangle - \langle\chi_{n,k}|\mathcal{H}_1|\chi_{n,k}\rangle a_{nk,nk}^{(1)}] + \cdots . \tag{H.54}$$

This expression defines the energy shifts order by order,

$$E_{n,k}(\lambda) \equiv E_n^0 + \lambda\Delta E_{n,k}^{(1)} + \lambda^2\Delta E_{n,k}^{(2)} + \lambda^3\Delta E_{n,k}^{(3)} + \cdots . \tag{H.55}$$

H.3.0.1 First order

Comparing Eqs. (H.54) and (H.55) we obtain for the first-order shift,

$$\Delta E_{n,k}^{(1)} = \langle \chi_{n,k} | \mathcal{H}_1 | \chi_{n,k} \rangle . \tag{H.56}$$

Note that this expression coincides with the expression for the first-order energy shift given by Eq. (H.22) from non-degenerate perturbation theory. The coupling to other manifolds does not enter the expression; only matrix elements diagonal in the quantum number n appear. In this sense, the quantum number n is *always* conserved to *first* order in perturbation theory and n is called a "good quantum number to *first* order in perturbation theory".

H.3.0.2 Second order

Proceeding to second order we obtain with the aid of Eq. (H.49)

$$\Delta E_{n,k}^{(2)} = \sum_{m,l} \langle \chi_{n,k} | \mathcal{H}_1 | \psi_{m,l}^0 \rangle \langle \psi_{m,l}^0 | \psi_{n,k}^1 \rangle - \langle \chi_{n,k} | \mathcal{H}_1 | \chi_{n,k} \rangle a_{nk,nk}^{(1)} . \tag{H.57}$$

This is the analogue of (H.24). Separating the term with $(m = n,\, l = k)$ from the summation it cancels against the subtracted term and we obtain

$$\Delta E_{n,k}^{(2)} = \sum_{m,l}{}' \langle \chi_{n,k} | \mathcal{H}_1 | \phi_{m,l} \rangle a_{ml,nk}^{(1)} , \tag{H.58}$$

where the prime indicates the summation constraint $(m \neq n,\, l \neq k)$, which implies $|\psi_{m,l}^0\rangle = |\phi_{m,l}\rangle$ as follows from (H.50). As Eq. (H.58) does not depend on $a_{nk,nk}^{(1)}$ we may choose $a_{nk,nk}^{(1)} = 0$.

To determine the other coefficients $a_{ml,nk}^{(1)}$ we consider Eq. (H.47) for $(\nu, \mu) = (m, l) \neq (n, k)$

$$E_{n,k}(\lambda) = E_m^0 + \frac{\lambda \langle \psi_{m,l}^0 | \mathcal{H}_1 | \chi_{n,k} \rangle + \lambda^2 \langle \psi_{m,l}^0 | \mathcal{H}_1 | \psi_{n,k}^1 \rangle + \cdots}{\lambda a_{ml,nk}^{(1)} + \lambda^2 a_{ml,nk}^{(2)} + \cdots} . \tag{H.59}$$

Replacing the l.h.s. by the expansion (H.55) and collecting the terms of first order in λ we obtain

$$(E_n^0 - E_m^0) a_{ml,nk}^{(1)} = \langle \psi_{m,l}^0 | \mathcal{H}_1 | \chi_{n,k} \rangle . \tag{H.60}$$

Here we distinguish two cases:

- for $(m = n,\, l \neq k)$ we have
$$\langle \chi_{n,l} | \mathcal{H}_1 | \chi_{n,k} \rangle = 0, \tag{H.61}$$
independent of the value of $a_{nl,nk}^{(1)}$; this provides us with the freedom to choose $a_{nl,nk}^{(1)} = 0$ for given n, k and all values of l.
- for $m \neq n$ we obtain, irrespective of the values of l and k,
$$a_{ml,nk}^{(1)} = \frac{\langle \phi_{m,l} | \mathcal{H}_1 | \chi_{n,k} \rangle}{E_n^0 - E_m^0} . \tag{H.62}$$

Substituting these expressions for $a_{ml,nk}^{(1)}$ the expression for the second order shift reduces to

$$\Delta E_{n,k}^{(2)} = \sum_{m,l}{}' \frac{|\langle \phi_{m,l} | \mathcal{H}_1 | \chi_{n,k} \rangle|^2}{E_n^0 - E_m^0} , \tag{H.63}$$

where the prime indicates the summation constraint $m \neq n$. Note that Eq. (H.63) is formally identical to the result (H.29) obtained with ordinary perturbation theory because summing over both m and l represents the summation over all states non-degenerate with the manifold $\{\chi_{n,k}\}$.

H.3.0.3 Relation with the unperturbed eigenstates

What remains to be done is to express the orthonormal set $\{|\chi_{n,k}\rangle\}$ in terms of the (known) set $\{|\phi_{n,k}\rangle\}$. From Eqs. (H.56) and (H.61) we find that this set diagonalizes the degenerate subspace,

$$\langle\chi_{n,l}|\mathcal{H}_1|\chi_{n,k}\rangle = \Delta E^{(1)}_{n,k}\delta_{kl}, \tag{H.64}$$

but in general not the rest of the Hilbert space - see (H.62),

$$\langle\phi_{m,l}|\mathcal{H}_1|\chi_{n,k}\rangle = (E^0_n - E^0_m)a^{(1)}_{ml,nk} \neq 0 \quad \text{for } m \neq n. \tag{H.65}$$

Hence, the orthonormal set $\{|\chi_{n,k}\rangle\}$ represents a basis for the degenerate subspace. Only when the operators $\mathcal{H}_0$ and $\mathcal{H}_1$ commute they can share a complete basis for the full Hilbert space (see Problem F.1). Thus, *restricting ourselves to the degenerate subspace* we may write

$$\mathcal{H}_1|\chi_{n,k}\rangle = \varepsilon_k|\chi_{n,k}\rangle \quad \text{for } k \in \{1, \cdots, g\}, \tag{H.66}$$

where we use the shorthand notation $\varepsilon_k \equiv \Delta E^{(1)}_{n,k}$. The eigenstates of the (unknown) set $\{|\chi_{n,k}\rangle\}$ and the (known) set $\{|\phi_{n,k}\rangle\}$ are related by a basis transformation,

$$|\chi_{n,k}\rangle = \sum_j c_{jk}|\phi_{n,j}\rangle \quad \text{with } \sum_j |c_{jk}|^2 = 1, \tag{H.67}$$

where the coefficients $c_{jk} \equiv \langle\phi_{n,j}|\chi_{n,k}\rangle$ are yet to be determined. These coefficients follow in three steps. First we project both sides of the Schrödinger equation for the degenerate subspace as given by Eq. (H.66) onto an arbitrary state $|\phi_{n,i}\rangle$ from the (known) set $\{|\phi_{n,k}\rangle\}$,

$$\langle\phi_{n,i}|\mathcal{H}_1|\chi_{n,k}\rangle = \varepsilon_k\langle\phi_{n,i}|\chi_{n,k}\rangle. \tag{H.68}$$

In the second step we replace $|\chi_{n,k}\rangle$ in Eq. (H.68) by the superposition (H.67) and obtain a set of g equations with g unknowns, one for each possible choice of $i \in \{1, \cdots, g\}$,

$$\sum_j c_{jk}[\langle\phi_{n,i}|\mathcal{H}_1|\phi_{n,j}\rangle - \varepsilon_k\delta_{ij}] = 0. \tag{H.69}$$

The third step is to solve these equations, which is possible if and only if the secular equation is satisfied,

$$\det\left|H'_{ij} - \varepsilon\,\delta_{ij}\right| = 0, \tag{H.70}$$

where

$$H'_{ij} = \langle\phi_{n,i}|\mathcal{H}_1|\phi_{n,j}\rangle \tag{H.71}$$

is the coupling matrix. Eq. (H.70) represents a polynomial of order g in powers of ε and is called the characteristic polynomial (secular equation) of the set of equations. The set of equations has g solutions for the set of coefficients $\{c_{1k}, \cdots c_{gk}\}$, each corresponding to one of the roots $\varepsilon_k \in \{\varepsilon_1, \cdots, \varepsilon_g\}$ of the characteristic polynomial. Finding these solutions amounts to the diagonalization of a $g \times g$ matrix eigenvalue problem, well-known from linear algebra. The roots correspond to the eigenvalues ε_k and the corresponding eigenvectors define the coefficients $\{c_{1k}, \cdots c_{gk}\}$. This completes our task of expressing $|\chi_{n,k}\rangle$ in terms of the (known) eigenstates $\{|\phi_{n,i}\rangle\}$ of the degenerate subspace of level n.

H.3.1 Reduction to ordinary first-order perturbation theory

Interestingly, in many practical cases involving degenerate levels we can calculate the perturbation shifts using the simple first-order expression form ordinary perturbation theory. In such cases it is said that "perturbation theory for degenerate levels *reduces to ordinary perturbation theory*". To introduce this topic we consider the special case in which we can select a basis $\{|\phi_{n,k}\rangle\}$ in which not only $\mathcal{H}_0$ but also $\mathcal{H}_1$ is diagonal,

$$\mathcal{H}_0|\phi_{n,k}\rangle = E_n^0|\phi_{n,k}\rangle \tag{H.72a}$$

$$\mathcal{H}_1|\phi_{n,k}\rangle = \Delta E_{n,k}|\phi_{n,k}\rangle. \tag{H.72b}$$

The existence of a basis $\{|\phi_{n,k}\rangle\}$ with the given properties is assured if the operators $\mathcal{H}_0$ and $\mathcal{H}_1$ commute (see Problem F.1). Obviously, in this case there is no need for explicit diagonalization because the full Hamiltonian is already diagonal, with the quantum number n associated with $\mathcal{H}_0$ and k with $\mathcal{H}_1$. This situation arises when, in the modeling of a physical system, a new feature is introduced; e.g., an external field. When the new feature involves a new degree of freedom (e.g., spin) the Hilbert space has to be expanded accordingly, turning the eigenstates $|\phi_n\rangle$ of $\mathcal{H}_0$ into degenerate manifolds $\{|\phi_{n,k}\rangle\}$ of $\mathcal{H}_0$. The energy shifts associated with the new feature are captured by the term $\mathcal{H}_1$ in the Hamiltonian. As the new Hamiltonian $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$ is already diagonal in the basis $\{|\phi_{n,k}\rangle\}$ we can restrict ourselves to calculating the energy shifts caused by the term $\mathcal{H}_1$,

$$\Delta E_{n,k} = \Delta E_{n,k}^{(1)} + \Delta E_{n,k}^{(2)} + \cdots \quad \text{for } k \in \{1, \cdots, g\}, \tag{H.73}$$

where

$$\Delta E_{n,k}^{(1)} = \langle\phi_{n,k}|\mathcal{H}_1|\phi_{n,k}\rangle \tag{H.74}$$

and

$$\Delta E_{n,k}^{(2)} = \sum_{m,l}{}' \frac{|\langle\phi_{m,l}|\mathcal{H}_1|\phi_{n,k}\rangle|^2}{E_n^0 - E_m^0} = 0. \tag{H.75}$$

Thus, the shift is given by the expression for "ordinary" *first* order perturbation theory because all terms of higher order ($\Delta E_{n,k}^{(2)}$, $\Delta E_{n,k}^{(3)}, \cdots$) vanish by orthogonality of the basis states $\{|\phi_{n,k}\rangle\}$, as follows directly by substituting (H.72b) into (H.75). This is of course as it should be because if the full Hamiltonian is already diagonal, we have the exact solution at our disposal and there is nothing that perturbation theory can add to this.

In view of the above, it is good practice to first check the commutation properties of $\mathcal{H}_0$ and $\mathcal{H}_1$ and search (in case $[\mathcal{H}_0, \mathcal{H}_1] = 0$) for a basis that is diagonal for all quantum numbers. Only if this search is unsuccessful we have to turn to perturbation theory for degenerate levels and diagonalize the perturbation matrix explicitly. Interestingly, in important cases (e.g., spin-orbit coupling) the commutation of the operators $\mathcal{H}_0$ and $\mathcal{H}_1$, although not exact, is close to perfect. In these cases we can take advantage of the rule that n is *always* a "good quantum number to *first* order in perturbation theory" and use the basis (H.50) for which $\mathcal{H}_1$ is only diagonal in the manifold of interest; i.e., $\mathcal{H}_1|\chi_{n,k}\rangle = \Delta E_{n,k}|\chi_{n,k}\rangle$. Whether first-order perturbation theory yields a sufficiently good approximation can only be justified by verifying that the higher-order shifts are sufficiently small. Fortunately, in many cases a rough estimate suffices to convince ourselves that this requirement is satisfied.

Comment A: the level splitting by the orbital Zeeman interaction $\mathcal{H}_Z$ is an example that can be analyzed with first-order perturbation theory in the basis of the Schrödinger Hamiltonian $\mathcal{H}_0$. This is possible because $[\mathcal{H}_0, \mathcal{H}_Z] = 0$. In this case the exact result coincides with the first-order expression from ordinary perturbation theory.

Comment B: the level shifts by the relativistic mass correction, the spin-orbit interaction and the hyperfine interaction are examples of cases where *full* commutation with the perturbation is absent but the Hamiltonian is close to diagonal in the basis of the subspace of interest because the principal quantum number is *always* a "good quantum number to *first* order in perturbation theory". Hence, this is a good approximation as long as the second- and higher-order contributions are negligible.

H.4 Example: the two-fold degenerate case

In this section we turn to the special case of a 2-fold *degenerate* manifold of two *bound states* $|a\rangle$ and $|b\rangle$ with primary quantum number n, unperturbed energy E_n^0 and satisfying the Schrödinger equation

$$\mathcal{H}_0 \left|\phi_{n,i}\right\rangle = E_n^0 \left|\phi_{n,i}\right\rangle \text{ with } \left|\phi_{n,i}\right\rangle \in \{|a\rangle, |b\rangle\}. \tag{H.76}$$

For simplicity the unperturbed states states $|a\rangle$ and $|b\rangle$ are presumed to be orthonormal. The full Hamiltonian $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$ satisfies the Schrödinger equation

$$\mathcal{H} \left|\psi_{n,k}\right\rangle = E_{n,k} \left|\psi_{n,k}\right\rangle, \tag{H.77}$$

where $E_{n,k} = E_n^0 + \Delta E_k$, with ΔE_k the energy shift of level $k \in \{1, 2\}$. In the perturbation theory for degenerate levels the first-order approximation to the level shift, $\Delta E_k \simeq \varepsilon_k$, is obtained by diagonalization of the degenerate subspace with respect to the perturbation $\mathcal{H}_1$. In this approximation the problem reduces to a two-level system as discussed in Appendix G.2.

I

Variational Methods

I.1 Introduction

I.1.1 Fundamental theorem

The fundamental theorem of variational calculus states that the equation $M(x) = 0$ holds if the integral $I = \int_a^b M(x)h(x)dx$ vanishes for *any* smooth function $h(x)$. By smooth we mean in this context that $h(x)$, $h'(x)$ and $h''(x)$ are continuous.

I.1.2 Extremal values of a continuous function

In physics we often have to search for a *maximum* or *minimum* (i.e., *extremum*) of some function. So, let $f(\mathbf{v})$ be a *continuous* function of n variables, represented by the *real* vector $\mathbf{v} = (v_1, \cdots, v_n)$, defined on the *domain* $\mathbf{v} \in V$. In general $f(\mathbf{v})$ will not be bounded; i.e., have no extreme points. Therefore, we shall suppose that $f(\mathbf{v})$ has an extreme value at the point $\mathbf{v}_0$. Three types of extrema can be distinguished: (a) *critical* points (if $\partial_{\mathbf{v}} f(\mathbf{v})|_{\mathbf{v}=\mathbf{v}_0} = 0$); (b) *singular* points (if $\partial_{\mathbf{v}} f(\mathbf{v})$ does *not* exist at $\mathbf{v}_0$); (c) *boundary* points (if $\mathbf{v}_0$ is a point at the boundary of the domain V). Here $\partial_{\mathbf{v}}$ represents the gradient with respect to the variables $v_1, \cdots, v_n$. Inversely, if we have identified a critical, singular or boundary point, this is not automatically an extreme point. The point can also be a *local* rather than a *global* extremum or some kind of *saddle* point. In such cases further inspection will be necessary and may be labor intensive.

I.1.3 Extremal values of a function in the presence of constraints - Lagrange multipliers

In important cases we have to search for the extrema of the function $f(\mathbf{v})$ in the presence of one or more *constraints*. Such a constraint can be an inequality (e.g., restricting the domain of definition of the function) or an equality (e.g., a normalization condition). In the latter case the number of independent variables is reduced, which means that in principle one of the variables could be eliminated but for many types of constraints this turns out to be impractical.

For *smooth* functions we can turn to the method of *Lagrange multipliers*. To introduce this method we consider again the function $f(\mathbf{v})$ of n variables, represented by the *real* vector $\mathbf{v} = (v_1, \cdots, v_n)$, defined on the *domain* $\mathbf{v} \in V$ and subject to $m \leq n-1$ equality constraints

$$g_j(\mathbf{v}) = 0, \quad (1 \leq j \leq m) . \tag{I.1}$$

If $f(\mathbf{v})$ has a *critical* point at $\mathbf{v}_0$ we can define a function of $n+m$ variables,

$$\mathcal{L}(\mathbf{v}, \lambda_1, \cdots \lambda_m) = f(\mathbf{v}) + \lambda_1 g_1(\mathbf{v}) + \cdots + \lambda_m g_m(\mathbf{v}), \tag{I.2}$$

with the property that $\mathcal{L}(\mathbf{v}, \lambda_1, \cdots \lambda_m)$ has a *critical* point at $(\mathbf{v}_0, \lambda_1, \cdots \lambda_m)$ provided $\partial_{\mathbf{v}} g_j(\mathbf{v}_0) \neq 0$ for $1 \leq j \leq m$. The function $\mathcal{L}(\mathbf{v}, \lambda_1, \cdots \lambda_m)$ is called the Lagrangian of the function $f(\mathbf{v})$ under the constraints $g_j(\mathbf{v}) = 0$ and $\lambda_1, \cdots \lambda_m$ are called *Lagrange multipliers.* Any critical point must satisfy the set of $n + m$ equations

$$\frac{\partial \mathcal{L}}{\partial v_i} = 0 \quad (1 \leq i \leq n) \tag{I.3}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_j} = 0 \quad (1 \leq j \leq m)\,. \tag{I.4}$$

I.2 Rayleigh-Ritz variational principle

In this section we introduce the variational principle for the determination of eigenvalues and eigenfunctions of a quantum mechanical system. According to this principle any state $|\psi\rangle$ in the Hilbert space of the Hamiltonian $\mathcal{H}$ for which the expectation value

$$E[\psi] = \frac{\langle\psi|\mathcal{H}|\psi\rangle}{\langle\psi|\psi\rangle} \tag{I.5}$$

is stationary under *arbitrary functional variations* of the state $|\psi\rangle$, is an eigenstate of $\mathcal{H}$ with eigenvalue $E[\psi]$.

To derive this theorem we consider the change of the energy functional $E[\psi] = \langle\psi|\mathcal{H}|\psi\rangle$ under an arbitrary variation $\psi \to \psi + \delta\psi$

$$\delta E = E[\psi + \delta\psi] - E[\psi] = \frac{\langle\psi + \delta\psi|\mathcal{H}|\psi + \delta\psi\rangle - E\langle\psi + \delta\psi|\psi + \delta\psi\rangle}{\langle\psi + \delta\psi|\psi + \delta\psi\rangle}. \tag{I.6}$$

Expanding this expression up to lowest order in the variation we obtain with the aid of Eq. (I.5)

$$\begin{aligned}\langle\psi|\psi\rangle\delta E + \cdots &= \langle\psi|\mathcal{H}|\psi\rangle + \langle\delta\psi|\mathcal{H}|\psi\rangle + \langle\psi|\mathcal{H}|\delta\psi\rangle - E\langle\psi|\psi\rangle - E\langle\delta\psi|\psi\rangle - E\langle\psi|\delta\psi\rangle + \cdots \\ &= \langle\delta\psi|\mathcal{H} - E|\psi\rangle + \langle\psi|\mathcal{H} - E|\delta\psi\rangle + \cdots\,. \end{aligned} \tag{I.7}$$

Hence, for $E[\psi]$ to be stationary ($\delta E = 0$) under real variations of the state vector we require

$$\langle\psi|\psi\rangle\delta E = \langle\delta\psi|\mathcal{H} - E|\psi\rangle + \langle\psi|\mathcal{H} - E|\delta\psi\rangle = 0. \tag{I.8}$$

As $\mathcal{H}$ is an hermitian operator this condition can be written in the form

$$\langle\delta\psi|\mathcal{H} - E|\psi\rangle + \langle\delta\psi|\mathcal{H} - E|\psi\rangle^* = 0. \tag{I.9}$$

This shows that the variation of the bra $\langle\psi|$ gives rise to a term which is the complex conjugate of the term arising from the variation of the ket $|\psi\rangle$, which has the convenient consequence that if the energy is stationary under variation of $\langle\psi|$, the condition for being stationary under variation of $|\psi\rangle$ is simultaneously satisfied. Hence, we can restrict the variational procedure to the variation of *either* $\langle\psi|$ *or* $|\psi\rangle$. To complete the derivation we note that the condition

$$\langle\delta\psi|\mathcal{H} - E|\psi\rangle = 0 \text{ under arbitrary variation of } \psi \tag{I.10}$$

can only be satisfied if $|\psi\rangle$ is a stationary state of the system with eigenvalue E,

$$\mathcal{H}|\psi\rangle = E|\psi\rangle. \tag{I.11}$$

Eq. (I.5) is sometimes referred to as the *Rayleigh ratio* as it can be identified with the Lagrange multiplier for the normalization constraint $\langle\psi|\psi\rangle = \mathcal{N}$ in the lagrangian variation of the functional $f[\psi] = \langle\psi|\mathcal{H}|\psi\rangle$ - see Problem I.1. The variational principle was introduced by Rayleigh in 1873 [91] for the identification of the eigenmodes of acoustical systems. The method was further developed by Ritz in 1909 [92].

Problem I.1. Show that the Rayleigh ratio (I.5) can be identified with the Lagrange multiplier for the normalization constraint $\langle\psi|\psi\rangle = \mathcal{N}$ of the functional $f[\psi] = \langle\psi|\mathcal{H}|\psi\rangle$.

Solution. The Lagrangian for the variation of $f[\psi] = \langle\psi|\mathcal{H}|\psi\rangle$ under the constraint $\langle\psi|\psi\rangle = 1$ is given by

$$\mathcal{L} = \langle\psi|\mathcal{H}|\psi\rangle + \lambda(\langle\psi|\psi\rangle - \mathcal{N}).$$

This Lagrangian is stationary if

$$\delta\mathcal{L} = \langle\psi + \delta\psi|\mathcal{H}|\psi + \delta\psi\rangle + \lambda\langle\psi + \delta\psi|\psi + \delta\psi\rangle = 0,$$

which is satisfied for

$$-\lambda = \frac{\langle\psi + \delta\psi|\mathcal{H}|\psi + \delta\psi\rangle}{\langle\psi + \delta\psi|\psi + \delta\psi\rangle} \underset{\delta\psi\to 0}{=} \frac{\langle\psi|\mathcal{H}|\psi\rangle}{\langle\psi|\psi\rangle} = E[\psi]. \qquad \square$$

I.2.1 Estimating the ground state energy

The "arbitrary" variations considered in the variational principle remain meaningless unless the eigenmodes of the system are known. So what to do if the eigenstates of a quantum mechanical system are unknown? In the present section we show that the ground state energy of a quantum mechanical system can be found or at least approximated by cleverly choosing a *trial* wavefunction in which we try to include all our knowledge (or intuition) about the system, for instance by imposing a known symmetry. To demonstrate the procedure we consider a physical system described by the Hamiltonian $\mathcal{H}$, for which a complete set of (unknown) eigenstates $\{|\psi_n\rangle\}$ satisfy the Schrödinger equation

$$\mathcal{H}|\psi_n\rangle = E_n|\psi_n\rangle. \tag{I.12}$$

Here $E_0, E_1, E_2, \cdots$ are the eigenvalues in order of growing energy, E_0 being the energy of the (non-degenerate) ground state $|\psi_0\rangle$. Let us denote the trial state of the system by $|\chi\rangle$. The corresponding trial wavefunction $\chi(\mathbf{r})$ is given a functional form (preferably norm conserving) involving on one or more coefficients $\alpha, \beta, \gamma, \cdots$ which serve as *variational parameters*. In general $|\chi\rangle$ will not be an eigenstate of $\mathcal{H}$ and is not necessarily normalized. Expressing $|\chi\rangle$ as a linear combination of the basis states $\{|\psi_n\rangle\}$ we have

$$|\chi(\alpha, \beta, \gamma, \cdots)\rangle = \sum_n a_n(\alpha, \beta, \gamma, \cdots)|\psi_n\rangle. \tag{I.13}$$

Since $E_n > E_0$, with $n > 0$, we can derive the following inequality for the expectation value of $\mathcal{H}$,

$$\langle\chi|\mathcal{H}|\chi\rangle = \sum_{mn} a_m^* a_n \langle\psi_m|\mathcal{H}|\psi_n\rangle = \sum_n |a_n|^2 E_n \geq E_0 \sum_n |a_n|^2. \tag{I.14}$$

With the aid of this inequality we obtain for the energy of the system

$$E(\alpha, \beta, \gamma, \cdots) = \frac{\langle\chi|\mathcal{H}|\chi\rangle}{\langle\chi|\chi\rangle} \geq E_0. \tag{I.15}$$

This expression is intuitively clear: an arbitrary state will in general not be the ground state and therefore have an energy larger than E_0. The better we succeed in capturing the properties of the ground state in the trial state the lower will be the energy and the better will be our estimate of E_0. Hence, this important theorem can be used to derive an upper bound for the ground-state energy of the system. By variation of the parameters $\alpha, \beta, \gamma, \cdots$ we can minimize the energy of our trial state. The minimum is obtained if the following equations are satisfied

$$\frac{\partial E(\alpha, \beta, \gamma, \cdots)}{\partial\alpha} = \frac{\partial E(\alpha, \beta, \gamma, \cdots)}{\partial\beta} = \frac{\partial E(\alpha, \beta, \gamma, \cdots)}{\partial\gamma} = \cdots = 0. \tag{I.16}$$

In view of the inequality (I.15) the minimal value must be larger or equal than E_0. Accordingly, simultaneously imposing the conditions (I.16) is not equivalent to an *arbitrary* variation of $|\psi_0\rangle$. In other words, minimizing the energy by variation of the trial wavefunction does not lead us necessarily to the exact state. However, the procedure neither excludes this possibility. In Problem I.2 we demonstrate the method by obtaining an upper bound for the ground state energy of a harmonic oscillator.

Problem I.2. The Hamiltonian of the 1D harmonic oscillator is given by

$$\mathcal{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2.$$

Determine an upper bound for the ground state energy of the 1D harmonic oscillator using the variational trial function $\chi(x) = \exp[-\frac{1}{2}\alpha x^2]$. Compare this value with the exact value.

Solution. Applying the inequality (I.15) the energy of the trial state $\chi(x)$ satisfies the relation

$$E(\alpha) = \frac{\int \exp[-\frac{1}{2}\alpha x^2]\mathcal{H}\exp[-\frac{1}{2}\alpha x^2]dx}{\int \exp[-\alpha x^2]dx} \geq E_0.$$

Substituting the expression for $\mathcal{H}$ and evaluating the integrals we obtain

$$E(\alpha) = -\frac{\hbar^2}{2m\alpha}\frac{\int_0^\infty \exp[-\frac{1}{2}x^2]\frac{d^2}{dx^2}\exp[-\frac{1}{2}x^2]dx}{\int_0^\infty \exp[-x^2]dx} + \tfrac{1}{2}m\omega^2\alpha\frac{\int_0^\infty x^2\exp[-x^2]dx}{\int_0^\infty \exp[-x^2]dx} = \frac{\hbar^2}{4m\alpha} + \tfrac{1}{4}m\omega^2\alpha \geq E_0.$$

In this case we have only one variational parameter. Varying $E(\alpha)$ with respect to α the stationary value is reached for

$$\frac{\partial E}{\partial \alpha} = -\frac{\hbar^2}{4m\alpha^2} + \tfrac{1}{4}m\omega^2 = 0,$$

which implies $\alpha = \hbar/m\omega$. It is easily verified that the stationary value corresponds to the *minimum* of the function $E(\alpha)$ and substituting the value $\alpha = \hbar/m\omega$ we obtain

$$E = \tfrac{1}{2}\hbar\omega.$$

This happens to the be the exact value because our trial function is of the same form as the true ground state. Any other choice of trial function will yield $E(\alpha) > \frac{1}{2}\hbar\omega$. □

I.3 Variational method for degenerate states

In many cases we have to deal with the phenomenon of *lifting of a degeneracy* by some form of symmetry breaking in the Hamiltonian $\mathcal{H}$. To discuss this phenomenon we consider a set with g degenerate states which we shall denote by $|\chi_i\rangle$, with $i \in \{1, \cdots, g\}$. The states $|\chi_i\rangle$ need not be orthogonal or normalized. Because the states are degenerate we have

$$E^0 = \frac{\langle\chi_i|\,\mathcal{H}_0\,|\chi_i\rangle}{\langle\chi_i|\chi_i\rangle} \text{ for } i \in \{1, \cdots, g\}, \tag{I.17}$$

where $\mathcal{H}_0$ is the Hamiltonian without symmetry breaking. To analyze how the degeneracy of the states $|\chi_i\rangle$ is lifted we choose a trial wavefunction in which g variational parameters $c_1, c_2, \cdots, c_g$ are chosen as the coefficients of a linear combination of the type

$$|\psi_k\rangle = \sum_{i=1}^{g} c_i\,|\chi_i\rangle\,. \tag{I.18}$$

The energy of the trial wave function is given by

$$E_k = \frac{\langle\psi_k|\,\mathcal{H}\,|\psi_k\rangle}{\langle\psi_k|\psi_k\rangle} = \frac{\sum_i\sum_j c_i^* c_j\,\langle\chi_i|\,\mathcal{H}\,|\chi_j\rangle}{\sum_i\sum_j c_i^* c_j\,\langle\chi_i|\,\chi_j\rangle}. \tag{I.19}$$

Using the rule that it suffices to vary either the state vector or its adjoint (see Section I.2) we differentiate this expression with respect to c_i^* and obtain

$$\frac{\partial E_k}{\partial c_i^*} = \frac{\left(\sum_i \sum_j c_i^* c_j\right) \sum_j c_j H_{i,j} - \left(\sum_i \sum_j c_i^* c_j H_{ij}\right) \sum_j c_j S_{ij}}{\left(\sum_{i,j} c_i^* c_j S_{ij}\right)^2} = \frac{\sum_j c_j (H_{ij} - E S_{ij})}{\sum_{i,j} c_i^* c_j S_{ij}}. \tag{I.20}$$

Here we defined

$$H_{ij} = \langle \chi_i | \mathcal{H} | \chi_j \rangle \tag{I.21}$$

as well as the quantity

$$S_{ij} = \langle \chi_i | \chi_j \rangle , \tag{I.22}$$

called the *overlap integral*. The energy E_k is stationary under variation of c_i^* when $\partial E/\partial c_i^* = 0$, which is the case for

$$\sum_j c_j (H_{ij} - E S_{i,j}) = 0. \tag{I.23}$$

Repeating this procedure for all c_i^* with $i \in \{1, \cdots, g\}$ we obtain g simultaneous equations for the g unknown variables c_j. This set of equations only has non-trivial solutions when the so-called *secular determinant* vanishes,

$$\det |H_{ij} - E S_{ij}| = 0. \tag{I.24}$$

This expression is called the *secular equation* and represents a polynomial of power g. The solutions define g energy levels, E_i with $i \in \{1, \cdots, g\}$, which are stationary with respect to small variations of the coefficients of the corresponding trial states.

I.3.1 Lifting of degeneracy by a small symmetry-breaking term

As a first special case we discuss the situation where the Hamiltonian $\mathcal{H}$ can be broken up in two hermitian parts,

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1, \tag{I.25}$$

where the Schrödinger equation $\mathcal{H}_0 |\phi_i\rangle = E^0 |\phi_i\rangle$, with $i \in \{1, \cdots, g\}$, defines a g-fold degenerate manifold of levels corresponding to an orthonormal set of g eigenstates $\{|\phi_i\rangle\}$. The term $\mathcal{H}_1$ is presumed to break the symmetry and is here taken to be *small*.

Using the formalism introduced above we set $|\chi_i\rangle = |\phi_i\rangle$ and since the $\{|\phi_i\rangle\}$ are eigenstates of $\mathcal{H}_0$ we have zero overlap between different states,

$$S_{ij} = \langle \phi_i | \phi_j \rangle = \delta_{i,j}. \tag{I.26}$$

Further, we have

$$H_{i,j} = \langle \phi_i | \mathcal{H}_0 | \phi_j \rangle + \langle \phi_i | \mathcal{H}_1 | \phi_j \rangle = E^0 \delta_{ij} + \mathcal{H}'_{ij}, \tag{I.27}$$

where we defined

$$\mathcal{H}'_{i,j} = \langle \phi_i | \mathcal{H}_1 | \phi_j \rangle . \tag{I.28}$$

Substituting Eqs. (I.26) and (I.27) into the secular equation (I.24) we obtain

$$\det \left| \mathcal{H}'_{i,j} - \varepsilon_1 \delta_{i,j} \right| = 0, \tag{I.29}$$

where $\varepsilon_1 = E - E^0$. This equation has g solutions, $\varepsilon_1 = \Delta E_i$ with $i \in \{1, \cdots, g\}$, which correspond to the energy shifts of the levels E_i with respect to the degenerate value E^0,

$$E_i = E^0 + \Delta E_i \text{ with } i \in \{1, \cdots, g\}. \tag{I.30}$$

Comment: This result coincides with first-order perturbation theory for a degenerate level and is further discussed in Section H.3.

I.3.2 Variation method applied to two degenerate states

As a second special case we discuss the situation of two normalized real states $|a\rangle$ and $|b\rangle$, not necessarily orthogonal, and of the same energy.

- Because the states are degenerate we have

$$\mathcal{H}_{00} \equiv \mathcal{H}_{aa} = \langle a|\, \mathcal{H}\, |a\rangle = \langle b|\, \mathcal{H}\, |b\rangle = \mathcal{H}_{bb}. \tag{I.31}$$

- Because the states are real and $\mathcal{H}$ is hermitian

$$V \equiv \mathcal{H}_{ab} = \langle a|\, \mathcal{H}\, |b\rangle = \langle b|\, \mathcal{H}^{\dagger}\, |a\rangle^{*} = \langle b|\, \mathcal{H}\, |a\rangle = \mathcal{H}_{ba}$$

- Because the states are not necessarily orthogonal we have in general

$$S \equiv S_{ab} = \langle a|b\rangle = \langle b|a\rangle = S_{ba} \neq 0$$

In search for a state of lower energy we construct a trial state

$$|\chi\rangle = a\,|a\rangle + b\,|b\rangle\,. \tag{I.32}$$

Applying the variational principle to the coefficients a and b we obtain the secular equation

$$\begin{vmatrix} \mathcal{H}_{00} - E & V - ES \\ V - ES & \mathcal{H}_{00} - E \end{vmatrix} = 0, \tag{I.33}$$

which can be written as $(\mathcal{H}_{00} - E)^2 = (V - ES)^2$. Solving the quadratic equation yields for the eigenvalues

$$E_{\pm} = \frac{(\mathcal{H}_{00} \pm V)}{(1 \pm S)}. \tag{I.34}$$

The coefficients fixing the eigenstates are found from

$$\langle a|\, \mathcal{H}\, |\chi\rangle = E_{\pm}\, \langle a|\chi\rangle \Leftrightarrow a_{\pm}\mathcal{H}_{00} + b_{\pm}V = E_{\pm}\,(a_{\pm} + b_{\pm}S)\,, \tag{I.35}$$

which can be written as

$$\frac{a_{\pm}}{b_{\pm}} = \frac{E_{\pm}S - V}{\mathcal{H}_{00} - E_{\pm}} = \pm 1. \tag{I.36}$$

As the overlap is non-zero the normalization condition is

$$1 = \langle \chi|\chi\rangle = |a|^2 + |b|^2 + 2abS \tag{I.37}$$

Hence, setting $|a|^2 = |b|^2 = |c|^2$ we find for the coefficients

$$|c_{\pm}|^2 = \frac{1}{2(1 \pm S)}. \tag{I.38}$$

Comment: This variational method is used to calculate the exchange splitting in two-electron atoms (in this case $S = 0$). It is also the basis for the description of bonding and anti-bonding orbitals in quantum chemistry (in this case $S \neq 0$).

J

Clebsch-Gordan coefficients

J.1 Relation with the Wigner $3j$ symbols

The Clebsch-Gordan coefficients $\langle j_1m_1; j_2m_2|JM\rangle$ for the coupling $(j_1 \times j_2)$ of two quantized angular momenta $\mathbf{j}_1$ and $\mathbf{j}_2$ into the total quantized angular momentum $\mathbf{J}$ are related to the Wigner $3j$ symbols,

$$\langle j_1m_1; j_2m_2|JM\rangle \equiv (-1)^{j_1-j_2+M}\sqrt{2J+1}\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix}, \tag{J.1}$$

where $|JM\rangle$ is a state of the coupled angular momentum and $|j_1m_1; j_2m_2\rangle \equiv |j_1m_1\rangle_1 \otimes |j_2m_2\rangle_2$ is a product state. Note that $(-1)^{j_1-j_2+M} = (-1)^{j_2-j_1-M}$ because $j_1 - j_2 + M$ is always an integer. By convention the CGCs with maximum projection of the total angular momentum are non-negative (*sign convention)*,

$$\langle j_1j_1; j_2(J-j_1)|JJ\rangle \geq 0, \tag{J.2}$$

and unity for the stretched state (*stretched-state convention*),

$$\langle j_1j_1; j_2j_2|J_{max}J_{max}\rangle = 1, \tag{J.3}$$

where $J_{max} = j_1 + j_2$.
The Wigner $3j$ symbols can be visualized by the vector diagram shown in Fig. J.1. The properties are listed below for the $3j$ symbol $\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$:

Reality:

- $3j$ symbols are real

Selection rules:

- $3j$ symbols are zero unless the *triangle inequality* $\Delta\,(j_1j_2j_3)$ holds; this is the case if

  $$|j_a - j_b| \leq j_c \leq j_a + j_b \tag{J.4}$$

  with a, b, c representing any cyclic permutation of $1, 2, 3$.

- $3j$ symbols are zero unless $j_1 + j_2 + j_3$ is *integral* and the angular momentum *projection is conserved*,

  $$j_1 + j_2 + j_3 \in \{0, 1, \cdots\} \quad \text{and} \quad m_1 + m_2 + m_3 = 0 \tag{J.5}$$

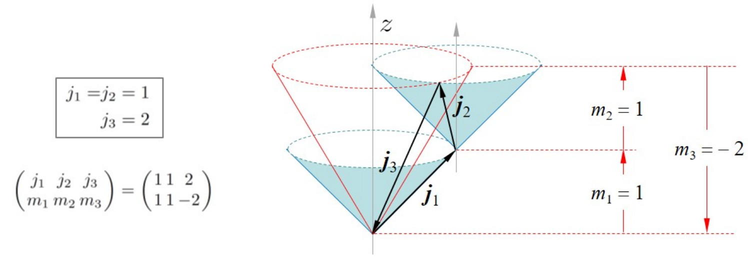


Figure J.1: To establish whether a $3j$ symbol is manifestly zero one can use a vector diagram, shown here for the example $j_1 = j_2 = 1$, $j_3 = 2$. The given $3j$ symbol is valid (*i.e.*, not manifestly zero) because the triangle is *closed* and the projections onto the z axis are quantized in integer values; hence, the triangle inequality is satisfied, $j_1 + j_2 + j_3$ is integral and the projections onto the z axis add up to zero.

Symmetries:

- invariant under cyclic permutation
- multiplied by $(-1)^{j_1+j_2+j_3}$ under interchange of two columns
- multiplied by $(-1)^{j_1+j_2+j_3}$ under simultaneous change of sign, $m_1, m_2, m_3 \to -m_1, -m_2, -m_3$

Orthogonality:

$$\sum_{m_1 m_2} (2j_3+1) \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3' \\ m_1 & m_2 & m_3' \end{pmatrix} = \delta_{j_3,j_3'} \delta_{m_3,m_3'} \tag{J.6}$$

$$\sum_{j_3=|j_1-j_2|}^{j_1+j_2} \sum_{m_3=-j_3}^{j_3} (2j_3+1) \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1' & m_2' & m_3 \end{pmatrix} = \delta_{m_1,m_1'} \delta_{m_2,m_2'} \tag{J.7}$$

In particular:

$$\sum_{m_1} (2j_3+1) \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & (m_3-m_1) & -m_3 \end{pmatrix}^2 = 1 \tag{J.8}$$

$$\sum_{J=0}^{2j} \sum_{M=-J}^{J} (2J+1) \begin{pmatrix} j & j & J \\ m_1 & m_2 & M \end{pmatrix}^2 = 1 \tag{J.9}$$

Even and odd summations:

$$\sum_{J=\text{even}}^{2j} \sum_{M=-J}^{J} (2J+1) \begin{pmatrix} j & j & J \\ m_1 & m_2 & M \end{pmatrix}^2 = \frac{1}{2} \tag{J.10a}$$

$$\sum_{J=\text{odd}}^{2j} \sum_{M=-J}^{J} (2J+1) \begin{pmatrix} j & j & J \\ m_1 & m_2 & M \end{pmatrix}^2 = \frac{1}{2} \tag{J.10b}$$

Racah formula

$$\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix} = (-1)^{j_1-j_2+M}\sqrt{\Delta(j_1j_2J)}$$
$$\times\sqrt{(j_1+m_1)!\,(j_1-m_1)!\,(j_2+m_2)!\,(j_2-m_2)!\,(J+M)!\,(J-M)!}$$
$$\times\sum_t \frac{(-)^t}{t!(J-j_2+t+m_1)!(J-j_1+t-m_2)!(j_1+j_2-J-t)!(j_1-t-m_1)!(j_2-t+m_2)!} \qquad \text{(J.11)}$$

where $t = 0, 1, \cdots$ and the triangle coefficient

$$\Delta(abc) \equiv \frac{(a+b-c)!(b+c-a)!(c+a-b)!}{(a+b+c+1)!} \qquad \text{(J.12)}$$

offers a logical measure for the *triangle inequalities*; these are satisfied for $\Delta(j_1j_2J) > 0$, whereas for $\Delta(j_1j_2J) = 0$ they are violated.

J.1.1 Special cases for given values of J

- J taking its maximum value $J = j_1 + j_2$

$$\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix} = \frac{(-1)^{j_1-j_2+M}}{\sqrt{2J+1}}\sqrt{\frac{(2j_1)!(2j_2)!}{(2J)!}}\sqrt{\frac{(J+M)!\,(J-M)!}{(j_1+m_1)!\,(j_1-m_1)!\,(j_2+m_2)!\,(j_2-m_2)!}} \qquad \text{(J.13)}$$

- the *stretched* case (J, M maximal); this implies $J = j_1 + j_2$ and $M = J$, with $m_1 = j_1$ and $m_2 = j_2$,

$$\begin{pmatrix} j_1 & j_2 & j_1+j_2 \\ m_1 & m_2 & -(j_1+j_2) \end{pmatrix} = \frac{(-1)^{2j_1}}{\sqrt{2(j_1+j_2)+1}}\,\delta_{m_1,j_1}\delta_{m_2,j_2}. \qquad \text{(J.14)}$$

  Since $2j_1$ is always integral, the corresponding Clebsch-Gordan coefficient is always unity,

$$\langle j_1j_1j_2j_2|(j_1+j_2)(j_1+j_2)\rangle = 1. \qquad \text{(J.15)}$$

- $J = 0$:

$$\begin{pmatrix} j & j & 0 \\ m & -m & 0 \end{pmatrix} = (-1)^{j-m}\frac{1}{\sqrt{2j+1}} \Leftrightarrow \langle jm00|j,m\rangle = 1 \qquad \text{(J.16)}$$

- $J = 1/2$ (note that $j + m$ is always an integer):

$$\begin{pmatrix} j & j+\frac{1}{2} & \frac{1}{2} \\ m & -m' & \pm\frac{1}{2} \end{pmatrix} = (-1)^{j-m'-1/2}\sqrt{\frac{j+1\pm m}{2\,(j+1)\,(2j+1)}}\,\delta_{m',m\pm\frac{1}{2}} \qquad \text{(J.17a)}$$

$$\begin{pmatrix} j & j-\frac{1}{2} & \frac{1}{2} \\ m & -m' & \pm\frac{1}{2} \end{pmatrix} = (-1)^{j-m-1}\sqrt{\frac{j\mp m}{2j\,(2j+1)}}\,\delta_{m',m\pm\frac{1}{2}} \qquad \text{(J.17b)}$$

- $J = 1$ (note that $m = 0$ can only occur for integral j and in that case $2j + 1 = \text{odd}$):

$$\begin{pmatrix} j & j & 1 \\ m & -m & 0 \end{pmatrix} = (-1)^{j-m} \frac{m}{\sqrt{j\,(j+1)\,(2j+1)}} \tag{J.18a}$$

$$\begin{pmatrix} j & j & 1 \\ m & -m' & \pm 1 \end{pmatrix} = (-1)^{j-\min(m,m')} \sqrt{\frac{j(j+1) - mm'}{2j(j+1)(2j+1)}}\, \delta_{m',m\pm 1} \tag{J.18b}$$

$$\begin{pmatrix} j+1 & j & 1 \\ m & -m-1 & 1 \end{pmatrix} = (-1)^{-j-m-1} \sqrt{\frac{(j-m)(j-m+1)}{2\,(2j+3)\,(j+1)\,(2j+1)}} \tag{J.18c}$$

$$\begin{pmatrix} j+1 & j & 1 \\ m & -m & 0 \end{pmatrix} = (-1)^{-j-m-1} \sqrt{\frac{(j+m+1)(j-m+1)}{(2j+3)\,(j+1)\,(2j+1)}} \tag{J.18d}$$

- $J = 2$:

$$\begin{pmatrix} j & j & 2 \\ m & -m & 0 \end{pmatrix} = (-1)^{j-m} \frac{3m^2 - j\,(j+1)}{\sqrt{j(j+1)(2j+3)\,(2j+1)\,(2j-1)}} \tag{J.19a}$$

$$\begin{pmatrix} j & j & 2 \\ m & -m' & \pm 1 \end{pmatrix} = (-1)^{j-m}\,(1 \pm 2m) \sqrt{\frac{3}{2}\frac{j(j+1) - mm'}{j(j+1)(2j+3)\,(2j+1)\,(2j-1)}}\, \delta_{m',m\pm 1} \tag{J.19b}$$

$$\begin{pmatrix} j & j & 2 \\ m & -m' & \pm 2 \end{pmatrix} = (-1)^{j-m} \sqrt{\frac{3}{2}\frac{(j^2 - 1 - mm')\,(j(j+2) - mm')}{j(j+1)(2j+3)\,(2j+1)\,(2j-1)}}\, \delta_{m',m\pm 2} \tag{J.19c}$$

- *Shell summation:*

$$\begin{pmatrix} l & k & l \\ 0 & 0 & 0 \end{pmatrix} \sum_{m=-l}^{l} (-1)^m \begin{pmatrix} l & k & l \\ -m & 0 & m \end{pmatrix} = \delta_{k,0} \tag{J.20}$$

J.1.2 Special cases for integer values of l

- $m_1 = m_2 = m_3 = 0$ and $l_1 + l_2 + l_3$ is *odd*

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} = 0 \tag{J.21}$$

- $m_1 = m_2 = m_3 = 0$ and $l_1 + l_2 + l_3$ is *even* (including zero)

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} = \frac{(-1)^p p! \sqrt{\Delta(l_1 l_2 l_3)}}{(p - l_1)!(p - l_2)!(p - l_3)!}, \tag{J.22}$$

where $p \equiv (l_1 + l_2 + l_3)/2$ and $\Delta(l_1 l_2 l_3)$ is given by Eq. (J.12).

- $l_3 = 0$ and $l, l' = \textit{integer}$, only nonzero if $l + l'$ is even; this implies $l' = l$ and $p \equiv (l + l')/2 = l$

$$\begin{pmatrix} l & l' & 0 \\ 0 & 0 & 0 \end{pmatrix} = (-1)^l \frac{1}{\sqrt{2l+1}}\, \delta_{l',l} \tag{J.23}$$

- $l_3 = 1$ and $l, l' =$ *integer*, only nonzero if $l + l' + 1$ is even; this implies $l' = l \pm 1$ and $p \equiv (l + l' + 1)/2 = \max(l, l')$

$$\begin{pmatrix} l \; l' \; 1 \\ 0 \; 0 \; 0 \end{pmatrix} = (-1)^{\max(l,l')} \sqrt{\frac{\max(l,l')}{(2l+1)(2l'+1)}}\, \delta_{l',l\pm1} \tag{J.24}$$

- $l_3 = 2$ and $l, l' =$ *integer*, only nonzero if $l + l' + 2$ is even; this implies $l' = l,\ l \pm 2$ and $p \equiv (l + l' + 2)/2 = \max(l, l')$

$$\begin{pmatrix} l \; l' \; 2 \\ 0 \; 0 \; 0 \end{pmatrix} = \begin{cases} (-1)^{l+1} \sqrt{\dfrac{l\,(l+1)}{(2l+3)\,(2l+1)\,(2l-1)}} & l' = l \\ (-1)^{l} \sqrt{\dfrac{\max(l,l')}{(2l+1)(2l'+1)}} \left(\dfrac{3}{4} \dfrac{l+l'}{l+l'+1} \right) & l' = l \pm 2. \end{cases} \tag{J.25}$$

Problem J.1. Use the exchange rule for $3j$ symbols to show that the symmetry of Clebsch-Gordan coefficients under exchange of the coupled angular momenta is given by

$$\langle j m_1 j_2 m_2 | JM \rangle = (-1)^{j_1 + j_2 - J} \langle j_2 m_2; j_1 m_1 | JM \rangle.$$

Solution. The relation between the Clebsch-Gordan coefficients and the $3j$ symbols is given by

$$\langle j_1 m_1; j_2 m_2 | JM \rangle \equiv (-1)^{j_1 - j_2 + M} \sqrt{2J+1} \begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix}.$$

Using the exchange rule for $3j$ symbols,

$$\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix} = (-1)^{j_1 + j_2 + J} \begin{pmatrix} j_2 & j_1 & J \\ m_2 & m_1 & -M \end{pmatrix},$$

we find

$$\langle j_1 m_1 j_2 m_2 | JM \rangle = (-1)^{3j_2 - j_1 + J} \langle j_2 m_2 j_1 m_1 | JM \rangle.$$

Since $(-1)^{4j_2} \equiv 1$ for all allowed values of j_2, and $J - j_1 - j_2$ is always integer this results in the desired expression for exchange. □

J.2 Derivation of the properties of $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ in the coupled representation

J.2.1 Commutation relations for $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$

It is not straightforward to calculate the properties of $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ in the coupled basis, $\{|(j_1 j_2) JM\rangle\}$ because $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ are defined with respect to the uncoupled basis, $\{|j_1 m_1; j_2 m_2\rangle\}$. Important insights can be obtained by deriving *selection rules* for the matrix elements of $j_{az}, j_{a\pm}$ in the coupled basis, with $a \in \{1, 2\}$.

Starting from the commutation relations (3.1) we find for the commutation relations between J_x, J_y, J_z and j_{ax}, j_{ay}, j_{az}

$$\begin{aligned} &[J_x, j_{ax}] = 0 && [J_y, j_{ay}] = 0 && [J_z, j_{az}] = 0 \\ &[J_x, j_{ay}] = i\hbar j_{az} && [J_y, j_{az}] = i\hbar j_{ax} && [J_z, j_{ax}] = i\hbar j_{ay} \\ &[J_x, j_{az}] = -i\hbar j_{ay} && [J_y, j_{ax}] = -i\hbar j_{az} && [J_z, j_{ay}] = -i\hbar j_{ax}, \end{aligned} \tag{J.26}$$

with $a \in \{1, 2\}$. Likewise, starting from the commutation relations (3.25) we find for the commutation relations between $J_\pm, J_z$ and $j_{a\pm}, j_{az}$, with $a \in \{1, 2\}$,

$$[J_\pm, j_{a\pm}] = 0, \quad [J_\pm, j_{a\mp}] = \pm 2\hbar j_{az}, \quad [J_z, j_{a\pm}] = \pm\hbar j_{a\pm}. \tag{J.27}$$

J.2.2 Selection rules for j_1 and j_2 in the coupled basis

From the commutation relations (J.26) and (J.27) we can derive the following *selection rules* for the matrix elements of $j_{az}, j_{a\pm}$, with $a \in \{1, 2\}$, in the coupled basis $\{|J, M\rangle\}$:

S1: *The matrix elements of* j_{az} *are diagonal in* M.

S2: *The operators* $j_{a\pm}$ *raise/lower* M *by one.*

S3: *The matrix elements of* j_{az}*, and* $j_{a\pm}$ *are zero for* $|J' - J| > 1$.

Problem J.2. Derive selection rule S1,

$$\langle J'M'|j_{az}|JM\rangle = 0 \quad \text{unless} \quad M' = M.$$

Solution. Using the hermiticity of J_z the selection rule S1 follows from the commutation of J_z and j_{az},

$$0 = \langle J'M'|[J_z, j_{az}]|JM\rangle = (M' - M)\hbar\langle J'M'|j_{az}|JM\rangle. \qquad \square$$

Problem J.3. Derive selection rule S2,

$$\langle J'M'|j_{a\pm}|JM\rangle = 0 \quad \text{unless} \quad M' = M \pm 1.$$

Solution. This follows from the commutation relation $[J_z, j_{a\pm}] = \pm\hbar j_{a\pm}$. Using the hermiticity of J_z we have

$$\langle J'M'|[J_z, j_{a\pm}]|JM\rangle = (M' - M)\hbar\langle J'M'|j_{a\pm}|JM\rangle = \pm\hbar\langle J'M'|j_{a\pm}|JM\rangle,$$

which simplifies to the condition

$$(M' - M \mp 1)\hbar\langle J'M'|j_{a\pm}|JM\rangle = 0. \qquad \square$$

Problem J.4. Derive selection rule S3,

$$\langle J'M'|j_{az}|JM\rangle = 0 \quad \text{and} \quad \langle J'M'|j_{a\pm}|JM\rangle = 0 \quad \text{for} \quad |J' - J| > 1.$$

Solution. Let us assume $J' \geq J$ and write $J' = J + n$, where $n \geq 0$. Using the commutation relation $[J_+, j_{a-}] = 2\hbar j_{az}$ we have $\langle J'M'|[J_+, j_{a-}]|J, M-1\rangle = 2\hbar\langle J'M'|j_{az}|J, M-1\rangle$. Since J_+ and J_- are hermitian conjugates this relation can be rewritten in the form

$$C_{J',-M'}\langle J'M' - 1|j_{a-}|J, M - 1\rangle - C_{J,M-1}\langle J', M'|j_{a-}|J, M\rangle = 2\langle J'M'|j_{az}|J, M - 1\rangle, \tag{a}$$

where the coefficients $C_{J,M}$ are given by Eq. (3.24b). This relation should hold for all M and M' satisfying the inequalities (b)

$$-J \leq M \leq J \quad \text{and} \quad -J' \leq M' \leq J' = J + n. \tag{b}$$

Since according to selection rule S2 we have $\langle J'M'|j_{a+}|JM\rangle = 0$ unless $M' = M + 1$ and the same holds for $\langle J'M' - 1|j_{a+}|JM - 1\rangle = 0$, we derive from the inequalities (b) the condition

$$M \leq J + n - 1 \leq J$$

to assure that Eq. (a) holds for nonzero matrix elements. As this is the case only for $0 \leq n \leq 1$, we find $\langle J'M'|j_{az}|JM\rangle = 0$ for $n > 1$. For $J' \leq J$ and for $\langle J'M'|j_{a\pm}|JM\rangle = 0$ the proof proceeds analogously. $\square$

J.2.3 Matrix elements of $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ in the coupled basis - phase convention for coupling

Before we can evaluate the matrix elements of $j_{1z}, j_{1\pm}$ and $j_{2z}, j_{2\pm}$ in the coupled basis $\{|J, M\rangle\}$, with $|j_1 - j_2| \leq J \leq j_1 + j_2$, we have to introduce a new phase convention. For matrix elements diagonal in J the conditions are familiar. The operators act within a single, $(2J+1)$-dimensional subspace in which *all eigenstates carry the same phase.* This is assured by the lowering operator J_- in combination with the Condon and Shortley phase convention - see Eqs. (3.24). For matrix elements off-diagonal in J the operators act between two of such subspaces and we have to settle the relative phase by a new convention.

To elaborate on the above we consider the matrix elements of j_{1z}. In view of the selection rule S1 these are of the type $\langle J'M|j_{1z}|JM\rangle$. For $J' = J$ it follows from the hermiticity of j_{1z} that the matrix element is equal to its complex conjugate,

$$\langle JM|j_{1z}|JM\rangle = \langle JM|j_{1z}|JM\rangle^*, \tag{J.28}$$

which shows that the value of $\langle j_{1z}\rangle$ is *real*, a well know property of hermitian operators. For $J' \neq J$ the operator j_{1z} acts between two subspaces of different J for which the relative phase phase is not defined. As we shall show below, the matrix elements $\langle J'M|j_{1z}|JM\rangle$ can always be chosen to be *real* but, depending on the relative phase of $|J, M\rangle$ and $|J', M\rangle$, the sign can be positive or negative. Clearly, a new phase convention is required to remove this ambiguity.

For the matrix elements of j_{1+} the conditions are similar. In view of the selection rule S2 the matrix elements are of the type $\langle J', M+1|j_{1+}|J, M\rangle$. Since a given raising operator is the hermitian conjugate of the corresponding lowering operator we have for $J' = J$

$$\langle J, M+1|j_{1+} + j_{2+}|J, M\rangle = \langle J, M+1|j_{1+} + j_{2+}|J, M\rangle^* = C_{J,M}. \tag{J.29}$$

This shows that (as a result of the Condon and Shortley phase convention) the *sum* of the matrix elements $\langle J, M+1|j_{1+}|J, M\rangle$ and $\langle J, M+1|j_{2+}|J, M\rangle$ is *real* and *positive.* This condition can only be satisfied for arbitrary J and M if the imaginary parts of these matrix elements always cancel, which is the case if these matrix elements are *real* individually. For $J' \neq J$ we shall find (in complete analogy with the matrix elements of j_z) that a new phase convention is required to uniquely define the matrix elements of j_{1+}.

So, it will be our task to extend the Condon and Shortley phase convention with a new element to assure that the phase of the matrix elements of j_{1z} and $j_{1\pm}$ is well-defined for $J' \neq J$ in the coupled basis. For this purpose we first derive two phase rules:

(*a*) *For given J and J', the matrix elements of j_{1+} carry a phase which is independent of M.* Recalling the commutation relations (J.27) we consider the equation

$$\langle J', M+2|[J_+, j_{1+}]|J, M\rangle = 0. \tag{J.30}$$

Since J_+ and J_- are hermitian conjugates this equation can be rewritten in the form

$$C_{J',-M-2}\langle J', M+1|j_{1+}|J, M\rangle = C_{J,M}\langle J', M+2|j_{1+}|J, M+1\rangle, \tag{J.31}$$

where the $C_{J,M}$ are shift coefficients - see Eq. (3.24b). Note that for $-J \leq M \leq J'-2$ the ratio $C_{J',-M-2}/C_{J,M}$ is a *positive real number* in the Condon and Shortley phase convention. Hence, for given J and J', we find by induction the same phase for all values of M.

(*b*) *For $J' \neq J$ the matrix elements of j_{1z} and j_{2z} are equal in magnitude but opposite in sign.*

$$\langle J', M|j_{2z}|J, M\rangle = -\langle J', M|j_{1z}|J, M\rangle. \tag{J.32}$$

Since $J_z = j_{1z} + j_{2z}$ we have

$$\langle J', M|j_{1z}|J, M\rangle + \langle J', M|j_{2z}|J, M\rangle = \langle J', M|J_z|J, M\rangle = M\delta_{J,J'}, \tag{J.33}$$

For $J' \neq J$ this leads to Eq. (J.32).

At this point we introduce the following *phase convention* [38, 28]:

For $J' \neq J$ the matrix elements of j_{1z} are chosen to be real and non-negative,

$$\langle J', M|j_{1z}|J, M\rangle \geq 0. \tag{J.34}$$

The justification of this convention requires a few steps. Our goal is to show by induction that in the *off-diagonal* case ($J' \neq J$) the matrix elements $\langle J \pm 1, M|j_{1z}|J, M\rangle$ carry the opposite/same phase as $\langle J \pm 1, M+1|j_{1+}|J, M\rangle$. The procedure is based on the commutation relations (J.27) by considering the following equation:

$$\langle J', M+1|[J_+, j_{1z}]|J, M\rangle = -\hbar\langle J', M+1|j_{1+}|J, M\rangle. \tag{J.35}$$

Since J_+ and J_- are hermitian conjugates this equation can be rewritten in the form

$$C_{J',-M-1}\langle J', M|j_{1z}|J, M\rangle + \langle J', M+1|j_{1+}|J, M\rangle = C_{J,M}\langle J', M+1|j_{1z}|J, M+1\rangle. \tag{J.36}$$

We now consider two cases:

(a) $J' = J - 1$: The ratio $C_{J',-M-1}/C_{J,M}$ is a positive real number in the Condon and Shortley phase convention (for all physically relevant values of M). It then follows from Eq. (J.36) that, *if the matrix element $\langle J', M|j_{1z}|J, M\rangle$ carries the same phase as $\langle J', M+1|j_{1+}|J, M\rangle$ this is also the case for $\langle J', M+1|j_{1z}|J, M+1\rangle$.* Hence, by induction this is the case for any relevant value of M. So, what remains to be done is show that for *one particular value* of J, J' and M, the matrix elements $\langle J', M|j_{1z}|J, M\rangle$ and $\langle J', M+1|j_{1+}|J, M\rangle$ carry the same phase. We claim that $M = -J' - 1$ is such a particular value. For this value we have $C_{J',-M-1} = 0$ and Eq. J.36) reduces to

$$C_{J,-J'-1}\langle J', -J'|j_{1z}|J, -J'\rangle = \langle J', -J'|j_{1+}|J, -J'-1\rangle. \tag{J.37}$$

Since the coefficient $C_{J,-J'-1} = C_{J,J'}$ is *positive* for $J' = J - 1$ the phase is the same on both sides, which completes the proof by induction.

At this point we can adopt the phase convention (J.34) and find as an immediate consequence that the matrix elements of j_{1+} are *real and non-negative* for $J' = J - 1$,

$$\langle J-1, M+1|j_{1+}|J, M\rangle \geq 0. \tag{J.38}$$

(b) $J' = J + 1$: For all relevant values of M the ratio $C_{J,M}/C_{J',-M-1}$ is real and positive in the Condon and Shortley phase convention. From Eq. (J.36) it then follows that if the phase of $\langle J', M+1|j_{1z}|J, M+1\rangle$ *is opposite to that of* $\langle J', M+1|j_{1+}|J, M\rangle$ this also applies to $\langle J', M|j_{1z}|J, M\rangle$. We claim that $M = J$ is the particular value for which this applies. Since $C_{J,M} = 0$ for $M = J$, the equation (J.36) reduces to

$$\langle J', J|j_{1z}|J, J\rangle = -(1/C_{J',-J-1})\langle J', J+1|j_{1+}|J, J\rangle. \tag{J.39}$$

Since the coefficient $C_{J',-J-1} = C_{J',J}$ is *positive* for $J' = J+1$ the matrix elements must have *opposite* phase.

At this point we can adopt the phase convention (J.34) and find for the case $J' = J + 1$ that the matrix elements of j_{1+} are *real and non-positive*,

$$\langle J+1, M+1|j_{1+}|J, M\rangle \leq 0. \tag{J.40}$$

In summary: the matrix elements of $j_{1z}, j_{1\pm}$ are *always real* in the coupled basis. By adopting the convention (J.34) we assure that for $J' \neq J$ the matrix elements $\langle J', M|j_{1z}|J, M\rangle$ are *always positive* (if not zero). The sign of the matrix elements $\langle J', M|j_{1+}|J, M\rangle$ cannot be chosen independently. Under the convention (J.34) it follows that the sign of $\langle J-1, M+1|j_{1+}|J, M\rangle$ is *positive* and that of $\langle J+1, M+1|j_{1+}|J, M\rangle$ *negative.* The convention (J.34) can be regarded as an extension of the Condon and Shortly phase convention for the off-diagonal case of the coupled basis.

J.2.4 Recurrence relations - the phase of Clebsch-Gordan coefficients

The orthonormality relations (3.98) leave us freedom to choose the phase of the CGCs. As it turns out, this freedom can be used to define the CGCs as *real* numbers, positive or negative. We start by showing this for the CGCs associated with stretched states. It follows from Eq. (3.91) that

- for given j_1 and j_2, the CGC of the *stretched state* is *unity* by convention,

$$\langle j_1 j_1; j_2 j_2 | J_{max}, J_{max} \rangle = 1, \tag{J.41}$$

where $J_{max} \equiv j_1 + j_2$. To analyze the other CGCs we make repeated use of the shift operators. To explain the procedure we start by deriving two recurrence relations:

1. The first recursion relation is based on the operator identity $J_- = j_{1-} + j_{2-}$,

$$\langle j_1, m_1; j_2, m_2 - 1 | J_- | JM \rangle = \langle j_1, m_1; j_2, m_2 - 1 | j_{1-} + j_{2-} | JM \rangle. \tag{J.42}$$

   Note that we added commas in the notation where this improved the readability. As raising and lowering are hermitian conjugate operations we obtain (using the symmetry relation $C_{j_2, m_2 - 1} = C_{j_2, -m_2}$) the following recursion relation for given j_1, j_2 and J,

$$\begin{aligned} C_{J,-M} \langle j_1, m_1; j_2, m_2 - 1 | J, M - 1 \rangle = C_{j_1, m_1} \langle j_1, m_1 + 1; j_2, m_2 - 1 | JM \rangle \\ + C_{j_2, -m_2} \langle j_1 m_1; j_2 m_2 | JM \rangle, \end{aligned} \tag{J.43}$$

   subject to the selection rule (3.94), $m_1 + m_2 = M$.

2. The second recursion relation is based on the operator identity $J_+ = j_{1+} + j_{2+}$,

$$\langle j_1, m_1; j_2, m_2 + 1 | J_+ | J, M \rangle = \langle j_1, m_1; j_2, m_2 + 1 | j_{1+} + j_{2+} | J, M \rangle. \tag{J.44}$$

   As raising and lowering are hermitian conjugate operations we obtain for given j_1, j_2 and J,

$$\begin{aligned} C_{J,M} \langle j_1 m_1; j_2, m_2 + 1 \, | J, M + 1 \rangle = C_{j_1, -m_1} \langle j_1, m_1 - 1; j_2, m_2 + 1 | J, M \rangle \\ + C_{j_2, m_2} \langle j_1 m_1; j_2 m_2 | J, M \rangle, \end{aligned} \tag{J.45}$$

   subject to the selection rule $m_1 + m_2 = M$.

To proceed we focus on two special cases:

- Recurrence relation (J.43) specialized to the case $m_1 = j_1$. Since $C_{j_1, m_1} = C_{j_1, j_1} = 0$, the relation reduces to

$$\langle j_1 j_1; j_2, m_2 - 1 | J, M - 1 \rangle = C_{j_2, -m_2} / C_{J,-M} \langle j_1 j_1; j_2 m_2 | JM \rangle. \tag{J.46}$$

  The ratio $C_{j_2, -m_2} / C_{J,-M}$ is *positive* in the Gordon and Shortley phase convention (for all physically relevant values of m_2 and M). Hence, for given j_1, j_2 and J, the phase of the coefficients $\langle j_1 j_1; j_2 m_2 | JM \rangle$ is *independent of M*.

- Recurrence relation (J.45) specialized to the case to $M = J$. Since $C_{J,M} = C_{J,J} = 0$, the relation reduces to

$$\langle j_1, m_1 - 1; j_2, m_2 + 1 | JJ \rangle = -C_{j_2, m_2} / C_{j_1, -m_1} \langle j_1 m_1; j_2 m_2 | JJ \rangle, \tag{J.47}$$

  The ratio $C_{j_2, m_2} / C_{j_1, -m_1}$ is *positive* in the Gordon and Shortley phase convention (for all physically relevant values of m_1 and m_2). Hence, for given j_1, j_2 and J, the phase of the coefficient $\langle j_1 m_1; j_2 m_2 | JJ \rangle$ is conserved up to a *sign*, which *alternates* when changing m_1 in unit steps.

With the above findings we established the following: if *one* of the coefficients $\langle j_1 m_1; j_2 m_2|J, M\rangle$ is real, all CGCs with the same combination of j_1, j_2 and J are *real* in the Condon and Shortley phase convention. Importantly, this condition can always be satisfied because, for given j_1, j_2 and J, we have the freedom to choose the CGCs with $m_1 = j_1$ and $M = J$ to be *real* and *positive* (*sign convention*) - see Problem J.5

$$\langle j_1 j_1; j_2 m_2|J, J\rangle \geq 0, \tag{J.48}$$

with $m_2 = J - j_1$. The use of the above rules is demonstrated in Problem J.9. As you will see the procedure is labor intensive. Therefore, it makes little sense to rederive these values every time we need them. Once we understand the procedure, it suffices to determine the values with *Mathematica*. For reference purposes some old-fashioned look-up tables are included in Appendix J.

Example: Spin-doublet system with half-integral total angular momentum

The case of coupling a spin $s = \frac{1}{2}$ to orbital angular momentum l is special since only two J values can emerge, $j = l \pm \frac{1}{2}$. This implies that j and m_J have half-integral values. Using the m_J selection rule (3.90) the Clebsch-Gordan decomposition (3.92) reduces to two terms,

$$|(l \pm \tfrac{1}{2}), m_J\rangle = \sum_{m_s=+1/2}^{-1/2} |l(m_J - m_s), sm_s\rangle\langle l(m_J - m_s), sm_s|(l \pm \tfrac{1}{2}), m_J\rangle. \tag{J.49}$$

This expression can be written in the form

$$|(l + \tfrac{1}{2}), m_J\rangle = \cos\alpha|l(m_J - m_s), \uparrow\rangle + \sin\alpha|l(m_J - m_s), \downarrow\rangle \tag{J.50}$$

$$|(l - \tfrac{1}{2}), m_J\rangle = \cos\beta|l(m_J - m_s), \uparrow\rangle + \sin\beta|l(m_J - m_s), \downarrow\rangle. \tag{J.51}$$

Note that the states are normalized,

$$\cos^2\alpha + \sin^2\alpha = 1 \quad \text{and} \quad \cos^2\beta + \sin^2\beta = 1. \tag{J.52}$$

As the states $|(l + \frac{1}{2}), m_J\rangle$ and $|(l - \frac{1}{2}), m_J\rangle$ form an orthonormal pair, the coefficients must satisfy the orthogonality condition

$$\cos\alpha\cos\alpha + \sin\alpha\sin\alpha = 0. \tag{J.53}$$

By defining $\sin\beta = \cos\alpha$ and $\cos\beta = -\sin\alpha$ both conditions are satisfied, and for angles in the interval $0 \leq \alpha \leq \pi/2$ the definition is also in accordance with the Condon and Shortly phase convention (J.48).

Problem J.5. Show that, for given j_1, j_2 and J, the Clebsch-Gordan coefficient with $m_1 = j_1$ and $M = J$ can be chosen to be *real* and *non-negative* - see Eq. (J.48).

Solution. We start with phase rule J.38 for the special case $M = J$,

$$\langle J+1, J+1|j_{1+}|J, J\rangle \leq 0,$$

which is based on the Condon and Shortley phase convention. We insert the closure relation (3.93) before and after the operator j_+. Applying the selection rules $J = m_1 + m_2$ and $J+1 = m_1' + m_2'$, the four-fold summation over m_1, m_2, m_1' and m_2' reduces to a two-fold summation over m_1 and m_1', with $m_2' = J + 1 - m_1'$. As m_2 is conserved by j_{1+} the summation over m_1' reduces to a single term, defined by $m_1' = J + 1 - m_2 = m_1 + 1$. This leaves us with the summation over m_1,

$$\sum_{m_1}\langle J+1, J+1|j_1, m_1+1; j_2, J-m_1\rangle\langle j_1, m_1+1|j_{1+}|j_1 m_1\rangle\langle j_1 m_1; j_2, J-m_1|J, J\rangle \leq 0.$$

Note that each term of the summation has to be negative (or zero) because, in view of J.46, all $2j_1 + 1$ terms carry the *same* phase. In particular, we have for the term with $m_1 = j_1 - 1$,

$$\langle J+1, J+1|j_1, j_1; j_2, m_2\rangle C_{j_1, j_1-1}\langle j_1, j_1-1; j_2, m_2|J, J\rangle \leq 0.$$

We now have to choose a phase convention such that this inequality is satisfied. This requires that the two CGCs be real and have opposite sign because $\langle j_1, m_1+1|j_{1+}|j_1 m_1\rangle = C_{j_1,m_1}$ is real and non-negative in the Condon and Shortley phase convention. As the two CGCs differ in m_1 by one unit of angular momentum, this is achieved by adopting the convention J.48. □

Problem J.6. Show that the CGCs with $M = J$ satisfy the following relation

$$\langle j_1, m_1-1; j_2, m_2+1|JJ\rangle = (-1)^{j_1-m_1} C_{j_2,m_2}/C_{j_1,-m_1}\langle j_1 m_1; j_2 m_2|JJ\rangle.$$

Solution. In view of the property J.48, we know that (for given j_1, j_2 and J) $\langle j_1 m_1; j_2, m_2|JJ\rangle$ is positive for $m_1 = j_1$ and $|j_1 - j_2| \leq J \leq j_1 + j_2$. Since the sign alternates with m_1 we find the desired relation by writing the sign in Eq. (J.47) in the form of the phase factor $(-1)^{j_1-m_1}$. □

Problem J.7. Show that when lowering M by lowering m_2 in the CGC of the stretched state we obtain

$$\langle j_1 j_1; j_2, j_2-1|J_{max}, J_{max}-1\rangle = \sqrt{j_2/(j_1+j_2)}\langle j_1 j_1; j_2 j_2|J_{max} J_{max}\rangle,$$

which *is real* and *positive.*

Solution. Specializing Eq. (J.46) to the stretched case ($m_1 = j_1$, $m_2 = j_2$, $M = J = J_{max} = j_1 + j_2$) we have $C_{j_2,-m_2} = C_{j_2,-j_2} = \sqrt{2j_2}$ and $C_{J,-M} = C_{J,-J} = \sqrt{2J}$, and obtain the desired relation. In view of Eq. (J.41) the result is real and positive. □

Problem J.8. Show that the CGC for antiparallel coupling of equal angular momenta is given by

$$\langle j, m; j, -m|00\rangle = \frac{(-1)^{j-m}}{\sqrt{2j+1}}.$$

Solution. Antiparallel coupling of two equal angular momenta ($j_1 = j_2 = j$) corresponds to $J = M = 0$. In this case all CGC are of the type $\langle j, m; j, -m|00\rangle$, with $-j \leq m \leq j$. These CGCs all have the same magnitude but differ in sign (see Problem J.6),

$$\langle j, m-1; j, -m+1|00\rangle = (-1)^{j-m}\langle j, m; j, -m|00\rangle.$$

Secondly, using the sum rule 3.99a we find

$$\sum_{m=-j}^{j} |\langle jm; j, -m\,|00\rangle\,|^2 = (2j+1)|\langle j, j; j, -j|00\rangle|^2 = 1.$$

Since $\langle j, j; j, -j|00\rangle \geq 0$ by convention (J.48), the two relations lead us to the desired result. □

Problem J.9. Derive the Clebsch-Gordan coefficients for coupling of the angular momenta $j_1 = \frac{3}{2}$ and $j_2 = \frac{1}{2}$ (see Table J.1). Note that this is an example of a doublet system with integral total angular momentum.

Solution. First the case $J = 2$. We start from the *stretched state convention*,

$$\langle \tfrac{3}{2}\tfrac{3}{2}; \tfrac{1}{2}\tfrac{1}{2}\,|22\rangle = 1. \tag{a}$$

Applying Eq. (J.46) and using the Gordon and Shortly phase convention (3.20) we calculate $C_{j_2,-m_2}/C_{J,-M} = C_{1/2,-1/2}/C_{2,-2} = \frac{1}{2}$ and obtain

$$\langle \tfrac{3}{2}\tfrac{3}{2}; \tfrac{1}{2} - \tfrac{1}{2}\,|21\rangle = \sqrt{1/4}. \tag{b}$$

Note that we made one step down and moved to the $M = +1$ block in in Table J.1. From here we make a step down within the $M = +1$ block by applying the recurrence relation (J.45) with the values $m_1 = \frac{3}{2}$, $m_2 = -\frac{1}{2}$ and $M = 1$. Calculating $C_{2,1} = 2$, $C_{3/2,-3/2} = \sqrt{3}$ and $C_{1/2,-1/2} = 1$ we find by substituting the values (a) and (b)

$$C_{2,1}\langle \tfrac{3}{2}\tfrac{3}{2}; \tfrac{1}{2}\tfrac{1}{2}\,|22\rangle = C_{3/2,-3/2}\langle \tfrac{3}{2}\tfrac{1}{2}; \tfrac{1}{2}\tfrac{1}{2}\,|21\rangle + C_{1/2,-1/2}\langle \tfrac{3}{2}\tfrac{3}{2}; \tfrac{1}{2} - \tfrac{1}{2}\,|21\rangle \Rightarrow \langle \tfrac{3}{2}\tfrac{1}{2}; \tfrac{1}{2}\tfrac{1}{2}\,|21\rangle = \sqrt{3/4}. \tag{c}$$

Table J.1: Clebsch-Gordan coefficients for coupling of the angular momenta $j_1 = \frac{3}{2}$ and $j_2 = \frac{1}{2}$. Only the non-zero coefficients are indicated. Note the selection rule $M = m_1 + m_2$.

$\langle j_1 j_2, m_1, m_2 \| JM\rangle$	$\|2,+2\rangle$	$\|2,+1\rangle$	$\|1,+1\rangle$	$\|2,0\rangle$	$\|1,0\rangle$	$\|1,-1\rangle$	$\|2,-1\rangle$	$\|2,-2\rangle$
$\langle j_1 j_2, +\frac{3}{2}, +\frac{1}{2}\|$	1							
$\langle j_1 j_2, +\frac{3}{2}, -\frac{1}{2}\|$		$\sqrt{1/4}$	$\sqrt{3/4}$					
$\langle j_1 j_2, +\frac{1}{2}, +\frac{1}{2}\|$		$\sqrt{3/4}$	$-\sqrt{1/4}$					
$\langle j_1 j_2, +\frac{1}{2}, -\frac{1}{2}\|$				$\sqrt{1/2}$	$\sqrt{1/2}$			
$\langle j_1 j_2, -\frac{1}{2}, +\frac{1}{2}\|$				$\sqrt{1/2}$	$-\sqrt{1/2}$			
$\langle j_1 j_2, -\frac{1}{2}, -\frac{1}{2}\|$						$\sqrt{1/4}$	$\sqrt{3/4}$	
$\langle j_1 j_2, -\frac{3}{2}, +\frac{1}{2}\|$						$-\sqrt{3/4}$	$\sqrt{1/4}$	
$\langle j_1 j_2, -\frac{3}{2}, -\frac{1}{2}\|$								1
$M = m_1 + m_2$	2	1		0		-1		-2

To move to the $M = 0$ block (next step down) we turn to the recurrence relation (J.43) with the values $m_1 = \frac{1}{2}$, $m_2 = \frac{1}{2}$, $M = 1$. Calculating $C_{2,-1} = \sqrt{6}$, $C_{3/2,1/2} = \sqrt{3}$ and $C_{1/2,-1/2} = 1$ we find by substituting the values (b) and (c)

$$C_{2,-1}\langle \tfrac{3}{2}\tfrac{1}{2}; \tfrac{1}{2} - \tfrac{1}{2}\,|20\rangle = C_{3/2,1/2}\langle \tfrac{3}{2}\tfrac{3}{2}; \tfrac{1}{2} - \tfrac{1}{2}\,|21\rangle + C_{1/2,-1/2}\langle \tfrac{3}{2}\tfrac{1}{2}; \tfrac{1}{2}\tfrac{1}{2}\,|21\rangle \Rightarrow \langle \tfrac{3}{2}\tfrac{1}{2}; \tfrac{1}{2}\tfrac{1}{2}\,|20\rangle = \sqrt{1/2}. \tag{d}$$

We make the step down within the $M = 0$ block by returning to the recurrence relation (J.45) but this time with the values with the values $m_1 = \frac{1}{2}$, $m_2 = -\frac{1}{2}$ and $M = 0$. Calculating $C_{2,0} = \sqrt{6}$, $C_{3/2,-1/2} = 2$ and $C_{1/2,-1/2} = 1$ we find by substituting (c) and (d)

$$C_{2,0}\langle \tfrac{3}{2}\tfrac{1}{2}; \tfrac{1}{2}\tfrac{1}{2}\,|21\rangle = C_{3/2,-1/2}\langle \tfrac{3}{2} - \tfrac{1}{2}; \tfrac{1}{2}\tfrac{1}{2}\,|20\rangle + C_{1/2,-1/2}\langle \tfrac{3}{2}\tfrac{1}{2}; \tfrac{1}{2} - \tfrac{1}{2}\,|20\rangle \Rightarrow \langle \tfrac{3}{2} - \tfrac{1}{2}; \tfrac{1}{2}\tfrac{1}{2}\,|20\rangle = \sqrt{1/2}.$$

Continuing this procedure we obtain all Clebsch-Gordan coefficients for $J = 2$.

We turn to the case $J = 1$. In this case we have no stretched state, so we have to determine the starting value in a different way. We start with the recurrence relation (J.47) with the values $m_1 = \frac{3}{2}$, $m_2 = -\frac{1}{2}$ and $M = 1$. Calculating $-C_{j_2,m_2}/C_{j_1,-m_1} = C_{1/2,-1/2}/C_{3/2,-3/2} = \frac{1}{2}$ and obtain

$$\langle \tfrac{3}{2}\tfrac{1}{2}; \tfrac{1}{2}\tfrac{1}{2}\,|11\rangle = -\sqrt{1/3}\,\langle \tfrac{3}{2}\tfrac{3}{2}; \tfrac{1}{2} - \tfrac{1}{2}\,|11\rangle\,. \tag{e}$$

Using the orthonormality relation

$$|\langle \tfrac{3}{2}\tfrac{1}{2}; \tfrac{1}{2}\tfrac{1}{2}\,|11\rangle\,|^2 + |\langle \tfrac{3}{2}\tfrac{3}{2}; \tfrac{1}{2} - \tfrac{1}{2}\,|11\rangle\,|^2 = 1$$

we find, using Eq. (e) and the *sign convention* (J.2)

$$\langle \tfrac{3}{2}\tfrac{3}{2}; \tfrac{1}{2} - \tfrac{1}{2}\,|11\rangle = \sqrt{3/4} \;\text{ and }\; \langle \tfrac{3}{2}\tfrac{1}{2}; \tfrac{1}{2}\tfrac{1}{2}\,|11\rangle = -\sqrt{1/4}.$$

From here on we can continue with the recurrence procedure as discussed for $J = 2$ until all values of the $J = 1$ manifold are determined. □

J.3 Symmetry properties of Clebsch-Gordan coefficients

J.3.1 Coupling order and exchange symmetry

Let us have a closer look at the coupling order introduced in Section 3.4.1. To reveal its signature we consider the matrix elements of j_{1z} for a given coupling order

$$\langle (j_1 j_2)J'M|j_{1z}|(j_1 j_2)JM\rangle. \tag{J.54}$$

Let us first focus on the *off-diagonal* case ($J' \neq J$). Adopting the convention (J.34), the off-diagonal matrix elements of j_{1z} are *real and non-negative*,

$$\langle (j_1 j_2)J'M | j_{1z} | (j_1 j_2)JM \rangle \geq 0. \tag{J.55}$$

Furthermore, recalling Eq. (J.32) we know that the off-diagonal matrix elements of j_{1z} and j_{2z} are equal in magnitude but have opposite sign,

$$\langle (j_1 j_2)J'M | j_{2z} | (j_1 j_2)JM \rangle = -\langle (j_1 j_2)J'M | j_{1z} | (j_1 j_2)JM \rangle. \tag{J.56}$$

Our task is to apply this *convention* to a system of two angular momenta, say $\boldsymbol{j}_a$ and $\boldsymbol{j}_b$. By assigning the values $\boldsymbol{j}_a \to \boldsymbol{j}_1$ and $\boldsymbol{j}_b \to \boldsymbol{j}_2$ we have for $J' = J - 1$

$$\langle (j_a j_b)(J-1), M | j_{az} | (j_a j_b)JM \rangle \geq 0 \quad \text{and} \quad \langle (j_a j_b)(J-1), M | j_{bz} | (j_a j_b)JM \rangle \leq 0. \tag{J.57a}$$

On the other hand, by assigning $\boldsymbol{j}_b \to \boldsymbol{j}_1$ and $\boldsymbol{j}_a \to \boldsymbol{j}_2$ we have

$$\langle (j_b j_a)(J-1), M | j_{bz} | (j_b j_a)JM \rangle \geq 0 \quad \text{and} \quad \langle (j_b j_a)(J-1), M | j_{az} | (j_b j_a)JM \rangle \leq 0. \tag{J.57b}$$

This shows that, by *changing the coupling order*, we *also change the sign* of the *off-diagonal* matrix elements of j_{az} (and j_{bz}),

$$\langle (j_a j_b)(J-1), M | j_{az} | (j_a j_b)JM \rangle = - \langle (j_b j_a)(J-1), M | j_{az} | (j_b j_a)JM \rangle. \tag{J.58}$$

First we consider the coupling order $(j_a j_b)$. What can we say about the phase of $|(j_a j_b)JM\rangle$? First of all, for given J, all angular momentum states carry the same phase since the phase is independent of M in the Condon and Shortley phase convention - see Eqs. (3.24). Extending this convention by adopting (J.34) we know that the matrix elements $\langle (j_a j_b)(J-1), M | j_{az} | (j_a j_b)JM \rangle$ are real and *positive* for any physically relevant value of M. So, if the phase of $|(j_a j_b)J, M\rangle$ is real and *positive* for one particular value of J, this is also the case for the state $|(j_a j_b)(J-1), M\rangle$ and (by induction) for all physically relevant values of J. This particular state exists because the stretched state $|(j_a j_b)J_{max}, J_{max}\rangle$, with $J_{max} = j_a + j_b$, has unit phase by convention - see Eq. (3.91).

What about the other coupling order? As in the previous case, for given J the states $|(j_b j_a)JM\rangle$ carry the same phase (independent of M). However, in view of Eq. (J.56) the matrix elements $\langle (j_b j_a)(J-1), M | j_{az} | (j_b j_a)JM \rangle$ are real and *negative*. So, if the phase of $|(j_b j_a)JM\rangle$ is real and *positive* for one particular value of J it has to be real and *negative* for $|(j_b j_a)(J-1), M\rangle$ and (by induction) *alternate between positive and negative* for all physically relevant values of J. Also in the present case this particular value exists because, recalling Eq. (3.81), the *stretched* state has unit phase, irrespective of the coupling order, $|(j_a j_b)J_{max}, J_{max}\rangle = |(j_a j_b)J_{max}, J_{max}\rangle$. Accordingly, the phase of the states $|(j_b j_a)J, M\rangle$ is real and alternates between positive and negative in such a way that it is positive for the stretched state. Thus, comparing both coupling orders, we arrive at

$$|(j_b j_a)J, M\rangle = (-1)^{j_a + j_b - J} |(j_a j_b)J, M\rangle. \tag{J.59}$$

About the *diagonal* case ($J' = J$) we can be short since the phase drops out the matrix element (J.54). Also the *off-diagonal* matrix elements with $|J' - J| > 1$ do not add new conditions on the phase. These can be disregarded as nonphysical as they violate the selection rule S3. Recalling the equivalence (3.97) we find from Eq. (J.59) for the symmetry relation of CGCs

$$\langle j_b m_b; j_a m_a | J, M \rangle = (-1)^{j_a + j_b - J} \langle j_a m_a; j_b m_b | J, M \rangle. \tag{J.60}$$

J.3.1.1 Exchange symmetry for identical particles

The case of equal spin, $j_1 = j_2 = j$, deserves special attention in view of its importance for the exchange symmetry of identical particles. In this case one cannot distinguish between the two coupling orderings and we have to add the amplitudes. This can be done in two independent ways. In particular in the form of a symmetric (+) and a antisymmetric (−) linear combination,

$$\begin{aligned}|jjJM\rangle_{\pm} &= \sqrt{\tfrac{1}{2}}\,[|(j_1j_2)J,M\rangle \pm |(j_2j_1)J,M\rangle] \\ &= \sqrt{\tfrac{1}{2}}\,\left[1 \pm (-1)^{2j-J}\right]|(j_1j_2)J,M\rangle. \qquad \text{(J.61)}\end{aligned}$$

Note that the second line vanishes for certain combinations of j and J. These combinations are excluded by the exchange symmetry. For the coupling of *integer* spins the value of $J_{max} = 2j$ is always *even*, whereas for *half-integer* spins it is always *odd*. As a consequence, for *integer* spins the *even* (*odd*) total spin states are *symmetric* (*antisymmetric*) whereas for *half-integer* spins this is the case for *odd* (*even*) total spin. In particular, for half-integer spins the stretched pair states are symmetric for odd J, whereas for integer spins this is the case for even J. These symmetry rules are known from some well-known examples. For instance, for two identical particles without spin the total spin is zero. In this case the spin of the pair is even and the corresponding state symmetric under exchange. For two identical spin-$\frac{1}{2}$ particles the total spin can be 0 or 1. In this case the spin is half-integer. For total spin zero (which is even) the corresponding spin state is antisymmetric under exchange; for $J = 1$ (which is odd) the spin state of the pair is symmetric.

Problem J.10. For antiparallel coupling of two equal angular momenta, $j_1 = j_2 = j$, the total angular momentum is zero, $J = M = 0$. Show that in this case the Clebsch-Gordan decomposition (3.92) can be written in the form

$$\sqrt{2j+1}|jj00\rangle = \sum_m (-1)^{j-m}|j,m;j,-m\rangle.$$

Solution. We start from the Clebsch-Gordan decomposition (3.92) for the special case $j_1 = j_2 = j$,

$$|jjJM\rangle = \sum_{m_1,m_2} |jm_1;jm_2\rangle\langle jm_1;jm_2|jjJM\rangle.$$

For antiparallel coupling the total angular momentum is zero, $J = M = 0$. Using the selection rule $M = m_1 + m_2$ the decomposition takes the form

$$|jj00\rangle = \sum_m |j,m;j,-m\rangle\langle j,m;j,-m|jj00\rangle.$$

Substituting

$$\langle j,m;j,-m|00\rangle = \frac{(-1)^{j-m}}{\sqrt{2j+1}}$$

(see Problem J.8) we arrive at the desired result. □

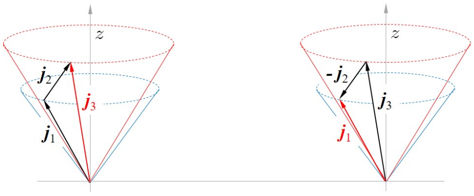
J.3.2 Time-reversal symmetry

Thus far we analyzed the addition of angular momenta, $\boldsymbol{j}_3 = \boldsymbol{j}_1 + \boldsymbol{j}_2$. What about subtraction? As suggested by the vector model (see Fig. J.2) we would like to define $\boldsymbol{j}_2 = \boldsymbol{j}_3 - \boldsymbol{j}_1$ by introducing a new angular momentum operator $\overline{\boldsymbol{j}} = -\boldsymbol{j}$, which has the properties of a "reversed" angular momentum. As it turns out, to properly define $\overline{\boldsymbol{j}}$ we have to study $\boldsymbol{j}$ under *time reversal.* A complicating factor is that the transformation $\boldsymbol{j} \mapsto \overline{\boldsymbol{j}}$ is found to be antilinear [97]. All this being said, we shall not enter into the properties of antilinear maps because, within the Condon and Shortley phase convention, the symmetry properties of the CGCs can be obtained algebraically from the recurrence relations J.43 and J.45.

In this section we derive the relation

$$\langle j_2,-m_2;j_3,m_3|j_1m_1\rangle = (-1)^{j_2+m_2}\left(\frac{2j_1+1}{2j_3+1}\right)^{1/2}\langle j_1m_1;j_2m_2|j_3m_3\rangle, \qquad \text{(J.62)}$$

where quantum numbers were chosen in compliance with the vector model. The prefactor will follow from the recurrence relations.



(a) (b)

Figure J.2: Vector addition diagrams; (a) diagram for $j_3 = \boldsymbol{j}_1 + \boldsymbol{j}_2$; (b) diagram for $j_1 = \boldsymbol{j}_3 + (-\boldsymbol{j}_2)$. The figure is very suggestive but the addition of a "reversed" angular momentum cannot be properly defined within the vector model. This model handles the quantum numbers but does not capture the phase rules.

We start by renaming $J \to j_3$ and $M \to m_3$ in the recurrence relations J.43 and J.45. With some rearrangement we bring these relations in the form

$$C_{j_1,m_1}\langle j_1, m_1+1; j_2, m_2-1|j_3m_3\rangle = -C_{j_2,-m_2}\langle j_1m_1; j_2m_2|j_3m_3\rangle \\ + C_{j_3,-m_3}\langle j_1m_1; j_2, m_2-1|j_3, m_3-1\rangle, \quad \text{(J.63a)}$$

$$C_{j_1,-m_1}\langle j_1, m_1-1; j_2, m_2+1|j_3m_3\rangle = -C_{j_2,m_2}\langle j_1m_1; j_2m_2|j_3m_3\rangle \\ + C_{j_3,m_3}\langle j_1m_1; j_2, m_2+1|j_3, m_3+1\rangle. \quad \text{(J.63b)}$$

Note that both equations satisfy the selection rule

$$m_1 + m_2 = m_3. \quad \text{(J.64)}$$

Now we relable the indices cyclicly, $(1,2,3) \to (2,3,1)$,

$$C_{j_1,-m_1}\langle j_2m_2; j_3, m_3-1|j_1, m_1-1\rangle = C_{j_2,m_2}\langle j_2, m_2+1; j_3, m_3-1|j_1m_1\rangle \\ + C_{j_3,-m_3}\langle j_2m_2; j_3m_3|j_1m_1\rangle. \quad \text{(J.65)}$$

Note that in this process the selection rule J.64 was replaced by $m_2 + m_3 = m_1$. As suggested by the vector model (see Fig. J.2), the selection rule J.64 can be restored by renaming the quantum number $m_2 \to -m_2$. However, we can do better. By renaming $m_2 \to -m_2 - 1$ and $m_3 \to m_3 + 1$ we not only recover the selection rule J.64 but find that $\langle j_2, -m_2; j_3, m_3|j_1m_1\rangle$ satisfies, *up to a sign*, the "other" recurrence relation, Eq. J.63b,

$$C_{j_1,-m_1}\langle j_2, -(m_2+1); j_3m_3|j_1, m_1-1\rangle = C_{j_2,m_2}\langle j_2, -m_2; j_3m_3|j_1m_1\rangle \\ + C_{j_3,m_3}\langle j_2, -(m_2+1); j_3m_3+1|j_1m_1\rangle. \quad \text{(J.66)}$$

Note that we used the symmetry rules $C_{j_2,-m_2-1} = C_{j_2,m_2}$ and $C_{j_3,-m_3-1} = C_{j_3,m_3}$. To obtain full equivalence, *including the sign*, we compare Eqs. (J.66) and (J.63b) term by term and add the appropriate phase factor,

$$-\langle j_1, m_1-1; j_2, m_2+1|j_3m_3\rangle = (-1)^{j_2+(m_2+1)}\alpha\langle j_2, -(m_2+1); j_3m_3|j_1, m_1-1\rangle \quad \text{(J.67a)}$$

$$\langle j_1m_1; j_2m_2|j_3m_3\rangle = (-1)^{j_2+m_2}\alpha\langle j_2, -m_2; j_3m_3|j_1m_1\rangle \quad \text{(J.67b)}$$

$$-\langle j_1m_1; j_2, m_2+1|j_3, m_3+1\rangle = (-1)^{j_2+(m_2+1)}\alpha\langle j_2, -(m_2+1); j_3m_3+1|j_1m_1\rangle. \quad \text{(J.67c)}$$

Here α is a proportionality constant, independent of m_1, m_2 and m_3. Since j_2+m_2 is always integral and the CGCs always real, α is a real number, which may still depend on j_1, j_2 and j_3.

To determine sign and magnitude of α, we start by applying the exchange rule (J.59) to bring Eq. (J.67b) in the form

$$\langle j_1 m_1; j_2 m_2 | j_3 m_3\rangle = (-1)^{j_2+m_2}(-1)^{j_2+j_3-j_1}\alpha\langle j_3 m_3; j_2, -m_2 | j_1 m_1\rangle. \tag{J.68}$$

First we shall demonstrate that the sign of α is *positive*. For given j_1, j_2, j_3, the relations (J.67) hold for all physically meaningful values of m_1, m_2, m_3. Therefore, we can determine the sign of α by considering a special case. For $m_1 = j_1$ and $m_3 = j_3$, Eq. (J.68) becomes

$$\langle j_1 j_1; j_2 m_2 | j_3 j_3\rangle = (-1)^{j_2+m_2}(-1)^{j_2+j_3-j_1}\alpha\langle j_3 j_3; j_2, -m_2 | j_1 j\rangle. \tag{J.69}$$

From the selection rule J.64 we infer that $m_2 = j_3 - j_1$ and since $j_2 + j_3 - j_1$ is always integral the Eq. (J.68) simplifies to

$$\langle j_1 j_1; j_2 m_2 | j_3 j_3\rangle = \alpha\langle j_3 j_3, j_2, -m_2 | j_1 j_1\rangle. \tag{J.70}$$

Hence, in view of the convention (J.48) we find $\alpha > 0$.

To determine the magnitude of α we compare the moduli of the two sides of Eq. (J.68),

$$|\langle j_1 m_1; j_2, m_3 - m_1 | j_3, m_3\rangle|^2 = |\alpha|^2 |\langle j_3 m_3; j_2, m_1 - m_3; | j_1 m_1\rangle|^2$$

Summing the l.h.s. over j_1 and the r.h.s. over j_3 we find by unitarity - see sum rule (3.99a),

$$\sum_{m_3=-j_3}^{j_3} \delta(j_1 j_2 j_3) = |\alpha|^2 \sum_{m_1=-j_1}^{j_1} \delta(j_1 j_2 j_3). \tag{J.71}$$

Since $\alpha > 0$ we find for the magnitude of α,

$$\alpha = \left(\frac{2j_3+1}{2j_1+1}\right)^{1/2}. \tag{J.72}$$

Substituting this value in Eq. (J.67b) we arrive at Eq. (J.68)

$$\langle j_2, -m_2; j_3, m_3 | j_1 m_1\rangle = (-1)^{j_2+m_2}\left(\frac{2j_1+1}{2j_3+1}\right)^{1/2}\langle j_1 m_1; j_2 m_2 | j_3 m_3\rangle. \tag{J.73}$$

Other symmetry relations can be derived in same way but rather than presenting these separately we introduce in the next section the Wigner $3j$ symbols. These are best suited to deal with the symmetry properties of the CGCs.

It is instructive to consider the stretched case, $m_1 = j_1$, $m_2 = j_2$ and $m_3 = j_3 = j_1 + j_2$,

$$\langle j_2, -j_2; j_3, j_3 | j_1 j_1\rangle = (-1)^{2j_2}\left(\frac{2j_3+1}{2j_1+1}\right)^{1/2}\langle j_1 j_1; j_2 j_2 | j_3 j_3\rangle. \tag{J.74}$$

This expression shows that the reversal is accompanied by a phase factor $(-1)^{2j_2}$, which is unity for integer j_2 but introduces a sign change in the half-integer case. This sign change is typical for time reversal of angular momenta without a classical analog [97].

J.3.3 Wigner $3j$ symbols

An inconvenience of CGCs are the non-intuitive symmetry properties, not to speak of the forest of competing notations in the literature.[1] This spoils the joy of rapid manipulation and assessment. An important simplification results by expressing the CGCs in terms of Wigner $3j$ symbols. These

---

[1] A compilation of notation conventions is given by Edmonds [38].

are obtained by treating the three angular momenta on equal footing. For this purpose we note that the kets of the basis $\{|j,m\rangle\}$ are related to the bras of the hermitian conjugate basis $\{\langle j,m|\}$ by the transformation

$$\langle j,m| = \sum_{m'} g_{m,m'} |j,m'\rangle, \tag{J.75}$$

where $g_{m,m'}$ is a metric tensor,

$$g_{m,m'} = \left(\begin{smallmatrix} j \\ m\; m' \end{smallmatrix}\right) = (-1)^{j+m}\delta_{m,m'}, \tag{J.76}$$

in the notation of Wigner. Note that $j+m$ is always an integer. The kets transform as contravariant vectors and the bras as covariant vectors. In tensor calculus the transformation (J.75) is known as *lowering* of the contravariant index. By lowering of the indices m_1 and m_2 the CGCs take the form

$$\langle j_1m_1; j_2m_2|j_3m_3\rangle = (-1)^{j_1+m_1}(-1)^{j_2+m_2}|j_1m_1\rangle \otimes |j_2m_2\rangle \otimes |j_3m_3\rangle. \tag{J.77}$$

Since the CGCs are real constants the tensor $|j_1m_1\rangle \otimes |j_2m_2 \otimes |j_3m_3\rangle$ must be a real scalar; i.e., be invariant under rotation. We denote this scalar by the expression

$$(-1)^{-2j_2}\sqrt{2j_3+1}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix}, \tag{J.78}$$

where the factor in brackets is the Wigner $3j$ symbol. Actually, at this point there is no reason to put the prefactor in front of the $3j$ symbol and/or the minus sign in front of m_3. However, we are free to do so and as we shall see this choice maximizes the symmetry. Using the projection rule $m_1+m_2=m_3$ the relation between the CGC and the $3j$ symbol becomes

$$\langle j_1m_1; j_2m_2|j_3m_3\rangle = (-1)^{j_1-j_2+m_3}\sqrt{2j_3+1}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix}, \tag{J.79}$$

or, equivalently,

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} = \frac{(-1)^{j_1-j_2+m_3}}{\sqrt{2j_3+1}}\langle j_1m_1; j_2m_2|j_3m_3\rangle. \tag{J.80}$$

Note that $j_1-j_2+m_3$ is always integral.

J.3.3.1 Sum rules

From the *orthonormality relations* (3.98) immediately follow the following *sum rules*:

$$\sum_{j_3=|j_1-j_2|}^{j_1+j_2}\sum_{m_3=-j_3}^{j_3}(2j+3)\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}^2 = 1 \tag{J.81a}$$

$$\sum_{m_1=-j_1}^{j_1}\sum_{m_2=-j_2}^{j_2}(2j+3)\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}^2 = 1. \tag{J.81b}$$

Note that the second summations can be eliminated by applying the projection rule.

Problem J.11. Derive the *symmetry property* 2*a*,

$$\begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}.$$

Solution. Converting Eq. (J.60) to $3j$ symbols we find

$$(-1)^{j_2-j_1-m_3}\sqrt{2j_3+1}\begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix} = (-1)^{j_1+j_2-j_3}(-1)^{j_1-j_2-m_3}\sqrt{2j_3+1}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}.$$

This can be rewritten in the form

$$\begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix} = (-1)^{4j_1-j_1-j_2-j_3}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}.$$

Since $4j_1$ is always *even* and $j_1+j_2+j_3$ always *integral* this corresponds to the desired property. □

Problem J.12. Derive the *symmetry property* $2b$,

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}.$$

Solution. Converting Eq. (J.62) to $3j$ symbols we find

$$(-1)^{j_2-j_3+m_1}\sqrt{2j_1+1}\begin{pmatrix} j_2 & j_3 & j_1 \\ -m_2 & m_3 & -m_1 \end{pmatrix} = (-1)^{j_2+m_2}\left(\frac{2j_1+1}{2j_3+1}\right)^{1/2}(-1)^{j_1-j_2+m_3}\sqrt{2j_3+1}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix}.$$

Using the *projection rule*, $m_3 = m_1 + m_2$, this can be rewritten in the form

$$\begin{pmatrix} j_2 & j_3 & j_1 \\ -m_2 & m_3 & -m_1 \end{pmatrix} = (-1)^{2m_2-2j_2}(-1)^{j_1+j_2+j_3}\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix}.$$

Note that $2m_2 - 2j_2$ is always *even*. So, after cyclic rotation of the l.h.s. and renaming $m_3 \to -m_3$ this corresponds to the desired property. □

J.3.3.2 Properties of $3j$ symbols

The $3j$ symbols inherit their *principal properties* from the CGCs. So, the $3j$ symbols are *zero unless*

1. The *triangle inequality* holds, $\Delta(j_1j_2j_3) > 0$ - see Eq. (J.12);
2. The *sum* of the angular momenta is *integral*, $j_1 + j_2 + j_3 \in \mathbb{Z}$.
3. The angular momentum *projection is conserved*, $m_1 + m_2 + m_3 = 0$.

The *symmetry properties* are given by two easy-to-memorize rules:

1. *Invariance* under *cyclic* permutations of the columns
2. *Phase changes by* $(-1)^{j_1+j_2+j_3}$
   a. under *exchange* of two columns
   b. under *inverted projection*, $m_1, m_2, m_3 \to -m_1, -m_2, -m_3$.

The symmetry properties $2a$ and $2b$ are derived in Problems J.11 and J.12.

J.3.3.3 Vector model for $3j$ symbols

In Fig. J.3 the $3j$ symbols are illustrated by diagrams in which the vectors add up to zero (note the difference with Fig. 3.2). With such diagrams the validity of the triangle inequality and the projection rule are easily established. As *most* combinations of the indices turn out to be nonphysical, rapid disposal of these combinations greatly enhances the efficiency of calculations. Only if a $3j$ symbol is *not* manifestly zero it has to be evaluated. Beware that not being *manifestly* zero does not exclude being zero. This happens in some high symmetry cases as a consequence of the symmetry rules. Many properties of $3j$ symbols are summarized in Appendix J.

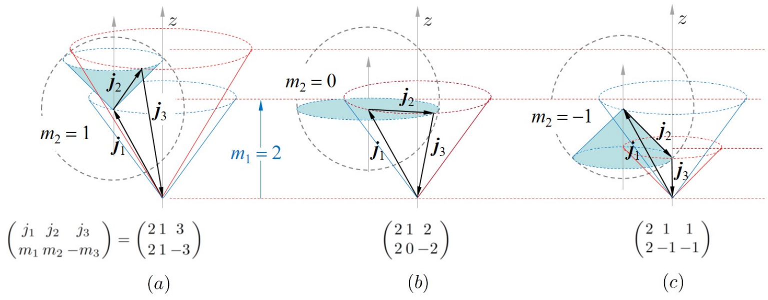


Figure J.3: Graphical representation of $3j$ symbols with $j_1 = 2$, $j_2 = 1$: a.) $j_3 = 3$; b.) $j_3 = 2$; c.) $j_3 = 1$. The $3j$ symbol is valid (i.e., not manifestly zero) if the triangle is closed and the projections onto the z axis are quantized in integer values. This implies that the triangle inequality is satisfied, $j_1 + j_2 + j_3$ is integral, and the projections onto the z axis add up to zero, $m_1 + m_2 + m_3 = 0$. The phase rules are not captured in these diagrams.

J.4 Addition of three angular momenta

In the case of three angular momenta, $\boldsymbol{j}_1$, $\boldsymbol{j}_2$ and $\boldsymbol{j}_3$, acting in three *independent* subspaces of the Hilbert space, the total angular momentum is given by

$$\mathbf{J} = \boldsymbol{j}_1 + \boldsymbol{j}_2 + \boldsymbol{j}_3.$$

In this case there is no unique addition procedure. Defining $\boldsymbol{j}_{12} = \boldsymbol{j}_1 + \boldsymbol{j}_2$, $\boldsymbol{j}_{23} = \boldsymbol{j}_2 + \boldsymbol{j}_3$ and $\boldsymbol{j}_{13} = \boldsymbol{j}_1 + \boldsymbol{j}_3$ we have

$$|(j_{12}j_3)JM\rangle = \sum_{m_{12}=-j_{12}}^{j_{12}} \sum_{m_3=-j_3}^{j_3} |j_{12}m_{12}; j_3m_3\rangle\langle j_{12}m_{12}; j_3m_3|(j_{12}j_3)JM\rangle, \tag{J.82a}$$

where $|j_{12}m_{12}; j_3m_3\rangle \equiv |(j_1j_2)j_{12}m_{12}\rangle \otimes |j_3m_3\rangle$ and

$$|j_{12}m_{12}\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |j_1m_1; j_2m_2\rangle\langle j_1m_1; j_2m_2|j_{12}m_{12}\rangle. \tag{J.82b}$$

Substituting (J.82b) into (J.82a) we obtain an expression for $|(j_{12}j_3)JM\rangle$. Likewise we obtain an expression for $(j_1j_{23})J'M'\rangle$.

$$|(j_{12}j_3)JM\rangle = \sum_{m_1,m_2,m_3,m_{12}} |j_1m_1; j_2m_2; j_3m_3\rangle\langle j_1m_1; j_2m_2|j_{12}m_{12}\rangle\langle j_{12}m_{12}; j_3m_3|(j_{12}j_3)JM\rangle, \tag{J.83a}$$

$$|(j_1j_{23})J'M'\rangle = \sum_{m_1,m_2,m_3,m_{23}} |j_1m_1; j_2m_2; j_3m_3\rangle\langle j_2m_2; j_3m_3|j_{23}m_{23}\rangle\langle j_1m_1; j_{23}m_{23}|(j_1j_{23})J'M'\rangle, \tag{J.83b}$$

where $|j_1m_1; j_2m_2; j_3m_3\rangle \equiv |j_1m_1\rangle \otimes |j_2m_2\rangle \otimes |j_3m_3\rangle$. Projection of the coupling scheme (J.83b) onto (J.83a) results in recoupling coefficient,

$$\langle(j_{12}j_3)JM|(j_1j_{23})J'M'\rangle = \sum_{\substack{m_1,m_2,m_3\\m_{12},m_{23}}} \langle j_1m_1; j_2m_2|j_{12}m_{12}\rangle\langle j_{12}m_{12}; j_3m_3|(j_{12}j_3)JM\rangle$$
$$\langle j_2m_2; j_3m_3|j_{23}m_{23}\rangle\langle j_1m_1; j_{23}m_{23}|(j_1j_{23})J'M'\rangle. \tag{J.84}$$

Note that the value of the recoupling coefficient is independent of M. So, aiming for maximum symmetry we may sum over M provided we divide by $2J + 1$ to correct for double counting. After

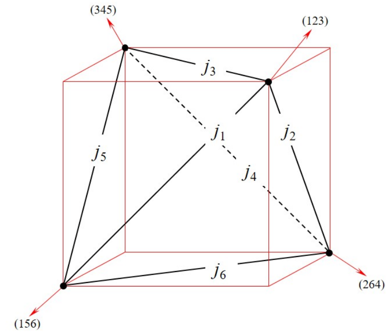


Figure J.4: Graphical representation of Wigner 6-j symbols in the form of a tetrahedron. Each corner vertex represents one of four $3j$ symbols. The $6j$ symbols are invariant under all symmetry operations, rotations and/or reflections, of the tetrahedral group (resulting in all 4! permutations of the corner points).

turning to $3j$ symbols the recoupling coefficient takes the form

$$\langle (j_{12}j_3)JM|(j_1j_{23})J'M'\rangle = \sum_{\substack{m_1,m_2,m_3\\ m_{12},m_{23},M}} (-1)^{\varphi}\sqrt{(2j_{12}+1)(2j_{23}+1)(2J'+1)/(2J+1)}$$
$$\times \begin{pmatrix} j_1 & j_2 & j_{12} \\ m_1 & m_2 & -m_{12} \end{pmatrix} \begin{pmatrix} j_{12} & j_3 & J \\ m_{12} & m_3 & -M \end{pmatrix} \begin{pmatrix} j_2 & j_3 & j_{23} \\ m_2 & m_3 & -m_{23} \end{pmatrix} \begin{pmatrix} j_1 & j_{23} & J' \\ m_1 & m_{23} & -M' \end{pmatrix}, \quad \text{(J.85)}$$

where

$$\varphi = j_1 - j_2 + m_{12} + j_{12} - j_3 + M + j_2 - j_3 + m_{23} + j_1 - j_{23} + M'. \quad \text{(J.86)}$$

J.4.1 Wigner $6j$ symbols

In view of the CGC selection rules the coupling coefficient (J.85) vanishes unless $m_{12} = m_1 + m_2$, $m_{23} = m_2 + m_3$, $M = m_{12} + m_3 = m_1 + m_2 + m_3 = m_1 + m_{23} = M'$. The latter condition can only be satisfied for all values of M if also $J = J'$. Wigner expressed this in the form

$$\langle (j_{12}j_3)J'M'|(j_1j_{23})JM\rangle = (-1)^{j_1+j_2+j_3+J}\sqrt{(2j_{12}+1)(2j_{23}+1)} \begin{Bmatrix} j_1 & j_2 & j_{12} \\ j_3 & J & j_{23} \end{Bmatrix} \delta_{J,J'}\delta_{M,M'}, \quad \text{(J.87)}$$

which includes a so-called $6j$ symbol. Note that $j_1 + j_2 + j_3 + J$ is always an integer. Like the $3j$ symbols also the Wigner $6j$ symbols offer optimized symmetry rules. By comparing Eq. (J.87) with Eq. (J.85) the $6j$ symbol can be expressed in the following form - see Problem J.13,

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{Bmatrix} = \sum_{\substack{m_1,m_2,m_3\\ m_4,m_5,m_6}} (-1)^{\sigma} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_1 & j_5 & j_6 \\ m_1 & m_5 & -m_6 \end{pmatrix} \begin{pmatrix} j_2 & j_6 & j_4 \\ m_2 & m_6 & -m_4 \end{pmatrix} \begin{pmatrix} j_3 & j_4 & j_5 \\ m_3 & m_4 & -m_5 \end{pmatrix}, \quad \text{(J.88)}$$

where $\sigma = j_4 + j_5 + j_6 + m_4 + m_5 + m_6$. A graphical representation is given in Fig. J.4.

Problem J.13. Derive Eq. (J.88).

Solution. We start from Eq. (J.87). Comparing this expression with Eq. (J.85) we obtain an expression for the $6j$symbol. To emphasize the symmetry we relabel the following quantum numbers: $(j_3, m_3) \to (j_4, m_4)$, $(j_{12}, m_{12}) \to (j_3, m_3)$, $(J, M) \to (j_5, m_5)$, $(j_{23}, m_{23}) \to (j_6, m_6)$. In this way we obtain

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{Bmatrix} = \sum_{\substack{m_1,m_2,m_3\\ m_4,m_5,m_6}} (-1)^{\varphi_1} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} \begin{pmatrix} j_3 & j_4 & j_5 \\ m_3 & m_4 & -m_5 \end{pmatrix} \begin{pmatrix} j_2 & j_4 & j_6 \\ m_2 & m_4 & -m_6 \end{pmatrix} \begin{pmatrix} j_1 & j_6 & j_5 \\ m_1 & m_6 & -m_5 \end{pmatrix},$$

where $\varphi_1 = 3j_1 + j_2 + j_3 - j_4 + j_5 - j_6 + m_3 + 2m_5 + m_6$ (note that $j_1 + j_2 + j_4 + j_5 \in \mathbb{Z}$). To arrive at the full symmetry of the $6j$ symbol we reorder the $3j$ symbols to match the arrangement indicated on the tetrahedron in Fig. J.4. For this purpose we swap the last two columns in both the third and the fourth $3j$ symbol and adapt the phase accordingly,

$$\left\{\begin{matrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{matrix}\right\} = \sum_{\substack{m_1,m_2,m_3 \\ m_4,m_5,m_6}} (-1)^{\varphi_2} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} \begin{pmatrix} j_3 & j_4 & j_5 \\ m_3 & m_4 & -m_5 \end{pmatrix} \begin{pmatrix} j_2 & j_6 & j_4 \\ m_2 & -m_6 & m_4 \end{pmatrix} \begin{pmatrix} j_1 & j_5 & j_6 \\ m_1 & -m_5 & m_6 \end{pmatrix},$$

with $\varphi_2 = 4j_1 + 2j_2 + j_3 + 2j_5 + j_6 + m_3 + 2m_5 + m_6$. Next we can replace the summations over m_3, m_4, m_5 and m_6 by summations over $-m_3$, $-m_4$, $-m_5$ and $-m_6$ adapting the phase accordingly,

$$\left\{\begin{matrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{matrix}\right\} = \sum_{\substack{m_1,m_2,m_3 \\ m_4,m_5,m_6}} (-1)^{\varphi_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_3 & j_4 & j_5 \\ -m_3 & -m_4 & m_5 \end{pmatrix} \begin{pmatrix} j_2 & j_6 & j_4 \\ m_2 & m_6 & -m_4 \end{pmatrix} \begin{pmatrix} j_1 & j_5 & j_6 \\ m_1 & m_5 & -m_6 \end{pmatrix},$$

with $\varphi_3 = 2j_1 - j_3 + 2j_5 + j_6 - m_3 - 2m_5 - m_6$ (since $2j_1 + 2j_2 + 2j_3$ is always even). To arrive at Eq. (J.88) we apply the projection rule to the second $3j$ symbol and obtain for the exponent of the phase $\varphi_4 = j_4 + j_5 + j_6 + m_4 + m_5 + m_6 + 2j_1 + 2j_5 - m_3 - m_4 - 3m_5 - 2m_6$. Since $2j_i \pm 2m_i$ is always *even* and $(-1)^{2m_i} = (-1)^{-2m_i}$ for $i \in \{1,2,3,4,5,6\}$ as well as the projection rules $m_3 + m_4 - m_5 = 0$ and $m_1 + m_5 - m_6 = 0$ we have

$$(-1)^{\varphi_4} = (-1)^{\sigma - 2m_1 - m_3 - m_4 - m_5 + 2m_6} = (-1)^{\sigma - 2m_1 - 2m_5 + 2m_6} = (-1)^{\sigma},$$

where $\sigma = j_4 + j_5 + j_6 + m_4 + m_5 + m_6$. □

J.4.1.1 Properties of $6j$ symbols

The $3j$ symbols inherit their *properties* from the CGCs. So, the $3j$ symbols are *zero unless*

1. The *triangle inequality* holds for all combinations $\Delta(j_k j_l j_m) > 0$ for $k, l, m \in \{1, \cdots 6\}$
2. The *sum* of the angular momenta is *integral*, $j_k + j_l + j_m \in \mathbb{Z}$ for $k, l, m \in \{1, \cdots 6\}$

Symmetry properties: $6j$ symbols *are invariant* under all symmetry operations of the tetrahedral group - see Fig. J.4. These result in invariance under

1. *permutations* of the columns; e.g.,

$$\left\{\begin{matrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{matrix}\right\} = \left\{\begin{matrix} j_2 & j_1 & j_3 \\ j_5 & j_4 & j_6 \end{matrix}\right\}$$

2. *simultaneous exchange* of *two* opposing elements of the rows; e.g.,

$$\left\{\begin{matrix} j_1 & j_2 & j_3 \\ j_4 & j_5 & j_6 \end{matrix}\right\} = \left\{\begin{matrix} j_4 & j_5 & j_3 \\ j_1 & j_2 & j_6 \end{matrix}\right\}.$$

J.5 Relation with the Wigner $6j$ symbols

The Wigner $6j$ symbols are defined by the relation

$$\langle (j_1 j_2) j_{12}, j_3; J'M' | j_1, (j_2 j_3) j_{23}; JM \rangle = (-)^{j_1+j_2+j_3+J} \sqrt{(2j_{12}+1)(2j_{23}+1)} \left\{\begin{matrix} j_1 & j_2 & j_{12} \\ j_3 & J & j_{23} \end{matrix}\right\} \delta_{JJ'} \delta_{MM'}. \tag{J.89}$$

and can be expressed as a sum over a product of four $3j$ symbols

$$\left\{\begin{matrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{matrix}\right\} = \sum_{\substack{m_1,m_2,m_3 \\ M_1,M_2,M_3}} (-1)^{\sigma} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} J_2 & J_3 & j_1 \\ M_2 & -M_3 & m_1 \end{pmatrix} \begin{pmatrix} J_3 & J_1 & j_2 \\ M_3 & -M_1 & m_2 \end{pmatrix} \begin{pmatrix} J_1 & J_2 & j_3 \\ M_1 & -M_2 & m_3 \end{pmatrix} \tag{J.90}$$

where $\sigma = J_1 + J_2 + J_3 + M_1 + M_2 + M_3$.

The Wigner $6j$ symbols $\begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix}$ have the following properties:

Reality:

The $6j$ symbols are all real.

Selection rules:

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix} = 0 \qquad \text{unless the triads } (j_1, j_2, j_3)\,(j_1, J_2, J_3)\,(J_1, j_2, J_3)\,(J_1, J_2, j_3) \tag{J.91}$$

- satisfy the triangular inequalities
- have an *integral* sum

Symmetry:

- invariant under column permutation,

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix} = \begin{Bmatrix} j_2 & j_1 & j_3 \\ J_2 & J_1 & J_3 \end{Bmatrix} \tag{J.92}$$

- invariant under simultaneous exchange of two elements from the first line with the corresponding elements from the second line,

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix} = \begin{Bmatrix} J_1 & J_2 & j_3 \\ j_1 & j_2 & J_3 \end{Bmatrix} \tag{J.93}$$

Orthogonality:

$$\sum_j (2j+1) \begin{Bmatrix} j_1 & j_2 & j \\ J_1 & J_2 & J \end{Bmatrix} \begin{Bmatrix} j_1 & j_2 & j \\ J_1 & J_2 & J' \end{Bmatrix} = \frac{\delta_{J'J}}{(2J+1)} \tag{J.94}$$

Special cases [38]

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_2 & j_1 & 0 \end{Bmatrix} = (-1)^s \frac{1}{\sqrt{(2j_1+1)(2j_2+1)}}, \tag{J.95a}$$

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_2 & j_1 & 1 \end{Bmatrix} = \frac{(-1)^s}{2} \frac{X}{\sqrt{j_1(j_1+1)(2j_1+1)j_2(j_2+1)(2j_2+1),}} \tag{J.95b}$$

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_2 & j_1 & 2 \end{Bmatrix} = \frac{(-1)^s}{2} \frac{3X\,(X+1) - 4j_1(j_1+1)j_2(j_2+1)}{\sqrt{(2j_1-1)j_1(j_1+1)(2j_1+1)(2j_1+3)(2j_2-1)j_2(j_2+1)(2j_2+1)(2j_2+3)}}, \tag{J.95c}$$

where $s = j_1 + j_2 + j_3$ and $X = j_3(j_3+1) - j_1(j_1+1) - j_2(j_2+1)$.

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_2-1 & j_1-1 & 1 \end{Bmatrix} = \frac{(-1)^s}{2} \sqrt{\frac{(j_1+j_2+j_3)(j_1+j_2+j_3+1)(j_1+j_2-j_3-1)(j_1+j_2-j_3)}{(2j_1-1)j_1(2j_1+1)(2j_2-1)j_2(2j_2+1)}}, \tag{J.96a}$$

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_2-1 & j_1 & 1 \end{Bmatrix} = \frac{(-1)^s}{2} \sqrt{\frac{(j_1+j_2+j_3+1)(j_1+j_2-j_3)(j_2+j_3-j_1)(j_1+j_3-j_2+1)}{j_1(j_1+1)(2j_1+1)(2j_2-1)j_2(2j_2+1)}}, \tag{J.96b}$$

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ j_2-1 & j_1+1 & 1 \end{Bmatrix} = \frac{(-1)^s}{2} \sqrt{\frac{(j_2+j_3-j_1-1)(j_2+j_3-j_1)(j_1+j_3-j_2+1)(j_1+j_3-j_2+2)}{(j_1+1)(2j_1+1)(2j_1+3)(2j_2-1)j_2(2j_2+1)}}. \tag{J.96c}$$

Sum rules:

- Summation over two 6j symbols

$$\sum_{J''}(-1)^{J+J'+J''}\,(2J''+1)\begin{Bmatrix} j_1 & j_2 & J \\ j_3 & j & J'' \end{Bmatrix}\begin{Bmatrix} j_2 & j_3 & J'' \\ j_1 & j & J' \end{Bmatrix} = \begin{Bmatrix} j_3 & j_1 & J' \\ j_2 & j & J \end{Bmatrix} \tag{J.97}$$

- summation over product of four 3j symbols

$$\sum_{q}\sum_{\substack{m_1 m_2 \\ m_1', m_2'}}(-1)^{\sigma}\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & M \end{pmatrix}\begin{pmatrix} j_2' & j_1' & J' \\ m_2' & m_1' & M' \end{pmatrix}\begin{pmatrix} j_1 & j_1' & k \\ -m_1 & m_1' & q \end{pmatrix}\begin{pmatrix} j_2 & j_2' & k \\ -m_2 & -m_2' & q \end{pmatrix} =$$

$$= \delta_{J,J'}\delta_{M,M'}\frac{1}{2J+1}\begin{Bmatrix} j_1 & j_2 & J \\ j_2' & j_1' & k \end{Bmatrix}, \tag{J.98}$$

  where $\sigma = j_1' + j_2' + J + m_1 + m_2 + q$.

- summation over the product of three 3j symbols

$$\sum_{m_1 m_2 m_3}(-1)^{\sigma}\begin{pmatrix} j_1 & j_2 & j \\ m_1 & -m_2 & m \end{pmatrix}\begin{pmatrix} j_2 & j_3 & j' \\ m_2 & -m_3 & m' \end{pmatrix}\begin{pmatrix} j_3 & j_1 & J \\ m_3 & -m_1 & M \end{pmatrix} =$$

$$= \begin{Bmatrix} j' & J & j \\ j_1 & j_2 & j_3 \end{Bmatrix}\begin{pmatrix} j' & J & j \\ m' & M & m \end{pmatrix}, \tag{J.99}$$

  where $\sigma = j_1 + j_2 + j_3 + m_1 + m_2 + m_3$. Note that $m_1 = m_2 - m$ and $m_3 = m_2 + m'$; *i.e.*, the sum runs over a single independent index (m_2).

- summation over product of two 3j symbols

$$\sum_{m}(-1)^{j+m}\begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & -m \end{pmatrix}\begin{pmatrix} j_1' & j_2' & j \\ m_1' & m_2' & m \end{pmatrix}$$

$$= (-1)^{-j_2-j_1'-j-J}\sum_{J,M}(-1)^{J+M}(2J+1)\begin{Bmatrix} j_1 & j_1' & J \\ j_2' & j_2 & j \end{Bmatrix}\begin{pmatrix} j_1 & j_1' & J \\ m_1 & m_1' & -M \end{pmatrix}\begin{pmatrix} j_2 & j_2' & J \\ m_2 & m_2' & M \end{pmatrix} \tag{J.100}$$

  Note that $M = m_1 + m_1' = -m_2 - m_2'$ and $m = m_1 + m_2 = -m_1' - m_2'$; *i.e.*, the summations over m and M both consist only of a single nonzero term.

Solution.

J.5.1 Tables of Clebsch-Gordan coefficients

Table J.2: The Clebsch-Gordan coefficients $\langle j_1 j_2 m_1 m_2 | jm\rangle$ for the coupling of two angular momenta $\mathbf{j}_1$ and $\mathbf{j}_2$ to the total angular momentum $\mathbf{J} = \mathbf{j}_1 + \mathbf{j}_2$ for the cases $(j_1 \times j_2) = (\frac{1}{2} \times \frac{1}{2})$, $(1 \times \frac{1}{2})$, $(\frac{3}{2} \times \frac{1}{2})$, $(2 \times \frac{1}{2})$.

$\vert m\vert = 1$	$\vert 1,1\rangle$	$\vert 1,-1\rangle$	$m = 0$	$\vert 1,0\rangle$	$\vert 0,0\rangle$
$\langle j_1j_2,\frac{1}{2},\frac{1}{2}\vert$	1	0	$\langle j_1j_2,\frac{1}{2},-\frac{1}{2}\vert$	$\sqrt{1/2}$	$\sqrt{1/2}$
$\langle j_1j_2,-\frac{1}{2},-\frac{1}{2}\vert$	0	1	$\langle j_1j_2,-\frac{1}{2},\frac{1}{2}\vert$	$\sqrt{1/2}$	$-\sqrt{1/2}$

$(\frac{1}{2} \times \frac{1}{2})$

$\vert m\vert = \frac{3}{2}$	$\vert\frac{3}{2},\frac{3}{2}\rangle$	$\vert\frac{3}{2},-\frac{3}{2}\rangle$			
$\langle j_1j_2,1,\frac{1}{2}\vert$	1	0			
$\langle j_1j_2,-1,-\frac{1}{2}\vert$	0	1			
$m = \frac{1}{2}$	$\vert\frac{3}{2},\frac{1}{2}\rangle$	$\vert\frac{1}{2},\frac{1}{2}\rangle$	$m = -\frac{1}{2}$	$\vert\frac{3}{2},-\frac{1}{2}\rangle$	$\vert\frac{1}{2},-\frac{1}{2}\rangle$
$\langle j_1j_2,1,-\frac{1}{2}\vert$	$\sqrt{1/3}$	$\sqrt{2/3}$	$\langle j_1j_2,0,-\frac{1}{2}\vert$	$\sqrt{2/3}$	$\sqrt{1/3}$
$\langle j_1j_2,0,\frac{1}{2}\vert$	$\sqrt{2/3}$	$-\sqrt{1/3}$	$\langle j_1j_2,-1,\frac{1}{2}\vert$	$\sqrt{1/3}$	$-\sqrt{2/3}$

$(1 \times \frac{1}{2})$

$\vert m\vert = 2$	$\vert 2,2\rangle$	$\vert 2,-2\rangle$			
$\langle j_1j_2,\frac{3}{2},\frac{1}{2}\vert$	1	0			
$\langle j_1j_2,-\frac{3}{2},-\frac{1}{2}\vert$	0	1			
$m = 1$	$\vert 2,1\rangle$	$\vert 1,1\rangle$	$m = -1$	$\vert 2,-1\rangle$	$\vert 1,-1\rangle$
$\langle j_1j_2,\frac{3}{2},-\frac{1}{2}\vert$	$\sqrt{1/4}$	$\sqrt{3/4}$	$\langle j_1j_2,-\frac{1}{2},-\frac{1}{2}\vert$	$\sqrt{3/4}$	$\sqrt{1/4}$
$\langle j_1j_2,\frac{1}{2},\frac{1}{2}\vert$	$\sqrt{3/4}$	$-\sqrt{1/4}$	$\langle j_1j_2,-\frac{3}{2},\frac{1}{2}\vert$	$\sqrt{1/4}$	$-\sqrt{3/4}$
			$m = 0$	$\vert 2,0\rangle$	$\vert 1,0\rangle$
			$\langle j_1j_2,\frac{1}{2},-\frac{1}{2}\vert$	$\sqrt{1/2}$	$\sqrt{1/2}$
			$\langle j_1j_2,-\frac{1}{2},\frac{1}{2}\vert$	$\sqrt{1/2}$	$-\sqrt{1/2}$

$(\frac{3}{2} \times \frac{1}{2})$

$\vert m\vert = \frac{5}{2}$	$\vert\frac{5}{2},\frac{5}{2}\rangle$	$\vert\frac{5}{2},-\frac{5}{2}\rangle$			
$\langle j_1j_2,2,\frac{1}{2}\vert$	1	0			
$\langle j_1j_2,-2,-\frac{1}{2}\vert$	0	1			
$m = \frac{3}{2}$	$\vert\frac{5}{2},\frac{3}{2}\rangle$	$\vert\frac{3}{2},\frac{3}{2}\rangle$	$m = -\frac{3}{2}$	$\vert\frac{5}{2},-\frac{3}{2}\rangle$	$\vert\frac{3}{2},-\frac{3}{2}\rangle$
$\langle j_1j_2,2,-\frac{1}{2}\vert$	$\sqrt{1/5}$	$\sqrt{4/5}$	$\langle j_1j_2,-1,-\frac{1}{2}\vert$	$\sqrt{4/5}$	$\sqrt{1/5}$
$\langle j_1j_2,1,\frac{1}{2}\vert$	$\sqrt{4/5}$	$-\sqrt{1/5}$	$\langle j_1j_2,-2,\frac{1}{2}\vert$	$\sqrt{1/5}$	$-\sqrt{4/5}$
$m = \frac{1}{2}$	$\vert\frac{5}{2},\frac{1}{2}\rangle$	$\vert\frac{3}{2},\frac{1}{2}\rangle$	$m = -\frac{1}{2}$	$\vert\frac{5}{2},-\frac{1}{2}\rangle$	$\vert\frac{3}{2},-\frac{1}{2}\rangle$
$\langle j_1j_2,1,-\frac{1}{2}\vert$	$\sqrt{2/5}$	$\sqrt{3/5}$	$\langle j_1j_2,0,-\frac{1}{2}\vert$	$\sqrt{3/5}$	$\sqrt{2/5}$
$\langle j_1j_2,0,\frac{1}{2}\vert$	$\sqrt{3/5}$	$-\sqrt{2/5}$	$\langle j_1j_2,-1,\frac{1}{2}\vert$	$\sqrt{2/5}$	$-\sqrt{3/5}$

$(2 \times \frac{1}{2})$

Table J.3: Clebsch-Gordan coefficients $\langle j_1 j_2 m_1 m_2 | jm\rangle$ for the coupling of two angular momenta for the cases $(j_1, j_2) = (\frac{5}{2} \times \frac{1}{2})$, $(3 \times \frac{1}{2})$.

$\|m\| = 3$	$\|3, 3)$	$\|3, -3)$
$\langle j_1 j_2, \frac{5}{2}, \frac{1}{2}\|$	1	0
$\langle j_1 j_2, -\frac{5}{2}, -\frac{1}{2}\|$	0	1

$m = 2$	$\|3, 2\rangle$	$\|2, 2\rangle$
$\langle j_1 j_2, \frac{5}{2}, -\frac{1}{2}\|$	$\sqrt{1/6}$	$\sqrt{5/6}$
$\langle j_1 j_2, \frac{3}{2}, \frac{1}{2}\|$	$\sqrt{5/6}$	$-\sqrt{1/6}$

$m = -2$	$\|3, -2\rangle$	$\|2, -2\rangle$
$\langle j_1 j_2, -\frac{3}{2}, -\frac{1}{2}\|$	$\sqrt{5/6}$	$\sqrt{1/6}$
$\langle j_1 j_2, -\frac{5}{2}, \frac{1}{2}\|$	$\sqrt{1/6}$	$-\sqrt{5/6}$

$m = 1$	$\|3, 1\rangle$	$\|2, 1\rangle$
$\langle j_1 j_2, \frac{3}{2}, -\frac{1}{2}\|$	$\sqrt{1/3}$	$\sqrt{2/3}$
$\langle j_1 j_2, \frac{1}{2}, \frac{1}{2}\|$	$\sqrt{2/3}$	$-\sqrt{1/3}$

$m = -1$	$\|3, -1\rangle$	$\|2, -1\rangle$
$\langle j_1 j_2, -\frac{1}{2}, -\frac{1}{2}\|$	$\sqrt{2/3}$	$\sqrt{1/3}$
$\langle j_1 j_2, -\frac{3}{2}, \frac{1}{2}\|$	$\sqrt{1/3}$	$-\sqrt{2/3}$

$m = 0$	$\|3, 0\rangle$	$\|2, 0\rangle$
$\langle j_1 j_2, \frac{1}{2}, -\frac{1}{2}\|$	$\sqrt{1/2}$	$\sqrt{1/2}$
$\langle j_1 j_2, -\frac{1}{2}, \frac{1}{2}\|$	$\sqrt{1/2}$	$-\sqrt{1/2}$

$(\frac{5}{2} \times \frac{1}{2})$

$\|m\| = \frac{7}{2}$	$\|\frac{7}{2}, \frac{7}{2}\rangle$	$\|\frac{7}{2}, -\frac{7}{2}\rangle$
$\langle j_1 j_2, 3, \frac{1}{2}\|$	1	0
$\langle j_1 j_2, -3, -\frac{1}{2}\|$	0	1

$m = \frac{5}{2}$	$\|\frac{7}{2}, \frac{5}{2}\rangle$	$\|\frac{5}{2}, \frac{5}{2}\rangle$
$\langle j_1 j_2, 3, -\frac{1}{2}\|$	$\sqrt{1/7}$	$\sqrt{6/7}$
$\langle j_1 j_2, 2, \frac{1}{2}\|$	$\sqrt{6/7}$	$-\sqrt{1/7}$

$m = -\frac{3}{2}$	$\|\frac{7}{2}, -\frac{5}{2}\rangle$	$\|\frac{5}{2}, -\frac{5}{2}\rangle$
$\langle j_1 j_2, 0, -\frac{1}{2}\|$	$\sqrt{6/7}$	$\sqrt{1/7}$
$\langle j_1 j_2, -1, \frac{1}{2}\|$	$\sqrt{1/7}$	$-\sqrt{6/7}$

$m = \frac{3}{2}$	$\|\frac{7}{2}, \frac{3}{2}\rangle$	$\|\frac{5}{2}, \frac{3}{2}\rangle$
$\langle j_1 j_2, 2, -\frac{1}{2}\|$	$\sqrt{2/7}$	$\sqrt{5/7}$
$\langle j_1 j_2, `, \frac{1}{2}\|$	$\sqrt{5/7}$	$-\sqrt{2/7}$

$m = -\frac{3}{2}$	$\|\frac{7}{2}, -\frac{3}{2}\rangle$	$\|\frac{5}{2}, -\frac{3}{2}\rangle$
$\langle j_1 j_2, 0, -\frac{1}{2}\|$	$\sqrt{5/7}$	$\sqrt{2/7}$
$\langle j_1 j_2, -1, \frac{1}{2}\|$	$\sqrt{2/7}$	$-\sqrt{5/7}$

$\|\mathrm{m}{=}\frac{1}{2}$	$\|\frac{7}{2}, \frac{1}{2}\rangle$	$\|\frac{5}{2}, \frac{1}{2}\rangle$
$\langle j_1 j_2, 1, -\frac{1}{2}\|$	$\sqrt{3/7}$	$\sqrt{4/7}$
$\langle j_1 j_2, 0, \frac{1}{2}\|$	$\sqrt{4/7}$	$-\sqrt{3/7}$

$m = -\frac{1}{2}$	$\|\frac{7}{2}, -\frac{1}{2}\rangle$	$\|\frac{5}{2}, -\frac{1}{2}\rangle$
$\langle j_1 j_2, 0, -\frac{1}{2}\|$	$\sqrt{4/7}$	$\sqrt{3/7}$
$\langle j_1 j_2, -1, \frac{1}{2}\|$	$\sqrt{3/7}$	$-\sqrt{4/7}$

$(3 \times \frac{1}{2})$

Table J.4: Clebsch-Gordan coefficients $\langle j_1 j_2 m_1 m_2 | jm\rangle$ for the coupling of two angular momenta for the cases $(j_1, j_2) = (\frac{7}{2} \times \frac{1}{2})$, $(4 \times \frac{1}{2})$.

$\vert m\vert = 4$	$\vert 4,4\rangle$	$\vert 4,-4\rangle$			
$\langle j_1 j_2, \frac{7}{2}, \frac{1}{2}\vert$	1	0			
$\langle j_1 j_2, -\frac{7}{2}, -\frac{1}{2}\vert$	0	1			
$m = 3$	$\vert 4,3\rangle$	$\vert 3,3\rangle$	$m = -3$	$\vert 4,-3\rangle$	$\vert 4,-3\rangle$
$\langle j_1 j_2, \frac{7}{2}, -\frac{1}{2}\vert$	$\sqrt{1/8}$	$\sqrt{7/8}$	$\langle j_1 j_2, -\frac{7}{2}, -\frac{1}{2}\vert$	$\sqrt{7/8}$	$\sqrt{1/8}$
$\langle j_1 j_2, \frac{5}{2}, \frac{1}{2}\vert$	$\sqrt{7/8}$	$-\sqrt{1/8}$	$\langle j_1 j_2, -\frac{5}{2}, \frac{1}{2}\vert$	$\sqrt{1/8}$	$-\sqrt{7/8}$
$m = 2$	$\vert 4,2\rangle$	$\vert 3,2\rangle$	$m = -2$	$\vert 4,-2\rangle$	$\vert 3,-2\rangle$
$\langle j_1 j_2, \frac{5}{2}, -\frac{1}{2}\vert$	$\sqrt{2/8}$	$\sqrt{6/8}$	$\langle j_1 j_2, -\frac{3}{2}, -\frac{1}{2}\vert$	$\sqrt{6/8}$	$\sqrt{2/8}$
$\langle j_1 j_2, \frac{3}{2}, \frac{1}{2}\vert$	$\sqrt{6/8}$	$-\sqrt{2/8}$	$\langle j_1 j_2, -\frac{5}{2}, \frac{1}{2}\vert$	$\sqrt{2/8}$	$-\sqrt{6/8}$
$m = 1$	$\vert 4,1\rangle$	$\vert 3,1\rangle$	$m = -1$	$\vert 4,-1\rangle$	$\vert 3,-1\rangle$
$\langle j_1 j_2, \frac{3}{2}, -\frac{1}{2}\vert$	$\sqrt{3/8}$	$\sqrt{5/8}$	$\langle j_1 j_2, -\frac{1}{2}, -\frac{1}{2}\vert$	$\sqrt{5/8}$	$\sqrt{3/8}$
$\langle j_1 j_2, \frac{1}{2}, \frac{1}{2}\vert$	$\sqrt{5/8}$	$-\sqrt{3/8}$	$\langle j_1 j_2, -\frac{3}{2}, \frac{1}{2}\vert$	$\sqrt{3/8}$	$-\sqrt{5/8}$
			$m = 0$	$\vert 4,0\rangle$	$\vert 3,0\rangle$
			$\langle j_1 j_2, \frac{1}{2}, -\frac{1}{2}\vert$	$\sqrt{1/2}$	$\sqrt{1/2}$
			$\langle j_1 j_2, -\frac{1}{2}, \frac{1}{2}\vert$	$\sqrt{1/2}$	$-\sqrt{1/2}$

$(\frac{7}{2} \times \frac{1}{2})$

$\vert m\vert = \frac{9}{2}$	$\vert \frac{9}{2}, \frac{9}{2}\rangle$	$\vert \frac{9}{2}, -\frac{9}{2}\rangle$			
$\langle j_1 j_2, 4, \frac{1}{2}\vert$	1	0			
$\langle j_1 j_2, -4, -\frac{1}{2}\vert$	0	1			
$m = \frac{7}{2}$	$\vert \frac{9}{2}, \frac{5}{2}\rangle$	$\vert \frac{7}{2}, \frac{7}{2}\rangle$	$m = -\frac{7}{2}$	$\vert \frac{9}{2}, -\frac{7}{2}\rangle$	$\vert \frac{7}{2}, -\frac{7}{2}\rangle$
$\langle j_1 j_2, 4, -\frac{1}{2}\vert$	$\sqrt{1/9}$	$\sqrt{8/9}$	$\langle j_1 j_2, -3, -\frac{1}{2}\vert$	$\sqrt{8/9}$	$\sqrt{1/9}$
$\langle j_1 j_2, 3, \frac{1}{2}\vert$	$\sqrt{8/7}$	$-\sqrt{1/9}$	$\langle j_1 j_2, -4, \frac{1}{2}\vert$	$\sqrt{1/9}$	$-\sqrt{8/9}$
$m = \frac{5}{2}$	$\vert \frac{9}{2}, \frac{5}{2}\rangle$	$\vert \frac{7}{2}, \frac{5}{2}\rangle$	$m = -\frac{5}{2}$	$\vert \frac{9}{2}, -\frac{5}{2}\rangle$	$\vert \frac{7}{2}, -\frac{5}{2}\rangle$
$\langle j_1 j_2, 3, -\frac{1}{2}\vert$	$\sqrt{2/9}$	$\sqrt{7/9}$	$\langle j_1 j_2, -2, -\frac{1}{2}\vert$	$\sqrt{7/9}$	$\sqrt{2/9}$
$\langle j_1 j_2, 2, \frac{1}{2}\vert$	$\sqrt{7/9}$	$-\sqrt{2/9}$	$\langle j_1 j_2, -3, \frac{1}{2}\vert$	$\sqrt{2/9}$	$-\sqrt{7/9}$
$m = \frac{3}{2}$	$\vert \frac{9}{2}, \frac{3}{2}\rangle$	$\vert \frac{7}{2}, \frac{3}{2}\rangle$	$m = -\frac{3}{2}$	$\vert \frac{9}{2}, -\frac{3}{2}\rangle$	$\vert \frac{7}{2}, -\frac{3}{2}\rangle$
$\langle j_1 j_2, 2, -\frac{1}{2}\vert$	$\sqrt{3/9}$	$\sqrt{6/9}$	$\langle j_1 j_2, -1, -\frac{1}{2}\vert$	$\sqrt{6/9}$	$\sqrt{3/9}$
$\langle j_1 j_2, 1, \frac{1}{2}\vert$	$\sqrt{6/9}$	$-\sqrt{3/9}$	$\langle j_1 j_2, -2, \frac{1}{2}\vert$	$\sqrt{3/9}$	$-\sqrt{6/9}$
$m = \frac{1}{2}$	$\vert \frac{9}{2}, \frac{1}{2}\rangle$	$\vert \frac{7}{2}, \frac{1}{2}\rangle$	$m = -\frac{1}{2}$	$\vert \frac{9}{2}, -\frac{1}{2}\rangle$	$\vert \frac{7}{2}, -\frac{1}{2}\rangle$
$\langle j_1 j_2, 1, -\frac{1}{2}\vert$	$\sqrt{4/9}$	$\sqrt{5/9}$	$\langle j_1 j_2, 0, -\frac{1}{2}\vert$	$\sqrt{5/9}$	$\sqrt{4/9}$
$\langle j_1 j_2, 0, \frac{1}{2}\vert$	$\sqrt{5/9}$	$-\sqrt{4/9}$	$\langle j_1 j_2, -1, \frac{1}{2}\vert$	$\sqrt{4/9}$	$-\sqrt{5/9}$

$(4 \times \frac{1}{2})$

K

Irreducible tensor operators

K.1 Definition

An *irreducible tensor operator* of *rank* k, denoted by $\mathrm{T}^{(k)}$, is defined as a set of $2k+1$ components, $T_{k,q} \in \{T_{k,-k}, \cdots, T_{k,k}\}$, which transform according to the $(2k+1)$-dimensional representation $\mathscr{D}^{k}_{q'q}(\alpha,\beta,\gamma)$ of the full rotation group (see Section 3.7.5),

$$T'_{kq} = P_R T_{kq} P_R^{\dagger} = \sum_{q'=-k}^{+k} T_{kq'} \mathscr{D}^{k}_{q'q}, \tag{K.1}$$

where

$$P_R \equiv P(\alpha,\beta,\gamma) = P_z(\alpha)P_y(\beta)P_z(\gamma) = e^{-i\alpha J_z/\hbar} e^{-i\beta J_y/\hbar} e^{-i\gamma J_z/\hbar}. \tag{K.2}$$

is the operator for a rotation of the physical system over the Euler angles α,β,γ (see Section 3.7.3). The components $T_{k,q}$ act in a $(2k+1)$-dimensional subspace of the basis $\{|\alpha JM\rangle\}$ of the standard representation $\{J^2, J_z\}$. The symbol α points to all other (including non-angular momentum) quantum numbers of the system. Alternatively, irreducible tensor operators can be defined by the following commutation relations of their standard components - see Problem K.1 [86],

$$[J_z, T_{k,q}] = q\,\hbar\, T_{k,q} \tag{K.3a}$$

$$[J_\pm, T_{k,q}] = \sqrt{(k \mp q)\,(k \pm q + 1)}\,\hbar\, T_{k,q\pm1} = \sqrt{k\,(k+1) - q\,(q \pm 1)}\,\hbar\, T_{k,q\pm1}. \tag{K.3b}$$

The $T_{k,q}$ of *integral* rank are called *spherical tensor* operators [93] because they transform like the spherical harmonics $Y_k^q(\hat{\mathbf{r}})$. For this reason the hermitian conjugate of $T_{k,q}$ is given by

$$T^{\dagger}_{k,q} = (-1)^q T_{k,-q} \;\text{ for } k = 0,1,2,\cdots. \tag{K.4}$$

This can be verified with the aid of the commutation relations (K.3).

Problem K.1. Derive the commutation relations K.3 from the transformation properties K.1

Solution. To deal with J_a we consider an infinitesimal rotation $\delta\varphi$ about the direction $\hat{\mathbf{a}}$,

$$P_a(\delta\varphi) T_{k,q} P_a^{\dagger}(\delta\varphi) = \sum_{q'} T_{k,q'} \langle k, q'|P_a(\delta\phi)|k,q\rangle.$$

Using Eq. (3.163) we have $P_R \to P_a(\delta\varphi) = 1 - i\delta\varphi J_a/\hbar$ and the l.h.s. becomes

$$P_a(\delta\varphi) T_{k,q} P_a^{\dagger}(\delta\varphi) = T_{k,q} - \frac{i}{\hbar}\delta\varphi J_a T_{k,q} + \frac{i}{\hbar}\delta\varphi T_{k,q} J_a + \frac{1}{\hbar^2}(\delta\varphi)^2 J_a T_{k,q} J_a.$$

On the r.h.s. the rotation matrix becomes

$$\langle k,q'|P_a(\delta\phi)|k,q\rangle = \delta_{q,q'} - \frac{i}{\hbar}\delta\varphi\langle k,q'|J_a|k,q\rangle.$$

Thus we obtain

$$T_{k,q} - \frac{i}{\hbar}\delta\varphi J_a T_{k,q} + \frac{i}{\hbar}\delta\varphi T_{k,q}J_a + \frac{1}{\hbar^2}(\delta\varphi)^2 J_a T_{k,q} J_a = T_{k,q} - \frac{i}{\hbar}\delta\varphi\sum_{q'}\langle k,q'|J_a|k,q\rangle.$$

Subtracting $T_{k,q}$ from both sides and dividing by $i\delta\varphi/\hbar$ this expression reduces in the limit $\delta\varphi \to 0$ to the form

$$[J_a, T_{k,q}] = \sum_{q'}\langle k,q'|J_a|k,q\rangle T_{k,q}.$$

In particular, this holds for $\hat{\mathbf{a}} \to \hat{\mathbf{x}}$, $\hat{\mathbf{a}} \to \hat{\mathbf{y}}$ and $\hat{\mathbf{a}} \to \hat{\mathbf{z}}$. For $\hat{\mathbf{a}} \to \hat{\mathbf{z}}$ the sum reduces to

$$[J_z, T_{k,q}] = q\,\hbar\, T_{k,q}.$$

Since $J_\pm = J_x \pm iJ_y$ we find

$$[J_\pm, T_{k,q}] = \sqrt{k\,(k+1) - q\,(q\pm 1)}\,\hbar\, T_{k,q\pm 1}.$$ □

K.1.1 Spherical tensor operators of rank zero - scalar operators

The simplest class of irreducible tensor operators are scalar operators $T_{0,0}$. These are irreducible tensor operator of rank 0; *i.e.*, $k = 0$ and $q = 0$, which means that scalar operators only have a single component. Any operator that commutes with J_+, J_- and J_z is a scalar operator in the vector space spanned by the basis $\{|\alpha JM\rangle\}$,

$$[J_z, T_{0,0}] = 0 \tag{K.5a}$$

$$[J_\pm, T_{0,0}] = 0. \tag{K.5b}$$

In other words, scalar operators are invariant under arbitrary rotations of the vector space in which they operate. An example of a scalar operator is the operator $\mathbf{L}^2$ in the vector space $\{|LM\rangle\}$. Operators like p_r and r or any function $f(r)$ of r are scalar operators within the vector space $\{|nlm\rangle\}$ of hydrogenic atoms with only principal structure.

K.1.2 Spherical tensor operators of rank 1 - vector operators

Vector operators are defined as operators with the same transformation properties as the radius vector $\mathbf{r}$ under proper rotations; *i.e.*, rotations in which the handedness of the coordinate system is conserved. In the *cartesian basis* an arbitrary real vector operator $\mathbf{T}$ can be written as

$$\mathbf{T} = \hat{\mathbf{x}}T_x + \hat{\mathbf{y}}T_y + \hat{\mathbf{z}}T_z \Leftrightarrow \begin{pmatrix} T_x \\ T_y \\ T_z \end{pmatrix}, \tag{K.6}$$

where T_x, T_y, T_z are called the cartesian components. Vector operators can be expressed as irreducible tensor operators of rank 1; i.e., $k = 1$ and $q = 0, \pm 1$ and we write $\mathbf{T}$ for $\mathrm{T}^{(1)}$. In the standard phase convention (*Condon and Shortley*) the *standard components* of $\mathbf{T}$ are given by s

$$T_{1,0} = T_z, \qquad T_{1,\pm 1} = \mp\sqrt{\tfrac{1}{2}}(T_x \pm iT_y). \tag{K.7}$$

Substituting the inverse relations

$$T_x = -\sqrt{\tfrac{1}{2}}(T_{1,+1} - T_{1,-1}) \quad \text{and} \quad T_y = i\sqrt{\tfrac{1}{2}}(T_{1,+1} + T_{1,-1}) \tag{K.8}$$

into Eq. (K.6) the standard decomposition takes the form

$$\mathbf{T} = \hat{\mathbf{u}}_{+1}^{*} T_{1,+1} + \hat{\mathbf{u}}_0 T_{1,0} + \hat{\mathbf{u}}_{-1}^{*} T_{1,-1}, \tag{K.9}$$

where the $\hat{\mathbf{u}}_q$, with $q = 0, \pm 1$, are the *spherical* unit vectors

$$\hat{\mathbf{u}}_{+1} = -\sqrt{\tfrac{1}{2}}(\hat{\mathbf{x}} + i\hat{\mathbf{y}}), \qquad \hat{\mathbf{u}}_{-1} = +\sqrt{\tfrac{1}{2}}(\hat{\mathbf{x}} - i\hat{\mathbf{y}}), \qquad \hat{\mathbf{u}}_0 = \hat{\mathbf{z}}. \tag{K.10}$$

The commutation relations (K.3) become

$$\begin{array}{ll} [J_z, T_{1,0}] = 0 & [J_z, T_{1,\pm1}] = \pm\hbar\, T_{1,\pm1} \\ {[J_\pm, T_{1,0}]} = \sqrt{2}\,\hbar\, T_{1,\pm1} & [J_\pm, T_{1,\mp1}] = \sqrt{2}\,\hbar\, T_{1,0}. \end{array} \tag{K.11}$$

In particular, the *standard components* J_{+1}, J_0, J_{-1} of the angular momentum operator $\mathbf{J}$ are defined as

$$J_0 = J_z; \qquad J_{\pm1} = \mp\sqrt{\tfrac{1}{2}}(J_x \pm iJ_y) = \mp\sqrt{\tfrac{1}{2}}J_\pm. \tag{K.12}$$

Using these relations the commutations take the form

$$\begin{array}{ll} [T_{1,0}, T_{1,0}] = 0 & [T_{1,0}, T_{1,\pm1}] = \pm\hbar\, T_{1,\pm1} \\ {[T_{1,\pm1}, T_{1,0}]} = \mp\hbar\, T_{1,\pm1} & [T_{1,\pm1}, T_{1,\mp1}] = \mp\hbar\, T_{1,0}. \end{array} \tag{K.13}$$

Example K.1. The simplest example of a vector operator is the radius vector *operator* $\underline{\mathbf{r}}$ with cartesian components $\underline{x}, \underline{y}, \underline{z}$ and standard components

$$T_{1,0} = \underline{z}, \qquad T_{1,\pm1} = \mp\sqrt{\tfrac{1}{2}}(\underline{x} \pm i\underline{y}), \tag{K.14}$$

which can be rewritten in the form

$$T_{1,0} = \underline{r}\cos\underline{\theta} = \sqrt{4\pi/3}\,\underline{r}Y_1^0(\underline{\hat{\mathbf{r}}}), \tag{K.15a}$$

$$T_{1,\pm1} = \mp\sqrt{\tfrac{1}{2}}\,\underline{r}(\sin\underline{\theta}\cos\underline{\phi} \pm i\sin\underline{\theta}\sin\underline{\phi}) = \sqrt{4\pi/3}\,\underline{r}Y_1^{\pm1}(\underline{\hat{\mathbf{r}}}), \tag{K.15b}$$

where $\underline{r} = |\underline{\mathbf{r}}|$ and $\underline{\hat{\mathbf{r}}}$ is the operator for the direction unit vector. Note that these standard components transform like the spherical harmonics,

$$T_{1,m} = \sqrt{4\pi/3}\,\underline{r}Y_1^m(\underline{\hat{\mathbf{r}}}). \tag{K.16}$$

Substituting these quantities into the commutation relations (K.11) we find

$$\begin{array}{ll} [L_z, Y_1^0(\underline{\hat{\mathbf{r}}})] = 0 & [L_z, Y_1^{\pm1}(\underline{\hat{\mathbf{r}}})] = \pm\hbar Y_1^{\pm1}(\underline{\hat{\mathbf{r}}}) \\ {[L_\pm, Y_1^0(\underline{\hat{\mathbf{r}}})]} = \sqrt{2}\hbar Y_1^{\pm1}(\underline{\hat{\mathbf{r}}}) & [L_\pm, Y_1^{\mp1}(\underline{\hat{\mathbf{r}}})] = \sqrt{2}\hbar Y_1^0(\underline{\hat{\mathbf{r}}}). \end{array} \tag{K.17}$$

Using the relations (1.53) and (1.55) we readily verify (with the aid of a test function) that these commutation relations are indeed satisfied.[1]

Example K.2. A second example of an irreducible vector operator is the angular momentum operator $\mathbf{L}$, which acts in the vector space spanned by the basis $\{|LM\rangle\}$ of the standard representation $\{L^2, L_Z\}$. The operator $\mathbf{L}$ is a (pseudo) vector with cartesian components L_x, L_y, L_z and standard components

$$T_{1,0} = L_z, \qquad T_{1,\pm1} = \mp\sqrt{\tfrac{1}{2}}(L_x \pm iL_y) = \mp\sqrt{\tfrac{1}{2}}L_\pm. \tag{K.18}$$

Substituting these definitions into the commutation relations (K.11) we find

$$\begin{array}{ll} [L_z, L_z] = 0 & [L_z, L_\pm] = \pm\hbar L_\pm \\ {[L_\pm, L_z]} = \mp\hbar L_\pm & [L_\pm, L_\mp] = \pm 2\hbar L_z, \end{array} \tag{K.19}$$

which coincide with the commutation relations derived in Chapter 1, as it should be.

---

[1] Note that a test function $\varphi(\mathbf{r})$ is a function of $\mathbf{r} = \{x, y, z\}$. The operator $\underline{\mathbf{r}}$ acting on this test function yields the variable $\mathbf{r}$.

K.1.3 Generalization - spherical tensor operators of rank 2

We now generalize the vector concept by introducing tensors in the cartesian space. The simplest tensor is a tensor of rank 0. This is a quantity which is invariant under proper rotations about the origin of the cartesian space. A tensor of rank 1 is a quantity with the same transformation properties under proper rotations as the fundamental vector

$$\mathbf{r} = \hat{\mathbf{x}}x + \hat{\mathbf{y}}y + \hat{\mathbf{z}}z \Leftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{K.20}$$

where $\hat{\mathbf{u}} = \partial_u \mathbf{r}$, with $u \in \{x, y, z\}$. A tensor of rank 2 is defined as a quantity with the same transformation properties under proper rotations as the fundamental tensor

$$\overleftrightarrow{\mathbf{r}} = \hat{\mathbf{x}}\hat{\mathbf{x}}xx + \hat{\mathbf{x}}\hat{\mathbf{y}}xy + \hat{\mathbf{x}}\hat{\mathbf{z}}xz + \hat{\mathbf{y}}\hat{\mathbf{x}}yx + \hat{\mathbf{y}}\hat{\mathbf{y}}yy + \hat{\mathbf{y}}\hat{\mathbf{z}}yz + \hat{\mathbf{z}}\hat{\mathbf{x}}zx + \hat{\mathbf{z}}\hat{\mathbf{y}}zy + \hat{\mathbf{z}}\hat{\mathbf{z}}zz \Leftrightarrow \begin{pmatrix} xx\;xy\;xz \\ yx\;yy\;yz \\ zx\;zy\;zz \end{pmatrix}, \tag{K.21}$$

where $\hat{\mathbf{u}}\hat{\mathbf{v}} = \partial_{uv} \overleftrightarrow{\mathbf{r}}$, with $u, v \in \{x, y, z\}$. It is straightforward to generalize this concept and define tensors of higher rank.

K.2 Wigner-Eckart theorem

Let $T_{k,q}$ be the standard components of an irreducible tensor operator in the angular momentum vector space defined by the standard representation $\{J^2, J_z\}$. The Wigner-Eckart theorem states that any matrix element of $T_{k,q}$ in this representation can be reduced to the form

$$\langle J'M'|T_{k,q}|JM\rangle = \frac{1}{\sqrt{2J'+1}} \langle J'||T_k||J\rangle (JM; kq|J'M'\rangle, \tag{K.22}$$

where $(JM, kq|J'M'\rangle$ is a Clebsch-Gordan coefficient and the factor $\langle J'||T_k||J\rangle$ stands for all properties of the matrix element that do not depend on M, M' and q. This factor is called the *reduced matrix element* of the tensor for given values of J and J'. In a way, the notation $\langle J'||T_k||J\rangle$ suggests more than it actually delivers because it only points to the *existence* of such a factor.[1] Its value has to be determined on a case by case basis by explicit calculation of one of the matrix elements $\langle J'M'|T_{k,q}|JM\rangle$. Importantly, since $\langle J'||T_k||J\rangle$ is independent of M, M' and q we can use the most convenient matrix element for this purpose. Some key examples are given in Section K.3.

In terms of $3j$ symbols, the theorem takes the form

$$\langle J'M'|T_{k,q}|JM\rangle = (-1)^{J-k+M'} \langle J'||T_k||J\rangle \begin{pmatrix} J & k & J' \\ M & q & -M' \end{pmatrix}. \tag{K.23}$$

This form is best suited for rearranging the quantum numbers (taking advantage of the convenient symmetry properties of the $3j$ symbols). For *integral* k the expression is equivalent to

$$\langle J'M'|T_{k,q}|JM\rangle = (-1)^{J'-M'} \langle J'||T_k||J\rangle \begin{pmatrix} J' & k & J \\ -M' & q & M \end{pmatrix}. \tag{K.24}$$

To arrive at this result we used the permutation rule for $3j$ symbols. Furthermore, since $J' + M'$ is always integral, we have $(-1)^{J'+M'} = (-1)^{-J'-M'}$. As k was taken to be integral also $J - J'$ is

---

[1] The presence of the prefactor $1/\sqrt{2J'+1}$ is a matter of taste. Being independent of M, M' and q, this prefactor can equally well be included in $\langle J'||T_k||J\rangle$ (as some authors do).

integral and $(-1)^{J-J'} = (-1)^{-J+J'}$. In preparation for the application of sum rules it is useful to rewrite the Wigner-Eckart theorem in the form

$$\langle J'M'|T_{k,q}|JM\rangle = (-1)^{J-M}\frac{1}{\sqrt{2k+1}}\langle J'||T_k||J\rangle(J'M';J,-M|kq\rangle. \tag{K.25}$$

K.2.1 Reduction of matrix elements

Properties of the reduced matrix elements can be derived with the aid of *sum rules.* As will be shown, these rules are most conveniently applied in the Dirac notation starting from Eq. (K.25). This expression can be inverted by multiplying both sides by $(-1)^{J-M}\langle k'q'|J',M;J,-M)$ and summing over M and M'. For the evaluation we return to the $3j$ notation

$$\langle J'||T_k||J\rangle = (2k+1)\sum_{M,M'}(-1)^{J'-M'}\begin{pmatrix} J' & k & J \\ -M' & q & M\end{pmatrix}\langle J'M'|T_{k,q}|JM\rangle. \tag{K.26}$$

Another example is to sum both sides of the Wigner-Eckart theorem (K.22) over M and q,

$$\begin{aligned}\sum_{M,q}|\langle J'M'|T_{k,q}|JM\rangle|^2 &= \frac{1}{2J'+1}|\langle J'||T_k||J\rangle|^2\sum_{M,q}\langle J'M'|JM,k,q)(JM,k,q|J'M'\rangle \\ &= \frac{1}{2J'+1}|\langle J'||T_k||J\rangle|^2.\end{aligned}$$

Next we sum over all $2J'+1$ possible values of M' and arrive at

$$\sum_{M',M,q}|\langle J'M'|T_{k,q}|JM\rangle|^2 = |\langle J'||T_k||J\rangle|^2, \tag{K.27}$$

which is manifestly independent of M, M' and q.

K.2.2 Proof of Wigner-Eckart theorem

The proof proceeds in four steps

1. First we show that the matrix elements satisfy the projection rule (J.5) of the $3j$ symbols. Using the commutation relation (K.3a) we have

   $$q\hbar\langle J'M'|T_{k,q}|JM\rangle = \langle J'M'|[J_z,T_{k,q}]|JM\rangle.$$

   Expanding the commutator and using the hermiticity of J_z this becomes

   $$q\langle J'M'|T_{k,q}|JM\rangle = (M'-M)\,\langle J'M'|T_{k,q}|JM\rangle. \tag{K.28}$$

   Thus we obtained the projection rule (J.5): $\langle J'M'|T_{k,q}|JM\rangle = 0$ unless $q = M'-M$, which can be expressed as

   $$\langle J'M'|T_{k,q}|JM\rangle = \lambda\langle J'M'|J_q|JM\rangle \tag{K.29}$$

   because also $\langle J'M'|J_q|JM\rangle$ is zero unless $q = M'-M$. Our conjecture is that the proportionality constant λ corresponds to the reduced matrix element that we are looking for,

   $$\lambda = (2J'+1)^{-1/2}\langle J'||T_k||J\rangle. \tag{K.30}$$

   However, for the time being it remains to be shown that in Eq. (K.29) the *same* value for λ can be used for any combination of M, M' and q; in other words, we have to show that the value of λ is *independent* of M, M' and q. For the special case of vector operators ($k=1$) we recognize in Eq. (K.29) the structure of the *vector projection rule* used in the addition of angular momenta; *e.g.*, for $\mathbf{J} = \mathbf{L}+\mathbf{S}$ we find for $T_{1,q} = L_q$

   $$\langle J'M'|L_q|JM\rangle = \lambda\langle J'M'|J_q|JM\rangle. \tag{K.31}$$

2. To prove the Wigner-Eckart theorem we have to show that

$$\langle J'M'|T_{k,q}|JM\rangle = \lambda \langle JM,k,q|J'M'\rangle, \tag{K.32}$$

where the differences between different combinations of M, M' and q are contained in the Clebsch-Gordan coefficient $\langle JM,k,q|J'M'\rangle$ and the prefactor λ is *independent* of these values. This suggests the strategy to search for two equivalent recursion relations for *given* values of J, J' and k, one for the matrix elements $\langle J'M'|T_{k,q}|JM\rangle$ and the other for the corresponding Clebsch-Gordan coefficients $\langle JM,k,q|J'M'\rangle$, both relating terms with different combinations of M, M' and q. For the matrix elements such a relation follows from the commutation relation (K.3b),

$$\sqrt{k(k+1)-q(q\pm 1)}\hbar\langle J'M'|T_{k,q\pm 1}|JM\rangle = \langle J'M'|[J_\pm, T_{k,q}]|JM\rangle \tag{K.33}$$

Expanding the commutator and using the property that J_+ and J_- are hermitian conjugates we find

$$\begin{aligned}\sqrt{k(k+1)-q(q\pm 1)}\langle J'M'|T_{k,q\pm 1}|JM\rangle &= \\ &= \sqrt{J'(J'+1)-M'(M'\mp 1)}\langle J'(M'\mp 1)|T_{k,q}|JM\rangle - \\ &\quad - \sqrt{J(J+1)-M(M\pm 1)}\langle J'M'|T_{k,q}|J(M\pm 1)\rangle.\end{aligned} \tag{K.34}$$

3. To find the equivalent recursion relation for the Clebsch-Gordan coefficients we use the addition of two angular momenta $\mathbf{J}' = \mathbf{J} + \mathbf{k}$ and apply Clebsch-Gordan decomposition of the *coupled* states $|J'M'\rangle$ onto the *uncoupled* basis $\{|JM,kq\rangle\}$. The properties of the coupling vector, $\mathbf{k}$, is determined by the rank, k, of the tensor $T_{k,q}$. If $T_{k,q}$ represents a vector operator $\mathbf{J}$ and $\mathbf{J}'$ can differ by at most one unit of angular momentum ($k=1$ and $-1 \le q \le 1$); for second-rank tensors by at most two units of angular momentum ($k=2$ and $-2 \le q \le 2$), etc.. To obtain the recursion relation we use the operator identity

$$\lambda J'_\pm = \lambda J_\pm + \lambda k_\pm, \tag{K.35}$$

where λ can be any constant. First, we operate the l.h.s. onto the state $|J'M'\rangle$ and find

$$\lambda J'_\mp|J'M'\rangle = \lambda\sqrt{J'(J'+1)-M'(M'\mp 1)}\hbar|J'(M'\mp 1)\rangle \tag{K.36}$$

and subsequently, after decomposition of $|J'(M'\mp 1)\rangle$ with respect to the basis $\{|JM,kq\rangle\}$, this becomes

$$\lambda J'_\mp|J'M'\rangle = \lambda\sqrt{J'(J'+1)-M'(M'\mp 1)}\hbar\sum_{Mq}|JM,kq\rangle\langle JM,kq|J'(M'\mp 1)\rangle. \tag{K.37}$$

Note that at this point we already obtained an expression that carries the same prefactor and quantum numbers as the first term on the r.h.s. of Eq. (K.34). Second, we start with the decomposition

$$|J'M'\rangle = \sum_{Mq}|JM,kq\rangle\langle JM,kq|J'M'\rangle \tag{K.38}$$

and subsequently operate the r.h.s. of the identity (K.35) onto the states $|JM,kq\rangle$,

$$\begin{aligned}\lambda J'_\mp|J'M'\rangle = \sum_{Mq}&\lambda\sqrt{J(J+1)-M(M\mp 1)}\hbar|J(M\mp 1),kq\rangle\langle JM,kq|J'M'\rangle \\ &+\sum_{Mq}\lambda\sqrt{k(k+1)-q(q\mp 1)}\hbar|JM,k(q\mp 1)\rangle\langle JM,kq|J'M'\rangle.\end{aligned} \tag{K.39}$$

Note that we can bring λ under the summation as long as it is independent of M, M' and q. Comparing this expression with the recursion relation (K.34) we note that the Clebsch-Gordan coefficients have to be shifted to serve our purpose. This is easily realized by relabeling of quantum numbers ($M \to M'' \pm 1$ and $q \to q'' \pm 1$) and summing over the new dummy indices (M'' or q'') which is allowed because the projection rule (J.5) assures that the summation total is not affected,

$$\lambda J'_{\mp}|J'M'\rangle = \sum_{M''q} \lambda\sqrt{J\,(J+1) - M''\,(M''\pm 1)}\hbar|JM'', kq\rangle\langle J(M''\pm 1), kq|J'M'\rangle$$
$$+ \sum_{Mq''} \lambda\sqrt{k\,(k+1) - q''\,(q''\pm 1)}\hbar|JM, kq''\rangle\langle JM, k(q''\pm 1)|J'M'\rangle. \quad \text{(K.40)}$$

After again renaming the dummy indices, $M'' \to M$ and $q'' \to q$, we equate the expansions (K.37) and (K.40) term by term and obtain the relation

$$\begin{aligned}\lambda\sqrt{k\,(k+1) - q\,(q\pm 1)}\langle JM, k\,(q\pm 1)\,|J'M'\rangle &= \\ = \lambda\sqrt{J'\,(J'+1) - M'\,(M'\mp 1)}&\langle JM, kq|J'\,(M'\mp 1)\rangle - \\ - \lambda\sqrt{J\,(J+1) - M\,(M\pm 1)}&\langle J\,(M\pm 1)\,, kq|J'M'\rangle, \end{aligned} \quad \text{(K.41)}$$

where we know that λ can depend on J, J' and k but certainly does not depend on M, M' and q; *i.e.*, this recursion relation holds for any value of of M, M' and q. Note that all square root prefactors coincide with those of Eq. (K.34). So, with Eq. (K.41), we succeeded in our goal to relate the matrix elements of Eq. (K.34) to Clebsch-Gordan coefficients multiplied by a coefficient λ which does not depend on the combination of M', M and q; comparing Eqs. (K.34) and (K.41) we identify

$$\left.\begin{aligned}\langle J'M'|T_{k,q\pm 1}|JM\rangle &= \lambda\langle JM, k\,(q\pm 1)\,|J'M'\rangle \\ \langle J'\,(M'\mp 1)\,|T_{k,q}|JM\rangle &= \lambda\langle JM, kq|J'\,(M'\mp 1)\rangle \\ \langle J'M'|T_{k,q}|J\,(M\pm 1)\rangle &= \lambda\langle J\,(M\pm 1)\,, kq|J'M'\rangle\end{aligned}\right\} \quad \text{(K.42)}$$

Using the appropriate change of the dummy indices ($q \pm 1 \to q$, $M' \mp 1 \to M'$, $M \pm 1 \to M$) these three equations all reduce to a single relation,

$$\langle J'M'|T_{k,q}|JM\rangle = \lambda\langle JM, kq|J'M'\rangle. \quad \text{(K.43)}$$

4. As this relation holds with the same value of λ for any combination of M, M' and q, we have proved the existence of a prefactor with the properties of the reduced matrix element introduced in Eq. (K.22) and we may adopt the notation

$$\lambda = (2J'+1)^{-1/2}\,\langle J'||T_k||J\rangle, \quad \text{(K.44)}$$

which completes the proof.

K.3 Examples

K.3.1 Reduced matrix elements and hermitian conjugation of spherical tensor operators

Consider the following two expressions based on the Wigner-Eckart theorem:

$$\langle J'M'|T^{\dagger}_{k,q}|JM\rangle = \langle JM|T_{k,q}|J'M'\rangle^* = (-1)^{J'-k+M}\langle J||T_k||J'\rangle^* \begin{pmatrix} J' & k & J \\ M' & q & -M \end{pmatrix} \quad \text{(K.45a)}$$

$$\langle J'M'|T_{k,-q}|JM\rangle = (-1)^{J-k+M'}\langle J'||T_k||J\rangle \begin{pmatrix} J & k & J' \\ M & -q & -M' \end{pmatrix}. \quad \text{(K.45b)}$$

Eliminating the $3j$ symbols and using the condition $q = M - M'$ these expressions combine to

$$\langle J'M'|T_{k,q}^{\dagger}|JM\rangle = \frac{(-1)^{J-J'+q}\langle J||T_k||J'\rangle^*}{\langle J'||T_k||J\rangle}\langle J'M'|T_{k,-q}|JM\rangle. \tag{K.46}$$

Since $\langle J'||T_k||J\rangle$ is independent of M, M' and q we can determine the relation between $\langle J'||T_k||J\rangle$ and $\langle J||T_k||J'\rangle^*$ by choosing the most convenient matrix element for this purpose. For *integer rank* we can choose $q = 0$ and obtain

$$\langle J'M'|T_{k,0}^{\dagger}|JM\rangle = \frac{(-1)^{J-J'}\langle J||T_k||J'\rangle^*}{\langle J'||T_k||J\rangle}\langle J'M'|T_{k,0}|JM\rangle. \tag{K.47}$$

Since, for *integer rank*, $T_{k,0}$ is invariant under rotation, we have $T_{k,0}^{\dagger} = T_{k,0}$ and the above equation reduces to

$$\langle J'||T_k||J\rangle = (-1)^{J-J'}\langle J||T_k||J'\rangle^*. \tag{K.48}$$

K.3.2 Reduced matrix elements for angular momentum operators

To determine the reduced matrix element $\langle J'||J||J\rangle$ the Wigner-Eckart theorem is written in the form

$$\langle J'||J||J\rangle = (-1)^{J'-M'}\langle J'M'|J_q|JM\rangle \begin{pmatrix} J' & 1 & J \\ -M' & q & M \end{pmatrix}^{-1}. \tag{K.49}$$

Since $\langle J'||T_k||J\rangle$ is independent of M, M' and q we can use the most convenient matrix element to determine its value. We demonstrate this by choosing $q = 0$ and $M = M' = J$ and obtain with the aid of Eq. (J.18a)

$$\langle J'||J||J\rangle = \langle J'J'|J_0|JJ\rangle \begin{pmatrix} J' & 1 & J \\ -J' & 0 & J \end{pmatrix}^{-1} = \sqrt{J(J+1)(2J+1)}\,\delta_{J,J'}. \tag{K.50}$$

Hence, for arbitrary q, M and M' the angular momentum matrix elements can be written in the form

$$\langle J'M'|J_q|JM\rangle = (-1)^{J+M'-1}\sqrt{J(J+1)(2J+1)} \begin{pmatrix} J & 1 & J \\ M & q & -M' \end{pmatrix} \delta_{J,J'}. \tag{K.51}$$

Using Eqs. (J.18) it is readily verified that this one equation is equivalent to the set of Eqs. (3.22b) and (3.24b).

K.3.3 Reduced matrix elements for the spherical harmonics

As a second example we discuss the determination of the reduced matrix elements of tensor operators of the spherical harmonic type. This is done by comparing the expression obtained by applying the Wigner-Eckart theorem with the result of the Gaunt integral (L.59).

$$\langle l'm'|Y_k^q(\hat{\mathbf{r}})|lm\rangle = \begin{cases} (-1)^{m'+l'}\langle l'||Y_k(\hat{\mathbf{r}})||l\rangle \begin{pmatrix} l' & k & l \\ -m' & q & m \end{pmatrix} \\ (-1)^{m'}\sqrt{\dfrac{(2l'+1)(2k+1)(2l+1)}{4\pi}} \begin{pmatrix} l' & k & l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l' & k & l \\ -m' & q & m \end{pmatrix}. \end{cases} \tag{K.52}$$

This comparison yields

$$\langle l'||Y_k(\hat{\mathbf{r}})||l\rangle = (-1)^{-l'}\sqrt{\frac{(2l'+1)(2k+1)(2l+1)}{4\pi}} \begin{pmatrix} l' & k & l \\ 0 & 0 & 0 \end{pmatrix}, \tag{K.53}$$

which is only nonzero for even values of $k+l+l'$.

Some special cases are:

- $k=0$: Substituting the $3j$ symbol (J.23) we find
$$\langle l'||\sqrt{4\pi}Y_0(\hat{\mathbf{r}})||l\rangle = \sqrt{2l+1}\,\delta_{l,l'}. \tag{K.54}$$
- $k=1$: Substituting the $3j$ symbol (J.24) we find
$$\langle l'||\sqrt{4\pi/3}Y_1(\hat{\mathbf{r}})||l\rangle = (-1)^{-l'+\max(l,l')}\sqrt{\max(l,l')}\,\delta_{l',l\pm1} = \pm\sqrt{\max(l,l')}\,\delta_{l',l\pm1}. \tag{K.55}$$
- $k=2$: Substituting the $3j$ symbol (J.25) we find
$$\langle l'||\sqrt{4\pi/5}Y_2(\hat{\mathbf{r}})||l\rangle = \begin{cases} -\sqrt{\dfrac{l\,(l+1)\,(2l+1)}{(2l+3)(2l-1)}} & l'=l \\ +\sqrt{\max(l,l')\left(\dfrac{3}{4}\dfrac{l+l'}{l+l'+1}\right)} & l'=l\pm2. \end{cases} \tag{K.56}$$

K.3.4 Reduction of matrix elements of vector operators

As a first example we consider the reduction of the matrix elements $\langle lsJ'M'|\,L_q\,|lsJM\rangle$ of the orbital angular momentum operator $\mathbf{L}$ expressed in the basis $\{|JM\rangle\}$ of the standard representation $\{J^2,J_z\}$ of the total angular momentum operator $\mathbf{J}=\mathbf{L}+\mathbf{S}$. Our aim is to show that
$$\langle lsJ\,(M\pm q)|\,L_q\,|lsJM\rangle = \langle lsJ||L||lsJ\rangle\,\langle lsJ\,(M\pm q)|\,J_q\,|lsJM\rangle \qquad \text{for } q=0,\pm1, \tag{K.57}$$
where L_q and J_q are the standard components of the vector operators $\mathbf{L}$ and $\mathbf{J}$, respectively, and $\langle lsJ||L||lsJ\rangle$ is a proportionality constant independent of M and q.

We start by applying the Wigner-Eckart theorem (K.24) to the standard components L_q of the rank 1 tensor $\mathbf{L}$,
$$\langle lsJM'|\,L_q\,|lsJM\rangle = (-)^{J'-M'}\langle lsJ||L||lsJ\rangle\begin{pmatrix} J & 1 & J \\ -M' & q & M\end{pmatrix} \qquad \text{with } q=0,\pm1. \tag{K.58}$$

These matrix elements are only non-zero for $M'=M+q$. We evaluate subsequently the cases $q=0,\pm1$ using the expressions (J.18a) and (J.18b) for $3j$ symbols,
$$\begin{aligned}\langle lsJM|\,L_z\,|lsJM\rangle &= (-)^{J-M}\langle lsJ||L||lsJ\rangle\begin{pmatrix} J & 1 & J \\ -M & 0 & M\end{pmatrix} \\ &= \frac{\langle lsJ||L||lsJ\rangle}{\sqrt{J\,(J+1)\,(2J+1)}}\,M\,\hbar \end{aligned}\tag{K.59a}$$

$$\begin{aligned}\langle lsJ\,(M\pm1)|\,L_\pm\,|lsJM\rangle &= (-)^{J-M\mp1}\langle lsJ||L||lsJ\rangle\sqrt{2}\,\hbar\begin{pmatrix} J & 1 & J \\ -(M\pm1) & \pm1 & M\end{pmatrix} \\ &= \frac{\langle lsJ||L||lsJ\rangle}{\sqrt{J\,(J+1)\,(2J+1)}}\sqrt{J(J+1)-M(M\pm1)}\,\hbar. \end{aligned}\tag{K.59b}$$

Absorbing the joint prefactor $\sqrt{J\,(J+1)\,(2J+1)}$ into the reduced matrix element (it is independent of M, M' and q) and using the definitions of the J_z and $J_\pm$ operators we can rewrite Eqs. (K.59a) and (K.59b) in the form
$$\langle lsJM|\,L_z\,|lsJM\rangle = \langle lsJ||L||lsJ\rangle\,\langle lsJM|\,J_z\,|lsJM\rangle \tag{K.60a}$$
$$\langle lsJ\,(M\pm1)|\,L_\pm\,|lsJM\rangle = \langle lsJ||L||lsJ\rangle\,\langle lsJ\,(M\pm1)|\,J_\pm\,|lsJM\rangle\,. \tag{K.60b}$$

Here we recognize the *vector projection rule* used for describing the atomic fine structure in the low-field limit. Using the definitions $J_0 = (J_z/\hbar)$ and $J_{\pm 1} = \sqrt{1/2}\,(J_\pm/\hbar)$ these equations reduce to Eq. (K.57).

K.3.5 Calculation of reduced matrix elements in coupled bases

The Wigner-Eckart theorem allows the reduction of matrix elements involving coupled angular momenta. In this section we illustrate this for two angular momentum operators, $\boldsymbol{j_1}$ and $\boldsymbol{j_2}$, acting in the two (independent) subspaces of the representations $\{\boldsymbol{j}_1^2, j_{1z}\}$ and $\{\boldsymbol{j}_2^2, j_{2z}\}$, respectively, and together spanning the angular momentum space $\{\mathbf{J}^2, J_z\}$, where $\mathbf{J} = \boldsymbol{j}_1 + \boldsymbol{j}_2$ is the coupled angular momentum. Let us consider matrix elements of the type $\langle j_1' j_2 J' M' | T_{1:k,q} | j_1 j_2 J M \rangle$, where $T_{1:k,q}$ is a spherical tensor operator acting in *one* of the subspaces, the subspace defined by the basis $\{|j_1, m_1\}$, but the matrix element is expressed in the basis $\{|JM\rangle\}$ of the standard representation $\{\mathbf{J}^2, J_z\}$. Applying the Wigner-Eckart theorem we obtain

$$\langle j_1' j_2 J' M' | T_{1:k,q} | j_1 j_2 J M \rangle = (-1)^{J'-M'} \langle j_1' j_2 J' || \mathrm{T}_1^{(k)} || j_1 j_2 J \rangle \begin{pmatrix} J' & k & J \\ -M' & q & M \end{pmatrix}. \tag{K.61}$$

In view of the properties of the $3j$ symbols this matrix element is nonzero *only* for $q = M' - M$. In other words, for given values of M and M' only a *single* component of the tensor $\mathrm{T}_1^{(k)}$ is nonzero. Furthermore, as $\mathrm{T}_1^{(k)}$ operates within the subspace $\{|j_1, m_1\}$ it leaves the quantum numbers j_2 and m_2 unaffected. Hence, only the matrix elements diagonal in j_2 and m_2 are nonzero and it is our task is to determine reduced matrix elements of the type $\langle j_1' j_2 J' || \mathrm{T}_1^{(k)} || j_1 j_2 J \rangle$. In Problem K.2 we show that these are given by the expression for *repeated reduction*

$$\langle j_1' j_2 J' || \mathrm{T}_1^{(k)} || j_1 j_2 J \rangle = (-1)^{J+j_1'+j_2+j_1} \langle j_1' || \mathrm{T}_1^{(k)} || j_1 \rangle \sqrt{(2J+1)(2J'+1)} \begin{Bmatrix} j_1' & J' & j_2 \\ J & j_1 & k \end{Bmatrix}, \tag{K.62}$$

where $\langle j_1' || \mathrm{T}_1^{(k)} || j_1 \rangle$ is the reduced matrix element of $\mathrm{T}_1^{(k)}$ within the subspace $\{\boldsymbol{j}_1^2, j_{1z}\}$ as given by (K.50). Note that the reduced matrix element is independent of magnetic quantum numbers.

Analogously, we can reduce the matrix elements of scalar invariants of the type

$$\mathrm{T}_1^{(k)} \cdot \mathrm{T}_2^{(k)} = \sum_{m=-k}^{k} (-1)^m T_{1:k,m} T_{2:k,-m}, \tag{K.63}$$

where $\mathrm{T}_1^{(k)}$ and $\mathrm{T}_2^{(k)}$ operate in the two (independent) subspaces of the representations $\{\boldsymbol{j}_1^2, j_{1z}\}$ and $\{\boldsymbol{j}_2^2, j_{2z}\}$, respectively. In this case the repeated reduction (K.62) leads to the following expression - see Problem K.3

$$\langle j_1' j_2' J' M' | \mathrm{T}_1^{(k)} \cdot \mathrm{T}_2^{(k)} | j_1 j_2 J M \rangle = (-1)^{j_1+j_2+j_{1'}} \langle j_1' || \mathrm{T}_1^{(k)} || j_1 \rangle \langle j_2' || \mathrm{T}_2^{(k)} || j_2 \rangle \begin{Bmatrix} j_1' & j_2' & J \\ j_2 & j_1 & k \end{Bmatrix} \delta_{J,J'} \delta_{M,M'}. \tag{K.64}$$

Example K.3. As a first example we consider the coupling of the orbital angular momentum $\mathbf{L}$ with the spin $\mathbf{S}$, to form the total electronic angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$. Let us consider matrix elements of the type $\langle l' s j' M' | \, T_{k,q} \, | l s j M \rangle$, where $T_{k,q}$ is a spherical tensor operator operating in the subspace $\{\mathbf{L}^2, L_z\}$ of the orbital angular momentum but the matrix element is expressed in the basis $\{|jM\rangle\}$ of the total electronic angular momentum standard representation $\{\mathbf{J}^2, J_z\}$. With $\langle l' || T_k || l \rangle$ given by (K.50), the expression for the repeated reduction is

$$\langle l' s j' || T_k || l s j \rangle = (-1)^{j+k+l'+s} \langle l' || T_k || l \rangle \sqrt{(2j+1)(2j'+1)} \begin{Bmatrix} l' & j' & s \\ j & l & k \end{Bmatrix}. \tag{K.65}$$

Example K.4. As a second example we consider the coupling of the total electronic angular momentum $\mathbf{J}$ with the nuclear spin $\mathbf{I}$, to form the total angular momentum $\mathbf{F} = \mathbf{J} + \mathbf{I}$. Let us consider matrix elements of the type $\langle j'IF'M'|\, T_{k,q}\, |jIFM\rangle$, where $T_{k,q}$ is a spherical operator operating in the subspace $\{\mathbf{J}^2, J_z\}$ of the orbital angular momentum but the matrix element is expressed in the basis $\{|FM\rangle\}$ of the total electronic angular momentum standard representation $\{\mathbf{F}^2, F_z\}$. With $\langle j'||T_k||j\rangle$ given by (K.50), the expression for the repeated reduction is

$$\langle j'IF||T_k||jIF\rangle = (-1)^{F+k+j'+I}\langle j'||T_k||j\rangle\sqrt{(2F+1)\,(2F'+1)}\begin{Bmatrix} j' & F' & I \\ F & j & k \end{Bmatrix}. \qquad \text{(K.66)}$$

Example K.5. Actually, in the previous example the total electronic angular momentum is itself composed of two angular momenta, $\mathbf{J} = \mathbf{L} + \mathbf{S}$. If the spherical tensor operator T_k only acts in one of the subspaces, say the subspace of $\{\mathbf{L}^2, L_z\}$ we can repeat the reduction by substituting Eq. (K.65) into Eq. (K.66). In this way we reduce the matrix element from the standard representation $\{\mathbf{F}^2, F_z\}$ to the standard representation $\{\mathbf{L}^2, L_z\}$. Using the relation $(-1)^{2k} = 1$ we find

$$\langle j'IF||T_k||jIF\rangle = (-1)^{F+j'+j+l'+s+I}\langle l'||T_k||l\rangle \times \sqrt{(2j+1)\,(2j'+1)\,(2F+1)\,(2F'+1)}\begin{Bmatrix} l' & j' & s \\ j & l & k \end{Bmatrix}\begin{Bmatrix} j' & F' & I \\ F & j & k \end{Bmatrix}. \qquad \text{(K.67)}$$

Problem K.2. Derive Eq. (K.62).

Solution. We start by decomposing Eq. (K.61) with respect to the uncoupled basis $\{|jmsm_s)\}$,

$$\langle j'sJ'M'|\, T_{k,q}\, |jsJM\rangle = \sum_{m,m',m_s} \langle J'M'|j'm'sm_s)\langle j'm'|T_{k,q}|jm\rangle(jmsm_s\,|JM\rangle\,. \qquad \text{(K.68)}$$

Expressing the Clebsch-Gordan coefficients in terms of Wigner $3j$ symbols using Eq. (J.1) we have

$$(jmsm_s|JM\rangle = (-1)^{j-s+M}\sqrt{2J+1}\begin{pmatrix} j & s & J \\ m & m_s & -M \end{pmatrix} \qquad \text{(K.69a)}$$

$$(j'm'sm_s|J'M'\rangle = (-1)^{j'-s+M'}\sqrt{2J'+1}\begin{pmatrix} j' & s & J' \\ m' & m_s & -M' \end{pmatrix}. \qquad \text{(K.69b)}$$

To evaluate $\langle j'm'|T_{k,q}|jm\rangle$ we use the Wigner-Eckart theorem,

$$\langle j'm'|T_{k,q}|jm\rangle = (-1)^{j'-m'}\langle j'||T_k||j\rangle\begin{pmatrix} j' & k & j \\ -m' & q & m \end{pmatrix}. \qquad \text{(K.70)}$$

If the reduced matrix element $\langle j'||T_k||j\rangle$ cannot be reduced any further (i.e., $\boldsymbol{j}$ is not a composite angular momentum itself) it can be determined by direct evaluation of the l.h.s. of Eq. (K.70), as was demonstrated for the $\langle j'||Y_k(\hat{\mathbf{r}})||j\rangle$ in Section K.3.3. To evaluate Eq. (K.68) we substitute Eqs. (K.69a), (K.69b) and (K.70). Collecting the phase factors we obtain

$$(-1)^{2j'+j-2s+M'+M-m'}. \qquad \text{(K.71)}$$

Using the properties of the $3j$ symbols we know that $M = m_s + m$ and $M' = m_s + m'$; hence, $M' + M - m' = 2m_s + m = -M' + 4m_s + m - m_s + m'$. Furthermore, since $4m_s$ is even we have $(-1)^{4m_s} = 1$. Likewise we have $(-1)^{-2s} = (-1)^{2s}$ because $2s$ is an integer. Hence, Eq. (K.68) can be written in the form

$$\langle j'sJ'M'|\, T_{k,q}\, |jsJM\rangle = \langle j'||T_k||j\rangle\sqrt{(2J+1)\,(2J'+1)}\sum_{m,m',m_s}(-1)^{2j'+j-2s-M'+m-m_s+m'} \times \begin{pmatrix} j & s & J \\ m & m_s & -M \end{pmatrix}\begin{pmatrix} j' & s & J' \\ m' & m_s & -M' \end{pmatrix}\begin{pmatrix} j' & k & j \\ -m' & q & m \end{pmatrix}. \qquad \text{(K.72)}$$

Next we bring the triple product of the $3j$ symbols in the order of triple-product summation formula (J.99). For this we first reorder the second and third $3j$ symbols,

$$\langle j'sJ'M'|\,T_{k,q}\,|jsJM\rangle = \langle j'||T_k||j\rangle\sqrt{(2J+1)\,(2J'+1)}\sum_{m,m',m_s}(-1)^{2j'+j+2s-M'+m+m_s+m'}$$
$$\times\begin{pmatrix} j & s & J \\ m & -m_s & -M \end{pmatrix}\begin{pmatrix} s & j' & J' \\ m_s & -m' & M' \end{pmatrix}\begin{pmatrix} j' & j & k \\ m' & -m & -q \end{pmatrix}. \qquad \text{(K.73)}$$

Note that we have replaced m_s by $-m_s$, which is allowed because we sum over all values of m_s. At this point we can apply the $3j$ triple-product summation formula and obtain (after using the sign rule for $3j$ symbols)

$$\langle j'sJ'M'|\,T_{k,q}\,|jsJM\rangle = (-1)^{J'+J+k}(-1)^{j'+s-M'}\langle j'||T_k||j\rangle$$
$$\times\sqrt{(2J+1)\,(2J'+1)}\begin{Bmatrix} J' & k & J \\ j & s & j' \end{Bmatrix}\begin{pmatrix} J' & k & J \\ -M' & q & M \end{pmatrix}. \qquad \text{(K.74)}$$

Comparing Eq. (K.74) with Eq. (K.61) we find

$$\langle j'sJ'||T_k||jsJ\rangle = (-1)^{J+k+j'+s}\langle j'||T_k||j\rangle\sqrt{(2J+1)\,(2J'+1)}\begin{Bmatrix} J' & k & J \\ j & s & j' \end{Bmatrix}. \qquad \text{(K.75)}$$

Rearranging the $6j$ symbols using the relations (J.92) and (J.93) we arrive at Eq. (K.62). □

Problem K.3. Derive Eq. (K.64).

Solution. Starting from the scalar invariant $\mathrm{T}_1^{(k)}\cdot\mathrm{T}_2^{(k)} = \sum_{m=-k}^{k}(-1)^m T_{1:k,m}T_{2:k,-m}$, we separate the matrix element into parts for $\mathrm{T}_1^{(k)}$ and $\mathrm{T}_2^{(k)}$, operating in the two (independent) subspaces of the representations $\{\boldsymbol{j}_1^2, j_{1z}\}$ and $\{\boldsymbol{j}_2^2, j_{2z}\}$, respectively,

$$\langle j_1'j_2'J'M'|\mathrm{T}_1^{(k)}\cdot\mathrm{T}_2^{(k)}|j_1j_2JM\rangle = \sum_{q=-k}^{k}(-1)^m\langle j_1'j_2'J'M'|T_{1:k,q}T_{2:k,-q}|j_1j_2JM\rangle$$
$$= \sum_{j_1''j_2''}\sum_{J'',M''}\sum_{q=-k}^{k}(-1)^q\langle j_1'j_2'J'M'|T_{1:k,q}|j_1''j_2''J''M''\rangle\langle j_1''j_2''J''M''|T_{2:k,-q}|j_1j_2JM\rangle.$$

Applying the Wigner-Eckart theorem to each of these parts we obtain

$$\langle j_1'j_2'J'M'|\mathrm{T}_1^{(k)}\cdot\mathrm{T}_2^{(k)}|j_1j_2JM\rangle = \sum_{J'',M''}\sum_{q=-k}^{k}(-1)^{q+J'-M'+J''-M''}$$
$$\langle j_1'j_2'J'||\mathrm{T}_1^{(k)}||j_1j_2'J''\rangle\langle j_1j_2'J''||\mathrm{T}_2^{(k)}||j_1j_2J\rangle\begin{pmatrix} J' & k & J'' \\ -M' & q & M'' \end{pmatrix}\begin{pmatrix} J'' & k & J \\ -M'' & -q & M \end{pmatrix}.$$

In view of the $3j$ projection rule we have $q = M'-M''$ and $(-1)^{q+J'-M'+J''-M''} = (-1)^{J'+J''-2M''}$. Adding and subtracting J'' in the exponent and noting that $2(J''-M'')$ is even, the matrix element becomes after some rearrangement of the left $3j$ symbol

$$\langle j_1'j_2'J'M'|\mathrm{T}_1^{(k)}\cdot\mathrm{T}_2^{(k)}|j_1j_2JM\rangle = \sum_{J''}(-1)^{J'-J''}$$
$$\langle j_1'j_2'J'||\mathrm{T}_1^{(k)}||j_1j_2'J''\rangle\langle j_1j_2'J''||\mathrm{T}_2^{(k)}||j_1j_2J\rangle\sum_{M''=-J''}^{J''}\sum_{q=-k}^{k}\begin{pmatrix} J'' & k & J' \\ -M'' & -q & M' \end{pmatrix}\begin{pmatrix} J'' & k & J \\ -M'' & -q & M \end{pmatrix}.$$

In this form we can apply the orthogonality relation (J.6) and simplify the expression,

$$\langle j_1'j_2'J'M'|\mathrm{T}_1^{(k)}\cdot\mathrm{T}_2^{(k)}|j_1j_2JM\rangle = \sum_{J''}(-1)^{J'-J''}$$
$$\langle j_1'j_2'J'||\mathrm{T}_1^{(k)}||j_1j_2'J''\rangle\langle j_1j_2'J''||\mathrm{T}_2^{(k)}||j_1j_2J\rangle\,\frac{\delta_{J,J'}\delta_{M,M'}}{2J+1}.$$

Using the repeated reduction relation (K.62) twice the expression for the matrix element becomes

$$\langle j_1'j_2'J'M'|\mathrm{T}_1^{(k)}\cdot\mathrm{T}_2^{(k)}|j_1j_2JM\rangle = \langle j_1'||\mathrm{T}_1^{(k)}||j_1\rangle\langle j_2'||\mathrm{T}_2^{(k)}||j_2\rangle$$
$$\sum_{J''}(-1)^{J'+J+2j_1+j_2+2j_2'+j_1'}\,(2J''+1)\begin{Bmatrix} j_1' & J' & j_2\\ J'' & j_1 & k\end{Bmatrix}\begin{Bmatrix} j_2 & J'' & j_1\\ J & j_2' & k\end{Bmatrix}\delta_{J,J'}\delta_{M,M'}.$$

Rearranging the $6j$ symbols for applying the sum rule (J.97) and arrive at the desired expression.

$$\langle j_1'j_2'J'M'|\mathrm{T}_1^{(k)}\cdot\mathrm{T}_2^{(k)}|j_1j_2JM\rangle = \langle j_1'||\mathrm{T}_1^{(k)}||j_1\rangle\langle j_2'||\mathrm{T}_2^{(k)}||j_2\rangle$$
$$(-1)^{J'+J+2j_1+j_2+2j_2'+j_1'}\sum_{J''}(2J''+1)\begin{Bmatrix} J & j_2 & j_1'\\ j_1 & k & J''\end{Bmatrix}\begin{Bmatrix} j_2 & j_1 & J''\\ J & k & j_2'\end{Bmatrix}\delta_{J,J'}\delta_{M,M'}.\quad\square$$

K.4 Composition and reduction of tensor operators

K.4.1 Composition of tensor operators

Let us return to the *Clebsch-Gordan transformation* for the states

$$|k,q\rangle = \sum_{m=-j_1}^{j_1}\sum_{m=-j_2}^{j_2}|j_1,m_1;j_2,m_2\rangle\langle j_1,m_1;j_2,m_2|k,q\rangle. \tag{K.76}$$

We recall from Section 3.7.5 that the standard components $T_{l,m}$ of the spherical tensor operator $\mathrm{T}^{(l)}$ transform under rotation like the $Y_l^m(\hat{\mathbf{r}}) = \langle\hat{\mathbf{r}}|l,m\rangle$. Generalizing the transformation properties to tensors of arbitrary rank, we find that starting from two irreducible tensor operators, $\mathrm{T}_1^{(j_1)}$ and $\mathrm{T}_2^{(j_2)}$, we can construct irreducible operators of rank k with the aid of the transformation[1]

$$[\mathrm{T}_1^{(j_1)}\otimes\mathrm{T}_2^{(j_2)}]_q^k = \sum_{m=-j_1}^{j_1}\sum_{m=-j_2}^{j_2}T_{1:j_1,m_1}T_{2:j_2,m_2}\langle j_1,m_1;j_2,m_2|k,q\rangle. \tag{K.77}$$

Here, the $[\mathrm{T}_1^{(j_1)}\otimes\mathrm{T}_2^{(j_2)}]_q^k$ represent the $(2k+1)$ standard components of the irreducible tensor operator $[\mathrm{T}_1^{(j_1)}\otimes\mathrm{T}_2^{(j_2)}]^{(k)}$. The operator $[\mathrm{T}_1^{(j_1)}\otimes\mathrm{T}_2^{(j_2)}]$ is called the *direct product* of the Tensor operators $\mathrm{T}_1^{(j_1)}$ and $\mathrm{T}_2^{(j_1)}$. In matrix notation, the operator $\mathrm{T}^{(\mathrm{j})}$ is represented by a $(2j+1)\times(2j+1)$ matrix; i.e., it acts in a $(2j+1)$-dimensional space. Turning to $3j$ symbols and using the projection rule (J.5) the decomposition (K.77) takes the form

$$[\mathrm{T}_1^{(j_1)}\otimes\mathrm{T}_2^{(j_2)}]_q^k = \sum_{m=-j_1}^{j_1}T_{1:j_1,m}T_{2:j_2,(q-m)}(-1)^{j_1-j_2+q}\sqrt{2k+1}\begin{pmatrix} j_1 & j_2 & k\\ m & q-m & -q\end{pmatrix}. \tag{K.78}$$

Note that the triangle inequality for $3j$ symbols restricts the rank of $[\mathrm{T}_1^{(j_1)}\otimes\mathrm{T}_2^{(j_1)}]^{(k)}$ to the interval

$$|j_1-j_2|\le k\le j_1+j_2. \tag{K.79}$$

The operator $[\mathrm{T}_1^{(j_1)}\otimes\mathrm{T}_2^{(j_2)}]$ acts in a $(2j_1+1)(2j_2+1)$-dimensional space and is represented by a $(2j_1+1)(2j_2+1)\times(2j_1+1)(2j_2+1)$ matrix, which is in index notation given by $T_{1:j_1,m_1}T_{2:j_2,m_2}$.

---

[1] Note that two irreducible tensor operators of the same rank have the same transformation properties; hence are the same up to a scalar multiplyer.

Tensor contraction - invariance

Specializing to the case $j_1 = j_2 = j$ we find that for every two irreducible tensor operator $T_1^{(j)}$ and $T_2^{(j)}$ we can construct a scalar operator ($k = q = 0$), which, like $Y_0^0(\hat{\mathbf{r}})$, is invariant under pure rotations,

$$[\mathrm{T}_1^{(j)} \otimes \mathrm{T}_2^{(j)}]_0^0 = \sum_{m=-j}^{j} T_{1:j,m} T_{2:j,-m} \begin{pmatrix} j & j & 0 \\ m & -m & 0 \end{pmatrix} = \sum_{m=-j}^{j} T_{1:j,m} T_{2:j,-m} \frac{(-1)^{j-m}}{\sqrt{2j+1}}. \tag{K.80}$$

In case $\mathrm{T}_1^{(j)} = \mathrm{T}_2^{(j)} = T^{(j)}$ this invariant is called the *Casimir invariant* of the tensor $T^{(j)}$.

K.4.2 Reduction of products of tensor operators

We now consider the *inverse* Clebsch-Gordan transformation for the states,

$$|j_1, m_1; j_2, m_2\rangle = \sum_{k=|j_1-j_2|}^{j_1+j_2} \sum_{q=-k}^{k} |k, q\rangle\langle k, q|j_1, m_1; j_2, m_2\rangle. \tag{K.81}$$

we find by following an analogous procedure that the tensor product of two irreducible tensor operators of rank j_1 and j_2 can be written as the sum of irreducible tensor operators of rank k, with k again restricted to the interval (K.79),

$$T_{1:j_1m_1} T_{2:j_2m_2} = \sum_{k=|j_1-j_2|}^{j_1+j_2} T_{3:k,q} (-1)^{j_1-j_2+q} \sqrt{2k+1} \begin{pmatrix} j_1 & j_2 & k \\ m_1 & m_2 & -q \end{pmatrix} \delta_{q,(m_1+m_2)}. \tag{K.82}$$

The $T_{1:j_1m_1} T_{2:j_2m_2}$ are the $(2j_1+1)(2j_2+1)$ tensor components of the tensor operator $[\mathrm{T}_1^{(j_1)} \otimes \mathrm{T}_2^{(j_2)}]$, which is the direct product of the operators $\mathrm{T}_1^{(j_1)}$ and $\mathrm{T}_2^{(j_1)}$. Eq. (K.82) shows that $[\mathrm{T}_1^{(j_1)} \otimes \mathrm{T}_2^{(j_1)}]$ can written as the *direct* sum of irreducible tensor operators,[1]

$$[\mathrm{T}_1^{(j_1)} \otimes \mathrm{T}_2^{(j_1)}] = [\mathrm{T}_1^{(j_1)} \otimes \mathrm{T}_2^{(j_1)}]^{(|j_1-j_2|)} \oplus \cdots \oplus [\mathrm{T}_1^{(j_1)} \otimes \mathrm{T}_2^{(j_1)}]^{(j_1+j_2)}. \tag{K.83}$$

This decomposition is called *reduction* of the $(2j_1+1)(2j_2+1)$ tensor components of $[\mathrm{T}_1^{(j_1)} \otimes \mathrm{T}_2^{(j_1)}]$ into the *direct* sum of the standard components of all irreducible tensor operators $[\mathrm{T}_1^{(j_1)} \otimes \mathrm{T}_2^{(j_1)}]^{(k)}$,

$$(2j_1+1)(2j_2+1) = \sum_{k=|j_1-j_2|}^{j_1+j_2} (2k+1). \tag{K.84}$$

In the matrix representation this implies that the $(2j_1+1)(2j_2+1) \times (2j_1+1)(2j_2+1)$ matrix representing the operator $[\mathrm{T}_1^{(j_1)} \otimes \mathrm{T}_2^{(j_1)}]$ is transformed into a block-diagonal form with $(2k+1) \times (2k+1)$ matrices along the diagonal, each representing one of the irreducible operators $[\mathrm{T}_1^{(j_1)} \otimes \mathrm{T}_2^{(j_1)}]^{(k)}$. The latter are called irreducible because they cannot be reduced into blocks of lower dimension.

K.4.3 Clebsch-Gordan transformation for the rotation matrices

Sandwiching the rotation operator $P_R(\alpha, \beta, \gamma)$ between the inverse Clebsch-Gordan transformation (K.81) and its hermitian conjugate we obtain two expressions for the rotation matrices. On the l.h.s. we have

$$\begin{aligned} \mathscr{D}^{j_1}_{m_1',m_1} \mathscr{D}^{j_2}_{m_2',m_2} &= \langle j_1, m_1'|P_R(\alpha, \beta, \gamma)|j_1, m_1\rangle\langle j_2, m_2'|P_R(\alpha, \beta, \gamma)|j_2, m_2\rangle \\ &= \langle j_1, m_1'; j_2, m_2'|P_R(\alpha, \beta, \gamma)|j_1, m_1; j_2, m_2\rangle. \end{aligned} \tag{K.85}$$

---

[1] The sum of *non-intersecting* maps is called the *direct* sum of these maps.

On the r.h.s we find

$$\sum_{k=|j_1-j_2|}^{j_1+j_2}\sum_{q,q'=-k}^{k}\langle j_1,m_1';j_2,m_2'|k,q'\rangle(k,q'|P_R(\alpha,\beta,\gamma)|k,q)\langle k,q|j_1,m_1;j_2,m_2\rangle =$$

$$= \sum_{k=|j_1-j_2|}^{j_1+j_2}\sum_{q,q'=-k}^{k}\langle j_1,m_1';j_2,m_2'|k,q'\rangle\mathscr{D}^k_{q',q}\langle k,q|j_1,m_1;j_2,m_2\rangle. \tag{K.86}$$

Equating the two sides we obtain the Clebsch-Gordan transformation for the rotation matrices

$$\mathscr{D}^{j_1}_{m_1',m_1}\mathscr{D}^{j_2}_{m_2',m_2} = \sum_{k=|j_1-j_2|}^{j_1+j_2}\sum_{q,q'=-k}^{k}\langle j_1,m_1';j_2,m_2'|k,q'\rangle\mathscr{D}^k_{q',q}\langle k,q|j_1,m_1;j_2,m_2\rangle. \tag{K.87}$$

Problem K.4. Derive the Gaunt integral (L.58).

Solution. We start from Eq. (3.184) and express the spherical harmonics in terms of the rotation matrices

$$Y^{m_1}_{j_1}(\hat{\mathbf{r}})Y^{m_2}_{j_2}(\hat{\mathbf{r}}) = \sqrt{\frac{2j_1+1}{4\pi}}\sqrt{\frac{2j_2+1}{4\pi}}\mathscr{D}^{j_1}_{0,m_1}\mathscr{D}^{j_2}_{0,m_2}.$$

Applying the Clebsch-Gordan transformation (K.87) to the rotation matrices the product can be written in the form

$$Y^{m_1}_{j_1}(\hat{\mathbf{r}})Y^{m_2}_{j_2}(\hat{\mathbf{r}}) = \sum_{k=|j_1-j_2|}^{j_1+j_2}\sum_{q=-k}^{k}\sqrt{\frac{2j_1+1}{4\pi}}\sqrt{\frac{2j_2+1}{4\pi}}\mathscr{D}^k_{0,q}\langle j_1,0;j_2,0|k,0\rangle\langle k,q|j_1,m_1;j_2,m_2\rangle.$$

Recognizing $\mathscr{D}^k_{0,q} = \sqrt{4\pi/(2k+1)}Y^q_k(\hat{\mathbf{r}})$ and expressing the CGCs in terms of $3j$ symbol the product becomes

$$Y^{m_1}_{j_1}(\hat{\mathbf{r}})Y^{m_2}_{j_2}(\hat{\mathbf{r}}) = \sum_{k=|j_1-j_2|}^{j_1+j_2}\sum_{q=-k}^{k}(-1)^q\sqrt{\frac{(2j_1+1)(2j_2+1)(2k+1)}{4\pi}}Y^q_k(\hat{\mathbf{r}})\begin{pmatrix} j_1 & j_2 & k\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} j_1 & j_2 & k\\ m_1 & m_2 & -q\end{pmatrix}.$$

Turning to the integral we can apply on the r.h.s. the orthonormality relation (L.51). Since only the term with $k=j_3$ and $q=-m_3$ is nonzero we arrive at the desired result,

$$\int Y^{m_1}_{j_1}(\hat{\mathbf{r}})Y^{m_2}_{j_2}(\hat{\mathbf{r}})Y^{m_3}_{j_3}(\hat{\mathbf{r}})\,d\hat{\mathbf{r}} = \sqrt{\frac{(2j_1+1)(2j_2+1)(2j_3+1)}{4\pi}}\begin{pmatrix} j_1 & j_2 & j_3\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} j_1 & j_2 & j_3\\ m_1 & m_2 & m_3\end{pmatrix}. \qquad \square$$

K.4.4 Composition of spherical tensor operators from vector operators

Specializing Eq. (K.78) to the case $j_1 = j_2 = j = 1$ we can construct spherical tensor operators of rank 0, 1 and 2,

$$T_{3:k,q} \equiv [\mathrm{T}^{(j)}_1 \otimes \mathrm{T}^{(j)}_2]^k_q = (-1)^q\sqrt{(2k+1)}\sum_{m=-1}^{1}T_{1:1,m}T_{2:1,(q-m)}\begin{pmatrix} 1 & 1 & k\\ m & q-m & -q\end{pmatrix}. \tag{K.88}$$

This expression provides the algebraic underpinning of the vector model for the addition of angular momenta as introduced in Section 3.4.1. Denoting the two rank 1 spherical tensor operators (vector operators) by $\mathbf{U}$ and $\mathbf{V}$ we can express the standard components of the spherical tensor operators $\mathrm{T}^{(0)}$ (rank 0), $\mathrm{T}^{(1)}$ (rank 1) and $\mathrm{T}^{(2)}$ (rank 2) in terms of the standard components of $\mathbf{U}$ and $\mathbf{V}$:

- *standard* component of $\mathrm{T}^{(0)}$ in terms of the *standard* components of $\mathbf{U}$ and $\mathbf{V}$

$$T_{0,0} = \sqrt{\tfrac{1}{3}}\,(U_{+1}V_{-1} - U_0V_0 + U_{-1}V_{+1}). \tag{K.89}$$

- *standard* components of $\mathrm{T}^{(1)}$ in terms of the *standard* components of **U** and **V**

$$T_{1,0} = \sqrt{\tfrac{1}{2}}\,(U_{+1}V_{-1} - U_{-1}V_{+1}) \tag{K.90a}$$

$$T_{1,\pm} = \pm\sqrt{\tfrac{1}{2}}\,(U_{\pm 1}V_0 - U_0V_{\pm 1}). \tag{K.90b}$$

- *standard* components of $\mathrm{T}^{(2)}$ in terms of the *standard* components of **U** and **V**

$$T_{2,0} = \sqrt{\tfrac{1}{6}}\,(U_{+1}V_{-1} + 2U_0V_0 + U_{-1}V_{+1}) \tag{K.91a}$$

$$T_{2,\pm 1} = \sqrt{\tfrac{1}{2}}\,(U_{\pm 1}V_0 + U_0V_{\pm}) \tag{K.91b}$$

$$T_{2,\pm 2} = U_{\pm 1}V_{\pm 1}. \tag{K.91c}$$

Alternatively, using Eq. (3.30) we can express the standard components of **U** and **V** in terms of the cartesian components of **U** and **V**. Using this transformation we obtain an expression for the *standard* components of the spherical tensor operators $\mathrm{T}^{(0)}$, $\mathrm{T}^{(1)}$ and $\mathrm{T}^{(2)}$ in terms of the *cartesian* components of **U** and **V**:

- *standard* component of $\mathrm{T}^{(0)}$ in terms of the *cartesian* components of **U** and **V**

$$T_{0,0} = -\sqrt{\tfrac{1}{3}}\,(U_xV_x + U_yV_y + U_zV_z). \tag{K.92}$$

- *standard* components of $\mathrm{T}^{(1)}$ in terms of the *cartesian* components of **U** and **V**

$$T_{1,0} = i\sqrt{\tfrac{1}{2}}\,(U_xV_y - U_yV_x) \tag{K.93a}$$

$$T_{1,\pm 1} = \tfrac{1}{2}\,[(U_zV_x - U_xV_z) \pm i(U_zV_y - U_yV_z)]. \tag{K.93b}$$

- *standard* components of $\mathrm{T}^{(2)}$ in terms of the *cartesian* components of **U** and **V**

$$T_{2,0} = \sqrt{\tfrac{1}{6}}\,(2U_zV_z - U_xV_x - U_yV_y) \tag{K.94a}$$

$$T_{2,\pm 1} = \mp\tfrac{1}{2}\,[(U_xV_z + U_zV_x) \pm i(U_yV_z + U_zV_y)] \tag{K.94b}$$

$$T_{2,\pm 2} = \tfrac{1}{2}\,[U_xV_x - U_yV_y \pm i(U_xV_y + U_yV_x)]. \tag{K.94c}$$

K.4.5 Composition of cartesian tensor operators from vector operators

Next we turn to the *cartesian* components of the spherical tensors $\mathrm{T}^{(0)}$, $\mathrm{T}^{(1)}$ and $\mathrm{T}^{(2)}$ in terms of the *cartesian* components of **U** and **V**. This offers the possibility to write these tensor components in vector notation

- *cartesian* component of $\mathrm{T}^{(0)}$ in terms of the *cartesian* components of **U** and **V**

$$T_0 = -\sqrt{\tfrac{1}{3}}\,(U_xV_x + U_yV_y + U_zV_z). \tag{K.95}$$

This shows that, in vector notation, $\mathrm{T}^{(0)}$ can be written as a dot product (i.e., remains invariant under rotation),

$$\mathrm{T}^{(0)} = -\sqrt{\tfrac{1}{3}}\,\mathbf{U}\cdot\mathbf{V}. \tag{K.96}$$

- *cartesian* components of $\mathrm{T}^{(1)}$ in terms of the *cartesian* components of $\mathbf{U}$ and $\mathbf{V}$. Inverting Eqs. (K.93), using Eqs. (K.8), the cartesian components become

$$T_x = i\sqrt{\tfrac{1}{2}}\,(U_yV_z - U_zV_y) \tag{K.97a}$$

$$T_y = -i\sqrt{\tfrac{1}{2}}\,(U_xV_z - U_zV_x) \tag{K.97b}$$

$$T_z = i\sqrt{\tfrac{1}{2}}\,(U_xV_y - U_yV_x). \tag{K.97c}$$

Using the index notation the can be compactly written in the form

$$T_i = i\sqrt{\tfrac{1}{2}}\varepsilon_{ijk}U_jU_k \tag{K.98}$$

which shows that, in vector notation, $\mathrm{T}^{(1)}$ represents a cross product and transforms like a vector,

$$\mathrm{T}^{(1)} = i\sqrt{\tfrac{1}{2}}\begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ U_x & U_y & U_z \\ V_x & V_y & V_z \end{vmatrix} = i\sqrt{\tfrac{1}{2}}\,(\mathbf{U}\times\mathbf{V}). \tag{K.99}$$

- *cartesian* components of $T^{(2)}$ in terms of the *cartesian* components of $\mathbf{U}$ and $\mathbf{V}$. In the index notation the reduction of the cartesian components T_{ij} becomes - see Problem K.5

$$T_{ij} = U_iV_j = \tfrac{1}{3}\mathbf{U}\cdot\mathbf{V}\delta_{ij} + \tfrac{1}{2}(U_iV_j - U_jV_i) + \tfrac{1}{2}(U_iV_j + U_jV_i) - \tfrac{1}{3}\mathbf{U}\cdot\mathbf{V}\delta_{ij}. \tag{K.100}$$

This expression shows that T_{ij} (i.e., any 3×3 matrix) can be written as the sum of a diagonal matrix,

$$\tfrac{1}{3}\mathbf{U}\cdot\mathbf{V}\delta_{ij} \tag{K.101}$$

(fixed by the trace of the matrix), an antisymmetric matrix,

$$A_{i,j} = \tfrac{1}{2}(U_iV_j - U_jV_i) \tag{K.102}$$

(fixed by 3 matrix elements: U_xV_y, U_xV_z and U_yV_z), and a symmetric matrix of zero trace,

$$S_{i,j} = \tfrac{1}{2}(U_iV_j + U_jV_i) - \tfrac{1}{3}\mathbf{U}\cdot\mathbf{V}\delta_{ij} \tag{K.103}$$

(fixed by 5 matrix elements, U_xV_y, U_xV_z, U_yV_z, U_xU_y and U_yU_y, with U_zU_z following from the constraint of zero trace).

Problem K.5. Derive Eq. (K.100), which shows that any 3×3 matrix can be written as the sum of a *diagonal matrix*, an *antisymmetric matrix* and a *zero-trace diagonal matrix.*

Solution. We start from the *inverse decomposition (K.82) for* the case $j_1 = j_2 = j = 1$,

$$T_{1,m_1}T_{1,m_2} = \sum_{k=0}^{2} T_{k,q}(-1)^q\sqrt{2k+1}\begin{pmatrix} 1 & 1 & k \\ m_1 & m_2 & -q \end{pmatrix}\delta_{q,(m_1+m_2)}. \tag{K.104}$$

From this expression we obtain (most conveniently with computer algebra) the reduction of the spherical tensor components $T_{1,m_1}T_{1,m_2} = U_{m_1}V_{m_2}$:

$U_{m_1}V_{m_2}$	V_{+1}	V_0	V_{-1}
U_{+1}	$T_{2,+2}$	$\sqrt{\tfrac{1}{2}}\,(T_{2,+1} + T_{1,+1})$	$\sqrt{\tfrac{1}{6}}\,(T_{2,0} + \sqrt{3}\,T_{1,0} + \sqrt{2}\,T_{0,0})$
U_0	$\sqrt{\tfrac{1}{2}}\,(T_{2,+1} - T_{1,+1})$	$\sqrt{\tfrac{1}{3}}\,(T_{2,0} - \sqrt{2}\,T_{0,0})$	$\sqrt{\tfrac{1}{2}}\,(T_{2,-1} + T_{1,-1})$
U_{-1}	$\sqrt{\tfrac{1}{6}}\,(T_{2,0} - \sqrt{3}\,T_{1,0} + \sqrt{2}\,T_{0,0})$	$\sqrt{\tfrac{1}{2}}\,(T_{2,-1} - T_{1,-1})$	$T_{2,-2}$

To obtain the reduction of the *cartesian* tensor components U_iV_j, with $i,j \in \{x,y,z\}$, we the express the U_iV_j in terms of the $U_{m_1}U_{m_2}$ with the aid of Eqs. (K.8). Subsequently, we substitute for the $T_{k,q}$, the decomposition in terms of the *cartesian* components of $\mathbf{U}$ and $\mathbf{V}$. We demonstrate this explicitly for T_{xx}, T_{xy} and T_{xz} :

$$\begin{aligned}
T_{xx} &= U_xV_x = -\sqrt{\tfrac{1}{3}}\,T_{0,0} + \tfrac{1}{2}T_{2,-2} - \sqrt{\tfrac{1}{6}}\,T_{2,0} + \tfrac{1}{2}T_{2,2} = \tfrac{1}{3}\mathbf{U}\cdot\mathbf{V} + U_xU_x - \tfrac{1}{3}\mathbf{U}\cdot\mathbf{V}\\
T_{xy} &= U_xV_y = -i(\sqrt{\tfrac{1}{2}}\,T_{1,0} - \tfrac{1}{2}T_{2,-2} + \tfrac{1}{2}T_{2,2}) = \tfrac{1}{2}(U_xV_y - U_yV_x) + \tfrac{1}{2}(U_xV_y + U_yV_x)\\
T_{xz} &= U_xV_z = -\tfrac{1}{2}(T_{1,-1} + T_{1,1} - T_{2,-1} + T_{2,1}) = \tfrac{1}{2}(U_xV_z - U_zV_x) + \tfrac{1}{2}(U_xV_z + UV_x).
\end{aligned}$$

These results all satisfy the desired expression. □

L

Properties of functions, series and integrals

L.1 Finite sums of powers

The best known example is the *arithmetic series*:

$$\sum_{k=1}^{n} a_k = \tfrac{1}{2}n(a_n + a_1), \text{ with } a_{n+1} = a_n + v, \tag{L.1}$$

where v is called the difference between successive terms. In particular, for $a_k = k$ this becomes

$$\sum_{k=1}^{n} k = \tfrac{1}{2}n(n+1). \tag{L.2}$$

Related finite sums are

$$\sum_{k=1}^{n} k^2 = \tfrac{1}{6}n(n+1)(2n+1) \tag{L.3}$$

$$\sum_{k=1}^{n} k^3 = \tfrac{1}{4}n^2(n+1)^2. \tag{L.4}$$

Sums of powers of (generally complex) numbers:

$$\sum_{k=1}^{n} k^n z^k = \left(z\frac{d}{dz}\right)^n \frac{z - z^{n+1}}{1-z}. \tag{L.5}$$

The best known example is the *geometric series*

$$\sum_{k=1}^{n} a_k = a_1 \frac{1-r^n}{1-r}, \text{ with } a_{n+1} = r a_n. \tag{L.6}$$

The factor r is called the ratio of the series; i.e., the ratio of two subsequent terms. Other examples of summations of this type are

$$\sum_{k=1}^{n} k z^k = z\frac{1-(n+1)z^n + nz^{n+1}}{(1-z)^2} \tag{L.7}$$

$$\sum_{k=1}^{n} k^2 z^k = z\frac{1+z-(n+1)^2 z^n + (2n^2+2n-1)z^{n+1} - n^2 z^{n+2}}{(1-z)^3}. \tag{L.8}$$

L.2 Gamma function

The gamma function is defined for the complex z plane excluding the non-positive integers

$$\int_0^\infty e^{-x} x^{z-1} dx = \Gamma(z). \tag{L.9}$$

For integer values $z - 1 = n = 0, 1, 2, \cdots$ the gamma function coincides with the factorial function,

$$\int_0^\infty e^{-x} x^n dx = \Gamma(n+1) = n! \tag{L.10}$$

Some special values are:

$$\begin{array}{lll} \Gamma(-1/2) = -2\sqrt{\pi} = -3.545, & \Gamma(1/2) = \sqrt{\pi} = 1.772, & \Gamma(1) = 1 \\ \Gamma(-3/2) = \frac{4}{3}\sqrt{\pi} = 2.363, & \Gamma(3/2) = \frac{1}{2}\sqrt{\pi} = 0.886, & \Gamma(2) = 1 \\ & \Gamma(5/2) = \frac{3}{4}\sqrt{\pi} = 1.329, & \Gamma(3) = 2 \\ & \Gamma(7/2) = \frac{15}{8}\sqrt{\pi} = 3.323, & \Gamma(4) = 6. \end{array} \tag{L.11}$$

Some related integrals are

$$\int_0^\infty e^{-x^2} x^{2n+1} dx = \frac{1}{2} n! \tag{L.12}$$

$$\int_0^\infty e^{-x^2} x^{2n} dx = \frac{(2n-1)!!}{2^{n+1}} \sqrt{\pi}. \tag{L.13}$$

A useful integral relation is

$$\int_0^\infty e^{-\ell x^n} x^m dx = \frac{1}{\ell^{(m+1)/n}} \int_0^\infty e^{-x^n} x^m dx. \tag{L.14}$$

L.3 Polylogarithm

The polylogarithm $\mathrm{Li}_\alpha(z)$ is a special function defined over the unit disk in the complex plane by the series expansion

$$\mathrm{Li}_\alpha(z) = \mathrm{PolyLog}[\alpha, z] \equiv \sum_{\ell=1}^\infty \frac{z^\ell}{\ell^\alpha} \qquad (|z| < 1), \tag{L.15}$$

where α is an arbitrary complex number. By analytic continuation the polylogarithm can be defined over a larger range of z. For z and α on the *real axis* and for $\alpha > 1$ the polylogarithm are given by the Bose-Einstein integrals

$$F_\alpha^{\mathrm{BE}}(z) = \frac{1}{\Gamma(\alpha)} \int_0^\infty \frac{x^{\alpha-1}}{z^{-1}e^x - 1} dx \qquad (z < 1) \tag{L.16}$$

and the Fermi-Dirac integrals

$$F_\alpha^{\mathrm{FD}}(z) = \frac{1}{\Gamma(\alpha)} \int_0^\infty \frac{x^{\alpha-1}}{z^{-1}e^x + 1} dx \qquad (z \geq -1). \tag{L.17}$$

Recurrence relations:

$$F_\alpha^{\mathrm{BE/FD}}(z) = z\frac{d}{dz} F_{\alpha+1}^{\mathrm{BE/FD}}(z) \quad \Leftrightarrow \quad F_\alpha^{\mathrm{BE/FD}}(e^u) = \frac{d}{du} F_{\alpha+1}^{\mathrm{BE/FD}}(e^u). \tag{L.18}$$

L.4 Bose-Einstein function

The Bose-Einstein (BE) integrals are defined for real z and $\alpha > 1$ as

$$F_{\alpha}^{\mathrm{BE}}(z) = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} \frac{x^{\alpha-1}}{z^{-1}e^x - 1} dx \qquad (z < 1). \tag{L.19}$$

The integrals can be expanded in powers of z on the interval $0 < z < 1$,

$$F_{\alpha}^{\mathrm{BE}}(z) = \frac{1}{\Gamma(\alpha)} \sum_{\ell=1}^{\infty} \int_0^{\infty} x^{\alpha-1} z^{\ell} e^{-\ell x} dx = \sum_{\ell=1}^{\infty} \frac{z^{\ell}}{\ell^{\alpha}} = g_{\alpha}(z) = \mathrm{Li}_{\alpha}(z), \tag{L.20}$$

where $\mathrm{Li}_{\alpha}(z)$ is the polylogarithm. For *non-integer* values of α the BE-integrals can also be expanded in the form[1]

$$F_{\alpha}^{\mathrm{BE}}(e^{-u}) = \Gamma(1-\alpha)\, u^{\alpha-1} + \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \zeta(\alpha - n) u^n, \tag{L.21}$$

where the expansion in powers of $u = -\ln z$ is valid on the interval $0 < u < 2\pi$. For *integer* values $\alpha = m \in \{2, 3, 4, \cdots\}$ the BE-integrals the expansion is

$$F_{m}^{\mathrm{BE}}(e^{-u}) = \frac{(-u)^{m-1}}{(m-1)!} \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m-1} - \ln u \right) u^{m-1} + \sum_{\substack{n=0 \\ \neq m-1}}^{\infty} \frac{\zeta(m-n)}{n!} u^n, \tag{L.22}$$

with convergence for $0 < u \leq 2\pi$.

L.5 Fermi-Dirac function

The Fermi-Dirac (FD) integrals are defined for real z and $\alpha > 1$ as

$$F_{\alpha}^{\mathrm{FD}}(z) = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} \frac{x^{\alpha-1}}{z^{-1}e^x + 1} dx \qquad (z \geq -1). \tag{L.23}$$

The integrals can be expanded in powers of z on the interval $0 < z \leq 1$,

$$F_{\alpha}^{\mathrm{FD}}(z) = \frac{-1}{\Gamma(\alpha)} \sum_{\ell=1}^{\infty} \int_0^{\infty} x^{\alpha-1} (-z)^{\ell} e^{-\ell x} dx = -\sum_{\ell=1}^{\infty} \frac{(-z)^{\ell}}{\ell^{\alpha}} = f_{\alpha}(z) = -\mathrm{Li}_{\alpha}(-z), \tag{L.24}$$

where $\mathrm{Li}_{\alpha}(z)$ is the polylogarithm.

L.6 Riemann zeta function

The Riemann zeta function is defined as a Dirichlet series

$$\lim_{z \to 1} g_{\alpha}(z) = \zeta(\alpha) = \sum_{\ell=1}^{\infty} \frac{1}{\ell^{\alpha}}. \tag{L.25}$$

Some special values are:

$$\zeta(1/2) = -1.460, \zeta(3/2) = 2.612, \qquad \zeta(5/2) = 1.341, \zeta(7/2) = 1.127,$$

$$\zeta(1) \to \infty, \qquad \zeta(2) = \pi^2/6 = 1.645, \zeta(3) = 1.202, \quad \zeta(4) = \pi^4/90 = 1.082.$$

---

[1] For a derivation see J.E. Robinson, Phys. Rev. 83, 678 (1951).

L.7 Selected integrals and rules

For $\gamma > 0$ and $\varepsilon > 0$

$$\int_0^{\varepsilon} \sqrt{x}\,(\varepsilon - x)^{\gamma-1}\,dx = \frac{\sqrt{\pi}\Gamma(\gamma)}{2\,\Gamma\,(3/2+\gamma)}\varepsilon^{1/2+\gamma} \tag{L.26}$$

The *Leibniz integral rule* for moving boundaries is given by

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(x,t)dx = \int_{a(t)}^{b(t)} \frac{\partial}{\partial t} f(x,t)dx + f(x,t)\frac{d}{dt}b(t) - f(x,t)\frac{d}{dt}a(t) \tag{L.27}$$

In particular, for definite integrals with fixed boundaries this integration rule becomes

$$\frac{d}{dt}\int f(x,t)dx = \int \frac{\partial}{\partial t} f(x,t)dx. \tag{L.28}$$

L.8 Commutator algebra

If A, B, C and D are four arbitrary linear operators the following relations hold:

$$\begin{aligned}
[A,B] &= -\,[B,A] && \text{(L.29a)}\\
[A,B+C] &= [A,B] + [A,C] && \text{(L.29b)}\\
[A,BC] &= [A,B]\,C + B\,[A,C] && \text{(L.29c)}\\
[AB,CD] &= A[B,C]D + AC[B,D] + [A,C]DB + C[A,D]B && \text{(L.29d)}\\
0 &= [A,[B,C]] + [B,[C,A]] + [C,[A,B]]\,. && \text{(L.29e)}
\end{aligned}$$

Commutators containing B^n:

$$\begin{aligned}
[A,B^n] &= \sum_{s=0}^{n-1} B^s\,[A,B]\,B^{n-s-1} && \text{(L.30a)}\\
[A,B^n] &= nB^{n-1}\,[A,B] \;\text{ if } B \text{ commutes with } [A,B]. && \text{(L.30b)}
\end{aligned}$$

The exponential operator is defined as:

$$e^A \equiv \sum_{n=0}^{\infty} \frac{A^n}{n!}. \tag{L.31}$$

Baker-Campbell-Hausdorff formula:

$$e^A e^B = e^C, \tag{L.32}$$

where

$$C = A + B + \tfrac{1}{2}[A,B] + \tfrac{1}{12}([A,[A,B]] + [[A,B],B]) + \cdots .$$

Special cases:

$$\begin{aligned}
e^A e^B &= e^{A+B+\frac{1}{2}[A,B]} \text{ if } A \text{ and } B \text{ commute with } [A,B] && \text{(L.33a)}\\
e^A B e^{-A} &= B + [A,B] + \tfrac{1}{2!}[A,[A,B]] + \tfrac{1}{3!}[A,[A,[A,B]]] + \cdots && \text{(L.33b)}\\
e^A B e^{-A} &= B + [A,B] \text{ if } A \text{ commutes with } [A,B] && \text{(L.33c)}\\
e^A B e^{-A} &= e^{\gamma} B \text{ if } [A,B] = \gamma B, \text{ with } \gamma \text{ a constant.} && \text{(L.33d)}
\end{aligned}$$

Trotter formula:

$$e^{A+B} = \lim_{N\to\infty} (e^{A/N} e^{B/N})^N. \tag{L.34}$$

Zassenhaus formula:

$$e^{A+B} = e^A e^B \prod_{n=2}^{\infty} e^{C_n}, \tag{L.35}$$

where

$$C_2 = -\tfrac{1}{2}[A,B]; \quad C_3 = \tfrac{1}{6}(2[A,[A,B]] - [[A,B],B]).$$

L.9 Legendre polynomials

The associated Legendre differential equation is given by,

$$\left[\left(1-u^2\right)\frac{d^2}{du^2} - 2u\frac{d}{du} - \frac{m^2}{1-u^2} + l(l+1)\right] P_l^m(u) = 0 \tag{L.36}$$

For $m = 0$ this equation is called the Legendre differential equation and its solutions are the Legendre polynomials, defined by the *Rodrigues formula*:

$$P_l(u) = \frac{1}{2^l l!}\frac{d^l}{du^l}(u^2-1)^l. \tag{L.37}$$

$P_l(u)$ is a real polynomial of *degree* $l \in \{0, 1, 2\cdots\}$, *parity*

$$P_l(-u) = (-1)^l P_l(u) \tag{L.38}$$

and having l zeros in the interval $-1 \leq u \leq 1$. The Legendre polynomials of lowest degree are

$$P_0(u) = 1, \quad P_1(u) = u, \quad P_2(u) = \tfrac{1}{2}(3u^2 - 1) \tag{L.39}$$

$$P_3(u) = \tfrac{1}{2}(5u^3 - 3u), \quad P_4(u) = \tfrac{1}{8}(35u^4 - 30u^2 + 3). \tag{L.40}$$

The Legendre polynomials are generated by expansion of the function

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \frac{1}{\sqrt{r^2 - 2urr' + r'^2}} = \frac{1}{r_>}\sum_l \left(\frac{r_<}{r_>}\right)^l P_l(u), \tag{L.41}$$

where $r_> = \max\{r, r'\}$ and $r_< = \min\{r, r'\}$ with $r = |\mathbf{r}|$ and $r' = |\mathbf{r}'|$; further $u = \hat{\mathbf{r}} \cdot \hat{\mathbf{r}}' = \cos\theta$, with θ the angle included by the unit vectors $\hat{\mathbf{r}}$ and $\hat{\mathbf{r}}'$. The expansion (L.41) is called the *multipole expansion.*

For $m \neq 0$ the solutions of Eq. (L.36) are called the associated Legendre functions $P_l^m(u)$. For $m = 0$ they coincide with the Legendre polynomials. For $m > 0$ the $P_l^m(u)$ are obtained by differentiation of the $P_l(u)$,[1]

$$P_l^m(u) = (-1)^m(1-u^2)^{m/2}\frac{d^m}{du^m}P_l(u). \tag{L.42}$$

Note that $P_l^0(u) \equiv P_l(u)$. The differentiation fixes the relative sign of the polynomials of different power. These functions consist of the product of a function $(-1)^m(1-u^2)^{m/2}$ and a polynomial of degree $(l-m)$, parity $(-1)^{l-m}$ with $(l-m)$ zeros in the interval $-1 \leq u \leq 1$. For crossing from positive to negative m the sign and normalization are fixed by convention,[2]

$$P_l^{-m}(u) = (-1)^m\frac{(l-m)!}{(l+m)!}P_l^m(u), \tag{L.43}$$

---

[1] Here the phase factor $(-1)^m$ is introduced to simplify the definition of the spherical harmonics. This convention is implemented in *Mathematica* but deviates from the convention used in most texts on quantum mechanics. Beware that differences in phase convention affect the recursion relations (L.48) and (L.49).

[2] Note that the phase factor oscillates like $(-1)^m$ for $m > 0\cdots$ but is always 1 for $m \leq 0$. This is the signature of the Condon and Shortley phase convention [28].

where $0 \leq m \leq l$. The $P_l^m(u)$ are nonzero only for $-l \leq m \leq l$; i.e., the index m can assume $2l+1$ possible values for a given value of l. Particular cases are

$$P_l^0(u) = P_l(u), \quad P_l^l(u) = (-1)^{-l}(2l-1)!!(1-u^2)^{l/2}. \tag{L.44}$$

The *orthogonality* of the P_l^m is expressed by

$$\int_{-1}^{1} P_l^m(u) P_{l'}^m(u) du = 0 \qquad (l \neq l') \tag{L.45a}$$

$$\int_{-1}^{1} P_l^m(u) P_l^{m'}(u) du = 0 \qquad (m \neq m'). \tag{L.45b}$$

The *normalization* of the P_l^m is expressed by

$$\int_{-1}^{1} \left[P_l^m(u)\right]^2 du = 2\int_{0}^{1} \left[P_l^m(u)\right]^2 du = \frac{2}{2l+1}\frac{(l+m)!}{(l-m)!}. \tag{L.46}$$

The following recursion *relations* hold for $-l \leq m \leq l$:

$$(2l+1)uP_l^m(u) = (l-m+1)P_{l+1}^m(u) + (l+m)P_{l-1}^m(u) \tag{L.47}$$

$$\sqrt{1-u^2}P_l^{m+1}(u) = -(l-m+1)(l+m)\sqrt{1-u^2}P_l^{m-1}(u) - 2muP_l^m(u) \tag{L.48}$$

and

$$(1-u^2)\frac{d}{du}P_l^m(u) = -luP_l^m(u) + (l+m)P_{l-1}^m(u) \tag{L.49a}$$

$$= (l+1)uP_l^m(u) - (l-m+1)P_{l+1}^m(u) \tag{L.49b}$$

$$= -muP_l^m(u) - \sqrt{1-u^2}P_l^{m+1}(u) \tag{L.49c}$$

$$= +muP_l^m(u) + (l-m+1)(l+m)\sqrt{1-u^2}P_l^{m-1}(u). \tag{L.49d}$$

L.9.1 Spherical harmonics $Y_l^m(\theta,\varphi)$

The *spherical harmonics* of *degree* l and *order* m are defined as the joint, *normalized* eigenfunctions of $\mathbf{L}^2$ and L_z. Their relation to the associated Legendre polynomials is given by

$$Y_l^m(\theta,\varphi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)e^{im\varphi}, \tag{L.50}$$

with $-l \leq m \leq l$. The *orthonormality relation* is

$$\int Y_l^{m*}(\hat{\mathbf{r}})Y_{l'}^{m'}(\hat{\mathbf{r}})d\hat{\mathbf{r}} = \delta_{l'l}\delta_{m'm}. \tag{L.51}$$

Using Eq. (L.43) we find that the *complex conjugation* is given by

$$Y_l^{m*}(\hat{\mathbf{r}}) = (-1)^m Y_l^{-m}(\hat{\mathbf{r}}). \tag{L.52}$$

The *parity* under space inversion, $\hat{\mathbf{r}} = (\theta,\varphi) \to -\hat{\mathbf{r}} = (\pi-\theta, \varphi+\pi)$, is given by

$$Y_l^m(-\hat{\mathbf{r}}) = (-1)^l Y_l^m(\hat{\mathbf{r}}). \tag{L.53}$$

Alternatively, the spherical harmonics can be defined using only associated Legendre functions of positive m,[1]

$$Y_l^m(\theta,\varphi) = \varepsilon\sqrt{\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}}P_l^{|m|}(\cos\theta)e^{im\varphi} \quad \text{with} \quad \varepsilon = \begin{cases} (-1)^m & \text{for } m \geq 0 \\ 1 & \text{for } m < 0. \end{cases} \tag{L.54}$$

Note that the spherical harmonics with even l have even parity and those with odd l have odd parity. The spherical harmonics of lowest degree are:

$$\left.\begin{array}{rll}
 & \sqrt{\tfrac{4\pi}{2l+1}}Y_l^m(\theta,\varphi) - \textit{spherical} & \sqrt{\tfrac{4\pi}{2l+1}}Y_l^m(x,y,z) - \textit{cartesian} \\
Y_0^0: & 1 & \Leftrightarrow \ 1 \\
Y_1^0: & \cos\theta & \Leftrightarrow \ z/r \\
Y_1^{\pm1}: & \mp\sqrt{\tfrac{1}{2}}\sin\theta\ e^{\pm i\varphi} & \Leftrightarrow \mp\sqrt{\tfrac{1}{2}}\left(x \pm iy\right)/r \\
Y_2^0: & \tfrac{1}{2}\left(3\cos^2\theta - 1\right) & \Leftrightarrow \ \tfrac{1}{2}\left(3z^2 - r^2\right)/r \\
Y_2^{\pm1}: & \mp\sqrt{\tfrac{3}{2}}\sin\theta\cos\theta\ e^{\pm i\varphi} & \Leftrightarrow \mp\sqrt{\tfrac{3}{2}}\left(x \pm iy\right)z/r^2 \\
Y_2^{\pm2}: & \sqrt{\tfrac{3}{2}\tfrac{1}{2}}\sin^2\theta\ e^{\pm 2i\varphi} & \Leftrightarrow \ \sqrt{\tfrac{3}{2}\tfrac{1}{2}}\left(x \pm iy\right)^2/r^2 \\
Y_3^0: & \tfrac{1}{2}\left(2\cos^3\theta - 3\cos\theta\sin^2\theta\right) & \Leftrightarrow \ \tfrac{1}{2}\left(5z^2 - 3r^2\right)z/r^3 \\
Y_3^{\pm1}: & \mp\sqrt{\tfrac{3}{16}}\left(4\cos^2\theta\sin\theta - \sin^3\theta\right)e^{\pm i\varphi} & \Leftrightarrow \mp\sqrt{\tfrac{3}{16}}\left(x \pm iy\right)\left(5z^2 - r^2\right)/r^3 \\
Y_3^{\pm2}: & \sqrt{\tfrac{15}{2}\tfrac{1}{2}}\cos\theta\sin^2\theta\ e^{\pm 2i\varphi} & \Leftrightarrow \ \sqrt{\tfrac{15}{2}\tfrac{1}{2}}\left(x \pm iy\right)^2 z/r^3 \\
Y_3^{\pm3}: & \mp\sqrt{\tfrac{5}{2}}\sqrt{\tfrac{1}{8}}\sin^3\theta\ e^{\pm 3i\varphi} & \Leftrightarrow \mp\sqrt{\tfrac{5}{2}}\sqrt{\tfrac{1}{8}}\left(x \pm iy\right)^3/r^3.
\end{array}\right\} \tag{L.55}$$

Note that

$$Y_l^m(\theta,\varphi)|_{\theta=0} = \sqrt{\frac{2l+1}{4\pi}}\delta_{m,0}.$$

Furthermore, the phase factor oscillates like $(-1)^m$ for $m = 1, 3, 5, \cdots 0$ but is always 1 for $m < 0$; this is the signature of the Condon and Shortley phase convention.

The *addition theorem* relates the angle θ_{12} between two directions $\hat{\mathbf{r}}_1 = (\theta_1, \varphi_1)$ and $\hat{\mathbf{r}}_2 = (\theta_2, \varphi_2)$ relative to a coordinate system of choice,

$$\frac{2l+1}{4\pi}P_l(\cos\theta_{12}) = \sum_{m=-l}^{l} Y_l^{m*}(\hat{\mathbf{r}}_1)Y_l^m(\hat{\mathbf{r}}_2). \tag{L.56}$$

The product of two spherical harmonics can be expressed in terms of Wigner $3j$ symbols

$$\begin{aligned} Y_l^m(\hat{\mathbf{r}})Y_{l'}^{m'}(\hat{\mathbf{r}}) = \sum_{L=|l-l'|}^{l+l'}\sum_{M=-L}^{L}(-1)^M\sqrt{\frac{(2l+1)(2l'+1)(2L+1)}{4\pi}} \\ \times\begin{pmatrix} l\ l'\ L \\ 0\ 0\ 0 \end{pmatrix}\begin{pmatrix} l & l' & L \\ m & m' & M \end{pmatrix}Y_L^{-M}(\hat{\mathbf{r}}), \end{aligned} \tag{L.57}$$

---

[1] Note that the phase factor $(-1)^m$ is only included for positive m. This is the signature of the Condon and Shortley phase convention [28].

An important relation is the *integral over three spherical harmonics* [49]

$$\int Y_{l_1}^{m_1}(\hat{\mathbf{r}})Y_{l_2}^{m_2}(\hat{\mathbf{r}})Y_{l_3}^{m_3}(\hat{\mathbf{r}})d\hat{\mathbf{r}} = \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}}\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}. \tag{L.58}$$

In Dirac notation this becomes

$$\langle l'm'|\, Y_k^q(\hat{\mathbf{r}})\,|lm\rangle = (-1)^{m'}\sqrt{\frac{(2l'+1)(2k+1)(2l+1)}{4\pi}}\begin{pmatrix} l' & k & l \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} l' & k & l \\ -m' & q & m \end{pmatrix}. \tag{L.59}$$

Some special cases are:

- $k=0$: The $3j$ symbols are zero unless $l'=l$ and $m'=m$; hence, we find
$$\langle l'm'|\sqrt{4\pi}\, Y_0^0(\hat{\mathbf{r}})\,|lm\rangle = \delta_{l',l}\delta_{m',m} \tag{L.60}$$
- $k=1$: The first $3j$ symbol is zero unless $l'=l\pm1$; hence, we find
$$\langle l'm'|\sqrt{4\pi/3}\, Y_1^q(\hat{\mathbf{r}})\,|lm\rangle = (-1)^{m'+\max(l,l')}\sqrt{\max(l,l')}\begin{pmatrix} l' & 1 & l \\ -m' & q & m \end{pmatrix}\delta_{l',l\pm1} \tag{L.61}$$
- $k=2$: The first $3j$ symbol is zero unless $l'=l$, $l\pm2$; hence, we find
$$\langle l'm'|\sqrt{4\pi/5}\, Y_2^q(\hat{\mathbf{r}})\,|lm\rangle = \begin{cases} (-1)^{m'+l+1}\sqrt{\dfrac{l\,(l+1)\,(2l+1)}{(2l+3)(2l-1)}}\begin{pmatrix} l & 2 & l \\ -m' & q & m \end{pmatrix} & l'=l \\ (-1)^{m'+l}\sqrt{\dfrac{3}{4}\dfrac{l+l'}{l+l'+1}}\sqrt{\max(l,l')}\begin{pmatrix} l' & 2 & l \\ -m' & q & m \end{pmatrix} & l'=l\pm2. \end{cases} \tag{L.62}$$

L.10 Hermite polynomials

The Hermite differential equation is given by

$$y''-2xy'+2ny=0. \tag{L.63}$$

For $n=0,1,2,\ldots$ the solutions satisfy the Rodrigues formula

$$H_n(x) = (-1)^n e^{x^2}\frac{d^n}{dx^n}(e^{-x^2}). \tag{L.64}$$

These solutions are polynomials of *degree n* known as *Hermite polynomials.* Examples:

$$\begin{array}{ll} H_0(x)=1 & H_4(x)=16x^4-48x^2+12 \\ H_1(x)=2x & H_5(x)=32x^5-160x^3+120x \\ H_2(x)=4x^2-2 & H_6(x)=64x^6-480x^4+720x^2-120 \\ H_3(x)=8x^3-12x & H_7(x)=128x^7-1344x^5+3360x^3-1680x\,. \end{array} \tag{L.65}$$

The generating function is

$$e^{2tx-t^2} = \sum_{n=0}^{\infty} H_n(x)\frac{t^n}{n!}. \tag{L.66}$$

Useful recursion relations are

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x) \tag{L.67}$$

$$H_n'(x) = 2nH_{n-1}(x) \tag{L.68}$$

and the orthogonality relation is given by

$$\int_{-\infty}^{\infty} e^{-x^2}H_m(x)\,H_n(x) = 2^n n!\sqrt{\pi}\delta_{mn}\,. \tag{L.69}$$

L.11 Laguerre polynomials

Generalized Laguerre polynomials satisfy the following differential equation

$$xy'' + (\alpha + 1 - x)y' + 2ny = 0. \tag{L.70}$$

For $n = 0, 1, 2, \ldots$ the solutions satisfy the Rodrigues formula[1]

$$L_n^\alpha(x) = \frac{1}{n!} e^x x^{-\alpha} \frac{d^n}{dx^n}(e^{-x} x^{n+\alpha}) \tag{L.71}$$

$$= \sum_{m=0}^{n} (-1)^m \binom{n+\alpha}{n-m} \frac{x^m}{m!}$$

$$= \sum_{m=0}^{n} \frac{\Gamma(\alpha+n+1)}{\Gamma(\alpha+m+1)} \frac{(-1)^m}{(n-m)!} \frac{x^m}{m!} \tag{L.72}$$

These polynomials are well-defined also for real α because the ratio of two gamma functions differing by an integer is well-defined, $(\beta)_n = \beta(\beta+1)(\beta+2)\cdots(\beta+n-1) = \Gamma(\beta+n)/\Gamma(\beta)$. The *Laguerre polynomials* of lowest *degree* are given by

$$L_0^\alpha(x) = 1, \quad L_1^\alpha(x) = \alpha + 1 - x, \quad L_2^\alpha(x) = \tfrac{1}{2}(\alpha+1)(\alpha+2) - (\alpha+2)x + \tfrac{1}{2}x^2. \tag{L.73}$$

Some special cases for $\alpha = 0$ and $\alpha = -n$ are

$$L_0(x) = 1, \quad L_1(x) = 1 - x, \quad L_2(x) = 1 - 2x + \tfrac{1}{2}x^2, \quad L_n^{-n}(x) = (-1)^n \frac{x^n}{n!}. \tag{L.74}$$

The generating function is

$$\frac{(-1)^m\, t^m}{(1-t)^{m+1}} e^{-x/(1-t)} = \sum_{n=m}^{\infty} L_n^m\,(x)\, \frac{t^n}{n!}. \tag{L.75}$$

The generalized Laguerre polynomials satisfy the orthogonality relation

$$\int_0^\infty x^\alpha e^{-x} L_n^\alpha(x) L_m^\alpha(x) dx = 0 \text{ for } m \neq n \text{ (orthogonality relation)} \tag{L.76}$$

$$\int_0^\infty x^\alpha e^{-x} L_n^\alpha(x) dx = \Gamma(\alpha+1)\delta_{0,n}\,. \tag{L.77}$$

Useful recursion relations are:

$$xL_n^\alpha(x) = (2n+\alpha+1)L_n^\alpha(x) - (n+\alpha)L_{n-1}^\alpha(x) - (n+1)L_{n+1}^\alpha \tag{L.78}$$

$$\frac{d}{dx}L_n^\alpha(x) = -L_{n-1}^{\alpha+1}(x) = -[1 + L_1^\alpha(x) + \cdots + L_{n-1}^\alpha(x)]. \tag{L.79}$$

Series expansions:

$$L_n^{\alpha+1}(x) = \sum_{m=0}^{n} L_m^\alpha(x) \tag{L.80a}$$

$$\frac{d}{dx}L_n^\alpha(x) = -\sum_{m=0}^{n-1} L_m^\alpha(x) \tag{L.80b}$$

$$\frac{d^2}{dx^2}L_n^\alpha(x) = \sum_{m=0}^{n-2} (n-m-1)\, L_m^\alpha(x). \tag{L.80c}$$

---

[1] Different definitions can be found in the literature. Here we adhere to the definition of the generalized Laguerre polynomials as used in the *Handbook of Mathematical functions* by Abramowitz and Stegun (Eds.), Dover Publications, New York 1965. This definition is also used by *Mathematica*.

Further, it is practical to introduce a generalized normalization integral

$$J_\nu(m,\alpha) = \int_0^\infty x^{\alpha+\nu} e^{-x} [L_m^\alpha(x)]^2 dx. \tag{L.81}$$

Some special cases are given by

$$J_0(m,\alpha) = \int_0^\infty x^{\alpha} e^{-x} [L_m^\alpha(x)]^2 dx = \frac{\Gamma(\alpha+m+1)}{m!} \tag{L.82}$$

$$J_1(m,\alpha) = \int_0^\infty x^{\alpha+1} e^{-x} [L_m^\alpha(x)]^2 dx = \frac{\Gamma(\alpha+m+1)}{m!}(2m+\alpha+1) \tag{L.83}$$

$$J_2(m,\alpha) = \int_0^\infty x^{\alpha+2} e^{-x} [L_m^\alpha(x)]^2 dx = \frac{\Gamma(\alpha+m+1)}{m!}[6m(m+\alpha+1)+\alpha^2+3\alpha+2] \tag{L.84}$$

$$J_{-1}(m,\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} [L_m^\alpha(x)]^2 dx = \frac{1}{\alpha}\int_0^\infty x^{\alpha} e^{-x} [L_m^\alpha(x)]^2 dx = \frac{\Gamma(\alpha+m+1)}{m!}\frac{1}{\alpha} \tag{L.85}$$

The integrals $J_\nu(m,\alpha)$ with $\nu > 0$ are obtained from Eq. (L.82) by repetitive use of the recursion relation (L.78) and orthogonality relation (L.76); integrals $J_\nu(m,\alpha)$ with $\nu < 0$ are obtained from Eq. (L.82) by partial integration and use of the recursion relation (L.79), the orthogonality relation (L.76) and the special integral (L.77).
Selected ratios for $m = n - l - 1$ and $\alpha = 2l + 1$:

$$J_4/J_1 = \frac{1}{n}[35n^2(n^2-1) - 30n^2(l+2)(l-1) + 3(l+2)(l+1)l(l-1)] \tag{L.86}$$

$$J_3/J_1 = 2\left[5n^2 + 1 - 3l(l+1)\right] \tag{L.87}$$

$$J_2/J_1 = \frac{1}{n}\left[3n^2 - l(l+1)\right] \tag{L.88}$$

$$J_1/J_1 = 1 \tag{L.89}$$

$$J_0/J_1 = \frac{1}{2n} \tag{L.90}$$

$$J_{-1}/J_1 = \frac{1}{2n}\frac{1}{2l+1} \tag{L.91}$$

$$J_{-2}/J_1 = \frac{1}{8}\frac{1}{(l+1)(l+1/2)l} \tag{L.92}$$

$$J_{-3}/J_1 = \frac{1}{32n}\frac{3n^2 - l(l+1)}{(l+3/2)(l+1)(l+1/2)l(l-1/2)}. \tag{L.93}$$

L.12 Bessel functions

L.12.1 Spherical Bessel functions and Hankel functions

The spherical Bessel differential equation is given by

$$x^2 y'' + 2xy' + \left[x^2 - l(l+1)\right] y = 0. \tag{L.94}$$

The general solution is a linear combination of two particular solutions, solutions $j_l(x)$, regular (as x^l) at the origin and known as spherical Bessel functions of the first kind, and solutions $n_l(x)$, irregular at the origin and known as spherical Bessel function of the second kind (also called Neumann functions). The spherical Bessel functions and the Neumann functions are real. The general solution can also be written as a linear combination of two complex functions called Hankel functions of the first (+) and second (−) type,

$$h_l^\pm(x) = n_l(x) \pm i j_l(x) \tag{L.95}$$

or, vice versa

$$j_l\,(x) = \frac{1}{2} i \left[h_l^-(x) - h_l^+(x) \right] \tag{L.96a}$$

$$n_l\,(x) = \frac{1}{2} \left[h_l^-(x) + h_l^+(x) \right]. \tag{L.96b}$$

The spherical Bessel functions and Neumann functions are real and Hankel functions can be expressed in the form

$$j_l\,(x) = R_l \frac{\sin x}{x} + S_l \frac{\cos x}{x} \tag{L.97}$$

$$n_l\,(x) = R_l \frac{\sin x}{x} - S_l \frac{\cos x}{x} \tag{L.98}$$

$$h_l^{\pm}(x) = (R_l \pm i S_l)\, \frac{e^{\pm ix}}{x}, \tag{L.99}$$

where R_l and S_l are polynomials in $1/x$ with real coefficients. In the case of R_l the polynomial is of degree l and has parity $(-1)^l$; the polynomial S_l is of degree $l-1$ and has parity $(-1)^{l-1}$. For real x the polynomial $(R_l \pm i S_l)$ is of the form

$$R_l \pm i S_l = \sum_{s=0}^{l} \frac{(\pm i)^{s-l}}{2^s s!} \frac{(l+s)!}{(l-s)!} \left(\frac{1}{x} \right)^s. \tag{L.100}$$

For real argument x the functions $j_l\,(x)$ and $n_l\,(x)$ are real and the Hankel functions satisfy the relation

$$h_l^{\pm}(x) = h_l^{\mp}(x)^*.$$

The parity relations for the spherical Bessel, spherical Neumann and spherical Hankel functions are

$$j_l(-z) = (-1)^l j_l(z) \tag{L.101}$$

$$n_l(-z) = (-1)^{l-1} n_l(z) \tag{L.102}$$

$$h_l^{\pm}(-z) = (-1)^{l-1} h_l^{\mp}(z). \tag{L.103}$$

An integral representation of the spherical Bessel function is

$$j_l\,(kr) = \tfrac{1}{2}(-1)^l \int_{-1}^{1} e^{ikr\cos\theta} P_l(\cos\theta) d\cos\theta \tag{L.104}$$

Some special cases are given by

- Lowest orders:

$$\begin{aligned} j_0\,(x) &= \frac{\sin x}{x}, & n_0\,(x) &= \frac{\cos x}{x}, & h_0^{\pm}(x) &= \frac{e^{\pm ix}}{x}, \\ j_1\,(x) &= \frac{\sin x}{x^2} - \frac{\cos x}{x}, & n_1\,(x) &= \frac{\cos x}{x^2} + \frac{\sin x}{x}, & h_1^{\pm}(x) &= \left(\frac{1}{x^2} \mp i \frac{1}{x} \right) \frac{e^{\pm ix}}{x}. \end{aligned} \tag{L.105a}$$

- Asymptotic forms for $x \to \infty$

$$j_l\,(x) \underset{x\to\infty}{\sim} \frac{1}{x} \sin(x - \tfrac{1}{2} l\pi) \tag{L.106a}$$

$$n_l\,(x) \underset{x\to\infty}{\sim} \frac{1}{x} \cos(x - \tfrac{1}{2} l\pi) \tag{L.106b}$$

$$h_l^{\pm}(x) \underset{x\to\infty}{\sim} \frac{e^{\pm i(x - \frac{1}{2} l\pi)}}{x} \left[1 \pm i \frac{l(l+1)}{2x} \right]. \tag{L.106c}$$

- Asymptotic forms for $x \to 0$

$$j_l(x) \underset{x\to 0}{\sim} \frac{x^l}{(2l+1)!!}\left[1 - \frac{x^2}{2(2l+3)} + \cdots\right] \tag{L.107a}$$

$$n_l(x) \underset{x\to 0}{\sim} \frac{(2l+1)!!}{(2l+1)}\left(\frac{1}{x}\right)^{l+1}\left[1 + \frac{x^2}{2(2l-1)} + \cdots\right]. \tag{L.107b}$$

- Orthogonality relations

$$\int_0^\infty j_l(k_1 r)\, j_l(k_2 r)\, r^2 dr = \frac{\pi}{2k_1^2}\delta(k_1 - k_2) \tag{L.108}$$

$$\int_{-\infty}^{\infty} j_l(x)\, j_{l'}(x)\, dx = \frac{\pi}{2l+1}\delta_{l'l}. \tag{L.109}$$

L.12.1.1 Relation to Riccati functions

The Riccati functions $\hat{\jmath}_l(x)$, $\hat{n}_l(x)$ and $\hat{h}_l^{\pm}(x)$ are defined as

$$\hat{\jmath}_l(x) = x j_l(x) \tag{L.110a}$$

$$\hat{n}_l(x) = x n_l(x) \tag{L.110b}$$

$$\hat{h}_l^{\pm}(x) = x h_l^{\pm}(x). \tag{L.110c}$$

L.12.1.2 Relation to Bessel functions

The spherical Bessel functions are related to half-integer Bessel functions

$$j_l(x) = \sqrt{\frac{\pi}{2x}} J_{l+\frac{1}{2}}(x) \text{ for } l = 0, 1, 2, \ldots \tag{L.111}$$

$$n_l(x) = (-)^l \sqrt{\frac{\pi}{2x}} J_{-l-\frac{1}{2}}(x) \text{ for } l = 0, 1, 2, \ldots \tag{L.112}$$

L.12.2 Bessel functions

The Bessel differential equation is given by

$$x^2 y'' + x y' + \left(x^2 - n^2\right) y = 0. \tag{L.113}$$

The general solution is a linear combination of two particular solutions

$$y = A J_n(x) + B J_{-n}(x) \text{ for } n \neq 0, 1, 2, \cdots \tag{L.114a}$$

$$y = A J_n(x) + B Y_n(x) \text{ for all integer } n \tag{L.114b}$$

where A and B are arbitrary constants and $J_{\pm n}(x)$ are Bessel functions, which are defined by

$$J_{\pm n}(x) = \sum_{p=0}^{\infty} \frac{(-1)^p (x/2)^{2p\pm n}}{p!\Gamma(1 + p \pm n)}. \tag{L.115}$$

The $Y_n(x)$ are Neumann functions and are defined by

$$Y_n(x) = \frac{J_n(x)\cos n\pi - J_{-n}(x)}{\sin n\pi} \text{ for } n \neq 0, 1, 2, \cdots \tag{L.116}$$

$$Y_n(x) = \lim_{p\to n} \frac{J_n(x)\cos p\pi - J_{-n}(x)}{\sin p\pi} \text{ for } n = 0, 1, 2, \cdots. \tag{L.117}$$

Extracting the leading term from the Bessel expansion (L.115) results in

$$J_{\pm n}(x) = \frac{(x/2)^{\pm n}}{\Gamma(1 \pm n)} \left(1 - \frac{(x/2)^2}{(1 \pm n)} + \cdots \right). \tag{L.118}$$

The generating function is of the form

$$e^{x(z-1/z)/2} = \sum_{n=-\infty}^{\infty} J_n(x) z^n, \tag{L.119}$$

in particular for $z = 1$

$$\sum_{n=-\infty}^{\infty} J_n(x) = 1.$$

Some differential recursion relations are for any $n \in \mathbb{Z}$

$$2J'_n(x) = J_{n-1}(x) - J_{n+1}(x) \tag{L.120a}$$

$$\frac{2n}{x} J_n(x) = J_{n+1}(x) + J_{n-1}(x) \tag{L.120b}$$

$$\frac{d}{dx}\left[x^n J_n(x)\right] = x^n J_{n-1}(x). \tag{L.120c}$$

Some integral recursion relations are for any $n \in \mathbb{Z}$

$$\int x^{n+1} J_n(x) dx = x^{n+1} J_{n+1}(x) \tag{L.121}$$

$$\int x^{-n+1} J_n(x) dx = -x^{-n+1} J_{n+1}(x). \tag{L.122}$$

Completeness relations for any $m \neq 0$

$$J_0^2(x) + 2\sum_{n=1}^{\infty} J_n^2(x) = 1 \tag{L.123}$$

$$\sum_{n=-\infty}^{\infty} J_{n+m}(x) J_n(x) = 0. \tag{L.124}$$

Addition relation

$$J_n(x+y) \sum_{m=-\infty}^{\infty} J_m(x) J_{n-m}(y) \tag{L.125}$$

Special cases:

Bessel functions with negative integer index

$$J_{-n}(x) = (-1)^n J_n(x) \text{ for } n = 0, 1, 2, \cdots$$
$$Y_{-n}(x) = (-1)^n Y_n(x) \text{ for } n = 0, 1, 2, \cdots.$$

Bessel function of $n = 1/4$

$$J_{1/4}(x) = \frac{(x/2)^{1/4}}{\Gamma(5/4)} (1 - \frac{(x/2)^2 \Gamma(5/4)}{\Gamma(9/4)} + \cdots) \tag{L.126}$$

$$J_{-1/4}(x) = \frac{(x/2)^{-1/4}}{\Gamma(3/4)} (1 - \frac{(x/2)^2 \Gamma(3/4)}{\Gamma(7/4)} + \cdots) \tag{L.127}$$

Asymptotic expansions:

$$J_n(x) \underset{x\to\infty}{\simeq} \sqrt{\frac{2}{\pi}\frac{1}{x}}\cos\left(x - n\frac{\pi}{2} - \frac{\pi}{4}\right) \tag{L.128}$$

$$Y_n(x) \underset{x\to\infty}{\simeq} \sqrt{\frac{2}{\pi}\frac{1}{x}}\sin\left(x - n\frac{\pi}{2} - \frac{\pi}{4}\right) \tag{L.129}$$

Integral expressions for $\mu + \nu + 1 > \lambda > 0$

$$\int_0^\infty \frac{1}{r^\lambda} J_\mu(kr) J_\nu(kr) dr = \frac{k^{\lambda-1}\Gamma(\lambda)\Gamma\left(\frac{\mu+\nu-\lambda+1}{2}\right)}{2^\lambda \Gamma\left(\frac{\mu-\nu+\lambda+1}{2}\right)\Gamma\left(\frac{\mu+\nu+\lambda+1}{2}\right)\Gamma\left(\frac{\nu-\mu+\lambda+1}{2}\right)}. \tag{L.130}$$

Special cases $2\mu + 1 > \lambda > 0$

$$\int_0^\infty \frac{1}{r^\lambda}\left[J_\mu(kr)\right]^2 dr = \frac{k^{\lambda-1}\Gamma(\lambda)\Gamma\left(\frac{2\mu-\lambda+1}{2}\right)}{2^\lambda\left[\Gamma\left(\frac{\lambda+1}{2}\right)\right]^2\Gamma\left(\frac{2\mu+\lambda+1}{2}\right)}. \tag{L.131}$$

L.12.3 Jacobi-Anger expansion and related expressions

The Jacobi-Anger expansions are is given by

$$e^{iz\cos\theta} = \sum_{n=-\infty}^{n=\infty} i^n J_n(z) e^{in\theta} \tag{L.132}$$

$$e^{iz\sin\theta} = \sum_{n=-\infty}^{n=\infty} i^n J_n(z) e^{in(\theta-\pi/2)} = \sum_{n=-\infty}^{n=\infty} J_n(z) e^{in\theta}, \tag{L.133}$$

where n assumes only integer values. Using $\int_{-\pi}^{\pi} e^{in\alpha} d\alpha = \delta_{n,0}$ this leads to the following integral representation of the Bessel function

$$\int_{-\pi}^{\pi} e^{iz\sin\theta} e^{-im\theta} d\theta = \sum_{n=-\infty}^{n=\infty} J_n(z)\int_{-\pi}^{\pi} e^{in\theta} e^{-im\theta} d\theta = J_n(z), \tag{L.134}$$

in particular

$$J_0(z) = \int_{-\pi}^{\pi} e^{iz\sin\theta} d\theta = \int_{-\pi}^{\pi} e^{iz\cos\theta'} d\theta'. \tag{L.135}$$

This relation can be rewritten in several closely related forms

$$e^{iz\sin\theta} = J_0(z) + \sum_{n=1}^{n=\infty} J_n(z)[e^{in\theta} + (-1)^n e^{-in\theta}] \tag{L.136}$$

$$\cos(z\sin\theta) = \Re(e^{iz\sin\theta}) = J_0(z) + 2\sum_{n=2,4,\cdots}^{n=\infty} J_n(z)\cos(n\theta) \tag{L.137}$$

$$\sin(z\sin\theta) = \Im(e^{iz\sin\theta}) = 2\sum_{n=1,3,\cdots}^{n=\infty} J_n(z)\sin(n\theta). \tag{L.138}$$

L.13 The Wronskian and Wronskian Theorem

Let us consider a second-order differential equation of the following general form

$$\chi'' + F(r)\chi = 0 \tag{L.139}$$

and look for some general properties of this eigenvalue equation. The only restrictions will be that $F(r)$ is bounded from below and continuous over the entire interval $(-\infty, +\infty)$. To compare full solutions of Eq. (L.139) with approximate solutions the analysis of their *Wronskian* is an important tool. The Wronskian of two functions $\chi_1(r)$ and $\chi_2(r)$ is defined as

$$W(\chi_1, \chi_2) \equiv \chi_1\chi_2' - \chi_1'\chi_2. \tag{L.140}$$

Problem L.1. If the Wronskian of two functions $\chi_1(r)$ and $\chi_2(r)$ is vanishing at a given value of r, then the logarithmic derivative of these two functions are equal at that value of r.

Solution. The Wronskian $W(\chi_1, \chi_2)$ is vanishing at position r if $\chi_1\chi_2' - \chi_1'\chi_2 = 0$. This can be rewritten as

$$\frac{d\ln\chi_1}{dr} = \frac{\chi_1'}{\chi_1} = \frac{\chi_2'}{\chi_2} = \frac{d\ln\chi_2}{dr}.$$

Hence, the logarithmic derivatives are equal. □

Problem L.2. Show that the derivative of the Wronskian of two functions $\chi_1(r)$ and $\chi_2(r)$, which are (over an interval $a < r < b$) solutions of two differential equations $\chi_1'' + F_1(r)\chi_1 = 0$ and $\chi_2'' + F_2(r)\chi_2 = 0$, is given by

$$dW(\chi_1, \chi_2)/dr = [F_1(r) - F_2(r)]\chi_1\chi_2.$$

This is the differential form of the Wronskian theorem.

Solution. The two functions $\chi_1(r)$ and $\chi_2(r)$ are solutions (over an interval $a < r < b$) of the equations

$$\chi_1'' + F_1(r)\chi_1 = 0 \tag{L.141}$$
$$\chi_2'' + F_2(r)\chi_2 = 0, \tag{L.142}$$

Multiplying the upper equation by χ_2 and the lower one by χ_1, we obtain after subtracting the two equations

$$dW(\chi_1, \chi_2)/dr = \chi_1\chi_2'' - \chi_2\chi_1'' = [F_1(r) - F_2(r)]\chi_1\chi_2.$$

In integral form this expression is known as the *Wronskian theorem,*

$$W(\chi_1, \chi_2)|_a^b = \int_a^b [F_1(r) - F_2(r)]\chi_1(r)\chi_2(r)dr. \tag{L.143}$$

The Wronskian theorem expresses the overall variation of the Wronskian of two functions over a given interval of their joint variable. □

Problem L.3. Show that the derivative of the Wronskian of two functions $\chi_1(r)$ and $\chi_2(r)$, which are (over an interval $a < r < b$) solutions of two differential equations $\chi_1'' + F_1(r)\chi_1 + f_1(r) = 0$ and $\chi_2'' + F_2(r)\chi_2 + f_2(r) = 0$, is given by

$$dW(\chi_1, \chi_2)/dr = [F_1(r) - F_2(r)]\chi_1\chi_2 + f_1(r)\chi_2 - f_2(r)\chi_1.$$

Solution. The two functions $\chi_1(r)$ and $\chi_2(r)$ are solutions (over an interval $a < r < b$) of the equations

$$\chi_1'' + F_1(r)\chi_1 + f_1(r) = 0 \tag{L.144}$$
$$\chi_2'' + F_2(r)\chi_2 + f_2(r) = 0, \tag{L.145}$$

Multiplying the upper equation by χ_2 and the lower one by χ_1, we obtain after subtracting the two equations

$$dW(\chi_1, \chi_2)/dr = [F_1(r) - F_2(r)]\chi_1\chi_2 + f_1(r)\chi_2 - f_2(r)\chi_1.$$

In integral form this expression becomes

$$W(\chi_1,\chi_2)|_a^b = \int_a^b [F_1(r)-F_2(r)]\chi_1\chi_2 dr + \int_a^b [f_1(r)\chi_2 - f_2(r)]\chi_1 dr. \tag{L.146}$$

The Wronskian theorem expresses the overall variation of the Wronskian of two functions over a given interval of their joint variable. □

Example. Consider the 1D-Schrödinger equation (in Rydberg a.u.)

$$\tilde{\chi}'' + [U(\rho)+\varepsilon]\,\tilde{\chi}_l = 0. \tag{L.147}$$

For two solutions on the interval $a < r < b$, denoted by $\chi_1(r,\varepsilon_1)$ and $\chi_2(r,\varepsilon_2)$ and corresponding to energies ε_1 and ε_2, the Wronskian Theorem takes the form

$$W(\chi_1,\chi_2)|_a^b = (\varepsilon_1-\varepsilon_2)\int_a^b \chi_1(r)\chi_2(r)dr. \tag{L.148}$$

Likewise, for two solutions, $\chi_1(r)$ and $\chi_2(r)$, of the same energy but for different potentials, $U_1(r)$ and $U_2(r)$, respectively, the Wronskian Theorem takes the form

$$W(\chi_1,\chi_2)|_a^b = \int_a^b [U_2(r)-U_1(r)]\,\chi_1(r)\chi_2(r)dr. \tag{L.149}$$

L.14 Total differentials and partial derivatives

In this section we consider the function u, which is function of the variables x, y and z,

$$u = u(x,y,z) \tag{L.150}$$

in the presence of a single constraint

$$g(x,y,z) = 0. \tag{L.151}$$

In thermodynamics the constraint is given by the equation of state of the system under consideration. In view of the constraint we can express x, y, z and u in terms of (y,z), (y,z) or (x,y), respectively,

$$z = z(x,y) \quad \text{or} \quad y = y(x,z) \quad \text{or} \quad x = x(y,z) \tag{L.152}$$

$$u = u(x,y) \quad \text{or} \quad u = u(x,z) \quad \text{or} \quad u = x(y,z). \tag{L.153}$$

L.14.1 Total differential

The *total differential* (also called *exact differential*),

$$dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x dy \tag{L.154}$$

can always be written in the form

$$dz = A(x,y)dx + B(x,y)dy, \tag{L.155}$$

where

$$A(x,y) \equiv \left(\frac{\partial z}{\partial x}\right)_y \quad \text{and} \quad B(x,y) \equiv \left(\frac{\partial z}{\partial y}\right)_x, \tag{L.156}$$

which implies

$$\left(\frac{\partial A}{\partial y}\right)_x = \left(\frac{\partial B}{\partial x}\right)_y = \frac{\partial^2 z}{\partial x \partial y}. \tag{L.157}$$

Properties:

Path invariance: The integral $\int A(x,y)dx + \int B(x,y)dy$ is independent of the integration path in the x, y plane but only depends on the value of the function z at the beginning and the end point,

$$\int_{x_1,y_1}^{x_2,y_2} A(x,y)dx + \int_{x_1,y_1}^{x_2,y_2} B(x,y)dy = \int_{z_1}^{z_2} dz = z_2 - z_1, \tag{L.158}$$

where $z_1 = z(x_1, y_1)$ and $z_2 = z(x_2, y_2)$.

Minus 1 rule: The variables x, y, z satisfy the relation (see Problem L.4)

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1. \tag{L.159}$$

Problem L.4. Derive the minus 1 rule.

Solution. This rule follows by substituting the total differential

$$dy = \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\frac{\partial y}{\partial z}\right)_x dz \tag{L.160}$$

into Eq. (L.154)

$$dz = \left(\frac{\partial z}{\partial x}\right)_y dx + \left(\frac{\partial z}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_z dx + \left(\frac{\partial z}{\partial y}\right)_x \left(\frac{\partial y}{\partial z}\right)_x dz. \tag{L.161}$$

Since

$$\left(\frac{\partial z}{\partial y}\right)_x \left(\frac{\partial y}{\partial z}\right)_x = 1 \tag{L.162}$$

the differential is only satisfied if

$$\left(\frac{\partial z}{\partial x}\right)_y + \left(\frac{\partial z}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_z = 0, \tag{L.163}$$

which can be rewritten in the form of the *minus 1 rule.* □

Inversely, the expression (L.155) is in general *not* a total differential. Interpreting $A(x,y)$ and $B(x,y)$ as partial derivatives,

$$A(x,y) \equiv \left(\frac{\partial z_A}{\partial x}\right)_y \quad \text{and} \quad B(x,y) \equiv \left(\frac{\partial z_B}{\partial y}\right)_x, \tag{L.164}$$

we find that in general

$$\left(\frac{\partial A}{\partial y}\right)_x \neq \left(\frac{\partial B}{\partial x}\right)_y. \tag{L.165}$$

For this case we use the notation:

$$đz = A(x,y)dx + B(x,y)dy. \tag{L.166}$$

Only if the condition (L.157) holds the functions $z_A(x,y)$ and $z_B(x,y)$ are equal (up to a constant) and Eq. (L.155) becomes a total differential.

M

Square matrices

M.1 Nomenclature and notation

In this section we summarize the nomenclature and properties of *complex* square matrices of *dimension (order) n*. An arbitrary $n \times n$ matrix A and its *inverse* are written as

$$A \equiv (a^i_j) \text{ and } A^{-1} \equiv (a^i_j)^{-1} \equiv (\bar{a}^i_j), \tag{M.1}$$

where a^i_j and $\bar{a}^i_j$ represent the elements of these matrices with the convention that the *upper* or *contravariant* index is the *row* index and the *lower* or *covariant* index is the *column* index. We define the index order, *upper precedes lower.* Conforming to the standard matrix notation, the *first* index is the *row* index of the matrix (a^i_j) and the *second* index the *column* index. The *transposed* matrix

$$A^T \equiv \tilde{A} \equiv (\tilde{a}^i_j) = (a^j_i) \tag{M.2}$$

is obtained by exchanging rows and columns, the *complex conjugate*

$$A^* \equiv (a^i_j)^* = (a^{i*}_j) \tag{M.3}$$

by taking the complex conjugate of all elements, and the *hermitian conjugate* or *conjugate transpose*

$$A^\dagger \equiv (a^i_j)^\dagger = (a^{j*}_i) = \tilde{A}^* \tag{M.4}$$

is obtained by complex conjugation of the transposed matrix. The trace of the matrix A is defined as

$$\mathrm{tr}A = \sum_{i=1}^{n} a^i_i. \tag{M.5}$$

The determinant of the matrix A is defined by the *Leibniz expansion*

$$|A| \equiv \det A = \det(a^i_j) = \sum_P (-1)^p a^{i_1}_1 \cdots a^{i_n}_n = \varepsilon_{i_1,\cdots i_n} a^{i_1}_1 \cdots a^{i_n}_n, \tag{M.6}$$

where the summation runs over all permutations P of the indices $(i_1, \cdots i_n) = (1, \cdots n)$ and $p \in \{\text{even}, \text{odd}\}$ is the order of the permutation; $\varepsilon_{i_1,\cdots i_n}$ is a generalized Levi-Civita symbol.

With each element a^i_j of the matrix we can associate a *cofactor*, cofactor(a^i_j), which is defined as the *minor* (sub-determinant) $\det A^i_j$ corresponding to the sub-matrix A^i_j of that element, including its sign $(-1)^{i+j}$. The determinant $\det A$ can be expanded in terms of its minors $\det A^i_j$, *c.q.* cofactors, with respect to row i,

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a^i_j \det A = \sum_{i=1}^{n} a^i_j \mathrm{cofactor}(a^i_j). \tag{M.7}$$

This is called the *Laplace expansion.* Replacing the elements of the transposed matrix by their cofactors we obtain the *adjoint* matrix, in which the elements are the cofactors of the *transposed* matrix

$$\mathrm{Adj}(a_j^i) = \mathrm{cofactor}(a_i^j) = (-1)^{i+j} \det A_i^j . \tag{M.8}$$

For square matrices we define the following special cases:

$$\begin{array}{rlrcl}
\textit{complex:} & \textit{hermitian} & A = A^{\dagger} & \Leftrightarrow & a_j^i = a_i^{j*} \\
 & \textit{unitary} & A^{-1} = A^{\dagger} & \Leftrightarrow & \bar{a}_j^i = a_i^{j*} \\
 & \textit{real} & A = A^{*} & \Leftrightarrow & a_j^i = a_j^{i*} \\
\textit{real:} & \textit{symmetrical} & A = \tilde{A} & \Leftrightarrow & a_j^i = a_i^j \\
 & \textit{orthogonal} & A^{-1} = \tilde{A} & \Leftrightarrow & \bar{a}_j^i = a_i^j
\end{array}
\begin{array}{r}
(\mathrm{M.9}) \\ (\mathrm{M.10}) \\ (\mathrm{M.11}) \\ (\mathrm{M.12}) \\ (\mathrm{M.13})
\end{array}$$

M.2 Properties

If the inverse exists, the matrix A satisfies the following properties:

- The determinant of an $n \times n$ matrix (a_i^j) is invariant under addition to a given column (or row) a scalar multiple of *another* column (or rows)

  $$\begin{vmatrix} a_1^1 \cdots a_1^n \\ \vdots \quad\;\; \vdots \\ a_n^1 \cdots a_n^n \end{vmatrix} = \begin{vmatrix} a_1^1 + k a_1^j \cdots a_1^n \\ \vdots \qquad\qquad \vdots \\ a_n^1 + k a_1^j \cdots a_n^n \end{vmatrix} = \begin{vmatrix} a_1^1 & \cdots & a_1^n \\ \vdots & & \vdots \\ a_n^1 + k a_i^1 & \cdots & a_n^n + k a_i^n \end{vmatrix} . \tag{M.14}$$

- The determinant of a *triangular* $n \times n$ matrix (a_i^j), *i.e.* $a_i^j = 0$ for $i > j$ (or $j > i$), is equal to the product of the diagonal elements,

  $$\begin{vmatrix} a_1^1 & \cdots & a_1^n \\ 0 & \ddots & \vdots \\ 0 & 0 & a_n^n \end{vmatrix} = \prod_{i=1}^{n} a_i^i . \tag{M.15}$$

  This follows from the Leibniz expansion because the only nonzero product of n matrix elements is the diagonal one.

- The determinant of the blockdiagonal $(n+m) \times (n+m)$ matrix of the $n \times n$ matrix A and the $m \times m$ matrix B is equal to the product of the determinants of the matrices A and B,

  $$\begin{vmatrix} A & C \\ 0 & B \end{vmatrix} = \begin{vmatrix} a_1^1 \cdots a_1^n & c_1^1 \cdots c_1^m \\ \vdots \qquad \vdots & \vdots \qquad \vdots \\ a_n^1 \cdots a_n^n & c_n^1 \cdots c_m^m \\ & b_1^1 \cdots b_1^m \\ 0 & \vdots \qquad \vdots \\ & b_m^1 \cdots b_m^m \end{vmatrix} = \det A \det B . \tag{M.16}$$

  This follows from the Leibniz expansion because the only nonzero product of $n+m$ matrix elements are the ones containing only elements of the submatrices A and B. Note that the r.h.s. is independent of C because the elements of the submatrix C appear in the determinant only in combination with elements of the zero block.

- The product rule for the determinants of two $n \times n$ matrices, A and B,

$$\det AB = \det A \,\det B. \tag{M.17}$$

- The determinant of A is non-zero

$$\det A = \det(a^i_j) = \Delta \neq 0. \tag{M.18}$$

Proof: Because the inverse exists, $(\bar{a}^i_j) = (a^i_j)^{-1}$, we have $(a^i_j)(\bar{a}^i_j) = \mathbb{1}$. Thus we obtain with the aid of the product rule $1 = \det[(a^i_j)(\bar{a}^i_j)] = \Delta \det(\bar{a}^i_j)$. This implies $\Delta \neq 0$.

- The determinant of A^{-1} equals the inverse of the determinant of A,

$$\det A^{-1} = \det(\bar{a}^i_j) = \Delta^{-1}. \tag{M.19}$$

Proof: Because the inverse exists we have (with the aid of the product rule) $1 = \det[(a^i_j)(\bar{a}^i_j)] = \Delta \det(\bar{a}^i_j)$ with $\Delta \neq 0$. This can be rewritten in the form (M.19).

- The determinant of A^* equals the complex conjugate of $\det A$,

$$\det A^* = \det(a^{i*}_j) = \det(a^i_j)^* = \Delta^*. \tag{M.20}$$

Proof: $\det(a^{i*}_j) = \varepsilon_{i_1,\cdots i_n} a^{i_1*}_1 \cdots a^{i_n*}_n = (\varepsilon_{i_1,\cdots i_n} a^{i_1}_1 \cdots a^{i_n}_n)^* = [\det(a^i_j)]^*$.

- The determinant of A is invariant under transposition

$$\det \tilde{A} = \det A. \tag{M.21}$$

Proof: $\det(a^i_j) = \varepsilon_{i_1,\cdots i_n} a^{i_1*}_1 \cdots a^{i_n*}_n = \varepsilon_{i_1,\cdots i_n} a^1_{i_1} \cdots a^n_{i_n} = \det(a^j_i)$.

- *Kronecker property:*

$$a^i_k \bar{a}^k_j = \delta^i_j \text{ and } \bar{a}^i_{i'} a^{i'}_j = \delta^i_j \tag{M.22}$$

Proof:[1] $\mathbb{1} = (a^i_j)(\bar{a}^i_j) = (a^i_k \bar{a}^k_j) \Rightarrow (a^i_k \bar{a}^k_j) = \delta^i_j$ and $\mathbb{1} = (\bar{a}^i_j)(a^i_j) = (\bar{a}^i_k a^k_j) \Rightarrow \bar{a}^i_k a^k_j = \delta^i_j$.

- The elements of the inverse matrix are given by

$$\bar{a}^i_j = \Delta^{-1}\mathrm{Adj}(a^i_j), \tag{M.23}$$

where $\mathrm{Adj}(a^i_j)$ is the *adjoint* matrix of (a^i_j), *i.e.* $\mathrm{Adj}(a^i_j) = \mathrm{cofactor}(a^j_i) = (-1)^{i+j} \det(A^j_i)$.
Proof: Let $B = (\alpha^i_j) = \mathrm{Adj}(a^i_j)$ be the *adjoint* matrix of (a^i_j). Then, the product of the matrices A and B is given by

$$AB = (a^i_j)(\alpha^i_j) = (a^i_k \alpha^k_j) = \sum_{i=1}^{n} a^i_k \mathrm{cofactor}(a^j_k) = (\delta^i_j) \det A = \mathbb{1}\,\Delta. \tag{M.24}$$

To arrive at the result we used Eq. (M.7) and further $\sum_{i=1}^n a^i_k \mathrm{cofactor}(a^j_k) = 0$ for $i \neq j$.[2] Rewriting Eq. (M.24) we obtain at $A^{-1} = B/\Delta$, which had to be proven.

---

[1] Note that $a^i_k b^k_j \neq a^k_i b^j_k$. Therefore we need a convention. In matrix multiplication we use the *column on row* summation convention $(c^i_j) = (a^i_j)(b^i_j) = (a^i_k b^k_j)$. This means that we sum over the column index of the left matrix and the row index of the right matrix.

[2] Note that this expression corresponds to a determinant with two identical columns.

- The trace of the commutator of two $(n \times n)$ matrices, A and B, is zero,
$$\operatorname{tr}[A,B] = 0 \tag{M.25}$$
Proof: The trace $\operatorname{tr} AB$ of the product of the matrices A and B is given by
$$\operatorname{tr} AB = \operatorname{tr}(a^i_k b^k_j) = \sum_{i,k=1}^{n} a^i_k b^k_i = \sum_{i,k=1}^{n} b^k_i a^i_k = \operatorname{tr}(b^k_i a^i_l) = \operatorname{tr} BA.$$
Since $[A,B] = AB - BA$ and $\operatorname{tr}(A+B) = \operatorname{tr}A + \operatorname{tr}B$ this completes the proof.
- The trace of the product of three $(n \times n)$ matrices A, B and C is invariant under cyclic permutation of these matrices,
$$\operatorname{tr} ABC = \operatorname{tr} CAB = \operatorname{tr} BCA \tag{M.26}$$
Proof: This follows by direct application of property M.25.

M.2.1 Unitary matrices

Let $U = (u_i^{i'})$ be a unitary $n \times n$ matrix, $U^{-1} = U^\dagger \equiv \tilde{U}^* \Leftrightarrow \bar{u}^i_{i'} = u_i^{i'*}$. The unitarity implies the following properties:

- *Kronecker property*: The Kronecker property takes the following form
$$u_i^{i'*} u^i_{j'} = \delta^{i'}_{j'} \text{ and } u^i_{i'} u_j^{i'*} = \delta^i_j. \tag{M.27}$$
Proof: Substituting the unitarity condition $\bar{u}^i_{i'} = u_i^{i'*}$ in Eq. (M.22) yields the desired result.
- *Determinant* - The determinant of a unitary $n \times n$ matrix U equals a complex number of unit norm,
$$\Delta \equiv \det U = e^{i\varphi}. \tag{M.28}$$
where $\varphi \in \mathbb{R}$. For the special case $\varphi = 0$ the matrix U is called the *special unitary matrix.*
Proof: Using the properties M.20, M.30 and M.19 we find $\Delta^* = \det U^* = \det U^\dagger = \det U^{-1} = 1/\Delta \to |\Delta|^2 = 1$, which implies (M.28).
- *Matrix inversion rule*: The elements of the inverse matrix are given by
$$u^{i*}_{i'} = e^{-i\varphi} \text{cofactor}(u_i^{i'}), \tag{M.29}$$
Proof: This follows from Eq. (M.23) by substitution of $\Delta = e^{i\varphi}$.
- *Determinant of hermitian conjugate*: The determinant of $U^\dagger$ follows with Eq. (M.21),
$$\det U^\dagger = \det \tilde{U}^* = \det U^* = e^{-i\varphi}. \tag{M.30}$$
Proof: This follows from Eq. (M.21) and substitution of $\Delta = e^{i\varphi}$.
- *Invariance of determinant under unitary transformation:* The determinant of an arbitrary $n \times n$ matrix A is invariant under unitary transformation
$$\det UAU^\dagger = \det A. \tag{M.31}$$
Proof: Using the properties M.17 and M.28 we find det $UAU^\dagger) = \det U \det A \det U^\dagger = e^{i\varphi} \det A \, e^{-i\varphi} = \det A$.
- *Invariance of trace under unitary transformation:* The trace of an arbitrary $n \times n$ matrix A is invariant under unitary transformation
$$\operatorname{tr} UAU^\dagger = \operatorname{tr}A. \tag{M.32}$$
Proof: Using the properties M.26 and M.10 we find $\operatorname{tr} UAU^\dagger = \operatorname{tr} U^\dagger UA = \operatorname{tr} U^{-1}UA = \operatorname{tr}A$.

N

Vector relations

N.1 Inner and outer products

$$(\mathbf{u},\mathbf{v},\mathbf{w}) = \mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = \mathbf{v}\cdot(\mathbf{w}\times\mathbf{u}) = \mathbf{w}\cdot(\mathbf{u}\times\mathbf{v}) \tag{N.1}$$

$$\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) = (\mathbf{u}\cdot\mathbf{w})\mathbf{v} - (\mathbf{u}\cdot\mathbf{v})\mathbf{w} \tag{N.2}$$

$$\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) = -\,(\mathbf{v}\times\mathbf{w})\times\mathbf{u} \tag{N.3}$$

$$(\mathbf{u}\times\mathbf{v})\cdot(\mathbf{w}\times\mathbf{z}) = (\mathbf{u}\cdot\mathbf{w})(\mathbf{v}\cdot\mathbf{z}) - (\mathbf{u}\cdot\mathbf{z})(\mathbf{v}\cdot\mathbf{w}) \tag{N.4}$$

N.2 Gradient, divergence and curl

N.2.1 Helmholtz theorem

Any vector $\mathbf{A}$ can be separated in an *irrotational* part $\mathbf{A}_{\parallel}$ and a *rotational (solenoidal)* part $\mathbf{A}_{\perp}$, defined by

$$\mathbf{A} = \mathbf{A}_{\perp} + \mathbf{A}_{\parallel} \text{ with } \boldsymbol{\nabla}\cdot\mathbf{A}_{\perp} = 0 \text{ and } \boldsymbol{\nabla}\times\mathbf{A}_{\parallel} = 0. \tag{N.5}$$

N.2.2 Vector identities with a single vector differential operator

$$\boldsymbol{\nabla}\cdot(\mathbf{A}\phi) = (\mathbf{A}\cdot\boldsymbol{\nabla}\phi) + \phi(\boldsymbol{\nabla}\cdot\mathbf{A}) \tag{N.6}$$

$$\boldsymbol{\nabla}\times(\mathbf{A}\phi) = -(\mathbf{A}\times\boldsymbol{\nabla}\phi) + \phi(\boldsymbol{\nabla}\times\mathbf{A}) \tag{N.7}$$

$$\boldsymbol{\nabla}\cdot(\mathbf{A}\times\mathbf{B}) = \mathbf{B}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) + \mathbf{A}\cdot(\boldsymbol{\nabla}\times\mathbf{B}) \tag{N.8}$$

$$\boldsymbol{\nabla}\times(\mathbf{A}\times\mathbf{B}) = (\mathbf{B}\cdot\boldsymbol{\nabla})\mathbf{A} - \mathbf{B}(\boldsymbol{\nabla}\cdot\mathbf{A}) - (\mathbf{A}\cdot\boldsymbol{\nabla})\mathbf{B} + \mathbf{A}(\boldsymbol{\nabla}\cdot\mathbf{B}) \tag{N.9}$$

$$\boldsymbol{\nabla}(\mathbf{A}\cdot\mathbf{B}) = (\mathbf{B}\cdot\boldsymbol{\nabla})\mathbf{A} + (\mathbf{A}\cdot\boldsymbol{\nabla})\mathbf{B} + \mathbf{B}\times(\boldsymbol{\nabla}\times\mathbf{A}) + \mathbf{A}\times(\boldsymbol{\nabla}\times\mathbf{B}) \tag{N.10}$$

Examples:

$$\boldsymbol{\nabla}\cdot\mathbf{r} = 3 \Leftrightarrow \partial_i r_i = 3 \tag{N.11}$$

$$\boldsymbol{\nabla}\times\mathbf{r} = \mathbf{0} \Leftrightarrow \varepsilon_{ijk}\partial_j r_k = \mathbf{0} \tag{N.12}$$

$$(\mathbf{A}\cdot\boldsymbol{\nabla})\mathbf{r} = \mathbf{A} \Leftrightarrow A_i\nabla_i r_k = A_k \tag{N.13}$$

$$\boldsymbol{\nabla}\cdot\dot{\mathbf{r}} = \boldsymbol{\nabla}\times\dot{\mathbf{r}} = (\mathbf{A}\cdot\boldsymbol{\nabla})\dot{\mathbf{r}} = \mathbf{0} \tag{N.14}$$

Combining Eqs. (N.10), (N.12) and (N.13) we find

$$\boldsymbol{\nabla}(\mathbf{r}\cdot\mathbf{A}) = \mathbf{A} + (\mathbf{r}\cdot\boldsymbol{\nabla})\mathbf{A} + \mathbf{r}\times(\boldsymbol{\nabla}\times\mathbf{A}). \tag{N.15}$$

Likewise we find by combining Eqs. (N.10) with Eq. (N.9)

$$\boldsymbol{\nabla}(\dot{\mathbf{r}}\cdot\mathbf{A}) = (\dot{\mathbf{r}}\cdot\boldsymbol{\nabla})\mathbf{A} + \dot{\mathbf{r}}\times(\boldsymbol{\nabla}\times\mathbf{A}) \tag{N.16}$$

$$\frac{d}{dt}\mathbf{A} = \frac{\partial}{\partial t}\mathbf{A} + (\dot{\mathbf{r}}\cdot\nabla)\mathbf{A}. \tag{N.17}$$

Special case:

$$\boldsymbol{\nabla}r^n = nr^{n-1}\left(\mathbf{r}/r\right) = nr^{n-1}\hat{\mathbf{r}}. \tag{N.18}$$

For $n \geq 0$ this expression is valid everywhere, including the origin. For $n < 0$ the expression is irregular in the origin. In particular we have

$$\boldsymbol{\nabla}\frac{1}{r} = -\frac{\mathbf{r}}{r^3}. \tag{N.19}$$

Since

$$\boldsymbol{\nabla}|\mathbf{r}-\mathbf{r}'| = \frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|} = -\boldsymbol{\nabla}'|\mathbf{r}-\mathbf{r}'| \tag{N.20}$$

(as is easily verified in cartesian coordinates) we can generalize Eq. (N.18) to

$$\boldsymbol{\nabla}|\mathbf{r}-\mathbf{r}'|^n = n|\mathbf{r}-\mathbf{r}'|^{n-1}\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|} = -\boldsymbol{\nabla}'|\mathbf{r}-\mathbf{r}'|^n. \tag{N.21}$$

N.2.3 Expressions with second derivatives

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\phi) = 0 \tag{N.22}$$

$$\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) = 0 \tag{N.23}$$

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{A}) = \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) - \boldsymbol{\nabla}^2\mathbf{A} \tag{N.24}$$

Expressions for the laplacian:

$$\Delta r^n = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}r^n = nr^{n-2}\left(\boldsymbol{\nabla}\cdot\mathbf{r}\right) + n\mathbf{r}\cdot\boldsymbol{\nabla}r^{n-2} = 3nr^{n-2} + n(n-2)r^{n-2}. \tag{N.25}$$

We distinguish three cases:

$$\Delta r^n = n(n+1)r^{n-2} \ \text{ is regular everywhere } \begin{cases}\text{for } n \geq 0 \ \textit{also} \text{ at the origin}\\ \text{for } n < 0 \ \textit{outside} \text{ at the origin}\end{cases} \tag{N.26}$$

Note that for $n = -1$ this expression yields $\boldsymbol{\nabla}^2(1/r) = 0$ everywhere outside the origin. The case $n = -1$ is special at the origin because the laplacian produces a flux out of any sphere centered around the origin. With the Gauss divergence theorem we calculate for this flux

$$-\int \boldsymbol{\nabla}\left(1/r^3\right)\mathbf{r}\cdot d\mathbf{r} = -\oint\left(1/r^3\right)\mathbf{r}\cdot\hat{\mathbf{r}}dS = -4\pi. \tag{N.27}$$

Thus, Δr^{-1} can be regarded as a distribution and we write

$$\boldsymbol{\nabla}^2(1/r) = -4\pi\delta(\mathbf{r}). \tag{N.28}$$

Bibliography

[1] Milton Abramowitz and Irene Stegun. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables* (Dover Publications, New York, 1965).

[2] Christine A. Aidala, Steven D. Bass, Delia Hasch, and Gerhard K. Mallot. *The spin structure of the nucleon. Rev. Mod. Phys.*, 85:655–691, 2013.

[3] T. Andersen, H. K. Haugen, and H. Hotop. *Binding Energies in Atomic Negative Ions: III. Journal of Physical and Chemical Reference Data*, 28(6):1511–1533, 1999.

[4] E. Arimondo, M. Inguscio, and P. Violino. *Experimental determinations of the hyperfine structure in the alkali atoms. Rev. Mod. Phys.*, 49:31–75, 1977.

[5] Lloyd Armstrong. *Theory of the Hyperfine Structure of Free Atoms* (Wiley-Interscience, New York, 1971).

[6] E. Back and S. Goudsmit. *Kernmoment und Zeemaneffekt von Wismut. Zeitschrift für Physik*, 47(3-4):174–183, 1928.

[7] L. C. Balling and F. M. Pipkin. *Gyromagnetic Ratios of Hydrogen, Tritium, Free Electrons, and Rb85. Phys. Rev.*, 139:A19–A26, 1965.

[8] A. Banerjee, D. Das, and V. Natarajan. *Absolute frequency measurements of the D 1 lines in 39 K, 85 Rb, and 87 Rb with 0.1 ppb uncertainty. EPL (Europhysics Letters)*, 65(2):172, 2004.

[9] D.R. Bates and Agnette Damgaard. *The Calculation of the Absolute Strengths of Spectral Lines. Phil. Trans. R. Soc. Lond. A*, 242:101–122, 1949.

[10] Gordon Baym. *Lectures on Quantum Mechanics* (Benjamin/Cummings Publishing Corporation, New York, 1969).

[11] A. Beckmann, K.D. Böklen, and D. Elke. *Precision measurements of the nuclear magnetic dipole moments of 6Li,7Li,23Na,39K and 41K. Zeitschrift für Physik*, 270(3):173–186, 1974.

[12] J. Benhelm, G. Kirchmair, U. Rapol, T. Körber, C. F. Roos, and R. Blatt. *Measurement of the hyperfine structure of the $S_{1/2} - D_{5/2}$ transition in* $^{43}\mathrm{Ca}^{+}$. *Phys. Rev. A*, 75:032506, 2007.

[13] H. A. Bethe. *The Electromagnetic Shift of Energy Levels. Phys. Rev.*, 72:339–341, 1947.

[14] H.A. Bethe and E.E. Salpeter. *Quantum Mechanics of One and Two-Electron Atoms* (Plenum Publishing Corporation, New York, 1977).

[15] S. Bize, Y. Sortais, M. S. Santos, C. Mandache, A. Clairon, and C. Salomon. *High-accuracy measurement of the 87 Rb ground-state hyperfine splitting in an atomic fountain. EPL (Europhysics Letters)*, 45(5):558, 1999.

[16] F. Bloch. *Nuclear Induction. Phys. Rev.*, 70:460–474, 1946.

[17] Niels Bohr. *On the constitution of atoms and molecules, Part I. Philosophical Magazine Series 6*, 26(151):1–25, 1913.

[18] Niels Bohr. *On the constitution of atoms and molecules, Part II: Systems containing only a single nucleus. Philosophical Magazine Series 6*, 26(153):476–502, 1913.

[19] Niels Bohr. *On the constitution of atoms and molecules, Part III: Systems containing several nuclei. Philosophical Magazine Series 6*, 26(155):857–875, 1913.

[20] Niels Bohr. *Der Bau der Atome und die physikalischen und chemischen Eigenschaften der Elemente. Zeitschrift für Physik*, 9(1):1–67, 1922.

[21] G. Breit. *Does the Electron Have an Intrinsic Magnetic Moment? Phys. Rev.*, 72:984–984, 1947.

[22] G. Breit and I. I. Rabi. *Measurement of Nuclear Spin. Phys. Rev.*, 38:2082–2083, 1931.

[23] Roger C. Brown, Saijun Wu, J. V. Porto, Craig J. Sansonetti, C. E. Simien, Samuel M. Brewer, Joseph N. Tan, and J. D. Gillaspy. *Quantum interference and light polarization effects in unresolvable atomic lines: Application to a precise measurement of the* $^{6,7}Li\ D_2$ *lines. Phys. Rev. A*, 87:032504, 2013.

[24] B. A. Bushaw, W. Nörtershäuser, G. W. F. Drake, and H.-J. Kluge. *Ionization energy of* ^{6,7}Li *determined by triple-resonance laser spectroscopy. Phys. Rev. A*, 75:052503, 2007.

[25] S. Chandrasekhar. *Some Remarks on the Negative Hydrogen Ion and its Absorption Coefficient. Astrophysical Journal*, 100:176, 1944.

[26] C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum Mechanics* (John Wiley & Sons, New York, 1977).

[27] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg. *Introduction to Quantum Electrodynamics* (John Wiley & Sons, New York, 1989).

[28] E.U. Condon and G.H. Shortley. *The Theory of Atomic Spectra* (Cambridge University Press, Cambridge, 1951).

[29] E.A. Davis and I.J. Falconer. *J.J. Thomson and the Discovery of the Electron* (Taylor & Francis, London, 1907).

[30] J. de Boer. *Construction operator formalism for many particle systems.* In *Studies in Statistical Mechanics, J. de Boer and G.E. Uhlenbeck (Eds.)*, volume III, p. 212 (North Holland, 1965).

[31] H. Dehmelt. *New continuous Stern-Gerlach effect and a hint of "the" elementary particle. Zeitschrift für Physik D Atoms, Molecules and Clusters*, 10(2):127–134, 1988.

[32] M. Diermaier, C. B. Jepsen, B. Kolbinger, C. Malbrunot, O. Massiczek, C. Sauerzopf, M. C. Simon, J. Zmeskal, and E. Widmann. *In-beam measurement of the hydrogen hyperfine splitting and prospects for antihydrogen spectroscopy. Nature Communications*, p. 15749, 2017.

[33] P. A. M. Dirac. *On the Theory of Quantum Mechanics. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 112(762):661–677, 1926.

[34] P. A. M. Dirac. *The Quantum Theory of the Emission and Absorption of Radiation. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 114(767):243–265, 1927.

[35] P. A. M. Dirac. *The Quantum Theory of the Electron. Proceedings of the Royal Society of London. Series A*, 117(778):610–624, 1928.

[36] P.A.M. Dirac. *The Priciples of Quantum Mechanics* (Oxford University Press, Oxford, 1958).

[37] Gordon W. F. Drake. *High Precision Theory of Atomic Helium. Physica Scripta*, T83:83, 1999.

[38] A.R. Edmonds. *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, 1996).

[39] B.-G. Englert. *The Semiclassical Atom* (Springer Verlag, Heidelberg, 1988).

[40] U. Fano and G. Racah. *Irreducible Tensorial Sets* (Acdemic Press, New York, 1959).

[41] E. Fermi. *Eine statistische Methode zur Bestimmung einiger Eigenschaften des Atoms und ihre Anwendung auf die Theorie des periodischen Systems der Elemente. Zeitschrift für Physik*, 48(1-2):73–79, 1928.

[42] Richard P. Feynman, Frank L. Vernon, and Robert W. Hellwarth. *Geometrical representation of the Schrödinger Equation for Solving Maser Problems. J. Appl. Phys.*, 28:49–52, 1957.

[43] V. Fock. *Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. Zeitschrift für Physik*, 61(1-2):126–148, 1930.

[44] V. Fock. *Konfigurationsraum und zweite Quantelung. Zeitschrift für Physik*, 75(9-10):622–647, 1932.

[45] Joseph Fraunhofer. *Über die Beziehungen der Eigenschaften zu den Atomgewichten der Elemente. Denkschriften der Königlichen Akademie der Wissenschaften zu München*, 5:193–226, 1814-1815.

[46] G. Gabrielse, D. Hanneke, T. Kinoshita, M. Nio, and B. Odom. *New Determination of the Fine Structure Constant from the Electron g Value and QED. Phys. Rev. Lett.*, 97:030802, 2006.

[47] T.F. Gallagher. *Rydberg Atoms.* Cambridge New Art History and Criticism (Cambridge University Press, 1994). ISBN 9780521385312.

[48] Ian R. Gatland. *Integer versus half-integer angular momentum. American Journal of Physics*, 74(3):191–192, 2006.

[49] J.A. Gaunt. *The Triplets of Helium. Phil. Trans. R. Soc. London. Series A*, 228:151–196, 1928.

[50] Walther Gerlach and Otto Stern. *Das magnetische Moment des Silberatoms. Zeitschrift für Physik*, 9(1):353–355, 1922.

[51] R. Gerritsma, G. Kirchmair, F. Zähringer, E. Solano, R. Blatt, and C.F. Roos. *Quantum simulation of the Dirac equation. Nature*, 463:68–71, 2010.

[52] H. Goldstein. *Classical Mechanics* (Addison-Wesley Publishing Company, Amsterdam, 1974).

[53] John L. Hall. *Nobel Lecture: Defining and measuring optical frequencies*. Rev. Mod. Phys.*, 78:1279–1295, 2006.

[54] D. Hanneke, S. Fogwell, and G. Gabrielse. *New Measurement of the Electron Magnetic Moment and the Fine Structure Constant. Phys. Rev. Lett.*, 100:120801, 2008.

[55] Theodor W. Hänsch. *Nobel Lecture: Passion for precision. Rev. Mod. Phys.*, 78:1297–1309, 2006.

[56] D. R. Hartree. *The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods. Mathematical Proceedings of the Cambridge Philosophical Society*, 24:89–110, 1928.

[57] D. R. Hartree. *The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion. Mathematical Proceedings of the Cambridge Philosophical Society*, 24:111–132, 1928.

[58] J. J. Hudson, D. M. Kara, I. J. Smallman, B. E. Sauer, M. R. Tarbutt, and E. A. Hinds. *Improved measurement of the shape of the electron. Nature*, 473:493–496, 2011.

[59] Friedrich Hund. *Zur Deutung verwickelter Spektren, insbesondere der Elemente Scandium bis Nickel. Zeitschrift für Physik*, 33:345–371, 1925.

[60] David Jacob, Lingxuan Shao, Vincent Corre, Tilman Zibold, Luigi De Sarlo, Emmanuel Mimoun, Jean Dalibard, and Fabrice Gerbier. *Phase diagram of spin-1 antiferromagnetic Bose-Einstein condensates. Phys. Rev. A*, 86:061601, 2012.

[61] Ulrich D. Jentschura, Svetlana Kotochigova, Eric-Olivier Le Bigot, Peter J. Mohr, and Barry N. Taylor. *Precise Calculation of Transition Frequencies of Hydrogen and Deuterium Based on a Least-Squares Analysis. Phys. Rev. Lett.*, 95:163003, 2005.

[62] P. Jordan and O. Klein. *Zum Mehrkörperproblem der Quantentheorie. Zeitschrift für Physik*, 45(11-12):751–765, 1927.

[63] P. Jordan and E. Wigner. *Über das Paulische Äquivalenzverbot. Zeitschrift für Physik*, 47(9-10):631–651, 1928.

[64] Walter Kohn. *Nobel Lecture: Electronic structure of matter¯wave functions and density functionals. Rev. Mod. Phys.*, 71:1253–1266, 1999.

[65] Tjalling Koopmans. *Uber the Zuordnung von Wellenfunctionen und Eigenwerten zu den Einzelnen Electronen eines Atomes. Physica*, 1:104–113, 1934.

[66] A. Kramida, Yu. Ralchenko, J. Reader, and NIST ASD Team. *NIST Atomic Spectra Database (ver. 5.2) [Online] http://physics.nist.gov/asd.* (National Institute of Standards and Technology, Gaithersburg, 2014).

[67] P. Kusch and H. M. Foley. *The Magnetic Moment of the Electron. Phys. Rev.*, 74:250–263, 1948.

[68] Willis E. Lamb and Robert C. Retherford. *Fine Structure of the Hydrogen Atom by a Microwave Method. Phys. Rev.*, 72:241–243, 1947.

[69] L.D. Landau and E.M. Lifshitz. *Mechanics* (Pergamon Press, Oxford, 1976).

[70] A. Landé. *Die absoluten Intervalle der optischen Dubletts und Tripletts. Zeitschrift für Physik*, 25(1):46–57, 1924.

[71] J. M. Leinaas and J. Myrheim. *On the theory of identical particles. Il Nuovo Cimento B (1971-1996)*, 37(1):1–23, 1977.

[72] Dimitry Mendeleev. *Über die Beziehungen der Eigenschaften zu den Atomgewichten der Elemente. Zeitschrift für Chemie*, 12:405–406, 1869.

[73] E. Merzbacher. *Single Valuedness of Wave Functions. American Journal of Physics*, 30(4):237–247, 1962.

[74] Eugen Merzbacher. *Quantum Mechanics* (John Wiley & Sons, Inc., New York, 1998).

[75] Albert Messiah. *Quantum Mechanics I* (North-Holland Publishing Company, Amsterdam, 1961).

[76] Albert Messiah. *Quantum Mechanics II* (North-Holland Publishing Company, Amsterdam, 1961).

[77] B. Odom, D. Hanneke, B. D'Urso, and G. Gabrielse. *New Measurement of the Electron Magnetic Moment Using a One-Electron Quantum Cyclotron. Phys. Rev. Lett.*, 97:030801, 2006.

[78] F. Paschen and E. Back. *Normale und anomale Zeemaneffekte. Annalen der Physik*, 344(15):897–932, 1912.

[79] W. Pauli. *Zur Frage der theoretischen Deutung der Satelliten einiger Spektrallinien und ihrer Beeinflussung durch magnetische Felder. Naturwissenschaften*, 12(37):741–743, 1924.

[80] W. Pauli. *Über den Einflusz der Geschwindigkeitsabhängigkeit der Elektronenmasse auf den Zeemaneffekt. Zeitschrift für Physik*, 31(1):373–385, 1925.

[81] W. Pauli. *Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren. Zeitschrift für Physik*, 31(1):765–783, 1925.

[82] W. Pauli. *Nobel Lecture: Exclusion Principle and Quantum Mechanics. Nobelprize.org. Nobel Media AB 2014. Web. 23 Feb 2017.*, 1945.

[83] Ruben Pauncz. *The Symmetric Group in Quantum Chemistry* (CRC Press, Boca Raton, Florida, 1995).

[84] I. I. Rabi, S. Millman, P. Kusch, and J. R. Zacharias. *The Molecular Beam Resonance Method for Measuring Nuclear Magnetic Moments. The Magnetic Moments of* ${}_3\mathrm{Li}^6$, ${}_3\mathrm{Li}^7$ *and* ${}_9\mathrm{F}^{19}$. *Phys. Rev.*, 55:526–535, 1939.

[85] I.I. Rabi. *Space Quantization in a Gyrating Magnetic Field. Phys. Rev.*, 51:652–654, 1937.

[86] Giulio Racah. *Theory of Complex Spectra. II. Phys. Rev.*, 62:438–462, 1942.

[87] Norman Ramsey. *Molecular beams* (Oxford University Press, Oxford, 1956).

[88] Norman F. Ramsey. *A New Molecular Beam Resonance Method. Phys. Rev.*, 76:996–996, 1949.

[89] Norman F. Ramsey. *The Atomic Hydrogen Maser. Metrologia*, 1(1):7, 1965.

[90] H. Rauch, A. Zeilinger, G. Badurek, A. Wilfing, W. Bauspiess, and U. Bonse. *Verification of coherent spinor rotation of fermions. Phys. Lett. A*, 54:425–427, 1975.

[91] Lord Rayleigh. *Some general theorems relating to vibrations. Proc. London Math. Soc.*, 4:357–68, 1873.

[92] Walter Ritz. *Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik. Journal für die Reine und Angewandte Mathematik*, 135:1–61, 1909.

[93] M.E. Rose. *Spherical Tensors in Physics. Proceedings of the Physical Society. Section A*, 67(3):239, 1954.

[94] M.E. Rose. *Elementary Theory of Angular Momentum* (John Wiley & Sons, Inc., New York, 1967).

[95] D.E. Rutherford. *Substitutional Analysis* (Dover Publications, New York, 2013).

[96] Ernest Rutherford. *LXXIX. The scattering of α and β particles by matter and the structure of the atom. Philosophical Magazine Series 6*, 21(125):669–688, 1911.

[97] J.J. Sakurai. *Lectures on Quantum Mechanics* (Adison-Wesley Publishing Company, New York, 1994).

[98] Craig J. Sansonetti, C. E. Simien, J. D. Gillaspy, Joseph N. Tan, Samuel M. Brewer, Roger C. Brown, Saijun Wu, and J. V. Porto. *Absolute Transition Frequencies and Quantum Interference in a Frequency Comb Based Measurement of the* ^{6,7}Li *D Lines. Phys. Rev. Lett.*, 107:023001, 2011.

[99] Philipp Schindler, Daniel Nigg, Thomas Monz, Julio T Barreiro, Esteban Martinez, Shannon X Wang, Stephan Quint, Matthias F Brandl, Volckmar Nebendahl, Christian F Roos, Michael Chwalla, Markus Hennrich, and Rainer Blatt. *A quantum information processor with trapped ions. New Journal of Physics*, 15(12):123012, 2013.

[100] E. Schrödinger. *Quantisierung als Eigenwertproblem. Annalen der Physik*, 385(13):437–490, 1926.

[101] E. Schrödinger. *Über die kräftefreie Bewegung in der relativistischen Quantenmechanik. Sitzungsber. Preuss. Akad. Wiss., Phys.-Math. Kl.*, 24:418–428, 1930.

[102] James Trefil Sharon McGrayne, George F. Bertsch. *Atom* (Encyclopedia Britannica, 2021).

[103] J. C. Slater. *The Theory of Complex Spectra. Phys. Rev.*, 34:1293–1322, 1929.

[104] J. C. Slater. *Note on Hartree's Method. Phys. Rev.*, 35:210–211, 1930.

[105] J. C. Slater. *A Simplification of the Hartree-Fock Method. Phys. Rev.*, 81:385–390, 1951.

[106] A. Sommerfeld. *Zur Quantentheorie der Spektrallinien. Annalen der Physik*, 356(17):1–94, 1916.

[107] Otto Stern. *Ein Weg zur experimentellen Prüfung der Richtungsquantelung im Magnetfeld. Zeitschrift für Physik*, 7(1):249–253, 1921.

[108] L. H. Thomas. *The Motion of the Spinning Electron. Nature*, 117:514–, 1926.

[109] L. H. Thomas. *The calculation of atomic fields. Mathematical Proceedings of the Cambridge Philosophical Society*, 23:542–548, 1927.

[110] L. H. Thomas. *I. The kinematics of an electron with an axis. Philosophical Magazine Series 7*, 3(13):1–22, 1927.

[111] M. Tinkham. *Group theory and quantum mechanics* (McGraw-Hill, San Francisco, 1964).

[112] Sin-Itiro Tomonaga. *The Story of Spin* (The University of Chicago Press, Chicago, 1997).

[113] G. E. Uhlenbeck and S. Goudsmit. *Ersetzung der Hypothese vom unmechanischen Zwang durch eine Forderung bezäglich des inneren Verhaltens jedes einzelnen Elektrons. Die Naturwissenschaften*, 13(47):953–954, 1925.

[114] G. E. Uhlenbeck and S. Goudsmit. *Spinning Electrons and the Structure of Spectra. Nature*, 117:264–265, 1926.

[115] Albrecht Unsöld. *Beiträge zur Quantenmechanik der Atome. Annalen der Physik*, 387(3):355–393, 1927.

[116] H. C. van de Hulst, C. A. Muller, and J. H. Oort. *The spiral structure of the outer part of the Galactic System derived from the hydrogen emission at 21 cm wavelength. Bull. Astr. Inst. Neth.*, 12:117–149. Provided by the SAO/NASA Astrophysics Data System, 1954.

[117] J. Vanier and C. Audoin. *The Quantum Physics of Atomic Frequency Standards* (Adam Hilger Publ., Bristol UK, 1989).

[118] J. von Neumann and E. Wigner. *Über das Verhalten von Eigenwerten bei adiabatischen Processen. Physikalische Zeitung*, 30:467–470, 1929.

[119] S. A. Werner, R. Colella, A. W. Overhauser, and C. F. Eagen. *Observation of the Phase Shift of a Neutron Due to Precession in a Magnetic Field. Phys. Rev. Lett.*, 35:1053–1055, 1975.

[120] E.P. Wigner. *Group Theory and its application to the quantum mechanics of atomic spectra* (Acdemic Press, New York, 1959).

[121] Frank Wilczek. *Quantum Mechanics of Fractional-Spin Particles. Phys. Rev. Lett.*, 49:957–959, 1982.

[122] Jun Ye, Peter Junger, and John L. Hall. *Hyperfine structure and absolute frequency of the* ^{87}Rb *5P3/2 state. Opt. Lett.*, 21(16):1280–1282, 1996.

[123] Pieter Zeeman. *The Effect of Magnetisation on the Nature of Light Emitted by a Substance. Nature*, 55:347, 1897.

Index